Mixed languages: a functional–communicative approach

YARON MATRAS
University of Manchester

It has been suggested that the structural composition of mixed languages and the linguistic processes through which they emerge are to some extent predictable, and that they therefore constitute a language "type" (e.g. Bakker and Mous, 1994b; Bakker and Muysken, 1995). This view is challenged here. Instead, it is argued that the compartmentalisation of structures observed in mixed languages (i.e. the fact that certain structural categories are derived from one "parent" language, others from another) is the result of the cumulative effect of different contact mechanisms. These mechanisms are defined in terms of the cognitive and communicative motivations that lead speakers to model certain functions of language on an alternative linguistic system; each mechanism will typically affect particular functional categories. Four relevant processes are identified: lexical re-orientation, selective replication, convergence, and categorial fusion. Different combinations of processes will render different outcomes, hence the diversity of mixed languages as regards their structure, function, and development.

Introduction

Recent collections by Bakker and Mous (1994a) and by Thomason (1997a) highlight growing interest in what have been termed "mixed languages" or "bilingual mixtures". Examples of mixed languages that have received considerable attention in recent literature are Michif, the Cree–French mixture spoken by the Métis, descendants of buffalo hunters in the Canadian Plains and North Dakota; Ma'a, a Bantu language in Tanzania that includes material from Cushitic languages; Copper Island Aleut, an Aleut–Russian admixture spoken by a mixed population off the coast of Kamchatka; Media Lengua, a Quechua–Spanish mixture of Ecuador; Angloromani and other so-called Para-Romani languages, which are Romani-derived vocabularies used by ethnic Gypsies within non-standard varieties of various European languages; and others.

Most languages are to some extent at least "mixed", in the sense that they have components that can be traced back to more than one source language as a result of a situation of contact in the language's earlier history. So when languages are referred to in the literature explicitly as "mixed" (as for example in Bakker and Mous, 1994b; or Bakker and Muysken, 1995), it is presumably in order to highlight that they go beyond the commonly attested patterns of mixture. Mixed languages are thus understood implicitly at least to breach conventional constraints on contact-induced language change. The point of interest for a theory of grammar is therefore in defining and explaining qualitative and quantitative differences between the effects of contact in mixed languages, as opposed to more conventional linguistic systems.

It appears however that few structural properties are shared by a majority of the languages classified as mixed in the literature. The split between lexicon and grammar, each derived from a different source language, is often highlighted as prototypical or even as intrinsically constitutive of mixed languages (see e.g. Bakker, 1994 and elsewhere), but such splits are rarely if ever consistent. The difficulty in finding clear structural criteria for mixed languages suggests that the contact phenomena that give rise to them are not uniform but varied. The question I wish to pursue here is whether there is any functional significance to the structural compartmentalisation in a given mixed system, one that can be accounted for in terms of the motivations to create or maintain such a mixture. My suggestion is that mixed languages differ from conventional cases of contact in the *density of different contact phenomena* and their cumulative effect on the overall structure of the system.

The present paper is structured as follows: first I survey current definitions of mixed languages and models explaining their emergence; next I identify four function-based mechanisms of contact. I then proceed to examine the discourse relevance of mixed utterances in non-self-contained mixed systems, I present the Functional Turnover hypothesis that ex-

Address for correspondence
Department of Linguistics, University of Manchester, Manchester M13 9PL, UK
E-mail: y.matras@man.ac.uk

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

plains selective copying of grammatical features from an ancestral language, and finally I discuss the cumulative effect of different contact mechanisms in so-called plain or self-contained mixed systems. In conclusion I argue that mixed languages are anything but a coherent or uniform language type, but that they offer opportunities to identify a complexity of processes of contact-induced change, and so can definitely help revise our overall notion of structural "borrowing".

Typologies of mixed languages

Definitions

Genetic affiliation and historical reconstruction play a crucial role in attempts to define mixed languages. Thomason and Kaufman (1988) consider a language as mixed if it does not offer opportunities for historical reconstruction of its parent language (see also Bakker and Mous, 1994b; Bakker and Muysken, 1995; Thomason, 1995, 1997b, e). Bakker (1994, 27) puts this in more positive terms, defining mixed languages as idioms that evoke identification on the part of speakers of two separate source languages. Contrasting with pidgins and creoles, mixed languages are assumed to be products of full bilingualism (cf. Thomason, 1997b, e). This is not entirely unproblematic, however, if one takes into account the role of language attrition and language loss in producing mixed systems (see e.g. Sasse, 1992, on Krekonika, a Greek–Albanian mixture; and Ma'a; Boretzky and Igla, 1994, on Para-Romani). Paradoxically, what distinguishes mixed languages from pidgins and creoles is the fact that they show continuity of significant portions of the grammar from the grammaticiser parent language (or from both parent languages), yet it is precisely the interrupted transmission or non-continuity of a significant portion of structural material that identifies mixed languages as opposed to cases of "normal" language development.

Mixed languages have been argued to be the outcome of mixed marriages giving rise to new ethno-cultural identities (Bakker, 1994, 1996, 1997; for case studies see also Hancock, 1970, 1992; Golovko, 1994; de Gruiter, 1994; Thomason, 1997c). Indeed some approaches view the emergence of a new ethnic group as the chief setting for the development of a mixed language (see especially Bakker, 1997, as well as Thomason and Kaufman, 1988, and Thomason, 1995, 1997b, e). There are, however, cases of acculturation leading to the emergence of mixed languages in situations where no population mixture is attested (cf. Muysken's (1997), account of Media Lengua). In fact, as pointed out by Smith (1995; cf. also Thomason, 1995), the cultural motivation behind the emergence of mixed varieties can just as well be retention, re-gaining, or re-definition of ethnic identity, defying pressure to change or assimilate. This is the case with a series of languages treated as mixed in the literature, such as Para-Romani (Boretzky and Igla, 1994), Shelta (also called Gammon, the secret language of Irish Travellers, cf. Grant, 1994), or Ma'a (Mous, 1994; cf. also Sasse, 1992). As regards the overall social context of emergence, an expression of *cultural defiance* might therefore seem a more accurate indicator.

Smith (1995) has suggested a distinction between "plain" mixed languages, which serve as everyday community languages, and "symbiotic" varieties, which are specialised varieties of a non-mixed language used in the same community, typically secret languages. Like pidgins, secret languages are not native languages and have a restricted functional scope. Like creoles, so-called "plain" mixed languages are full-fledged native community languages that reflect a newly emerged ethnic identity (cf. Thomason, 1997e). One must, however, take into account that the great majority of mixed languages attested so far are spoken in communities alongside at least one of their parent or source languages, and we only rarely find them in isolation. This may add to their controversial status as mixed *registers* (cf. e.g. Mous, 1994 on Ma'a) rather than full-fledged languages. As far as the "languageness" (Thomason, 1997b) is concerned, then, varieties discussed in the literature as "mixed" are obviously spread along a rather wide continuum.

Models of emergence

Explanatory models of mixed languages can be grouped along three main points of focus: (1) Language maintenance and language shift, (2) unique and pre-determined processes ("intertwining"), and (3) conventionalisation of language mixing patterns. The first approach highlights the *substitution of major components* of a given language – for example entire morphological paradigms or typological features – through material (alternative structures) from a different source language. According to Thomason and Kaufman (1988) mixed languages arise in situations where language shift is partly resisted, and so parts of the original language are retained while significant portions are replaced by the pressure-exerting language. Sasse (1992) assumes re-lexification of the target language *after* a shift has taken place, with younger speakers having only partial access to the vocabulary of their ancestral language. Such cases

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

are considered by Sasse not as re-vitalisation (as proposed for Para-Romani by Boretzky and Igla, 1994; see also Boretzky, 1998), but as processes of language renewal through which *new* languages emerge. A variation is offered by Muysken (1981, 1997) under the heading of "relexification": Media Lengua lexical entries show Spanish phonological forms, but are argued to continue to conform functionally to the original Quechua entry.

A radically different approach is taken by Bakker (1997, 192–213) with his model of *language intertwining* (see also Bakker and Muysken, 1995, and Bakker and Mous, 1994b). Bakker questions the substitution of either lexical or grammatical components of a language. Instead he attributes the emergence of mixed varieties to one single and predetermined process involving mixed populations,[1] by which the "grammar" (sometimes specified as morphosyntax, or just as bound morphemes) of one language, typically that spoken by native women, combines with the "lexicon" of another, usually a colonial language spoken by men. Deviations from the lexicon/grammar split are explained by Bakker (1997) as the result of language-specific typological constraints. Bakker insists on the equal status of the two participating source languages,[2] but stresses nevertheless that women provide the grammatical input due to their influence on children as well as their native status. Bakker and Muysken (1995, 50) even argue that the grammar of a mixed language is derived from the language which its first generation of speakers know best. Thus, even the intertwining model recognises a functional hierarchy of some sort among the contributing languages.

Overlapping with both notions – of substitution and of a merger of components – is the view that attributes mixed languages to *deliberate or conscious creations*. This is more easily arguable in cases of registers that still function as secret languages and display various manipulative constructions such as camouflaging affixes or phoneme substitution (cf. for instance Grant, 1994 on Shelta, and see Bakker, 1998), though a conscious and deliberate creation is often assumed for other mixed languages as well. For Copper Island Aleut, Golovko (1994, 117) hypothesises that language mixture began as a game among adult Aleuts learning Russian, was then used as a secret code, and later conventionalised. A similar idea is expressed by Thomason (1995, 29) with regard to the general phenomenon of abruptly emerging mixed languages, and Bakker (1997, 213) similarly admits that intertwined languages are created "more or less consciously". Mous (1994) assumes deliberate creation of Ma'a as a register of Mbugu by speakers aiming to set themselves apart from Bantu speakers.

The third type of approach to the emergence of mixed languages views them as cases of a *conventionalisation of codeswitching patterns*. This view has most recently found a supporter in Auer (1998a, b), who assumes a gradual loss of the conversational function of language alternation as a means of expressing contrast.[3] Drawing on observations of non-meaningful or unmarked language alternation – classified as "monolectal codeswitching" by Meeuwis and Blommaert (1998) – Auer defines an intermediate stage on a continuum (Language Mixing). For the final stage (Fused Lects), Auer is inspired mainly by the attestation of an emergence of mixed styles as markers of group identity (cf. de Rooij, 1996; Maschler, 1997, 1998; Oesch Serra, 1998).[4] Another supporter of the view that mixed languages represent fossilised patterns of mixing is Myers-Scotton (1992, 1998), who traces the emergence of languages like Ma'a (Cushitic "lexicon", Bantu "grammar") to a change in the social roles and status of the participating languages, resulting in a "flipping" of matrix and embedded language; this is referred to as the "Matrix Language Turnover" model. Recent work within Myers-Scotton's "Matrix Language Frame" model, notably by Bolonyai (1998) and by Jake (1998), postulates the emergence of a "composite matrix language" in more conventional situations of language contact (language acquisition and shift). Unlike mixed languages, however, the composite matrix language consists here of surface material from just one language, combined only with abstract mapping rules from another. The striking feature of mixed languages is that they neither follow this latter pattern, nor always behave consistently as regards the separation, by source language, of content and system morphemes.

While it is evident that mixed languages are the

[1] Cf. Bakker and Muysken (1995, 50): "The way in which intertwined languages are formed appears to be highly uniform".

[2] Cf. Bakker (1997, 210): "Neither component is more important than the other; they both have the same weight".

[3] Such gradual transition from code switching to mixed languages is disputed by Muysken (1997) and by Bakker (1997), whose comparisons of mixed languages with code mixing involving the same source languages fail to show a continuum of mixing ranging from "average" to extreme. On similar grounds, Bakker and Mous (1994b, 5) argue against a continuum of lexical borrowing that leads to mixed languages. Borrowing is said to affect roughly 45% of lexical material, most of it peripheral, while in mixed languages up to 90% is affected, including the basic lexicon (see also Muysken, 1997, 378).

[4] It is not insignificant however that these mixed styles typically involve a conventionalisation of the use of borrowed discourse markers, an issue to which I shall return when discussing the notion of "fusion" below (see also Matras 1998).

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

product of social motivations to maintain a unique linguistic system in a community, no concentrated attempt has so far been made to relate the *purpose* behind the creation of a mixed variety to the actual compartmentalisation of structures that it displays. My concern here is to investigate the relation between social motivation and communicative needs, and the functional properties of linguistic categories that are affected by language admixture.

Structures as mental processing instructions

In attributing processing functions to linguistic elements, I am partly inspired by the work of the Bühlerians in the tradition of Functional Pragmatics (see Ehlich, 1986; Brünner and Graefen, 1994; Bührig and Matras, 1999; Redder and Rehbein, 1999). Already Bühler's (1934) view of language recognises a compartmentalisation of the linguistic system, assigning mental activities that are triggered by linguistic signals to a "symbolic field" where meaning is encoded, and a "deictic field" that encodes gestures of reference to components of the situation or by analogy to a mental map based on verbalised representations. Bühler's general approach has been extended in Functional Pragmatics to include such domains as the operative, expeditive, relational, and expressive fields, activation of which allows, respectively, the processing of elements of speech (in the case of inflection and conjunctions), to direct the listener's attention (interjections and particles), to convey internal categorisations performed on other linguistic elements (adpositions and local relations), or to portray real-world occurrences (onomatopoeic representation). While I will not follow this model in detail, I shall nevertheless assume that linguistic systems display such natural compartmentalisation of functions: structural categories are intrinsically responsible for triggering *linguistic–mental procedures* or processing instructions, i.e. they represent distinct cognitive–mental activities in which speaker and hearer engage during linguistic interaction. Compartmentalisation within a mixed language might constitute, then, a way of designating separate processing functions to each of the source languages.

A consistent grammar/lexicon split, for example, might be interpreted as a transposition of the symbolic field of language, where meaning and reference to real-world experience is conveyed, into the domain of an intruding language (in the case of colonial settings for mixed languages such as Michif or Media Lengua). Such *partial* transposition is well attested in conventional cases of languages in contact – consider for instance the wholesale import of names for mainstream institutions into minority or immigrant languages, the import of technical vocabulary, and so on. What this means for active bilinguals is not only that gaps in the native language are filled, which is how structure-based analyses regard the phenomenon of "cultural loans" (cf. Myers-Scotton, 1992), but that certain classes of meaning are negotiated outside the native language system, or rather that a compartmentalisation of meaning is achieved, with certain domains merging with representations from the contact language. The *wholesale* adoption of lexical entries, as assumed for the "ideal" mixed language, is a tendency to shift conceptual representation *per se* into the sphere of the contact language. Such a move will be motivated not by a sense of "equal contribution" from each language (cf. Bakker 1997), but rather by an unbalanced distribution of functions, speakers taking the "point of view" that negotiation of meaning in L2 is more attractive or advantageous than that in L1. There are two main reasons why such a grammar/lexicon split might arise in a speech community. The first is strategic, and pertains to the deliberate attempt to conceal meaning. The second is connected to a more permanent cultural adaptation, and has to do with adopting the L2-culture's realm of experience and representation of collective knowledge.

Contact mechanisms in functional perspective

In the following sections I wish to explore the connection between the motivation to mix and conventionalise patterns of mixing, the structural compartmentalisations that appear in some mixed languages, and the natural compartmentalisation of linguistic functions. I distinguish four contact mechanisms that appear to be involved, separately or in combination, in the creation of mixed languages: lexical re-orientation, selective replication, convergence, and categorial fusion.

LEXICAL RE-ORIENTATION is the conscious shifting of the linguistic field that is responsible for encoding meaning or conceptual representations away from the language in which linguistic interaction is normally managed, organised, and processed: speakers adopt in a sense one linguistic system to express *lexical meaning* (or symbols, in the Bühlerian sense of the term) and another to organise the relations among lexical symbols, as well as within sentences, utterances, and interaction. The result is a split, by source language, between lexicon and grammar. Re-orientation usually assumes that an underlying representation of similar value already exists in L1 (the underlying or receiving language). Thus re-orientation does not refer to the import of so-called cultural loans (new words for new objects or con-

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

cepts), but rather to the substitution of existing labels or symbols (cf. also Muysken, 1997, on re-lexification). In this, lexical re-orientation differs from prototypical cases of lexical borrowing. A further major difference is that, where lexical borrowing targets vocabulary for which counterpart expressions already do exist in a language, it does so only selectively and partially. Lexical re-orientation, on the other hand, is *wholesale*. The most obvious example of lexical re-orientation are secret languages. Here, lexical material is inserted into mixed utterances in order to conceal meaning. Consider the case of *Lekoudesch*, the secret language of Jewish cattle-traders in southwest Germany (and other German-speaking regions):

(1) D'r *guj veroumelt lou*.
"The man doesn't understand."

The italicised items indicate lexical elements that are recruited from a fixed inventory of secret lexical material, mainly of Ashkenazic Hebrew origin. The grammatical framework is dialectal German. From the purpose of such mixed utterances – to conceal meaning from outsiders – it is rather obvious which elements are natural candidates for lexical re-orientation: those, namely, that encode meaning – lexical roots of nouns, verbs, sentence adverbs and adjectives, numerals (especially in secret trade languages), as well as negation markers and existential verbs. The latter are "grammatical" in a sense; but unlike other grammatical devices – passive constructions, tense markers, complementisers, and others – existential expressions and negation markers can form the backbone of propositions, signifiying all the difference between being and not being, or a state of affairs and its absence. The notion of lexical re-orientation is intended to capture the inherent link between the motivation to mix – the marking out of propositional *meaning* on a consistent and wholesale basis, and the structures that are targeted – those elements, namely, that convey meaning. I return to discussing lexical re-orientation in more detail below.

SELECTIVE REPLICATION is a process which accounts for the continuous presence of structures from an earlier community language. In discussing interrupted language transmission across generations, Thomason and Kaufman (1988, 101) emphasise the relevance of hearing a structure in context to its acquisition by the younger generation. Selective replication assumes a motivation to reproduce salient elements of a language, without relying on grammatical operations – i.e. on language-internal processing instructions – in this language. The process differs from lexical re-orientation in that it does not, either specifically or exclusively, target conceptual representations (meaning), but aims rather at reproducing contextually relevant actions of speech as speakers try and re-activate impressions of an ancestral language. In addition to lexical vocabulary, deictic elements and interjections that are high on the scale of situative saliency are likely candidates for replication. Frozen or fossilised operative items (e.g. inflection) may accompany reproduced material but will usually remain non-productive. The main motivation for selective replication involves a turnover of the functions assigned to a linguistic system – from that of an everyday community language, to a specialised variety or register. The idea is that an older generation of speakers still uses the ancestral language as an all-purpose language, but also as an in-group secret language. The younger generation shifts to a different language for all-purpose communication, but has an interest in maintaining a secret in-group language. It reduces the ancestral language to a secret language, with far-reaching structural consequences. A full elaboration of the functional turnover hypothesis is given below.

CONVERGENCE: it is often impossible to assign certain portions of the "grammar" component of mixed languages to either one or the other contributing language, as they appear to be a hybrid product of the two. We might find material from one language taking over syntactic and semantic features of partly corresponding elements in another. Such processes are not confined to mixed languages, but can be observed in a variety of language contact situations where speakers synchronise operations that are responsible for the organisation of an utterance in two or more languages, while at the same time maintaining the material–structural autonomy of the linguistic systems.

I define convergence as the adaptation of an internal element in Language A to match the scope and distribution of an element in Language B that is perceived as its functional counterpart. Convergence thus assigns a particular structure or item in Language A the function of trigger for the same grammatical operation that is activated by a semantic or structural counterpart in Language B. The distinction between languages A and B is, however, for definition purposes only; the process is often mutual, triggering changes in both languages. Such harmonisation of grammatical operations may typically affect the layout of propositional content at the sentence level (i.e. word order and patterns of clause combining); an exemplary case is the loss of the infinitive in the Balkan languages and the use of finite complements, same-subject modal constructions, and purpose clauses. Convergence may also affect lexical semantics, as well as inflectional mor-

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

phology. Thus, in Turoyo, a NeoAramaic (northwest Semitic) language of southeastern Anatolia, the Kurmanji (northwest Iranian) ergative construction is imitated using inherited morphological material. A distinction is introduced between past-tense intransitives, where subject concord is marked through the set of affixes also employed for the present tenses of all verbs (*azzi-no* "I went"; Kurmanji *ez çû-m*), and transitives, where the subject concord markers derive from indirect object markers (*mīd-li* "I took"; Kurmanji *min girt*).

Convergence may be seen as a compromise between merging patterns and retention of structural autonomy. In this regard it may be indicative of the social positions and functions associated with the participating languages in a multilingual community. In eastern Anatolia, convergence also appears in the layout for present-tense verbs (cf. Chyet, 1995, 240): in the indicative, a prefix expressing progressivity precedes the verb stem, while person affixes follow (Turoyo *ko-ḥoz-eno*, Kurmanji *di-bîn-im*, Western Armenian *gə-desn-em*; "I see"). The pattern even appears in a larger area, which includes Persian (*mī-bīn-æm*) and Arabic (although here person marking also precedes the verb stem: *b-a-šūf-ø*). In all these languages, the subjunctive is expressed by a zero-prefix, or in some languages (namely Kurmanji, and other Kurdish varieties, and Persian) through variation between a zero-prefix and a subjunctive prefix (Turoyo *ḥoz-eno*, Kurmanji *bi-/bîn-im*, Persian *be-/bīnæm*, Western Armenian *desn-em*, Arabic *a-šūf-ø*).

CATEGORIAL FUSION: contrary to convergence, where mapping rules merge but the material autonomy of the systems is retained, fusion implies that speakers do not differentiate systems while carrying out certain linguistic processing operations, but instead draw on the resources of just one single system for a particular class of functions. Fusion is thus the wholesale non-separation of languages in both forms and functions of a given category or class of expressions. The typical cases of fusion are discourse-regulating elements – discourse particles, conjunctions, interjections, and hesitation markers – as well as phasal adverbs and focus particles. Cases of languages in contact that have replaced this class of items on a wholesale basis are widely attested. In Matras (1998) I discussed the predictability of the process of fusion. In terms of semantics, there is diachronic evidence that fusion tends to follow the hierarchy "contrast, change, restriction > addition, continuation, elaboration". In structural terms, elements that are lexically less analysable and more gesture-like, such as hesitation markers, are more likely to undergo fusion at a relatively early stage of contact than items that are lexical, or overlap with cognate lexical or deictic items, e.g. discourse markers such as "you know" or "then".

How can one use such data to construct an explanatory model of fusion? From a discourse-interactional viewpoint, "contrast" can be interpreted as a linguistic–mental gesture that aims at an intensified capturing of the listener's attention, indeed at direct intervention with hearer-sided processing of propositional content amid potential interactional disharmony (e.g. disbelief, as a result of the breaking of causal chains; cf. Rudolph, 1996). The high ranking of contrast and gesture-like expressions on the contact-susceptibility hierarchy suggests a cognitive motivation for fusion: it is triggered by cognitive pressure to reduce the mental processing load and so eliminate the choice between two competing systems while carrying out highly automaticised, gesture-like operations that may result in high conversational tension (monitoring-and-directing operations). Next-ranking positions on the fusion-susceptibility hierarchy show non-factuality and epistemic qualification of propositions, where the speaker's authority is potentially at stake and the task of "selling" an assertion to the listener entails a high "risk factor".

This interpretation of fusion as a cognitive process is supported by evidence from spontaneous speech production errors by bilinguals, which frequently target similar grammatical elements (see Matras, 1998, forthcoming). Further evidence comes from my own observations on early child bilingualism. In a situation where each parent consistently speaks their own language, and the parents live apart, language separation at the age of 1:11–2:7 is consistent. Mixing only takes place (a) where new vocabulary has been acquired in a specific context, and the matching expression in the other language is not yet known; or (b) in the few hours following a change in setting from the care of one parent to that of the other, where function words of the monitoring-and-directing class (interjections, presentatives, phasal adverbs, focus particles) from the foregoing situation are carried forward into the language of the "new" setting. (A detailed discussion of the data is beyond the scope of this paper.)

The orientation target for fusion is the *pragmatically dominant* language: the language which, in a given moment of discourse interaction, is granted maximum mental effort by speakers. This may be the speaker's first language, or one that is dominant for a particular domain of linguistic interaction, or one that exerts pressure due to its overall role as the majority language that is culturally prestigious or economically powerful (for a more detailed discussion of the notion of a pragmatically dominant language, and evidence from speech production, see

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

Matras, 1998, forthcoming). Fusion differs from lexical re-orientation and selective replication in that it is not a deliberate or conscious process, and that it targets items that are high on the scale of automaticised processing functions, rather than on the scale of referential or situative saliency.[5] Crucial to the relevance of fusion to the present discussion is the fact that, given what I would call "sociolinguistic admissibility" – no constraints on intelligibility and no loss of prestige – fusion as an occasional process that occurs with individual speakers in some situations, may give rise to long-term language change, by which the entire system of monitoring-and-directing operations, and possibly other functional categories in language too, are replaced on a wholesale basis. Again I refer to earlier work for a more elaborate discussion (Matras, 1998, forthcoming; and see also Salmons, 1990).

Lexical re-orientation in non-self-contained mixtures

Motivations

I return to Smith's (1995) sub-categorisation of mixed languages, and in particular to his use of the term "symbiotic" to designate those mixed varieties that are spoken alongside a related non-mixed language. Structural approaches tend to view symbiotic varieties as genuine mixed languages, or at least as representing underlying mixed languages (for somewhat conflicting views on Para-Romani varieties cf. Bakker and Van der Voort, 1991; Bakker, 1994; and Bakker, 1998; and see discussion below). A great proportion of so-called symbiotic mixed languages in fact involve rules and conventions that are not used for the actual production of fluent conversation, but merely for the production of single *mixed utterances*. I therefore prefer to call them "non-self-contained mixtures".

A possible motivation for the creation of non-self-contained mixtures may be, in the case of secret languages at least, to by-pass the norms of mainstream communicative interaction in a community while avoiding the sanctions normally associated with such a breach of social conventions. Mixing is intended to make the offence against the rules of communicative interaction as unapparent as possible.[6] The targeting of content words here is due to the trivial yet crucial fact that content words convey meaning. It is meaning that reveals thoughts and intentions, and so meaning that is sanctionable, and is therefore sought to be concealed. Once established, the in-group habit of disguising meaning might of course also function as a token in its own right, stressing or even serving to ascertain membership in the group. Some examples follow below.

The discourse position of mixed utterances

Lekoudesch was the secret language used until the early 1930s by Jewish cattle-traders in southwest Germany. It is closely related to other varieties with cognate designations – derived form the euphemistic *lošn-koudeš* (Ashkenazic Hebrew for "sacral language", and the term used to refer to written Hebrew) – which are attested in Germany and the Netherlands (cf. Meisinger, 1902; Moormann, 1920, 1922; Matras, 1991; Klepsch, 1996). Lekoudesch survived until the late 1980s among non-Jewish farmers in a number of communities that had had a significant pre-war Jewish population of cattle traders.[7] The examples presented here are taken from interviews with non-Jewish informants in the Swabian villages Rexingen and Buttenhausen, as well as with Jewish cattle traders originating from the same area who emigrated to Palestine in 1938.

Lekoudesch is listed by Smith (1995, 367) as a symbiotic mixed language, and indeed a formal inspection of isolated sentences renders the impression of a mixed system:

(2) *Lou dibra*, d'r *guj schäfft*!
"Don't speak, the man is (there)!" [=a stranger is listening]
(3) Alle *gimmel doff*.
"All three (are) good." [=about cattle]
(4) Die *goja* isch *haggel doff*, dia kennt-m'r *lekächa*.
"The woman is very pretty, one could take her."

If we were to follow the conventions that are common in descriptive accounts of mixed languages (cf. Bakker and Mous, 1994a; Thomason, 1997a), we would derive from (1)–(4) the following characterisation of features by source language for Lekoudesch:

(5) Structural compartmentalisation by source language in Lekoudesch:
Ashkenazic Hebrew: nouns, verbs, adjectives, adverbs, numerals, negation markers, copula *schäff-* (< Hebrew "to sit").

[5] That fusion may act as a separate process is supported by Muysken's (1997, 402) observation that Spanish-derived coordinating conjunctions in Media Lengua are exceptions to his re-lexification hypothesis, as they are not modelled in any way on the underlying Quechua system of semantic–functional categories.

[6] The use of specific structures to exclude outsiders from the interaction has therefore been termed "bystander deixis" by Rijkhoff (1998).

[7] For surveys of the traces left by Jewish traders' jargon in German dialects see Meisinger (1901), Althaus (1963), Weinberg (1969), Röll (1986), Post (1992).

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

German (Swabian/Franconian): articles and determiners, verb inflection, copula, pronouns, word order.

In addition, the lack of a copula in (2) could be interpreted (again, if we were to adopt a strictly descriptive viewpoint) as either simplification of the German sentence model, or perhaps as a replication of the Hebrew nominal clause structure, which lacks a present-tense copula. Further structural features encountered in Lekoudesch are listed in (6):

(6) a. Creation of new lexical items by adding a German agentive suffix to Hebrew roots: *duh-gemmer* "fisherman"
b. Creation of new lexical items based on German rules for lexical composition, and partly as loan translations from German secret languages (Rotwelsch and Jenisch varieties): *schochamajim* "coffee" (< "black water'), *seifelbajis* "toilet" (< "shit-house"), *kassirrosch* "pig-head"
c. Generalisation of *lou* as an all-purpose negator
d. Integration of a number of verbs of Romance origin: *disemma* "whisper", *dormena* "sleep"

Lekoudesch thus constitutes not an arbitrary insertion of lexical items into a German framework but rather a rule-governed, consistent, and conventionalised system. Yet its composition already reveals much about the purpose for its creation: it is essentially a lexical reservoir used to conceal meaning by manipulating key propositional items in key utterances. Basic vocabulary from Hebrew sources serves as the basis for this lexical reservoir. It is enriched through productive lexical composition as well as through borrowings from other neighbouring secret languages such as Jenisch – as seen with Romance-derived items (cf. discussion in Wexler, 1988, 139ff.).

The function of Lekoudesch is nicely illustrated by the following two examples from Jewish speakers. In the first, (7), a woman characterises the use of Hebrew-derived items – referred to here significantly first as "the Hebrew language", then as "the counting [=trade] language" – as strictly a secret trade jargon that was understood by members of the community but not actively used outside the domain of cattle trade:[8]

(7) Die hebräische Sprache, die/ die/ die Zahlasprache es is nur gewesa wenn m'r ainem zwaita net wissa lassa wollt wie hoch der Preis isch, aber sonscht kaine/ oder se hawwe saga kenne "was *schuckt* die Kuh?", "wieviel kostet die Kuh?" Des/ des kann ich mich erinnern.
"The Hebrew language, the/ the/ the trade language it was only when you didn't want the other person to know what the price is, but otherwise none/ or they could say 'how much is the cow worth?' That/ that I can remember."

In example (8), the use of Lekoudesch to conceal meaning (and its ironic outcome, as the bystander turns out to have been familiar with the secret code) is reconstructed in an anecdote:

(8) Es is zu ihm gekommen eine alte Kundin, [...] Und ... als die Frau gekauft hat, sagt er zu mainer Großmutter: "Hått die *Goja* auch ebbes *meschullemt* am *Bajiss*?" [...] Nå hat die aber das verschtaãnda, nå hått sie zu mainer Großmutter gesagt: "*Kenn*, *kenn*, aber nur a *Mattle*" [...] Un die Christin hat das gekannt, von früher, von de Viehhändler her.
"An old customer came to him [...] And ... as the woman was buying something he says to my grandmother: 'Has the woman paid anything [before] in the house [=shop]?' [...] But she understood, so she said to my grandmother: 'Yes, yes, but only a little' [...] And this Christian woman knew it, from earlier times, from the cattle traders."

A further example of a non-self-contained mixed variety comes from roughly the same region. *Jenisch* is used by populations of peripatetics or commercial nomads (showpeople, traders, craftsmen) in southwest Germany to designate a variety of secret lexicons. They typically draw on a pool of lexical items that contains elements of older Rotwelsch (camouflaged derivations of German lexical roots) and more recent loans mostly from Romance, Ashkenazic Hebrew, and Romani. My examples derive from Unterdeufstetten, an itinerant base-community in the Franconian district of Schwäbisch Hall.[9] It is one of several communities in the immediate vicinity in which itinerants began to settle in the second half of the 18th century (for a historical and ethnographic description, see Dürr, 1961). The term *Jenisch*, which speakers use to refer to their secret language, also appears as an adjective denoting culural particularities, as in *a "Rudel" isch a jenischer Wåche* "'Rudel' means a Jenisch cart", as well as an ethnonym: *Manische send Zigeuner, mir send Jenische* "Manisch people are Gypsies, we are Jenisch people". This self-designation is shared with similar groups within the

[8] Cf. also the statement of a non-Jewish farmer: *Dia håand lekoudesch dibra daß dia andra's it v'rschtãnda håand* "They spoke ledkoudesch so that the others would not understand".

[9] Fieldwork was conducted together with Thomas Jauch, in September 1996.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

region but also in more remote areas, and the Jenisch people of Unterdeufstetten are conscious of linguistic diferences and affinities among the various Jenisch varieties. They are also aware of historical links with Romani settlers in their community, and indeed there are plenty of traces of Romani material and spiritual culture, including for example the belief in the returning spirits of the dead (see Dürr, 1961, 105–106), referred to in the local Jenisch as *mulo* (Romani *mulo* "dead", but also "ghost, spirit").

Jenisch satisfies a series of criteria for mixed languages: it appears to have arisen in a situation of population mixture and is used to mark group identity in defiance of mainstream cultural norms. A look at single isolated utterances will tempt descriptivists into concluding that non-German lexicon is combined here with German (Franconian) grammar:

(9) I *nasch* in *Gatschemm* un *schwäch* ein *Blamm*.
"I go to the pub and drink a beer."

(10) Als Generaldirektor muscht du au *schenegla* sonscht *schiwsch tschi Lowe*
"As general director you need to work too otherwise you get no money."

However, looking at a brief conversation where Jenisch elements are employed, our impression of an actual "language" gives way to the observation that special lexicon is inserted only sporadically, consciously, and deliberately in order to shift meaning, and so settings, from a serious, straightforward, or factual interaction to one that challenges social and conversational taboos, and so might be less committing or even humorous:

(11) A: Wenn man müed isch?
B: *Nasch* ins *Tschiben*. Aber muscht *spanna*, daß a *schuggere Tschai* håscht, nå kåsch/ daß nei*durma* kåsch.
A: A *schuggere Tschai*, zum *durma*?
B: *Kenn*. Bisch miad, etzt *nascht* ins *Tschiben*, nemmsch *a schuggere Tschai* mit, und dei *Moss* dia *guftet* dann an de Ohra na.
A: "If you're tired?"
B: "Go to bed. But you need to see that you have a pretty girl, so you can/ so you can sleep."
A: "A pretty girl, to sleep?"
B: "Yes. You're tired, so you go to bed, you take a pretty girl with you, and then your wife will slap your ears."

Lexical re-orientation thus aims at manipulating conceptual representation with the intention of disguising key slots in discourse that are likely to cause friction, since they fail to conform to communicative norms and constraints on content regulation. The product of the process consists of scattered mixed utterances and does not constitute a self-contained communicative system. Hence there is no re-lexification, and strictly speaking no intertwining of grammar and lexicon to produce a "language". What exists are rather conventions that allow speakers to shift meaning or content away from the domain of common everyday speech, motivated by the need to sustain and express a separate system of values beyond the control reach of mainstream society norms. This implies a conscious orientation towards alternative sources of conceptual representation, including creative production of lexical items.

The Functional Turnover Hypothesis

Lexical re-orientation accounts for the conscious replacement of elements that carry key propositional representations: content words, negation, copula. But non-self-contained mixed languages may show additional structural categories as well. These typically enter a mixed system through selective replication, as part of what I call *Functional Turnover*: a language is used by an older generation of speakers as an everyday community language, though it is also kept secret from outsiders, thus also enabling a kind of "conspiratory" communication among members of the community in the presence of strangers. As linguistic assimilation proceeds in the community, the younger generation shifts largely to the outside or mainstream language for the purposes of everyday communication even within the community. The ancestral community language is then reduced to its function as a secret language, a function which, due to the nature of social relations between the minority community and majority society, is still required. The shift in the balance of the overall communicative functions assigned to the ancestral language – from an all-purpose language with secretive usages, to an exclusively secret language – explains the motivation to maintain portions of it (those that allow it to serve its purpose by concealing meaning), while gradually neglecting other components, notably the bulk of grammatical structures. The process is indeed reminiscent of the development of secret languages through Lexical Re-orientation, the difference being that the Functional Turnover Hypothesis assumes the gradual decline of a language once spoken in the community as an all-purpose language, rather than deliberate recruitment of lexical material to substitute for key lexical items in the everyday language. I call the attempt to keep an ancestral language alive for the purposes of secret in-group communication *selective replication*. I

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

devote this section to a discussion of Para-Romani varieties as examples of selective replication.

"Leftover" grammar

Contrary to the impression conveyed in the literature on mixed languages, there exists no authentic or at least no reliable documentation of a Para-Romani *language*. Existing texts primarily testify to the ability of authors – or rather their informants – to construct mixed sentences drawing on Romani vocabulary when asked to do so.[10] But there is little indication of the extent to which such mixing strategies are or were in use in actual conversation beyond the production of occasional mixed utterances.

Two exceptional sources that provide transcripts of tape-recorded interviews with Gypsies where Romani-derived items were used, are Acton (1971) on Angloromani (Romani "lexicon", English "grammar"), and Leigh (1998) on Caló (Romani "lexicon", Spanish "grammar"). Acton (1971, 121), however, reports that the presence of a tape-recorder made the Gypsies' conversation "slightly more formal in tone", and that "except in utterances which I was usually unable to catch, they stopped using Romani phrases". The transcript still shows frequent use of the word *Gaujo*, *Gauji* "non-Gypsy", which is also a common Romani loan into non-Romani based secret jargons and slang, as well as the tag *mush* "man, mate". Another frequently occurring Romani item was *moxadi* "polluted", a cultural term closely associated with the topic of the conversation, which was norms of hygiene and ritual purity among Gypsies and non-Gypsies. Thus there is attestation of both a symbolic and an emblematic usage of Romani-derived items, but not one that justifies the label of a mixed language. Of special interest to our discussion is the following interaction between Acton and several Gypsy men (Acton, 1971, 124–125):

(12) T: . . . you go up to our caravans, you don't see big old dogs round about up there, all hairs all over everywhere. You go in these houses, the door, an' the fust thing that comes up to the door is a great ol' dog run an' jump at the door, claw the door. [. . .]
Acton: I must confess we do that, we've got a dog at home.
T: Now you've got to speak as you find it, know what I mean.
M (embarrassed): I mean you got to speak as you find it, haven't you?
Acton: Yes.
T: Now I've been in . . .
W (very quickly): Don't *ker* agen, *mush*. D'man do-knows 'is *piyamengri*.
Acton: Sorry . . . er . . .
M: Now I used to be mates with a *Gaujo* fellow . . .

In the excerpt, T is comparing Gypsy cleanliness attitudes with those of the non-Gypsies, exemplifying them through the presence of hair-shedding dogs in non-Gypsy houses. Acton's confession of guilt embarrasses T and M, who try to elicit from the guest reassurance of not being offended. As T then wishes to continue, W intervenes, evidently in an attempt to avoid further offence. His words are directed at T and intended to by-pass Acton's conversational scrutiny. They are spoken very quickly, and key items are encoded in Romani: *ker* "do", *piyamengri* "tea" (a nominal derivation from "drink", and so most likely a conscious though perhaps not spontaneous creation). To use Rijkhof's (1998) terms (cf. also Clark and Carlson, 1982), Romani serves here as a negative bystander deixis, aiming at excluding the bystander – Acton – from accessing meaning and communicative intention.

In (12), then, the Romani component is functionally equivalent to the non-German components of Lekoudesch and Jenisch discussed above. The extraordinary feature of Para-Romani varieties, however, is the preservation of frozen and some semi-productive inflectional features, as well as of structural categories such as pronouns and demonstratives (cf. Bakker, 1998; Boretzky, 1998). Leigh (1998, 275) cites Spanish Gypsies using the phrase *mansa camelo tuque* "I love you", where *mansa* reflects the underlying Romani sociative case of the first-person pronoun, *cam-el-o* combines Romani third-person singular present-tense inflection with the Spanish conjugation ending, and *tuque* represents the Romani dative case of the second-person pronoun. Such frozen grammatical elements testify to earlier access to an inflected variety.

The question of how inflected Romani may have given way to the Para-Romani mixed varieties (attested mainly in Britain, Iberia, and Scandinavia) has been the subject of controversy in Romani linguistics. Hancock (1970, 1984, 1992) attributes the creation of Romani-based mixed varieties to population mixtures drawing deliberately on Romani lexicon for secretive purposes at the initial stage, then preserving the new mixed in-group variety to re-inforce Gypsy ethnic identity while abandoning inflected Romani. An alternative scenario proposed by Kenrick (1979) sees the decline of Romani proper as a process *leading* to the emergence of an in-group Gypsy mixed

[10] For a discussion of the authenticity of Para-Romani texts see Bakker (1998, 84–87) and Boretzky (1998, 110).

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

ethnolect, rather than resulting from it. Boretzky (1998) maintains on the basis of the lexical resources available in Para-Romani, that the "creators" of these varieties will have had direct access to a productive, full-fledged inflected Romani proper, but were not themselves active speakers of inflected Romani (see also Boretzky and Igla, 1994).

A case to be considered is a secret language used in the Hessean town of Giessen in central Germany, referred to by its speakers as *Jenisch* but by outsiders as *Manisch* or "Gypsy"[11] (Lerch, 1976). Giessen Jenisch or Manisch shares a core lexicon of diverse origins with other Jenisch varieties, such as Unterdeufstetten Jenisch considered above. In addition, it shows several hundred Romani-derived roots that are not shared with other secret vocabularies, rendering the impression of an underlying Jenisch variety that has more recently incorporated a significant structural layer from Romani. Alongside Romani-derived content words and negation markers, Manisch has some semi-productive noun derivational morphology from Romani (agentive suffixes and abstract nominalisation markers) as well as unbound grammatical morphemes, mainly expressions of time and location, quantifiers, interjections, and deictics. Significantly, the third-person pronouns *job, joi* appear in deictic–situational usages only (Lerch, 1976, 245):

(13) *Wer iš mado, də tšabo odər di tšāij? job iš mado*
"Who is drunk, the man or the woman? He is drunk."

The possessive pronouns *miro* "my" and *tiro* "your" are, according to Lerch, familiar to speakers, but are rarely used. Manisch also appears to be the only German-based secret language that employs Romani numerals. All this suggests a stage in which a population of German speakers had contact with active speakers of Romani, and copied elements of their speech without actually being able to process sentences in Romani, i.e. with no grammatical competence in the language. This stage is now transmitted in a fossilised form, and so we have in essence a replication of a copy. The selective copying of inflected Romani structures is best seen in lexical items typical of Manisch which are not found in the Romani compoent of other German secret languages, such as the verb "to come". Its infinitive *aben* in Manisch is derived from the Romani root *av-*, while the present-tense form *wild* "he/she comes" derives from the Romani present *vēla* "he/she comes". This "suppletivoid" usage indicates we are not dealing with an actual structural *system* of a language, but with the selective replication of single linguistic forms, based on the situation-specific contexts of their original usage by speakers of Romani. Apparently, members of the speech community came to treat Romani and Jenisch as functionally equivalent, and consequently allowed the two to merge. As Romani lost its role as an everyday community native language and a marker of ethnicity, its replication became subordinated to the communicaive purposes of Jenisch as a non-self-contained secret variety.

[11] Derived from *manuš*, an internal ethnonym in some Romani dialects of the German-speaking area, and a general term for "Gypsies" in German secret languages.

Transition and the functional continuum

One might define the breaking point away from Romani proper as a *turnover of functions*, where the need to retain a special variety is stronger than the ability to transmit a coherent linguistic–grammatical system. Normal language transmission then gradually gives way to selective replication of linguistic material. The Functional Turnover Hypothesis allows us to account for why, as Boretzky (1998, 98–99) points out, lexicon is retained, while in normal situations of language attrition lexicon is easily compromised: since the motivation for selective replication is, as with lexical re-orientation, primarily to by-pass mainstream communicative norms, priority is given to material that is pragmatically most salient for this purpose, namely to items that encode meaning. Speakers of Romani varieties with diminishing grammatical competence will have had a functional–communicative interest in preserving vocabulary, an interest that overrides the constraints of structural development patterns observed in normal situations of language attrition. Unlike lexical re-orientation, however, selective replication also allows for a partial retention of frozen grammatical components.

The Functional Turnover Hypothesis accounts for such retention without needing to postulate a gradual borrowing of L2 grammar into Romani, or re-grammaticisation. Consider the following data from Smart (1862–1863, 80) on a now extinct variety of English Romani, widely accepted in the literature as being the forerunner of (mixed) Angloromani. The examples illustrate how the Romani and English copula forms may be used interchangeably within the same corpus:

(14) *Dik, savo see?* A *gorgio?*
"Look, who is [that]? A stranger?"
(15) *Covvo Moosh* is a *gryengro*
"This man is a horse-dealer"

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

In the same or a closely related variety, the inflection of lexical verbs appears to have been variable too. Consider the following data, taken from notes on English Romani collected by T. W. Norwood in Cheltenham, Gloucestershire in April 1863:[12]

(16) *Kanna shummus tarno,* I used *to jiv*
when was.1SG young live.Ø
kerrasty
house.LOC
"When I was young, I used to live in a house"

(17) We shall *jassa* *kallako*
go.1PL.FUT tomorrow
"We shall go tomorrow"

These examples show Romani inflection (person and tense on verbs; locative case on a noun) and independent clause and phrase syntax (Pro-drop in the first clause of (16), and lack of indefinite article). At the same time, some tense–aspect forms are English, carrying with them English personal pronouns. This in turn results in a hybrid construction in (16), where, by analogy with English, the Romani root *jiv* "to live" is employed as an infinitive; and in the doubling in (17) of tense and person marking, which appear both in English and as inflections on the Romani lexical verb.

The inevitable conclusion from (14)–(17) is that the notion of a "grammaticiser" language is ambiguous here, while the "lexifier language" is clearly Romani. This means that, in transitional stages, it is difficult to argue eiher for a complete substitution of Romani grammar through the grammar of the contiguous language, *or* for a plain insertion of Romani lexical items into the strict framework of the contiguous language. It is, however, plausible that the speakers who produced these sentences were making a conscious choice in favour of Romani lexical vocabulary. On this basis, so-called Para-Romani might be defined as a process of *diminishing grammatical competence* coinciding with a deliberate effort to *maintain lexical competence*. The partial retention of grammar is a by-product of the effort to maintain a separate code, or selective replication. This in turn is motivated by the turnover of functions, the process by which the language becomes restricted to specific functions and discourse positions.

In the case of Para-Romani, there is room to consider the background for functional turnover already within underlying, western dialects of Romani proper that gave rise to mixed varieties (Iberian, British, German, and Scandinavian Romani). To begin with, we find among speakers of these dialects a reluctance to share the language and its structures with outsiders. This view of the language as secretive suggests of course that a form of negative accommodation or bystander deixis belongs to its primary functions. This is also supported by structural evidence. Where central and eastern European Romani generally adopt loanwords rather freely, western dialects show a preference for internal compositions and derivations, as well as euphemistic and cryptolectal formations. Liebich (1863, 90–92) has already noted the creation of numerous cryptic–interpretative placenames in German Romani (Sinti) that are strongly reminiscent of compositions in Rotwelsch, Jewish cattle traders jargon, and other secret varieties, such as *xamaskero foro* literally "the eating-town" for "Breslau", based on *bres-* > *freß-* German "to eat". Günther (1915, 16–19) points out similarities between German Romani and Jenisch in drawing on existing lexical–semantic resources for creative, euphemistic lexical compositions, as in Romani *muleskro khe*r lit. "dead man's house", Jenisch *Begerkittle* of the same composition, for "coffin" (see also Wagner, 1937). Typical of western dialects is also the use of Romani genitive derivations especially for the creation of words relating to human beings, professions, and economic resources such as animals, food, or agricultural terms – all reminiscent of lexical re-orientation strategies observed in non-self-contained secret languages. The same cluster of dialects also show a preference for group-specific autononyms – *kale* "blacks", *manuš* "persons", or the names *romaničel*, *sinte* – over the general ethnonym *rom* (which survives in the word "husband/wife" as well as in the name of the language, *romanes*), a sign of social and ethnic isolation, which will have brought speakers of these dialects closer to itinerants of non-Romani origin. All this suggests a gradual shift of balance between the everyday, bystander-neutral communicative function of Romani, and the in-group secretive function it shares with artificial secret languages, thus strengthening the impression of a process that may ultimately result in functional turnover.

The case for selective replication in Ma'a

This brings me to another controversial case of a mixed language, that of Ma'a. Mous (1994) considers Ma'a an in-group register of (Bantu) Mbugu, and even labels it "Inner Mbugu" as opposed to "Normal Mbugu". Inner Mbugu or Ma'a is characterised by occasional lexical insertions into a frame that is on the whole compatible with Normal Mbugu (Mous, 1994, 177):

[12] Source: Scott Macfie Collection 4.1–5, University of Liverpool Special Archives (Gypsy Collections). See also Grosvenor (1910, 217–219).

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

(18) *aní* ní mborá nitété jangu já
1sg cop girl 1sg:say:sbj 10:mine 10:con
uborá [. . .]
14.2:girl
"I was a girl, should I talk about my childhood
[. . .]

halí ya ku*yó*?
other 9:con 15:speak
[Is there] more to say?"

Mous suggests that Ma'a never had a different function (in our terms, it was never self-contained), but was a deliberate creation motivated by the need to highlight social–economic and ethnic separateness. While it seems that his synchronic characterisation of the language and its functions has not at all been challenged, alternative suggestions have been made with respect to the process of emergence. Thomason and Kaufman (1988, 223–228) and Thomason (1997c) view Ma'a as a Cushitic language that has borrowed Bantu grammar. Among the arguments put forth by Thomason are syntactic features – Ma'a has a Cushitic copula and possessive transitive verb – as well as the existence of non-Bantu vocabulary that is not Cushitic (a sizeable number can apparently be traced back to Maasai; see Mous, 1994), which, it is argued, contradicts the assumption of an underlying Bantu language (Mbugu or Pare) that has borrowed from Cushitic. Sasse's (1992) approach seems to integrate the two points of view, arguing in favour of a Cushitic language that borrowed from Bantu, was then abandoned in favour of Bantu, and finally served as a source of re-lexification material on the basis of which a new language emerged.

Let us consider the categories that are specifically reserved for Inner Mbugu or Ma'a. With the exception of lexical items, most features are said to derive from Cushitic, although no single Cushitic ancestor has so far been identified. The overview in (19) is based mainly on Mous (1994):

(19) Non-Bantu components in Ma'a:
Lexical items (also non-Cushitic), non-productive causative affixes, non-productive and optional nominal suffix, numerals between 1–5 and 10, demonstratives and possessives accompanying non-Bantu lexical items, copula and possessive verb, personal pronouns, adaptation pattern for the inflection of some verbs, partial exemption from Bantu noun class agreement.

Selective replication allows us to account for the retention, alongside Cushitic vocabulary, of frozen derivational features, and of deictic noun phrase modifiers (pronouns, demonstratives, possessives), as well as of the most basic and frequently occurring numerals and the copula and possessive verb. Thus the bulk of Cushitic material fits the profile of pragmatic, situative saliency typical of the retention of features following a functional turnover. This is also consistent with what is known about the history of Ma'a and the gradual loss of Cushitic-derived grammar over the past generations. The retention of a separate identity motivates selective replication of Cushitic material from the original Ma'a, which undergoes a turnover of functions from a full-fledged community language into a lexical register. It is still defined by speakers as a "language" because it still constitutes a form of speech that identifies them as members of the community and allows in-group as well as ritual communication (in much the same way that users of Romani-derived lexicon in varieties of English and Swedish continue to call these registers *Romany* or *Rommani*).

This historical scenario for the emergence of Ma'a does not differ significantly from some of those proposed elsewhere in the literature (especially Mous, 1994 and Thomason, 1997c); indeed I do not pretend to offer any new historical evidence to explain the state of affairs in Ma'a. The contribution I wish to make pertains rather to identifying selective replication as a contact mechanism in its own right, one that is tightly connected to the process of a turnover of the functions assigned to the language of a community. The motivation to retain a register in order to flag ethnic distinctness amid linguistic assimilation will target components of the ancestral language in a way that is not arbitrary. In this sense, Ma'a cannot be viewed simply as a case where Bantu (Mbugu or Pare) grammar has simply been borrowed, gradually replacing the original Cushitic structure, but rather as a case of language shift with subsequent deliberate (albeit diminishing) replication of Cushitic material.

The cumulative effect of contact mechanisms

So far we have identified lexical re-orientation in Lekoudesch and Jenisch, selective replication in Para-Romani and Ma'a, and lexical re-orientation supplemented by selective replication in Manisch. Our small sample already shows how diverse mixed languages are as regards their structural composition and compartmentalisation. In this final section I examine some of the more challenging cases of mixed languages reported on in the literature.

Functional compartmentalisation

The discussion will presuppose a natural function-based compartmentalisation of linguistic structures, the following domains being of immediate relevance:

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

(1) MEANING is encoded through content elements (lexical items) that constitute symbols for conceptual representations, and which in turn may evoke associations of mental images that are culture-specific.[13] (2) Alongside content elements, there exists a series of grammatical structures that are crucial to meaning in the sense that they are, at the most elementary level, CONSTITUTIVE OF PROPOSITIONS. These include at least negation and existential verbs (possession possibly being a derived existential relation). (3) Operative procedures[14] carry out the PROCESSING of propositional contents including the relations among them and those among their constituents (represented through inflection, subordinating conjunctions, word order, and more). (4) Certain grammatical items are involved in MONITORING AND DIRECTING interaction, which includes assessing the listener's involvement, and anticipating and processing the listener's response to propositions (represented through discourse markers, interjections, focus particles, and more). (5) Finally, some classes of grammatical items are inherently situative, as they are involved in pointing out components of the speech action or speech situation. They are often formative of stereotypical situation-bound actions of speech, and can in general be characterised as COMMUNICATIVELY SALIENT and, from a language learner's perspective, EASILY RETRIEVABLE. Typical grammatical items in this domain are deictics, anaphora, possessors, quantifiers such as basic numerals, and inflected forms of semantically salient verbs of motion.

It was suggested above that a natural function-based compartmentalisation of linguistic structures is reflected in distinct mechanisms of language change triggered by contact, and I will very briefly review the argument. In doing so, I refer to the 5 categories outlined in the preceding paragraph. Lexical re-orientation aims at shifting MEANING into a separate system, and so it targets categories 1 and 2, which are formative of conceptual–propositional representation. Selective replication seeks to reproduce actions of speech in a language without actually needing to rely on processing or directing operations in this language, and so it targets, alongside meaning, salient components that are easily retrievable (categories 1 and 5). Convergence allows harmonisation of systems while maintaining their surface-level or material autonomy. It targets the more abstract components involved in processing propositional content (category 3), but may similarly affect the arrangement of symbolic representations or lexical semantics (category 1). Finally, fusion is the wholesale merger of processing operations by functional category. Since it is motivated by the attempt to reduce the "cost of production" of speech in favour of a pragmatically dominant language, fusion will typically begin with the more complex and interactively intensive mental tasks outlined as category 4, and will only then infiltrate the more proposition-internal processing functions, under 3, beginning with those that are listener-oriented.[15]

Extreme convergence and fusion

Extreme convergence pertains to cases where convergent structures dominate a language's grammatical operations. To argue from the point of view of historical reconstruction, only a small portion of operative procedures can be accounted for through direct and straightforward descent or "inheritance". The Dutch interlanguages Javindo, documented by de Gruiter (1994), and Petjo, as discussed by van Rheeden (1994) – I risk some simplification and deal with both together, and the reader is referred to the sources for a detailed discussion – display on the whole structures that are derived from Dutch, though the mapping rules (or the rules for grammatical organisation, as it were) are still Malay. For example, there is a shift toward a purely analytical tense and aspect system. This is clearly an outcome of continuing abstract processing in Malay, which was not abandoned but reinforced through a language acquisition process. Interestingly, a number of operations show fusion with the underlying blueprint language, Javanese and Malay respectively, notably the use of relativisers and of passive and causative morphology; here is a Petjo example (van Rheeden, 1994, 226):

(20) kleren *njang di*-wassen door die frouw
clothes rel pass-wash by that woman
"the clothes that were washed by that woman"

To judge by the information that is available, fusion in these cases does not actually follow its normal course, but targets selected operations without the wholesale incorporation of utterance modifiers, conjunctions, etc. Indeed, it is not entirely clear that Malay in fact has the role of the pragmatically dominant language which would constitute the target for fusion. Some evidence, however, might be sought

[13] I use "Meaning" here in the narrow sense to refer to the specific type of symbolic–conceptual meaning represented by lexical entities.

[14] I use the term coined in the Functional Pragmatics tradition (Ehlich, 1986), which adopts Bühler's (1934) communicative–psychological approach to linguistic fields as reflections of mental processing tasks. See above.

[15] The historical evidence for the hierarchisation of structures that are susceptible to fusion is reviewed in Matras (1998) and cannot be repeated here.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

in the fact that Malay provides expeditive elements – interjections and terms of address.

A further case for convergence and fusion – this time fusion being the "extreme" mechanism at work – can be made for Copper Island Aleut. My source of information on the language is mainly Golovko (1994), and partly Thomason and Kaufman (1988), Thomason (1997d) and Vakhtin (1998). Aleut provides (1) content words, (2) noun and verb derivation, (3) noun inflection, and (4) converbs. From this it is already clear that no grammar/lexicon split can be claimed for Copper Island Aleut. Russian supplies (5) a selection of content words, most of which appear to be common in all non-mixed Aleut dialects as well and so may be considered a case for conventional borrowing, (6) conjunctions and connective particles, (7) subjunctive and conditional markers, (8) word order, (9) the finite verb inflection, (10) negation markers, (11) personal pronouns, when used in conjunction with past-tense verbs and the impersonal future auxiliary, and (12) possessive and object pronouns. An illustration is provided by the following, from Golovko (1994,115):

(21) *yesli by oni* ukaala-ag'aa-*li* huzu-um *by*
if subj they here-move-pst.pl all-refl subj
txichi qala-chaa-*l*
refl.pl be-glad-caus-pst
"If they came, everybody would be glad"

Though it is clear that Copper Island Aleut came into being in a setting of mixed households (Aleut women, Russian men), there is some dispute whether it is the product of imperfect acquisition of Aleut by Russians (Thomason and Kaufman, 1988), or of Russian influence on the speech of Aleuts (Golovko, 1994). Vakhtin (1998) proposes a scenario according to which Aleuts had shifted to Russian, then re-acquired Aleut into a Russian framework. However, the fact that Aleut converbs and the complex nominal inflection are retained tends to point in the direction of continuation of processing operations in Aleut and so against selective replication, in the sense defined above; the failure to restructure the lexicon on a Russian model shows that Russian impact, however substantial, does not qualify as lexical re-orientation, either.

Components 1–4 can therefore safely be assumed to be inherited, representing continuation of Aleut, and we can devote our attention to the Russian impact, items 5–11. The first of those, 5, as mentioned, can be dismissed as common to all Aleut dialects and so as not characteristic of a mixed variety. Item 6 is explainable as the usual scenario for fusion with a pragmatically dominant language. To this I also attribute item 7: fusion extends to cover the subjective evaluation of propositional content or the domain of modality and non-factuality. From a conversation-interactional and language-processing point of view, subjunctive and conditional markers can be said to occupy a position in between monitoring and directing operations, and proposition-internal processing operations: they process the hearer's potential attitudes and anticipated reactions to what is said. This position is reflected in their place on the continuum for fusion. Item 8 – word order – reflects partial convergence with Russian in the domain of the organisational display of propositional content. This pertains especially to the loss of Aleut SOV structure and so to the position of the verb, which leads us to the most puzzling feature of Copper Island Aleut – the adoption of Russian finite verb inflection.

Speakers of Copper Island Aleut have made a significant step toward a shift to Russian in merging those processing operations that are the essence of message structure: the initiation of the predication and so the anchoring of propositional content is achieved in Russian. It is therefore not a coincidence that word order rules conform to the language that provides finite verb inflection. Moreover, once we have established the principle of predication organisation in Russian as a case of fusion of operative procedures, it is easier to explain the remaining Russian features: negation markers serve as essentials of the predication, being part of the overall propositional representation. Personal pronouns, finally, appear with past-tense verbs where Russian inflection does not distinguish person, as well as with the future auxiliary *bud* which apparently has been stripped of its Russian conjugation and assumes instead an impersonal form (cf. Vakhtin, 1998, 320), and so they can be considered an integral part of verb inflection, allowing differentiation of person consistently throughout the tense paradigm.

In conclusion for these two cases, then, one might argue that once we have identified the type of contact mechanism at work, a significant portion, though by no means all of the actual structural composition of the mixed language becomes predictable. Convergence is unilateral in Javindo and Petjo, using a Malay blueprint with Dutch material. Fusion in Copper Island Aleut encompasses classes of items that are universal candidates for fusion in language contact situations. Unpredictable and perhaps even unique are the cases of fusion involving Malay passive morphology, and Russian finite verb inflection.

Lexical re-orientation, fusion, and convergence

The combination of these three mechanisms of contact – lexical re-orientation, fusion, and conver-

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

gence – characterises two languages that have received much attention in recent discussion contexts on mixed languages: Media Lengua and Michif. It is these two that conform most closely to the image of a grammar/lexicon split. Muysken (1997) offers an elaborate discussion of the grammatical composition of Media Lengua, which I follow and interpret here. Quechua, the so-called "grammaticiser" language, supplies the following categories: (1) word order, (2) noun and verb inflection, and (3) sentential complementation and connective structures (clitics). Spanish, the "lexifier" language, contributes (4) lexical roots, (5) interrogatives, (6) frozen morphology imported with Spanish content morphemes, such as the past participle ending *-do* in adjectives, (7) Spanish affixes that are also borrowed into Quechua: diminutive *-itu*, present participle *-ndu*, clitic *-tan* from Spanish *-también*, (8) the forms for deictics and personal pronouns, (9) coordinating conjunctions and discourse particles, (10) complementisers and the structure of embedded wh-questions, and (11) the forms for comparative and reflexive markers. An illustration is provided by the following example from Muysken (1997, 377):

(22) *Isi*-ga *asi nustru barrio*-ga *asi*
this-top thus our community-top thus
kostumbri-n *abla*-na
accustomed-3 talk-nom
"In our community we are accustomed to talking this way"

The historical evidence – there is no attestation of a population mixture, and all speakers appear to be fluent in Quechua as well as, more recently perhaps, in Spanish – allows us to assume continuation of a form of Quechua that has absorbed a significant Spanish component, or, to use Muysken's term, has been re-lexified. There are two challenges posed by the nature of the compartmentalisation in Media Lengua: first, to explain the wholesale adoption of content words from Spanish, and second, to account for the split in processing operations.

The answer to the first question was hinted at above: Media Lengua represents a case of lexical reorientation in a system that has become self-contained. The motivation of speakers to mark their status in-between the two communities – rural Quechua and urban Spanish (Muysken, 1997) – has led them to shift the system of meaning representations, and thereby adopt the socio-cultural associations triggered by the use of Spanish lexicon. Lexical re-orientation in this case is therefore both symbolic and expressive.

The split in processing operations may be explained firstly by the postulation of a process of fusion of directing operations with those of the pragmatically dominant language, accounting for items 9–11, with wh-complementisers and comparatives representing relations between propositions and actors respectively and so figuring at lower levels, yet still on the continuum of epistemic qualification that is potentially sensitive to fusion. Items 6 and 7 require little attention, as they fall within the domain of well-attested and conventional borrowing, and are not distinctive features of the mixed variety of Quechua under discussion. There remains therefore the problem of Spanish interrogatives, deictics, and pronouns. All three group together as representing actor constituents, and so it is tempting to view them as extensions of the lexical domain, which is indeed the approach which Muysken (1997) takes. It is possible that speakers' grammatical system-awareness – a kind of abstraction from "grammatical intuition" – marks out the inconsistency in noun phrase representation and leads them to try and extend lexical reorientation to cover lexical placeholders, thereby triggering a process similar to selective replication, by which forms are copied but the operational system that governs their usage is lost. This means that the actual functions of the Spanish deictic system are not reproduced in Media Lengua. To compensate for the breakdown in the operation of deictics, this portion of the linguistic system undergoes a process of convergence through which Spanish forms are taken to represent the underlying Quechua functional paradigms, and partly merge with Quechua markers to create a hybrid system.

Our final case-study, Michif, has been considered by Thomason and Kaufman (1988, 228–233) as a form of Cree with wholesale replacement of native Cree nominal structures by French noun phrases. I follow Bakker (1997) and Bakker and Papen (1997) in listing the contribution by category from the two source languages. Cree contributes to Michif (1) verbs, including verb inflection and converbs, (2) word order, (3) interrogatives, pronouns, and deictics, (4) the majority of adverbial particles, (5) the obviative suffix on nouns, (6) some conjunctions and connectives. The French component comprises (7) nouns, (8) articles, numerals, indefinites, possessives, (9) most negation markers, (10) some adjectives, (11) some conjunctions and connectives, and (12) forms for adverbial conjunctions, relativisers, and impersonal modality markers. The category of adpositions is mixed, drawing frequently on combinations of French prepositions and Cree postpositions.

The core of the Michif predication is clearly based on Cree input, and again we see that the language that contributed verb inflection also provides word order rules. Unlike Media Lengua, the substitution

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

of lexical material does not affect placeholders and category 3 remains intact. That Cree-derived processing of propositional content is productive is seen in the extension of the obviative affix to French nouns (item 5). It is the French component that appears contradictory. We see a move toward fusion of operative procedures in the domain of clause combining, represented by item 12, which, however, appears to have come to a halt half-way, and resulted instead in internal grammaticalisation processes based on French-derived material. Item 11 is somewhat puzzling, for if we are to assume that French constituted the pragmatically dominant language, then a much clearer effect of fusion might be expected at the level of discourse management. Consider the following example, adapted from Bakker (1997, 5–6):

(23) *un vieux* ana ayi/ *un vieux* opahikêt
an old this uh an old trapper
ê-nôhcihcikêt, *you see*, êkwa
trapped and
ayi/ *un matin* êkwaniskât ahkosiw, *but*
uh one morning woke-up be-sick
kêyapit ana
still this-one
wî-nitawi-wâpahtam *ses pièges*.
want-go-see his traps
"An old this uh/ an old trapper was trapping, you see, and uh/ one morning he woke up sick, but he still wanted to go and see to his traps."

French connectives appear alongside both Cree-derived and English forms. The impact of French as a pragmatically dominant language appears to have been rather short-lived, an observation which seems consistent with the overall history of the language and its rather abrupt emergence during a short period of intense French–Cree contacts (see Bakker, 1997 for details). French was later replaced altogether by English as the pragmatically dominant language, which is reflected in (23) by its infiltration into directing or discourse-managing operations – the seeds of fusion.

The short period of French influence might be the key to understanding the present composition of Michif. What may have started out as a process of lexical re-orientation similar to the one observed in Media Lengua, encompassed in Michif only nouns. Bakker (1997) offers a structural explanation for this, suggesting that the bound nature of the Cree verb and the non-availability of isolated verbal roots excluded substitution of Cree verbal entities through French counterparts. In addition, though, the French input might also have been subjected to a natural hierarchy of functional prominence, with nouns – being the most stable conceptual representations in time and space – figuring at the very top, followed by adjectives and finally by verbs. This corresponds exactly to the state of affairs in Michif as regards the French element in lexical classes. The striking feature of Michif is that lexical re-orientation – however constrained by these structural and perhaps also functional factors – is accompanied by fusion of the noun phrase grammar.

A key to understanding both Media Lengua and Michif is therefore the motivation toward lexical re-orientation and the shift it entails of meaning representation into a second language. The process is coupled in both cases (albeit to different degrees) with other contact mechanisms, namely fusion (represented along its natural continuum from interaction management to epistemic qualification) and convergence (as a compromise strategy). In both cases, key processing operations pertaining to the predication remain in the native language – represented foremostly by verb inflection and clause-level word order rules. The extraordinary feature of both languages is the stabilisation of lexical re-orientation and its conventionalisation within a self-contained system. This invites a characterisation as a *reversed turnover of functions*: what might have started off as manipulated mixed utterances serving to flag speakers' orientation toward a colonial culture eventually gave rise to an all-purpose native-community idiom.

Conclusion

Two central methodological questions have been addressed in this paper. The first has to do with the way we interpret functions of linguistic material – words and grammatical formations – especially as regards the cognitive and processing operations which they trigger, and the implementation of such interpretation in an explanatory model of language change through contact. The second is about the power of the model presented here to actually predict the mechanisms involved in the emergence of mixed languages and their structural outcome.

Approaches in cognitive linguistics that rely on an interpretation of structures and their usage by speakers, rather than on experimental data, face a general problem of scientific replicability and so of persuasiveness: in trying to reconstruct the internal functions of linguistic elements, the processing operations that they trigger, and what goes on in the subconscious planning mechanism that is responsible for the production of utterances, we rely primarily on an interpretation of data and their position in discourse. Such is the case when discussing the discourse function of mixed utterances in secret languages, but

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

also in interpreting the relevance of monitoring-and-directing operations to the fusion hierarchy (as discussed in Matras, 1998). As with many attempts to explain, rather than just describe, occurrences in language, we risk speculating about the causes of processes. However, such speculations once formulated can indeed be put to the test. It has been claimed that discourse operators are positioned at the head of a chain of grammatical items which are subjected to different contact mechanisms than, say, lexical items, and that keeping an ancestral language partially alive for specific functions will target a different selection of structures than consciously re-orienting oneself toward a new or neighbouring language. These claims derive from observations on the behaviour of the relevant classes of elements in contact situations. They involve relating those observations to what we know about the function and structure of linguistic categories. The overall message is that we must draw on the trivial but crucial realisation that words are not simply words, but are divided into function-oriented categories. When items display a similar behaviour in contact situations, it is not only legitimate but necessary to search for their common denominator, however unapparent it may seem from a structural viewpoint. It is on this basis that the hierarchy of fusion has been postulated, and that an explanation has been suggested that relates the susceptibility of wholesale replacement to cognitive overload surrounding automaticised monitoring-and-directing operations.

My point of departure in this contribution was that mixed languages do not offer a uniform language type, but are outcomes of diverse mechanisms and combinations of mechanisms. It is therefore impossible in my view to make general predictions about the linguistic outcome of extreme cases of population mixtures, as has been attempted in the literature (e.g. Bakker and Mous, 1994b, 1995). But can we use the explanatory model outlined here to make alternative predictions? While I have argued above that mixed languages are neither uniform nor predictable, I have at the same time maintained that, given the types and combinations of contact mechanisms which give rise to them, their structural compositions are not entirely arbitrary either. Thus we can on the whole predict the course of fusion, which will target monitoring-and-directing operations in a universal order. We can also explain why certain types of contact constellations are likely to trigger fusion – those for instance, where full bilingual competence involving linguistic processing operations will lead to literal interference of one system with another. We can further predict the types of social–communicative circumstances in which lexical re-orientation might occur, or where selective replication is likely to be encountered. We can explain why lexical re-orientation will target items that express meaning, and leave the grammar component of the language largely intact. We can finally predict that where preference is given to maintaining the material autonomy of a particular language, contact might be reflected through the synchronisation of the triggers of grammatical operations with their counterparts in the other language, resulting in what I offered above as a narrow definition of convergence.

What we cannot predict, however, is the combination of processes, or the full extent to which they will affect the linguistic systems involved in them. We can only assume in hindsight that cases such as that of Malay passive morphology in Javindo and Petjo, of French noun phrase morphology in Michif, or of Russian finite verb inflection in Copper Island Aleut are the result of specific structural constraints on developments such as Dutch–Malay convergence in the first instance, lexical re-orientation towards French in the second, and a shift to a Russian-based predication and sentence production in the final. A catalogue of possible structural constraints, with more predictive power, is indeed a desirable endeavour, but will depend on future availability of more first-hand data on mixed language systems.

The fact that motivations for mixing are diverse helps explain why mixed languages are not uniform. Moreover, it is precisely the fact that mixed languages are products of such complex density of motivations, resulting in diverse contact mechanisms, that makes them both relatively rare and perhaps also short-lived. The argument line adopted here was mainly based on the assumption that an appreciation of the structural compartmentalisation in mixed languages must firstly take into account the natural function-based and cognitive compartmentalisation of linguistic structures. It is here that I attempted to venture beyond some of the more general postulations of a split between lexicon, grammar, and function words. In providing us with various clusterings of linguistic features and categories, mixed systems pose a special challenge to grammatical theory and models of grammar, even beyond the issue of contact: they offer opportunities to identify natural alliances among structures, allowing testing of our hypotheses regarding compartmentalisation in grammar and language processing functions.

References

Acton, T. A. (1971). The functions of the avoidance of Moxadi Kovels. *Journal of the Gypsy Lore Society*, 3rd series, 50, 108–136.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

Althaus, H.-P. (1963). Jüdisch-hessische Sprachbeziehungen. *Zeitschrift für Mundartforschung* 30, 104–156.

Auer, P. (1998a). Introduction: bilingual conversation revisited. In P. Auer (ed.), *Code-switching in conversation: languages, interaction and identity*, pp. 1–24. London: Routledge.

Auer, P. (1998b). *From code-switching via language mixing to fused lects: toward a dynamic typology of bilingual speech*. (Interaction and Linguistic Structures 6). Freiburg: Deutsches Seminar I, Universität Freiburg.

Bakker, P. (1994). Michif, the Cree–French mixed language of the Métis buffalo hunters in Canada. In P. Bakker & M. Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, pp. 13–33. Amsterdam: IFOTT.

Bakker, P. (1996). Language intertwining and convergence: typological aspects of genesis of mixed languages. *Sprachtypologie und Universalienforschung* 49 (1), 9–20.

Bakker, P. (1997). *A language of our own: the genesis of Michif – the mixed Cree–French language of the Canadian Métis*. New York: Oxford University Press.

Bakker, P. (1998). Pararomani languages and secret languages. In Y. Matras (ed.), *The Romani element in non-standard speech*, pp. 69–96. Wiesbaden: Harrassowitz.

Bakker, P. & Mous, M. (1994a) (eds.). *Mixed languages: 15 case studies in language intertwining*. Amsterdam: IFOTT.

Bakker, P. & Mous, M. (1994b). Introduction. In P. Bakker & M. Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, pp. 1–11. Amsterdam: IFOTT.

Bakker, P. & Muysken, P. (1995). Mixed languages and language intertwining. In J. Arends, P. Muysken & N. Smith (eds.), *Pidgins and creoles: an introduction*, pp. 41–52. Amsterdam: Benjamins.

Bakker, P. & Papen, R. A. (1997). Michif: a mixed language based on Cree and French. In S. G. Thomason (ed.), *Contact languages: a wider perspective*, pp. 295–363. Amsterdam: Benjamins.

Bakker, P. & Van der Voort, H. (1991). Para-Romani languages: an overview and some speculations on their genesis. In P. Bakker & M. Cortiade (eds.), *In the margin of Romani: Gypsy languages in contact*, pp. 16–44. Amsterdam: Institute for General Linguistics.

Bolonyai, A. (1998). In-between languages: language shift/maintenance in childhood bilingualism. *International Journal of Bilingualism* 2, 21–43.

Boretzky, N. (1998). Der Romani-Wortschatz in den Romani-Misch-Dialekten. In Y. Matras (ed.), *The Romani element in non-standard speech*, pp. 97–132. Wiesbaden: Harrassowitz.

Boretzky, N. & Igla, B. (1994). Romani mixed dialects. In P. Bakker & M. Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, pp. 35–68. Amsterdam: IFOTT.

Brünner, G. & Graefen, G. (1994) (eds.). *Texte und Diskurse: Methoden und Forschungsergebnisse der Funktionalen Pragmatik*. Opladen: Westdeutscher Verlag.

Bühler, K. (1934 [1982]. *Sprachtheorie*. Stuttgart: Fischer.

Bührig, K. & Matras, Y. (1999) (eds.). *Sprachtheorie und sprachliches Handeln: Festschrift für Jochen Rehbein zum 60. Geburtstag*. Tübingen: Stauffenburg.

Chyet, M. (1995). Neo-Aramaic and Kurdish: an interdisciplenary consideration of their influence on each other. In S. Izreel & R. Drory (eds.), *Israel Oriental Studies XV: language and culture in the Near East*, pp. 219–252. Leiden: Brill.

Clark, H. H. & Carlson, T. B. (1982). Hearers and speech acts. *Language* 58, 332–373.

de Gruiter, M. (1994). Javindo, a contact language in prewar Samarang. In P. Bakker & M. Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, pp. 151–159. Amsterdam: IFOTT.

de Rooij, V. (1996). *Cohesion through contrast: discourse structure in Shaba Swahili/French conversations*. Amsterdam: IFOTT.

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

Ehlich, K. (1986). Funktional–pragmatische Kommunikationsanalyse: Ziele und Verfahren. In W. Hartung (ed.), *Untersuchungen yur Kommunikation: Ergebnisse und Perspektiven*, pp. 15–40. Berlin: Akademie der Wissenschaften.

Golovko, E. V. (1994). Copper Island Aleut. In P. Bakker & M. Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, pp. 113–121. Amsterdam: IFOTT.

Grant, A. (1994). Shelta: the secret language of Irish Travellers viewed as a mixed language. In P. Bakker & M. Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, pp. 123–150. Amsterdam: IFOTT.

Grosvenor, Lady A. (1910). A pilgrim's progress. *Journal of the Gypsy Lore Society*, new series 3, 204–224.

Günther, L. (1915). Vorbemerkung (Preface), Wittich, Engelbert. Die jenische Sprache. *Archiv für Kriminal-Anthropologie und Kriminalistik* 63 1–22.

Hancock, I. (1970). Is Anglo-Romanes a creole? *Journal of the Gypsy Lore Society*, third series, 49, 41–44.

Hancock, I. (1984). Romani and Angloromani. In P. Trudgill (ed.), *Language in the British Isles*, pp. 367–383. Cambridge: Cambridge University Press.

Hancock, I. (1992). The social and linguistic development of Scandoromani. In E. H. Jahr (ed.), *Language contact: theoretical and empirical studies*, pp. 37–52. Berlin: Mouton de Gruyter.

Jake, J. L. (1998). Constructing interlanguage: building a composite matrix language. *Linguistics* 36, 333–382.

Kenrick, D. (1979). Romani English. In I. Hancock (ed.), *Romani sociolinguistics* (*International Journal of the Sociology of Language* 19), 111–120.

Klepsch, A. (1996). Das Lachoudische: eine jiddische Sondersprache in Franken. In K. Siewert (ed.), *Rotwelsch-Dialekte*, pp. 81–93. Wiesbaden: Harrassowitz.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

Leigh, K. (1998). Romani elements in present-day Caló. In Y. Matras (ed.), *The Romani element in non-standard speech*, pp. 243–282. Wiesbaden: Harrassowitz.

Lerch, H.-G. (1976). *Das Manische in Giessen. Die Geheimsprache einer gesellschaftlichen Randgruppe, ihre Geschichte und ihre soziologischen Hintergründe*. Giessen: Anabas-Verlag.

Liebich, R. (1863). *Die Zigeuner in ihrem Wesen und ihrer Sprache*. Leipzig: Brockhaus.

Maschler, Y. (1997). Emergent bilingual grammar: the case of contrast. *Journal of Pragmatics* 28, 279–313.

Maschler, Y. (1998). On the transition from code-switching to a mixed code. In P. Auer (ed.), *Code-switching in conversation: languages, interaction and identity*, pp. 125–149. London: Routledge.

Matras, Y. (1991). Zur Rekonstruktion des jüdischdeutschen Wortschates ehemaliger "Judendörfer" in Südwetdeutschland. *Zeitschrift für Dialektologie und Linguistik* 58, 268–293.

Matras, Y. (1998). Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36, 281–331.

Matras, Y. (forthcoming). Fusion and the cognitive basis for bilingual discourse markers. *International Journal of Bilingualism*.

Meeuwis, M. & Blommaert, J. (1998). A monolectal view of code-switching: layered code-switching among Zairians in Belgium. In P. Auer (ed.), *Code-switching in conversation: languages, interaction and identity*, pp. 76–98. London: Routledge.

Meisinger, O. (1901). Die hebräischen Fremdwörter der Rappenauer Mundart. *Zeitschrift für hochdeutsche Mundarten* 2, 73–75.

Meisinger, O. (1902). Lotekhôlisch. Ein Beitrag zur Kenntnis der fränkischen Händlersprache. *Zeitschrift für hochdeutsche Mundarten* 3, 121–127.

Moormann, J. (1920). "Louter Lekoris": Een levende geheimtaal. *Tijdschrift voor taal en letteren* 8, 235–239, 306–333.

Moormann, J. (1922). Losche Nekôdesch: Een Limburgsche geheimtaal. *Tijdschrift voor taal en letteren* 10, 26–43, 68–87.

Mous, M. (1994). Ma'a or Mbugu. In P. Bakker, P. & M. Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, pp. 175–200. Amsterdam: IFOTT.

Muysken P. (1981). Halfway between Quechua and Spanish: the case for relexification. In A. Highfield & A. Valdman (eds.), *Historicity and variation in creole studies*, pp. 52–78. Ann Arbor: Karoma.

Muysken P. (1997). Media Lengua. In S. G. Thomason (ed.), *Contact languages*, pp. 365–426. Amsterdam: John Benjamins.

Myers-Scotton, C. (1992). Codeswitching as a mechanism of deep borrowing, language shift, and language death. In M. Brenzinger (ed.), *Language death: factual and theoretical explorations with special reference to East Africa*, pp. 31–58. Berlin: Mouton de Gruyter.

Myers-Scotton, C. (1998). A way to dusty death: the Matrix Language turnover hypothesis. In L. A. Grenoble & L. J. Whaley (eds.), *Endangered languages: language loss and community response*, pp. 289–316. Cambridge: Cambridge University Press.

Oesch Serra, C. (1998). Discourse connectives in bilingual conversation: the case of an emerging Italian–French mixed code. In P. Auer (ed.), *Code-switching in conversation: language, interaction, and identity*, pp. 101–122. London: Routledge.

Post, R. (1992). Jüdisches Sprachgut in den pfälzischen und südhessischen Mundarten. In A. H. Kuby (ed.), *Pfälzisches Judentum gestern und heute*, pp. 177–258. Neustadt: Verlag Pfälzische Post.

Redder, A. & Rehbein, J. (1999). Zusammenhänge von Grammatik und mentalen Prozessen. In A. Redder & J. Rehbein (eds.), *Grammatik und mentale Prozesse*, pp. 1–12. Tübingen: Stauffenburg.

Rijkhof, J. (1998). Bystander deixis. In Y. Matras (ed.), *The Romani element in non-standard speech*, pp. 51–67. Wiesbaden: Harrassowitz.

Röll, W. (1986). Bestandteile des deutschen Gegenwartwortschatzes jiddischer oder hebräischer Herkunft. In W. Röll & H.-P. Bayerdörfer (eds.), *Auseinandersetzungen um jiddische Sprache und Literatur*, pp. 54–62. Tübingen: Niemeyer.

Rudolph, E. (1996). *Contrast*. Berlin: de Gruyter.

Salmons, J. (1990). Bilingual discourse marking: code switching, borrowing, and convergence in some German–American dialects. *Linguistics* 28, 453–480.

Sasse, H.-J. (1992). Theory of language death. In M. Brenzinger (ed.), *Language death: factual and theoretical explorations with special reference to East Africa*, pp. 7–30. Berlin: Mouton de Gruyter.

Siewert, K. (1996) (ed.). *Rotwelsch-Dialekte*. Wiesbaden: Harrassowitz.

Smart, B. C. (1862/63). *The dialect of the English Gypsies*. (Appendix to Transactions of the Philological Society for 1862–1863). Berlin Asher.

Smith, N. (1995). An annotated list of creoles, pidgins, and mixed languages. In J. Arends, P. Muysken & N. Smith (eds.), *Pidgins and creoles: an introduction*, pp. 331–374. Amsterdam: Benjamins.

Thomason, S. G. (1995). Language mixture: ordinary processes, extraordinary results. In C. Silva-Corvalán (ed.), *Spanish in four continents: studies in language contact and bilingualism*, pp. 15–33. Washington, DC: Georgetown University Press.

Thomason, S. G. (1997a) (ed.). *Contact languages*. Amsterdam: John Benjamins.

Thomason, S. G. (1997b). Introduction. In S. G. Thomason (ed.), *Contact languages*, pp. 1–7. Amsterdam: Benjamins.

Thomason, S. G. (1997c). Ma'a (Mbugu). In S. G. Thomason (ed.), *Contact languages*, pp. 469–487. Amsterdam: Benjamins.

Thomason, S. G. (1997d). Mednyj Aleut. In S. G. Thomason (ed.), *Contact languages*, pp. 449–468. Amsterdam: Benjamins.

Thomason, S. G. (1997e). A typology of contact languages. In A. Spears & W. Donald (eds.), *The structure and*

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

status of pidgins and creoles, pp. 71–88. Amsterdam: Benjamins.

Thomason, S. G. & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

Vakhtin, N. (1998). Copper Island Aleut: a case of languages "resurrection". In L. A. Grenoble & L. J. Whaley (eds.), *Endangered languages: language loss and community response*, pp. 317–327. Cambridge: Cambridge University Press.

van Rheeden, H. (1994). Petjo: the mixed language of the Indos in Batavia. In P. Bakker & M. Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, pp. 223–237. Amsterdam: IFOTT.

Wagner, M. L. (1937). Stray notes on Spanish Romani, chapter 2: Cryptolalic formations in other Romani dialects. *Journal of the Gypsy Lore Society*, third series, 16, 27–32.

Weinberg, W. (1969). *Die Reste des Jüdischdeutschen*. Stuttgart: Kohlkammer.

Wexler, P. (1988). *Three heirs to a Judean–Latin legacy: Judeo-Ibero-Romance, Yiddish, and Rotwelsch*. Wiesbaden: Harrassowitz.

Received August 31, 1999 Revision accepted December 15, 1999

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

PEER COMMENTARIES

Mixing is functional (but what are the functions?)

PETER AUER
Universität Freiburg, Deutsches Seminar I, Institut für deutsche Sprache und ältere Literatur, Werthmannplatz 3, 79085 Freiburg, Germany
E-mail: auerp@sun2.ruf.uni-freiburg.de

Yaron Matras' paper makes an important contribution to research on so-called mixed languages by drawing our attention to the diversity of the varieties subsumed under this heading, both in sociolinguistic and in grammatical-(–functional) terms. It provides a healthy counterpart to the sweeping generalisations occasionally found in the literature on mixed languages. One of these is that in mixed languages, one of the ancestral languages provides the grammar and the other the lexicon (Bakker 1997 etc.). Matras argues very convincingly that such a view is not valid and cannot be upheld; it does not even apply to the most frequently mentioned (prototypical?) mixed languages such as Michif or Media Lengua (both of which include a considerable amout of grammatical material from the language which also acts as the lexifier). Instead mixed languages arise out of very different needs and in very different sociolinguistic contexts, usually in order to display a certain social affiliation with a given group (or disaffiliation with another, usually main-stream or majority group); these different needs and contexts then result in very heterogeneous mechanisms of language contact. I could not agree more with this point of view, and I will not go into the details here which would only repeat what Matras puts forward very convincingly. Instead I will focus on those aspects of this very rich paper in which his claims raise some questions or are even (in my opinion) contestable.

A preliminary note on terminology and on the extensional definition of "mixed languages" is indispensable. The term is surely somewhat ambiguous if only for the fact that many researchers on bilingualism use "mixing" in a much wider sense in order to describe a way of speaking which includes (frequent) alternation between two languages or varieties (as identified by the linguist) beyond discursive code-switching. If the term *mixed language* is used it should be kept in mind that it refers to more than (code-)*mixing* in this latter sense. In Auer (1999) I argue, citing Matras' earlier work, that (code-)mixing may result in "fusion" (roughly corresponding to the (self-contained) mixed languages of the literature) as soon as it becomes obligatory, i.e., as soon as it is no longer an option for the speaker. In other words, fusion requires some kind of structural sedimentation. Putting it more shortly, one might say that a fused variety (or: a self-contained "mixed language") is a variety in which language alternation has intruded into the grammatical system.

Given the need to separate code-mixing as a stylistic device from fusion/mixed languages (albeit on a continuum), it may be asked if all of the examples of "mixed languages" in Matras' paper satisfy the requirement of structural sedimentation. The question of how variable or obligatory the contact elements in a "mixed language" are, is unfortunately rarely addressed in the literature; however, it is decisive for delimiting it against the code-mixing styles that have been reported from so many bilingual communities. As a consequence, the empirical basis for determining structural sedimentation is sometimes rather weak. My impression is nonetheless that some of Matras' examples are not "mixed languages" in this sense; particularly in the case of the Lekoudesch, Jenisch and Para-Romani, data he mentions (i.e., in his examples for "lexical re-orientation" and "selective replication"), stylistic mixing instead of grammatical fusion seems to be at stake. In fact the very notion of what he calls "non-self contained mixtures" and others "symbiotic mixed languages" questions the issue of structural sedimentation, and *eo ipso* the description as sedimentated "languages" or varieties.[1] Arguably, the description of the Jenisch examples (9) and (10) as the result of a "special lexicon [which] is inserted only sporadically, consciously, and deliberately in order to shift meaning, and so settings, from a serious, straightforward, or factual interaction to tone that challenges social and conversational taboos", suggests that we are dealing with a stylistic, not a grammatical type of language alternation, which might better be called code-mixing, or, in this particular case, even (discourse-functional) code-switching. The key to an understanding of these ways of speaking is given by Matras himself when he describes Michif and Media Lengua as examples of "the stabilization of lexical re-orientation and its conventionalisation within a self-contained system ... what might have started off as manipulated mixed utterances serving to flag speakers' orientation toward a colonial culture eventually gave rise to an all-purpose native community idiom". This is also the analysis compatible with my mixing fusion continuum (as suggested in Auer, 1999).

The more far-reaching claim made by Matras is that his four types of contact (lexical re-orientation, selective replication, convergence and fusion) can be explained by their differential targeting of five "functional compartmentalisa-

[1] I do not see any problem in calling a style or a variety "mixed" or "fused" even though the type of alternation it displays is clearly unbalanced; i.e. when it is perfectly possible to identify a "matrix language" which provides the grammatical skeleton into which elements of the other language are inserted. This is frequently the case in code-mixing styles, and it seems to be by far the prevalent if not the only pattern of "fusion". And whether what we linguists are dealing with is "a language" or not is not to be decided by us anyway: rather, it is exclusively a matter of attitudes brought to bear by the speakers on their way of speaking.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

tions'' of language, i.e.: (1) meaning-encoding in content elements (lexical nouns, verbs, adjectives, etc.), (2) propositional encoding, (3) operative processing of propositional elements, (4) monitoring and directing in discourse and (5) deixis (in a comprehensive sense of the word). This ''functional compartmentalisation'' goes back to an attempt by K. Ehlich (1986, 33ff.) to elaborate and extend Bühler's distinction of the ''symbolic'' (*Symbolfeld*) and the ''deictic'' (*Zeigfeld*) field of language. Since the theory is not spelled out in full, it is hard to apply it to any field, including that of language contact.[2]

Matras suggests the following correspondences and thus functional explanations: *lexical re-orientation* is basically an attempt to conceal meaning from co-participants; it can therefore be expected to target the functional domains 1 and 2 (concepts plus propositional operators such as negation). This is indeed the case of various ''secret codes'' such as Jenisch or Lekoudesch. Of course, lexical borrowing is not restricted to secret codes but occurs in many (other) bilingual contexts as well, although for other reasons. (The need to name new objects – as in ''cultural loans'' – is only one of them, and not at all able to explain any lexical borrowing in bilingual communities.) In order to capture lexical re-orientation, a sociolinguistic trait is most distinctive: only one of the ''ancestral languages'' (i.e., in this case, German) is available at all and used monolingually, i.e., lexical re-orientation is not linked to bilingualism.

The process of *selective replication* has roughly the same function. The main difference according to Matras is that it occurs in a situation of language death in which a given generation of speakers has shifted to a new language of interaction (for instance, from Romani to English), but continues to use the language of the preceding generation for ''symbolic and emblematic usage'', i.e., both ''ancestral languages'' are spoken or have been spoken as monolingual varieties. Since one of the main functions of this archaic language is secretive, the same functional domains of language should be affected as in lexical re-orientation, but Matras also includes the functional domain 5 (deixis). The secretive function of selective replication nicely explains why lexical material from the dying language is retained (setting Para-Romani aside from other sociolinguistic contexts of language death). On the other hand, it is not clear how the retention of Romani components in the deictic field can be accounted for, given the closely related or even identical social–communicative functions of Para-Romani and ''secret codes'' such as Jenisch. In addition, Matras' example for Para-Romani (*Giessen Manisch*) includes non-deictic (and non-propositional) Romani elements as well (such as interjections, which are otherwise subsumed under the ''monitoring and directing''-domain (4)).[3] An alternative, and possibly more straightforward, explanation of ''selective replication'' suggests itself, i.e., that the remnants of inflectional morphology are due to the retention of more or less frozen code-switched utterances from a previous stage of bilingualism.

Convergence is the contact mechanism I found most difficult to get a precise idea of in Matras' model. It should target (in addition to conceptual representation), (1) the domain of ''processing of propositional contents'' (3), and is contrasted with *fusion*, which starts in the domain of ''monitoring and directing'' (4) but may affect other domains (particularly 3) as well. Fusion will therefore, in its final stage, represent the most intimate ''intertwining'' of linguistic systems. According to Matras, it is motivated by reducing the processing costs of dealing with two linguistic systems simultaneously, a claim which needs more space to be discussed than is available here.[4] To me, the critical question is whether convergence and fusion can be separated at all, or if one should not rather refer to only one mechanism of (''fusional'') language contact for which the first stages can be located in Matras' domains 1 and 4. This seems all the more plausible since it seems to be difficult to find instances of pure convergence, i.e., a type of language contact exclusively targeting domain 3 (in addition to omnipresent domain 1); all examples discussed by Matras involve at least fusion in addition to convergence, sometimes also lexical re-orientation.

This, of course, would seriously weaken the functional explanation based on the Ehlich model since we would end up with only two functionally based contact types, one of them defined negatively: non-self-contained varieties affecting the encoding of concepts and propositional operators (and possibly some deictic elements) could be singled out by their secretive function, while all the other candidates for ''mixed languages'' could affect all of Matras' functional domains of language, although very likely not in random order. In fact, the most convincing merits of Matras' paper may be that it gives a clear and functionally as well as socially motivated analysis of so-called non-self-contained mixed languages, showing that in many respects, they do not deserve this label at all, given the profound differences which separate them from self-contained mixed languages.

[2] One of the many questions Matras' short sketch raises is the inclusion of quantifiers and basic numerals in the deictic domain.

[3] Matras also mentions English Romani in connection with this, but presumably not as a case of ''selective replication''. The examples, particularly (17), suggest a more intimate intertwining of English and Romani; note, in particular, the double marking of the future tense in this hybrid construction, typical of ''fusion'' (in my and presumably also Matras' meaning).

[4] It should be mentioned, however, that while the ''processing'' explanation has some plausibility in the case of fillers, discourse particles and connectives, it seems difficult to apply to modality, subjunctive or conditional markers, which are also included under ''monitoring and directing''.

References

Auer, P. (1999). From code-switching via language mixing to fused lects: toward a dynamic typology of bilingual speech. *International Journal of Bilingualism* 3 (4), 309–332.

Bakker, P. (1995). Language intertwining and convergence: typological aspects of the genesis of mixed languages. *STUF* 49, 9–20.

Bakker, Peter (1997). A language of our own: the genesis of Michif, the mixed Cree – French language of the Canadian Métis. New York/Oxford: Oxford University Press.

Ehlich, K. (1986). Funktional–pragmatische Kommunkationsanalyse: Ziele und Verfahren. In W. Hartung (ed.), *Untersuchungen zur Kommunikation: Ergebnisse und Perspektiven*, pp. 15–40. Berlin: Akademie der Wissenschaften.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

Insertional codeswitching in an immigrant language: "just" borrowing or lexical re-orientation?

AD BACKUS
Babylon-Studies of Multilingualism in the Multicultural Society, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands
E-mail *A.M.Backus@kub.nl*

In this paper I will explore the usefulness of Yaron Matras' typology of contact mechanisms for the type of contact setting I am most familiar with myself: that of a typical immigrant language, in my case the variety of Turkish that is spoken in the Netherlands.

Of the mixed languages (MLs) discussed by Matras, Michif and Media Lengua look most like Dutch Turkish. However, even with these languages there are big differences. Does that mean that Dutch Turkish is not an ML? The answer is probably that indeed it is not, but a comparison is still likely to uncover interesting similarities. This is because all MLs arise out of a contact setting, as did Dutch Turkish. The mechanisms discussed by Matras are presumably typical of language contact in general, not idiosyncratic features of the languages he happens to discuss. That means that the same contact mechanisms are likely to be found in Dutch Turkish, just less frequently so. As Matras says, it's the density that makes the difference.

One of the most salient consequences of language contact is the phenomenon of insertional codeswitching (CS), of which we speak when a clause in one language, the base or matrix language, contains one or more elements, normally content words, from the other language (referred to as the guest or embedded language). Matras does refer to CS as relevant to the genesis of MLs, but has the type in mind that involves foreign discourse markers. CS data indeed often contain bilingual utterances in which the propositional content is all in one of the languages, but the discourse markers and utterance modifiers which frame this content are predominantly from the other language. It is this type of CS which often leads to what Matras refers to as a case of fusion.

In this commentary, I would like to explore the relevance of the more common type of CS, insertional CS, to the genesis of MLs. I must say at the outset that I see insertional CS as the synchronic equivalent of lexical borrowing, which I define in diachronic terms. Matras explicitly contrasts lexical borrowing with lexical re-orientation, and concludes that the former differs from the latter because it mainly concerns new cultural concepts (that is: lexical addition) and never leads to wholesale replacement of the matrix language lexicon. As is well known, some MLs are characterized by wholesale lexical borrowing, and the subsequent lexicon/grammar split that is so typical of MLs as a category. In what follows, I will present some results of an empirical investigation into the degree of vocabulary replacement in the immigrant variety of Turkish spoken in Holland (cf. Backus, 1992, and Backus, 1996, for a detailed presentation of the CS practised in this community).

On first inspection, MLs and insertional CS seem to have a lot in common. In both cases, the grammatical frame of one language hosts lexical material from another. An often-held notion is that MLs, such as Michif, must be fossilized CS (Myers-Scotton, 1998, 314). On the other hand, those who study MLs claim quite consistently that, although CS may have something to do with the genesis of MLs, it cannot be the mechanism that directly leads to it. The main reason for this disbelief is the total nature of the mixture in most documented MLs. In Michif, for example, 90% of the content words are from French, the presumed Embedded Language in insertional CS with Cree as the matrix language, but CS data never show such near-complete replacement of the stock of content words.

Even though individual examples of CS and sample sentences from Michif or Media Lengua do look very similar, there are some features that are disturbingly different. Examples (1a) and (1b) have a lot in common. In Media Lengua, all content words are Spanish. Similarly, in the Turkish–Dutch example in (1b), all content words are Dutch, embedded in a Turkish grammatical frame. However, the example in (1c) is more typical of Turkish–Dutch CS than (1b). Just one of the content words in the clause is Dutch. The density of CS varies from conversation to conversation and from informant to informant, but it is never as spectacular as figures for the replacement of content word vocabulary reached by MLs.

(1) a. *todabia no byen aprendi*-naku-n *porke*
still not well learn-pl-3 because
eskwela-bi *anda*-naku-n
school-LOC go-pl-3
(Media Lengua; Muysken, 1997, 401)
"They *do not learn well yet because* they *go to school*"

b. şöyle *hoek*-li, *schuin vorm*-lu
such corner-ADJ slanted shape-ADJ
(Turkish–Dutch CS; Backus, 1996, 102)
"with a *corner* like this, with a *slanted shape*"

c. ben *plaat*-lar-ı ne bileyim?
I record-pl-POSS what know-OPT-1sg
(Turkish–Dutch CS; Backus, 1996, 103)
"what do I know about his *record*s?"

In addition to this difference regarding the rate of vocabulary replacement, there are two more differences which space does not permit discussing here. One lies in the influence from the embedded language on the matrix language. Whenever CS becomes very intense, structural influence from the language that is supplying the content

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

words follows (cf., for instance, Hill and Hill, 1986; Johanson, 1992; Silva-Corvalan, 1994), while MLs show little structural interference. A comparison of MLs and CS lects also unveils some differences in their sociolinguistic characteristics.

In the light of all this, is the hypothesis that MLs constitute fossilized CS still tenable? In and of themselves, none of these differences invalidate the hypothesis; they do, however, call for evidence which, to the best of my knowledge, has not been brought forward so far. For vocabulary replacement, my topic here and the only one I've done the required work on, at least the following demonstration is needed for the hypothesis to be accepted: CS lects should show progressive vocabulary replacement in which the gap between dense CS and the 90% figure of some MLs is being closed.

Elsewhere (Backus, forthcoming), I have given a detailed analysis of this research question. Though I can say outright that the required demonstration could not be provided, studying the progress of vocabulary replacement in a CS lect might be worth our while anyway, as it can give us clues as to how likely it is that it could ever proceed further towards that magical 90%, and thus whether it should be included as an important contact mechanism in its own right.

Let's say there is a putative ML *in statu nascendi*, called Dutch Turkish. All Dutch elements occurring in the Turkish of bilingual Turks in Holland must be seen as belonging to its lexicon. A first question is what kind of words these are. Roughly half of the Dutch insertions belong to expected semantic domains, such as education, bureaucracy and Dutch social life. A second question is how pervasive Dutch lexical items are in bilingual speech. A detailed analysis of the fragment of conversation in my data with the highest density of insertional CS, a discussion about a hospital training program that one of the informants had been attending not long before the recording was made, yielded results that are nowhere near what we are accustomed to from MLs. Of all the semantic domains investigated in the data, it is here that vocabulary replacement has proceeded the furthest; if Turkish–Dutch CS does not look like an ML in this fragment, than it won't resemble one anywhere. It turns out that the Turkish clauses in this fragment still contain twice as many Turkish content words than Dutch ones (31 versus 15). However, all but one of the hospital-related words are Dutch, while most of the 31 Turkish content words belong to basic vocabulary. Most have general meaning; examples include words for "go on holiday", "fall in love", "girl", "month", "to go", etc.

This gives us no evidence for large-scale vocabulary replacement in Immigrant Turkish in general. There just seems to be a large influx of Dutch vocabulary in certain semantic fields, which means that we are talking about vocabulary addition rather than about vocabulary replacement. Most of the Dutch words have highly specific semantics, with strong connotations of life in Holland, and thus do not have adequate equivalents in Turkish. Basic vocabulary remains overwhelmingly Turkish in Turkish clauses (the reader should note that all fragments of bilingual speech also contain many Dutch clauses, since alternational CS, the regular alternation of sentences and clauses in the two languages, is at least as frequent as insertional CS). Bilinguals seem to follow a maxim of Bilingual Economy that says that, when engaging in insertional CS, you only take from the other language what you need.

Application of this maxim leads to a mixed lect that is not an ML because everything that the language already possesses in the way of useful vocabulary will remain. There is borrowing of specific vocabulary, but there is no good reason for borrowing basic vocabulary. Incidental gaps notwithstanding, the base language already has perfectly good words for the concepts encoded by such basic vocabulary. Thus, an immigrant to the United States is more likely to import the word *highschool* than the word *school*, because the former has a narrower, more specific, referent. If no basic vocabulary gets borrowed, the 90%-or-so replacement figure can never be reached.

Basic vocabulary belongs to the clausal frame in much the same way as suffixes, word order and other morphosyntactic features do. This makes across-the-board vocabulary replacement of the sort MLs seem to have undergone, and as is presumed by the hypothesis that MLs are cases of fossilized CS, unlikely.

In conclusion, we can say the following:

- Turkish–Dutch CS, and presumably CS in general, is characterized by heavy addition of new vocabulary, as in heavy borrowing, not by the addition of a second set of words for familiar concepts, as in lexical re-orientation.
- Fossilization of the present state of Turkish–Dutch CS would lead to a language quite unlike documented MLs, with much alternation between the two contributing source languages at sentence and clause levels, as well as to retention of much Turkish lexical material.
- Codeswitchers have widespread experience with code alternation. Though this is not found in data from MLs, it does keep the door open for balanced bilingualism and subsequent "language planning" (e.g. creation of an ML through lexical re-orientation and/or Selective Replication).
- Purely structurally speaking, CS lects and MLs look similar, but the match is far from perfect. CS is more erratic, with both lexicons open for use, while MLs have a relatively fixed, or crystallized, lexicon. MLs have a fairly defined distribution of material from the two languages, while CS lects show up irregular numbers of insertions, switched constituents, and shifts of matrix language.

These differences seem serious enough to justify Matras' choice not to include insertional CS among the contact mechanisms that play a role in the genesis of MLs. Whatever role it does play can be subsumed under the heading of lexical re-orientation, although it differs from typical cases of that phenomenon in one crucial aspect: insertional CS

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

does not normally have the conscious functional motivations that typify lexical re-orientation, except in isolated cases when learned words from the Embedded Language are used because that language signifies modernity. This relative lack of functional motivations is probably exactly what keeps mainstream immigrant populations from actively upgrading their lexical borrowing from a case of fairly superficial insertional CS and vocabulary addition to one of creating a new, mixed, possibly secret, language, built through the mechanism of wholesale vocabulary replacement.

References

Backus, A. (1992). *Patterns of language mixing: a study in Turkish–Dutch bilingualism.* Wiesbaden: Otto Harrassowitz.

Backus, A. (1996). *Two in one: bilingual speech of Turkish immigrants in the Netherlands.* Tilburg: Tilburg University Press.

Backus, A. (forthcoming) The link that is missing between codeswitching and mixed languages: vocabulary addition in Dutch Turkish. In P. Bakker & F. Field (eds.), *SPRAU: Working Papers of the University of Aarhus (Proceedings of the Workshop on Mixed Languages).*

Hill, J. & Hill, K. (1986). *Speaking Mexicano: dynamics of syncretic language in Central Mexico.* Tucson: University of Arizona Press.

Johanson, L. (1992). *Strukturelle Faktoren in Türkischen Sprachkontakten.* Stuttgart: Franz Steiner Verlag.

Myers-Scotton, C. (1998). A way to dusty death: the Matrix Language Turnover Hypothesis. In L. Grenoble & L. Whaley (eds.), *Endangered languages: language loss and community response*, pp. 289–316. Cambridge: Cambridge University Press.

Muysken, P. (1997). Media Lengua. In S. Thomason (ed.), *Contact languages: a wider perspective*, pp. 345–426. Amsterdam: Benjamins.

Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles.* Oxford: Clarendon.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

 Bilingualism: Language and Cognition 3 (2), 2000, 106–108 © 2000 Cambridge University Press

Social and communicative approaches to mixed languages

PETER BAKKER
Department of Linguistics, Aarhus University, Jens Chr. Skous Vej 7, 8000 Aarhus C, Denmark
E-mail: linpb@hum.au.dk

Yaron Matras' paper is a significant contribution to the study of mixed languages. Its innovative points are the attempt to compartmentalise the process of their genesis into a number of separate components, and its function-based orientation embedded in a communicative orientation. However, I do not think that this compartmentalisation explains more than earlier theories.

The paper's abstract style and the small selection of sample languages and sentences, coupled with somewhat idiosyncratic terminology which is not always followed throughout the paper, hide some of its inconsistencies and weaknesses.

I will first discuss some general problems with Matras' four processes. Then I sketch an alternative approach. After that I discuss serious problems in applying the compartmentalisation to one of the cases, Michif. In an additional case study, I show better (but admittedly less elaborate) predictive power for the intertwining model.

When Matras mentions descriptivists and structuralists, I should probably include myself among those. Matras' source of inspiration is his recourse to a communicative theory of the genesis of mixed languages. The ideas that I have supported or developed are theories of genesis as well, albeit that the explanatory motives given by me were social rather than communicative. The communicative approach cannot replace the social approach, at most it can complement it.

Of the four processes invoked by Matras two affect the lexicon almost exclusively: lexical reorientation and selective replication. These two processes affect languages in different ways, but the results can be superficially similar: a dichotomy of content words and grammatical system. In both cases the innovations with regard to the source languages are often overwhelming. There are clear communicative motives for their emergence, based on the functions of the new languages (identity marker versus secrecy). They explain many differences (including some structural ones) between languages which have been called "mixed". The concepts build on earlier studies, using new terminology. I think these are useful concepts for secret languages like Lekoudesch, Yenisch and Manisch.

But does "selective replication" also apply to the other cases discussed as resulting from this process? How "selective" should the vocabulary be? In the three best documented varieties of Para-Romani (Scandinavia, England, Spain) roughly an almost equal number of Romani roots are attested as in regular Romani dialects such as Bugurdži. Boretzky (1998) has calculated that out of 750 inherited Romani words, Kalderash Romani and Bugurddži Romani ("normal" Romani) had preserved 600 and 570 respectively, and the Para-Romani varieties of Spain, England and Scandinavia 580, 520 and 500 respectively. One can hardly call that "selective".

The other two processes, convergence and fusion, affect mostly grammar and pragmatic elements. In many of the cases these are just additive processes, combined with others. There are indeed cases of across-the-board convergence (some discussed in Bakker 1996, 2000a) yielding mixed languages, but I do not think this is the case for Petjo or Javindo. The presence of passive morphology in Petjo can easily be explained in the intertwining model, since it is one of the few morphological operations in the source variety of Malay, and syntactically it behaves as expected (see below).

The term "fusion" relates to a process which does not affect languages so radically, but still visibly. Discourse markers and conjunctions tend to be in the pragmatically dominant language, either temporarily or permanently, through fusion. Fusion therefore mostly involves discourse markers, interjections and word order. It is the most problematic of all. Even though it is said to concern mostly discourse markers, it is not understandable why it should be the "extreme" (main?) mechanism at work in Copper Island Aleut, with its intricate mixture on many levels. Also, in Michif, fusion departs from Matras' definition.

Despite Matras' argumentation, I would still maintain that all of the languages discussed here, except two, show a basic dichotomy of grammar (phonology, morphology, syntax) versus lexicon. We find variation almost exclusively in the free grammatical morphemes. The intertwining model is associated with new ethnic groups that are (often) the result of massive marriages of men from one group and women from another, and also with certain languages such as Para-Romani (but not necessarily secret languages such as Yenisch, Manisch and Lekoudesch). There are two structural exceptions to this pattern: Michif and Copper Island Aleut. I have given typological arguments for the deviant nature of Michif (Bakker, 1997), but I have no solution for Copper Island Aleut – but neither has anyone else, including Matras.

The intertwining model has been sketched in Bakker (1997, Chapter 7) and is elaborated in Bakker (2000b), where a much broader database is investigated of more than two dozen languages. Sociohistorical, diachronic and psycho-linguistic arguments are presented there, which I cannot repeat here. I see no reason to revise my view that what I call intertwined languages (which does not include all mixed languages) constitute a single type – albeit with some variation, of course. It is in the free grammatical morphemes (pronouns, copulas, demonstratives, negative markers, etc.), and only there, that we find both intralinguistic and interlinguistic variation, as also alluded to in Matras' paper.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

The assignment of a separate communicative function to negation and existential verbs, just because these appear to behave unexpectedly in some languages, is not justifiable. In most intertwined languages there is variation in the source language of negators and copulas, as predicted. Furthermore, equation of existential verbs with copulas, hinted at here and there in the text, does not appear to fit the data: in Matras' Lekoudesch and Angloromani data the locative function (cross-linguistically closely related to existential) is expressed in the respective lexifier languages (Hebrew, Romani), whereas the equative and identificational functions are from the respective grammaticiser languages (German, English). This is in line with other cases of mixed languages (e.g. Michif), and should be interpreted as a warning against categorial generalisations such as "copula". On the other hand, it confirms Matras' statement that the most lexical or deictic category of a set of meanings (here those expressed like copulas), will be the one from the lexicon language.

I would like to take issue with the communicative approach by focussing on the language I know best: Michif. Matras' hypothesis of its genesis is based on an assumption of the short-lived influence of French, referring to my work. Both the reference and the claim are incorrect. It is actually much easier to find French speakers among the Michif speakers than it is to find Cree speakers. Of the many dozens of Michif speakers I have worked with in Canada and the United States, only one could speak Cree, and perhaps one in three could speak French. Hence, French has lasted longer than Cree, and Matras' explanation for the compartmentalisation is based on a wrong assumption and is therefore invalid.

Matras believes that Michif is a case in which three of his four processes are combined: it was a case of lexical reorientation – apparently of Cree towards French – combined with "fusion of the noun phrase grammar" and convergence (unspecified where). Lexical reorientation is supposed to be the transfer of meaning, i.e. content elements, for the purpose of shifting meaning to a different and culture-specific system. This is impossible to reconcile with the facts of Michif. Lots of concepts relating to Cree culture can only be expressed in French (e.g. tipi, local food resources, plants and animals, sun dance, bow and arrow, kinnikinnik, etc.), whereas other, typically European, concepts can only be expressed in Cree (e.g. to shoot with a gun, to sell, to drive a car, to plough, to arrest). Furthermore, there would be no conceivable communicative reason to limit such a communication-based process process to nouns – Matras invokes the often observed fact that nouns are more prone to borrowing than adjectives and verbs, but this is irrelevant: what has happened here has nothing to do with lexical borrowing in "normal" situations of contact.

When Matras refers to fusion of noun phrase grammar it is not clear what he means. Perhaps he alludes to the fact that noun phrase word order is purely French (Métis French, that is), including the inherited variable order of French adjectives, rather than Cree as should be expected on the basis of his (somewhat inaccurate) item 2) regarding Michif. He offers no explanation for this. The little Cree influence on the noun phrase predates the genesis of Michif (Bakker 1997).

Further, the source languages of coordinating conjunctions contradict a scenario in which French elements replace Cree ones: "but" is never from French (but rather Cree and indeed English) and "and" is more often from French than from Cree whereas "or" is always from French. Assuming that Cree is the "L1", the "native language" or "inherited", this contradicts Matras' own (1998) universals of borrowing. Here his "fusion" concept does not apply, and reversal of Cree and French as L1 and L2 does not help here (see also below).

Fusion is also invoked as influencing the "continuum from interaction-management to epistemic qualification" in the case of Michif, but here again a concrete example is lacking. It appears that in Michif almost all of the interjections (equivalent to "Wow!", "Ah!", and the like) derive from Cree, only some swear words and interjections ("look!", "Oh my God!") come from French.

If we look at the deictic elements in Michif, it appears that some locative adverbs are from Cree (here, there, gone, far away) and some from both French and Cree (everywhere, outside, inside). Time adverbs are from French (yesterday, last night), from French and Cree (tomorrow, now), or from Cree (long time ago, at the same time, today, soon, by and by). Other types of adverbs show a similar, seemingly arbitrary distribution of French and Cree elements. No communicative motivation plays a role here. Matras' account fails.

If I may summarise Matras' view in Table 1, it is easy to see that it does not work for Michif. Content elements, relating to meaning (a), are about equally derived from the two languages. Communicative category (b), I have argued above, is an arbitrary and ad-hoc category in my view. Communicative goal (c) again shows both Cree and French linguistic elements, and there is no communicative reason for having noun phrases (with zero-inflection?) and NP word order from one language but sentential order and verb inflection from another. Linguistic elements relating to monitoring (d) are Cree rather than French, so the predicted fusion is rather shallow, and the French and Cree conjunctions even contradict Matras. In the final category (e), a communicative approach cannot explain why deictic elements and quantifiers are from both languages, why some anaphora are from Cree and why possessives are mostly from French – Cree nouns have Cree possessive markers. In short, Matras' approach is inconsistent with the Michif facts in all of his categories.

Here Matras falls into the trap of assuming that all languages, including mixed languages, have a "base" into which other items can be introduced (borrowed, or taken over by convergence in the sense of Matras). This is something I would like to take issue with, since I think it is the basis of some unnecessary and factually incorrect claims.

Concerning Michif, for instance, it is impossible to speak of L1 and L2, or a "native" language versus an "intruding" language. These are misleading terms in cases where fluent bilinguals, with social motives, attempt to

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

Table 1. *Matras' communicative approach*

communicative goal	*linguistic means*	*affected part of language*	*relevant contact processes*
(a) conveyance of meaning	content elements	lexicon	lexical orientation, selective replication
(b) value of proposition	some grammatical elements	negation, locative/existential copulas	lexical reorientation
(c) formulation and processing of proposition	language-internal relations	inflection, word order	convergence, fusion
(d) monitoring and directing	relations between speakers	discourse markers, interjections, focus markers	fusion
(e) convey speech-situation-bound information	communicatively salient elements	deictics, anaphora, possessives, quantifiers, motion verbs	selective replication

create and speak a new language based on the available resources: their mother tongues. It has nothing to do with cultural or communicative motives, nor with a colonial situation.

The communicative approach does not do more than make the simple observation that different processes are at work. It does not predict how these are combined under what circumstances (except for the two lexical processes when they are exclusive processes). The communicative goals do not affect the connected language categories any more or less than the others. A "half-way" explanation makes no sense either.

Does the model fare better with the other languages? I have no space here to discuss those in detail, but I do not think it does. Matras' explanations for the distribution of elements in the other languages often seem ad hoc and post hoc. Furthermore, for all non-secret languages Matras mentions remaining puzzles and problematic elements.

In short, I do not think that the communicative approach is superior to the social, structuralist or intertwining approach. Compartmentalisation is not able to predict more, or more accurately, the mixed languages facts than the social approach. As an example we can take the language of the Peranakan Chinese of Eastern Java, which has not figured in the recent literature on mixed languages, but which fits the predictions of the intertwining model exactly. The Peranakans descend from Malay-speaking Chinese traders and native Javanese women. The verbal and nominal roots are from Malay/Indonesian (90%), the fathers' language, and the morphology (and presumably word order and phonology) from Javanese (88%), the mothers' language. Function words are distributed equally from the two languages (53% Javanese), and so are the adjectives and adverbs (together 46% Javanese; all percentages from Dreyfuss and Oka, 1979, 251). Nevertheless, this example also shows a weakness of the intertwining approach: there is currently no explanation for the source languages of the individual function words. Probably here Matras' processes and communicative motivations can help in providing an answer.

The compartmentalisation is at the same time a strong and a weak point of the approach. It is strong, because complex processes can be explained. It is weak, because it can explain anything. There is hardly any predictive power in the crucial domains. More seriously, it is not compatible with the facts of the non-secret languages.

References

Bakker, Peter. (1996). Language intertwining and convergence: typological aspects of the genesis of mixed languages. *Sprachkontakt und Grammatikalisierung*, Nicole Nau & Martin Haase (eds.). Special Issue of *STUF. Sprachtypologie und Universalienforschung* 49 (1), 9–20.

Bakker, P. (1997). *"A language of our own": the genesis of Michif – the mixed Cree–French language of the Canadian Métis*. New York: Oxford University Press.

Bakker, P. (2000a). Rapid language change: creolization, intertwining, convergence. In C. Renfrew, A. McMahon and R. L. Trask, *Time depth in Historical Linguistics* (Papers in the Prehistory of Languages), pp. 575–610. Cambridge: McDonald Institute for Archaeological Research.

Bakker, P. (2000b). *Language intertwining: Structure and genesis of mixed languages*. Unpublished ms.

Boretzky, N. (1998). Der Romani-Wortzschatz in den Romani-Misch-Dialekten (Pararomani). In Y. Matras (ed.), *The Romani element in non-standard speech*, pp. 97–132.

Dreyfuss, G. R. & D. Oka. (1979). Chinese Indonesian: a new kind of language hybrid? In *Papers in Pidgin and Creole Linguistics* 2 (Pacific Linguistics, Series A – No. 57), pp. 247–274. Canberra: Australian National University.

Matras, Y. (1998). Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36, 281–331.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

Speech genres and other questions on fusion

JAN BLOMMAERT
Department of African Languages and Literatures, University of Ghent, Rozier 44, 9000 Gent, Belgium
E-mail: *jan.blommaert@rug.ac.be*

Yaron Matras offers us one of those rare papers that are genuinely stimulating and challenging to a variety of audiences and in a variety of domains. I am sure that other commentators will go into the grammatical arguments he develops; still others will discuss his views on language and cognition, and some will take issue with the case descriptions he offers. My own contribution to this debate will be restricted to what I believe Matras tells me and other linguistic anthropologists and sociolinguists regarding the interface between linguistic structure and the social use of linguistic resources. Matras' paper, though generally oriented towards linguistic questions and issues, also raises some wider questions, I believe. I will try to spell these out, begging Matras to indulge arguments and questions of mine that may seem to be off course at first glance.

The set of hypotheses offered in the paper is appealing: it offers us a linguistically grounded model of different ways of mixing and shifting, sensitive to interactional and communicative contexts. The contact mechanisms specified in the paper are interesting and the functional turnover hypothesis is a thought-provoking sociolinguistic explanation of particular forms of linguistic shifting. It is from these passages in the paper that I depart, for they are closest to my concern.

My point of departure is the presentation of examples in Matras' paper. There are two "types" of examples, I believe: traditional linguistic single-sentence or single-utterance examples (a majority) and examples that cover longer stretches of discourse or conversational sequences: examples 7, 8, 11 and 12. The latter examples, when inspected more closely, seem to me to raise questions of *genre*. In the sections in which these examples occur, Matras discusses motivations for creating what he calls non-self-contained mixtures, mentioning deviance of communicative norms and negative bystander deixis as verifiable functions. The occurrence of switches into the "secret" varieties is discourse-functional and therefore heavily dependent on specific locations in discourse and conversations. The shifts occur in places where from a conversation-analytical viewpoint they would fulfill major recontextualizing functions such as shifting the participation framework of the speech situation at hand. Going back to Bakhtin's definition of speech genre ("each sphere in which language is used develops its own relatively stable types of ... utterances" (1986, 60)) we see that one may as well approach the use of Lekoudesch and the Romani variants in terms of genres, i.e. specific speech styles that are used in specific speech situations: whenever an interactional shift in speech situation is required, the mixed speech style comes in as a way of reinforcing ingroup identity, of interactionally excluding uninvited overhearers, and of making sense. What happens then is the interactional management of genres, seen in terms of matching (but not necessarily fixed) sets of linguistic-communicative resources and socioculturally anchored spheres of interaction and conduct.

So far this is just a reformulation of what Matras says, for instance with regard to the social motivations for functional turnover. But it helps us to formulate another question. It could be interesting to explore the particular ways in which this nexus of linguistic patterning and conversational/discursive patterning works. The presentation of the examples raises the issue of whether the interactional work that is going on here (and hence, the social motivations for such shifting and for the creation of mixed varieties) is dependent on grammatical instruments only? That is: in defining the matching sets of linguistic-communicative resources and social spheres of interaction and conduct, why should we restrict the potential complex of shifting and mixing features to grammatical features, and not look into paralinguistic (intonational, prosodic, gestural) and metapragmatic (genre conventions, stylistic and normative awareness) features as ingredients of e.g. functional turnover? In sum, when we think about a functional-communicative approach to mixed languages, could we move beyond grammar so as to include for instance generic expectations, narrative structuring, transfer of metapragmatic features and so forth? We know that languages come with far more than grammar and lexicon, yet a lot of the literature on language contact seems to restrict contact phenomena to those "visible" categories (exemplified in this paper by Matras' "function-based compartmentalization of linguistic structures").

I will try to illustrate the question by means of two small examples: one dealing with narrative structure and one with written style. In both cases the data come from L2 users with restricted or partial linguistic competence in the L2. There is no linguistic "mixing" going on in the examples, yet they are language contact data.

The first example is a fragment from the beginning of an interview with a Somali woman, H in the transcript (Blommaert, 1999a). I shall first give the field transcript:

Example 1
H: I'm from *Somalia and my name is Habiba Mohammed and I=I have *five childrens and I coming here before the children are coming=when I was euh when I=I'm arrive in Belgium I was *alone\
A: ah\
H: yeah\in sake of the war=the war of Somalia\
A: uhuh\
H: And. I w=I'm. Twen'*thirty five years old\
A: uhuh\
H: and euh I was working in Somalia ICRC International Red Cross
A: that's

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

H: ICRC *Red Cross\
A: ah OK OK jaja
B + H: [acknowledge]
H: and I was euh office assistant\
A: ja
H: yeah. So Somalia is starting war *nineteen ninety one
A: uhuh\
H: so until ninety one to ninety five I was in Somalia
A: uhuh\
H: and [baby starts crying] wa [laughs] and I have *four children at that time and euhm ... My husband comes from euh *north Somalia
A: uhuh
H: and I *south Somalia is fighting north at=at south is fighting\
A: uhuh
H: so my=my husband and my children have no. *safety for their lives
A: uhuh

The Somali woman displays considerable problems with all sorts of grammatical structures in English, most critical of which from a narrative viewpoint is her struggle with tense and aspect marking in verbs, creating problems with outlining the sequential structure of events in a narrative and thus jeopardizing cohesion and coherence in the story. But what we see is that the narrator succeeds in bringing a considerable amount of narrative structure to her story. Looking at the woman's utterances from a viewpoint of ethnopoetic analysis (e.g. Hymes, 1998), we see structure emerge by means of the skillful use of two English discourse markers: "and" and "so". The patterning and distribution of these markers creates two parts in the narrative (a change in dominant discourse marker introduces a topic change) and organizes subordination within these parts (graphically represented by indentation):

PART I
1. I'm from *Somalia
2. **and** my name is Habiba Mohammed
3. **and** I=I have *five childrens
4. **and** I coming here before the children are coming
   {clarification}when I was euh when I=I'm arrive in Belgium I was *alone\
   yeah\in sake of the war=the war of Somalia\
6. **and**. I w=I'm. Twen'*thirty five years old\
7. **and** euh I was working in Somalia ICRC International red Cross
8. **and** I was euh office assistant\

PART II
1. **So** Somalia is starting war *nineteen ninety one
2. **so** until ninety one to ninety five I was in Somalia
3. **and** wa' [laughs] and I have *four children at that time
4. **and** euhm.. My husband comes from euh *north Somalia
5. **and** I *south Somalia is fighting north at=at south is fighting\
6. **so** my=my husband and my children have no. *safety for their lives

What we see here is how in an L2-learner situation specific parts of grammar seem to acquire functions not usually assigned to them. The connectives used here seem to have crucial narrative-structuring functions dividing the narrative into sequentially organized and argumentatively patterned lines and parts and so accounting for most of the "coherence" and storyline sequentiality in the narrative. The issue is: this is a contact phenomenon in which *narrative patterning* seems to perform important "grammatical" operations. Where do we place such phenomena in models of mixing and fusion?

The second example is about generic models and expectations. What follows is a brief fragment from a handwritten autobiography of a man from Shaba (Congo) (Blommaert, 1996; 1999b). The text is 17 pages long, and most of it is in Swahili. Towards the end, the text suddenly shifts into French. The shift is indexical of a genre shift: the end of the Swahili part is also the conclusion of his life history, and the French part is in fact a letter addressed to his sponsor, a Belgian lady. Let us take a look at a fragment from the French part of the text (a "transcript" of the handwritten version, keeping the lines and some graphic features of the manuscript), followed by an English translation:

Example 2
Cette lumière ça n'a pas illumine que moi qui a
été votre boy, la production aidera les vieux
et les jeunes gens, surtout les jours de Fetes des
Mariages et des deuils. Déjà une bonne somme d'
argent qui étaient destine à moi, c'étaient tombé
dans les mains d'autres pauvres et la Malle des
Habits aux missionnaires, cela ne vous a pas choquer
mais vous me parliez que Heureusement c'étaient
tombé dans les mains d'autres pauvres.

Translation
This light has not only illuminated me, who
was your houseboy, the production will help the old
and young people, especially on holidays and
weddings and funerals. Already a good sum of
money destined for me, it has fallen in the
hands of other poor people and the Suitcase with
Clothes for the missionaries, you were not shocked
but you spoke to me that Fortunately it had
fallen in the hands of other poor people.

What we see here, I believe, is again a contact phenomenon in which the allocation of linguistic resources takes on a peculiar shape. The text is monolingual – highly exceptional in the deeply "mixing" environment of urban Shaba (as elsewhere in urban Africa). So was the Swahili part of the text, though traces of codeswitching could be found, indicating that the monolingual variety was a generically marked variety requiring considerable effort from the author (Blommaert, 1999b). The "non-mixed" character of the text, in other words, has to be placed in a continuum of mixed varieties of French and Swahili, generically regimented, and of which monolingual French is one extreme, monolingual Swahili another. Both monolingual varieties are therefore deeply and fundamentally "mixed" varieties, paradoxical though that may seem. A second feature of the text is the frequency of errors. Clearly, the author has based himself on a spoken and colloquial variety of French in writing this text. The "seriousness" of the text genre does not seem to presuppose "full" grammatical and stylistic access to the language. Spoken varieties (by lack of others) can pass as "correct" and even status varieties without endangering the successful realization of the genre.

Two things need to be emphasized, I believe. First, in a multilingual environment such as this one, *no* visible mixing

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

can actually be mixing for what is being "mixed" is a generic model of a particular type of text (in this case: "serious" autobiographical narrative and formal letter to a European woman). Second, the fact that generic models and blueprints are borrowed does not presuppose the borrowing of all resources typically associated with the model. Cross-language and cross-cultural genre transfer allow for considerable "bricolage" in linguistic and stylistic terms. Despite these two qualifications, there certainly has been "fusion".

Yaron Matras has offered us a refined view of the dynamics of mixing in language contact. In doing so he pushes us further in specifying (and investigating) what exactly can and will be "mixed" in specific contact situations and with specific functions or outcomes. The linguistic side of this could (and should) be complemented with other dimensions of language and communication – the "invisible" or not-so-grammatical stuff in discourse – and I have tried to show that issues of speech genre in contact situations could be fruitful areas of research. Far more is now to be said about this, and for this we should thank Matras.

References

Bakhtin, M. M. (1986). *Speech genres and other late essays*. Austin: University of Texas Press.

Blommaert, J. (1996). A Shaba Swahili life history: text and translation. *Afrikanistische Arbeitspapiere (AAP)* 47 (*Swahili Forum* 3), 31–62.

Blommaert, J. (1999a). Investigating narrative inequality: home narratives by African asylum seekers in Belgium. *LPI Working paper nr. 1*. Gent: Department of African Languages and Cultures.

Blommaert, J. (1999b). Reconstructing the sociolinguistic image of Africa: grassroots writing in Shaba (Congo). *Text* 19, 175–200.

Hymes, D. H. (1998). When is oral narrative poetry? Generative form and its pragmatic conditions. *Pragmatics* 8, 475–500.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

 Bilingualism: Language and Cognition 3 (2), 2000, 112–114 © 2000 Cambridge University Press

The tortoise and the hare: distinguishing processes and end-products in language contact

PENELOPE GARDNER-CHLOROS
School of Languages, Linguistics and Culture, Birkbeck College, University of London, 43 Gordon Square, London WC1H OPD, UK
E-mail: p.gardner-chloros@ bbk.ac.uk

Matras's principal claim is that the variation in the composition of Mixed Languages (MLs) can be understood with reference to "natural function-based compartmentalisation of linguistic structures". As there are close connections between MLs and other types of language contact, in particular code-switching (CS), it is important that we examine this claim in that broader context. I will begin by considering some of the difficulties in specifying what constitutes an ML and some of the features they share with CS.

At a linguistic level, their supposedly characteristic split between grammar (from one language), and lexicon (from another) is said to be "rarely if ever consistent"; Matras points out that few of their structural properties are shared. In Michif, for example, the grammar itself is split, the internal structure of the nominal phrase being essentially French, and that of the verb phrase essentially Cree (Bakker 1994).

Models seeking to explain the *emergence* of MLs are said to be crucial to their definition; but as in the case of pidgins and creoles, such models do not provide a consistent picture either. Some claim that such languages represent a (manipulated) form of language shift; some are based on the idea that they arose in similar (unusual) historical circumstances, some attribute them to a process of conscious creation and some to a gradual conventionalization of CS patterns. Matras casts some doubt on Thomason's (1997) contention that MLs arise in societies where there is full bilingualism, referring to the important role, in their emergence, of language attrition and loss.

The search for a clear-cut definition is further hampered by the fact they are sometimes the principal means of communication in a community but more commonly auxiliary or "secret" languages; finally they are said to be both rare and, Matras suggests, probably short-lived – hence many of the examples we have are based on historical texts rather than contemporary sociolinguistic material. The description of the disagreements surrounding the emergence of Ma'a provides an illustration of the difficulties involved in assigning elements of the language to particular origins, with scholars holding opposite views as to the basic typology of this Bantu/Cushitic mixture and as to how each of these strands contributed to the final product.

At first sight it is therefore surprising that, in spite of these variations and the uncertainty with respect to definitions and origins, Matras is prepared to argue that "Once we have identified the type of contact mechanism at work, a significant portion, though by no means all of the actual structural composition of the ML becomes predictable". It seems inherently unlikely that *prediction* should be possible where *reconstruction* is so problematic. This situation has a parallel in the study of CS, in which the discovery of systematic underlying rules has so far proved elusive.

The fossilization of CS is presented here as one of the possible routes to an ML; an alternative view is that the relationship is more in the nature of a process/product distinction. Others clearly agree that MLs originate in CS (e.g. Myers-Scotton, 1993; Bakker and Mous, 1994), but it is less easy to see exactly at which point, and on what basis, it makes sense to put MLs into a separate category. One relevant issue discussed in the CS literature is whether the term CS should be reserved for purely alternational language mixtures, as opposed to other categories such as borrowing or code-mixing where convergence occurs (Clyne, 1987, 740–741; Hamers and Blanc, 2000, 309–310). I myself have argued that understanding how the continuum between these phenomena operates is more important than setting up such a taxonomy (Gardner-Chloros, 1995). On this view, the fact that the term ML is seen as designating particular, identifiable varieties, whereas CS is thought of in terms of a process, is partly a terminological matter in an area where everyone seems to draw the lines at different points (Milroy and Muysken, 1995, 12). Although CS does not always crystallize into an ML, the differences can to some extent be regarded as analogous to using a video camera or a still camera to view the same events.

Though one can focus on differences, there are many similarities. Like MLs, which are said to be short-lived, CS is often considered a transitional state of affairs. According to Gumperz, it occurs whenever two languages come together under conditions of rapid social change (1982, 64). For example, it is attested in Gal's well-known study of German–Hungarian bilingualism in Oberwärt, where language shift is underway from Hungarian to German (1979). Gumperz and Wilson (1971) showed that after many decades of language contact in Kupwar in India, a situation was reached where several varieties had converged to the point where distinctiveness between them was only preserved by lexical differences. Although it is discussed in terms of CS, there seems no reason why the resulting variety should not qualify as an ML.

The question of whether MLs arise in situations of "full bilingualism" is equally relevant for CS. The answer depends on what is meant by "full bilingualism", since in most "bilingual communities" there are intergenerational differences in the bilinguality of individuals, and in the extent to which they can use the contributing varieties monolingually. Although there are situations where CS is a stable variety within a bilingual majority community

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

(Poplack, 1980), such situations are the exception among studies of CS. Similarly, according to Matras, the majority of MLs are auxiliary to one of their "parent" varieties.

At a functional level, both CS and MLs frequently represent *compromise* speech forms (Scotton, 1976), and provide a means of preserving the ancestral language and the cultural values associated with it while adapting to the demands of modernity (Rindler-Schjerve, 1998). Linguistically speaking, both may include a phase where elements from the two contributing varieties may be used interchangeably. Matras quotes data from Smart on the interchangeable use of the Romani and English copula in English Romani; similar examples from the CS literature include Backus (1999, 275) and Gardner-Chloros (1991, 156).

Just as CS often shows considerable variation within the same community, depending on age, education, and network structure (Bentahila and Davies, 1991; Li Wei, 1998), an example such as that of Jenisch suggests that ML may not be much different: "The product of the process consists of scattered mixed utterances and does not constitute a self-contained communicative system". Matras' description of the sporadic insertion of Romani into a German base for e.g. humorous or convention-challenging purposes bears more than a passing similarity to the Creole insertions in English in the variety described by Sebba (1993) under the heading of CS.

As Matras says, most languages are mixed to some extent, and although MLs represent a particular type of focusing, the processes underlying them are not fundamentally different from those operating in other forms of contact. English can be seen as an ML, consisting of an Old English grammatical base with considerable relexification from Norman French. "At first those who spoke French were those of Norman origin, but soon through intermarriage and association with the ruling class numerous people of English extraction must have found it to their advantage to learn the new language, and before long the distinction between those who spoke French and those who spoke English was not racial but largely social" (Baugh, 1951, 135). In line with Bakker's (1994) account of ML formation, there were numerous marriages between Norman men and English women (Baugh, 1951, 141), which is consistent with Norman being the lexifier and Old English the principal provider of grammatical structure. Most striking of all is the evidence that the introduction of French words into English is closely correlated with the progressive adoption of English by the upper classes. Jespersen (1928, 94) shows that slightly over ten thousand French words entered the English language between 1250 and 1400 (calculation based on the words' first recorded usage). This corresponds exactly to the period when the upper classes were adopting English (these words represent 40% of all the French words in English, of which about 75% are still in current use). From the fifteenth century onwards, when English had become the language of the majority within the ruling classes, there was a sharp drop in borrowings from French (Jespersen, 1928, 94; Baugh, 1951, 214).

A recognition of the continuity of language contact processes is compatible with Auer's (1998) proposal that language contact progresses from CS, where alternation between two varieties is used meaningfully, to Language Mixing, where they are used complementarily and some distinctiveness is lost, to Fused Lects, where there is no longer any discourse mileage to be derived from the contrast. This pattern clearly represents an idealization. As Auer acknowledges, many contact situations do not lead to a completion of the process and various "structural domains" of language may progress along the CS–LM–FL cline at different rates (Auer, 1999, 324). Furthermore, in many communities all three stages are observed synchronically among different groups, or in different contexts (Bentahila and Davies, 1991). Another major theory of CS, Myers-Scotton's Matrix Language Frame Model (1993), fits uneasily, according to Matras, with the idea that there is a progression from CS to ML: the latter do not behave consistently as regards the separation of content and system morphemes, which is a cornerstone of the MLF theory.

All in all, in the absence of any evidence for discontinuity between CS and ML, it seems undeniable that in certain cases the first leads to the second. Although not all CS will lead to an ML, there seems no reason to postulate further mechanisms and it is unlikely that an ML would emerge without a prior CS stage.

This brings us to Matras's explanation of what underlies the different configurations in MLs. He claims that "we must draw on the trivial but crucial realisation that words are not simply words, but are divided into function-oriented categories". His claim is that these "cognitive" categories map onto the traditional grammatical categories in a complex manner, which explains the basis on which different grammatical sub-classes are transferred into a "separate system". For example, "pragmatically salient" components of the ancestral language (which may for instance consist of a package including frozen derivational features, basic numerals, some noun phrase modifiers, the copula, and the possessive verb), can be the subject of "selective replication" in the ML, although the motivations for grouping together these features is not obvious at first sight.

This proposal shows parallels with Azuma's proposals (1998) in relation to CS. Azuma's point of departure is that CS obeys the constraints of the MLF model, but he feels the need to offer an explanation as to why certain constituents are switched as a "chunk" in CS rather than on a word-by-word basis, and secondly as to why certain closed-class items are switched, whereas the theory states that this can only occur when they are part of an EL (embedded language) island, i.e. in a group. He suggests that any chunk/segment which can "meaningfully stand alone in the speaker's mind" can be switched, alongside closed-class items which have particular discourse functions or a "rich semantic/discourse function". Thus Azuma gives the example of "the" being the only English word used in a Japanese sentence, to convey the message that what "the" refers to is unique. His explanation for "that" (complementizer) being switched is that it highlights the relationship between the two parts of the sentence (i.e. CS with a discourse function).

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

The problem with this explanation is its ad hoc nature and its potential circularity. No external method for verifying which chunks "stand alone" in the speaker's mind is proposed, nor how we can decide that a function word has a "rich semantic/discourse function", except through the fact of its being switched. Matras's "natural" function-based compartmentalization of linguistic structures runs the same risk, unless some objective method of corroboration can be found. How can we decide which are, for example, the "salient components that are easily retrievable" (Matras's categories 1 and 5) without knowing the linguistic history of the particular individual? As often in language contact, attractive explanations which are not carefully situated as to the level of generalization – individual or group – cannot be used for prediction and so lose their force.

The study of MLs, and the search for a common rationale underlying their formation, adds a serious – and worthwhile – challenge to the study of language contact. First, it adds the methodological challenge of reconstruction – and the advantages of hindsight – to the synchronic assessment of norm formation in contact situations. Secondly, it adds a further dimension to the notion of focusing, showing how similar processes can lead to different outcomes. We should be wary, however, of resorting to explanations such as Matras's cognitively based categorization may draw us into, which are extremely difficult to test. The more conventional correlation of different language contact outcomes with detailed sociolinguistic description, and the comparison between these, may be the method of the tortoise rather than the hare, but, providing we live long enough, it may provide us with a more satisfactory basis for assessing the regularities in contact languages.

References

Auer, P. (1998a). Introduction: bilingual conversation revisited. In P. Auer (ed.), *Code-switching in conversation: languages, interaction and identity*, pp. 1–24. London: Routledge.

Auer, P. (1999). From codeswitching via language mixing to fused lects: toward a dynamic typology of bilingual speech. *International Journal of Bilingualism* 3, 309–332.

Azuma, S. (1998). Meaning and form in code-switching. In R. Jacobson (ed.), *Codeswitching worldwide*, pp. 109–125. Berlin: Mouton.

Backus, A. (1999). The intergenerational codeswitching continuum in an immigrant community. In G. Extra & L. Verhoeven (eds.), *Bilingualism and migration* (Studies in Language Acquisition 14), pp. 261–279. Berlin: Mouton de Gruyter.

Bakker, P. (1994). Michif, the Cree-French mixed language of the Métis buffalo hunters in Canada. In P. Bakker & M. Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, pp. 13–33. Amsterdam: IFOTT.

Bakker, P. & Mous, M. (1994) (eds.). *Mixed languages: 15 case studies in language intertwining*. Amsterdam: IFOTT.

Baugh, A. C. (1951). *A history of the English language*. London: Routledge & Kegan Paul.

Bentahila, A. & Davies, E. E. (1991). Constraints on code-switching: a look beyond grammar. *Papers for the symposium on code-switching in bilingual studies: theory, significance and perspectives*, pp. 369–405. Barcelona, March 1991: European Science Foundation.

Clyne, M. (1987). Constraints on code-switching: how universal are they? *Linguistics* 25, 739–764.

Gal, S. (1979). *Language shift: social determination of linguistic change in bilingual Austria.* New York: Academic Press.

Gardner-Chloros P. (1991). *Language selection and switching in Strasbourg*. Oxford: Clarendon.

Gardner-Chloros P. (1995). Code-switching in community, regional and national repertoires: the Myth of the discreteness of linguistic systems. In L. Milroy & P. Muysken*, One speaker two languages: cross-disciplinary perspectives on code-switching*, pp. 68–89. Cambridge: Cambridge University Press.

Gumperz, J. (1982). *Discourse strategies*. Cambridge, Cambridge University Press.

Gumperz, J. J. & Wilson, R. D. (1971). Convergence and creolization: a case from the Indo-Aryan–Dravidian border. In D. Hymes (ed.), *Pidginization and creolization of languages*, pp. 151–168. Cambridge, Cambridge University Press.

Hamers, J. and Blanc, M. (2000; 2nd edn.). *Bilinguality and bilingualism*. Cambridge: Cambridge University Press.

Jespersen, O. (1928; 4th edn.). *Growth and structure of the English language*. Leipzig: Teubner.

Li Wei (1998). Banana split? Variations in language choice and code-switching patterns of two groups of British-born Chinese in Tyneside. In R. Jacobson (ed.), *Codeswitching worldwide*, pp. 153–177. Berlin: Mouton.

Milroy, L. & Muysken, P. (1995). Introduction: code-switching and bilingualism research. In L. Milroy & P. Muysken*, One speaker two languages*: *cross-disciplinary perspectives on code-switching*, pp. 1–14. Cambridge: Cambridge University Press.

Myers-Scotton, C. (1993). *Duelling languages.* Oxford: Clarendon.

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: toward a typology of code-switching. *Linguistics* 18, 581–618.

Rindler-Schjerve R. (1998). Codeswitching as an indicator for language shift? Evidence from Sardinian–Italian bilingualism. In R. Jacobson (ed.), *Codeswitching worldwide*, pp. 221–249. Berlin/New York: Mouton.

Scotton C. M. (1976*).* Strategies of neutrality: language choice in uncertain situations*. Language* 52, 919–941.

Sebba, M. (1993). *London Jamaican*. London: Longman.

Thomason, S. G. (1997) (ed.). *Contact languages*. Amsterdam: John Benjamins.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

Selective replacement is extreme lexical re-orientation

MAARTEN MOUS
Department of African Linguistics, Leiden University, postbus 9515, 2300 RA Leiden, Netherlands
E-mail: *m.mous@let.leidenuniv.nl*

There are a number of aspects of Matras' approach to mixed languages that I consider to be sound and laudable. These include his search for functional explanations for the structural compartmentalisation in mixed languages; the fact that he does not assume that mixed languages form one uniform category, nor that they have all undergone the same processes; and his search for processes of change that are not limited to the emergence of mixed languages. Other researchers in the field have a similar attitude, as is apparent from the title of Thomason's (1995) article "Language mixtures: ordinary processes, extraordinary results". Matras' proposals relate the socio-historical scenarios which give rise to mixed languages (or language mixtures) with functionally motivated, interrelated linguistic changes. However promising his observations may be, the details of the nature of such a relationship do not become completely clear in the present article. In particular I will discuss the relation between his selective replication and functional turnover.

My comments, some of them critical, others more complementary in character, relate to five issues: What are the actual functions that can be correlated with certain linguistic changes or phenomena in mixed languages, and in particular, how prominent is secrecy as the function of certain varieties? What is the nature of the concept of self-containedness of the languages in question? What is the nature of selective replication and its link with functional turnover? How should processes of language change be portrayed, from the point of view of the language system or from that of the speaker? Finally, I add some comments on the factual details of Ma'a.

As examples of lexical re-orientation in non-self-contained mixtures, Matras discusses several forms of what he calls secret languages. At several points he acknowledges that these languages tend to function as markers of identity as well as secrecy but when he correlates the linguistic facts to the functions, he limits himself to the secrecy function, which according to him would explain the need to replace lexical items, numerals and negation markers. In my experience with the literature on cant, argot, secret languages, and trade languages, the functional distinction between secrecy and identity is never clear, for which reason instead of positing a category of secret languages, I only use argot, giving primacy to the function of identity, while at the same time distinguishing between those lects that have a limited scope of communication and those that are suitable for every-day communication. The function of identity is two-sided and can be a positive association with a certain cultural identity, or it can be a negative association with the dominant identity in the area concerned. Such a distinction allows an explanation of lexical re-orientation towards one particular language, or rather away from the dominant language, in which case any deviant elements are suitable to act as replacements. Other functions of lexically re-oriented non-self-contained mixtures are respect or fear, based on the concept of the power of the word. Such is the case in the various registers of respect which entail avoiding the name of the father-in-law and any word resembling it. Lexical re-orientation in these lects often makes use of the derivational possibilities in the dominant language, a means which is less successful if the function is to mark a different or a specific identity, or secrecy. Similarly, youth in the big city often develop registers that, apart from identity-marking, fulfil the function of linguistic competition. Here language games such as syllable inversion, insertion, or substitution may serve that purpose. Hence, while Matras' search for functional explanations is promising, research into the relationship between the function of a mixed language variety and its linguistic properties should be deepened and expanded.

Matras proposes the term "non-self-contained" mixtures for those symbiotic mixed languages that are not used for the production of fluent conversation but merely for the production of single mixed *utterances*. I agree that this is an important distinction. Matras seems to equate his "non-self-contained" mixtures with Smith's "symbiotic" mixtures. However, there is a category of self-contained symbiotic mixed languages and Ma'a is a case in point. Ma'a is symbiotic with the "pure" Bantu language Mbugu but conversations that are entirely in Ma'a are common; almost every Mbugu lexical root is replaced by an equivalent, including possessives, demonstratives, and pronouns. The example that Matras quotes from my work in (18) gives a wrong impression. This particular example served to show that people *can* use different registers within one conversation. The first sentence in (18) is in the symbiotic Bantu language Mbugu with one Ma'a root as intruder. The second line is by another person, later in the conversation and is an example of the Ma'a register. The publication from which the example is taken contains another sequence of a more typical Ma'a text.

Matras proposes that foreign structural properties other than negation and copula enter a mixed language through "selective replication" as part of a "functional turnover". I find the idea that speakers copy structural elements from another language (their ancestors' language) for emblematic functions both attractive and plausible. One of the possible reasons to do so is to bring the lost language to life, that is to say, in a situation of functional turnover. However, if selective replication refers to copying certain structural elements from another language, this language could also be a language other than the

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

former language. Indeed, Matras' examples of Jenisch show that there is no need to have a command of the grammar of the language copied from, as he observes that "[a]ll this suggests a stage in which a population of German speakers had contact with active speakers of Romani, and copied elements of their speech without actually being able to process sentences in Romani, i.e. with no grammatical competence in the language". The inherent connection between a functional turnover and selective replication is postulated but not argued for in the course of the article. This issue is important for his discussion of Ma'a as an example of selective replication since the necessity of a functional turnover for selective replication is the only linguistic basis on which such a turnover can be postulated in the case of Ma'a. In fact there are indications that foreign structural elements can enter a non-self-contained mixture independent of a functional turnover. For example, the register of respect among the Yemsa of Ethiopia includes deviant pronouns, and also verbal inflection for the second and third person (Aklilu Yilma, 1992, 5). In Matras' approach it is not grammatical transfer as such that is an instance of selective replication, since the transfer of copula and negation markers is attributed to lexical re-orientation, but rather it is that the grammatical transfer is from an ancestral language, with the purpose of keeping that language alive and using it for secret in-group communication.

He equates process and purpose. It may be preferable to study the processes as such and then relate them to functions. Sometimes Matras portrays selective replication as the next step in lexical re-orientation, but at other times the two processes are presented as though they exclude each other, whereby selective replication is characterised as "the gradual decline of a language ... *rather than* deliberate recruitment of lexical material". Matras proposes that the salient components that are easily retrievable are the elements that are replicated. These include the elements that are inherently situative, pointing out components of the speech action or speech situation, namely those that are communicatively salient and, for the language learner, easily retrievable. Concretely, deictics, possessors, quantifiers such as basic numerals, inflected forms of semantically salient verbs of motion are typically replicated. With the exception of the last category, Ma'a does indeed show parallel forms for these categories. In my view, the replacement of the forms (but not the categorisation) of closed sets of free-standing function words such as personal pronouns, possessives, and demonstratives is in essence not all that different from lexical re-orientation, and can still be characterised as deliberate recruitment, at least in origin, though not in actual use. Once a lect is self-contained, whether symbiotic or not, the need for its own forms for closed sets of function words is high; or, to put this the other way around, once the function words can be taken from the re-oriented lect, it comes close to being self-contained. Thus, what is essential is the communicative function of the lect, i.e., the outcome of the functional turnover in Matras' terms, but not the turnover itself. For this reason the mere presence of signs of selective replication does not warrant the assumption of a previous functional turnover. In the case of Ma'a where we find a limited number of elements that are not common for a Bantu language and are different from the symbiotic Bantu language, the thesis that these derive from their former language is based on the assumption that such selective replication is by necessity connected to a functional turnover. For example, the fact that Ma'a has a "possessive" verb is assumed to be a feature of a former Cushitic language. However, the root in question, *lo*, cannot be traced to a candidate former Cushitic language. The Cushitic language Iraqw, which is closest to the largest source of Cushitic lexical material in Ma'a, has a verb *koom* "to have" which also has the senses of "be with, be together", as in the "Bantu" pattern "be with" for possession. Incidentally, various Bantu languages such as Kirundi and Icibemba, have a possessive verb. Hence there is no evidence of a turnover here. Finally, at several points Matras mentions that functional turnover is a gradual process. It remains unclear whether this is an assumption necessitated by the model, or based on observation. In my view, the linguistic correlate of functional turnover, selective replication, does not need to be a gradual process.

Matras formulates lexical re-orientation as transfer of meaning, "to shift conceptual representation", or "shift meaning or content". Thus his perspective is from the language as a system. From the view of the actual language user there is no transfer of meaning but rather transfer of form to the dominant grammatical system. In an approach that aims at relating language change to function, it is more appropriate to start from the point of view of the language user in presenting the change.

I feel obliged to make some remarks on the details of the Ma'a facts, even if they are not crucial for the essence of Matras' account. I have already commented on the inappropriateness of example (18) and the argument of the possessive verb. Matras mentions the copula as a non-Bantu component in Ma'a. This is a reference to remarks by Thomason, quoting Ehret, on the fact that Ma'a uses a copula in constructions where Bantu languages would not. The actual copula concerned is, however, Bantu. Moreover, there is no difference in use of the copula between the Ma'a register, its symbiotic Bantu language Mbugu, and its closest relative Pare. Finally, Matras mentions an "unclear adaptation pattern for the inflection of some verbs" as a non-Bantu element, but unfortunately I do not know what he is referring to. In conclusion, the present paper forms a very interesting and stimulating approach to mixed language varieties and offers a promising direction for future research.

References

Aklilu Yilma (1992). The linguistic etiquette of Yemsa. *Journal of Ethiopian Studies* 25, 1–14.

Thomason, S.G. (1995). Language mixture: ordinary processes, extraordinary results. In C. Silva-Corvalán (ed.), *Spanish in four continents: studies in language contact and bilingualism*, pp. 15–33. Washington, DC: Georgetown University Press.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

Some considerations for explaining mixed languages

MELISSA G. MOYER
Departament de Filologia Anglesa, Edifici B, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain
Department of Linguistics, New York University
E-mail: melissa.moyer@uab.es

Yaron Matras' account of the existing structural disparity among mixed languages within a functional approach is a welcome proposal and it constitutes an important effort to give a systematic and theoretically unified explanation. Several questions, however, are raised by his approach.

The main argument of the article is based on the assumption that one can arrive at an explanatory account of mixed languages by relating lexical items and other surface phenomena of mixed languages (i.e. word order and inflectional morphemes) with specific language contact mechanisms which are cognitively motivated and allegedly responsible for the production of those mixed structures. Furthermore, those individual lexical items are grouped into five "natural communicative function" categories and they can be related to one or more of the language contact mechanisms proposed. This rather fixed functional compartmentalisation of the lexicon turns out to be an important limitation of the approach. Communicative functions in language rely on more than just a fixed relation between form and function. Also, functional meaning in many instances rests on structural units larger than the descriptive elements proposed by researchers of mixed languages.

Matras adopts the assignment by other researchers of various functional elements to different languages and then he proceeds to assign them without further question to four types of cognitively based language contact processes. The cognitive relevance of a description of mixed languages that does not recognize grammatical constituents such as phrases or conventionalized expressions is questionable. An example from the article which illustrates this problem comes from the Caló expression *mansa camelo tuque* (English: I love you). Matras uses this example to argue for the functional compartmentalisation of items such as pronouns, demonstratives and semi-productive inflectional features. While a mixed language may have existed when Romani gave way to Para-Romani and Caló, the example above might very well have come down to us as a whole expression used to express a special in-group intimacy. Therefore, the relevant communicative function might be tied to the entire expression rather than to individual compartmentalised structures such as those proposed by Matras.

The clear-cut compartmentalisation of structures by language is often not so neat as it might first appear. A given lexical item which may occur in both of the participating languages of a mixed system can be used to elicit the problem which arises from assigning an element with a specific function and contact mechanism (i.e. lexical reorientation) to both a grammaticiser language and a lexifier language. The scarce data available on some of the mixed languages discussed and the lack of language tagging of the few examples included in the paper make it hard to ascertain the degree of variation which might exist for the mixed languages discussed and how well the compartmentalisation proposed holds.

An instance in the article where such an inconsistency appears is with the classification of the copular verb in Leukoudesch and Para-Romani. In example (2) from Leukoudesch *Lou dibra, d'r guj* ***schäfft*** (English: Don't speak, the man is (there)) Matras classifies the Ashkenazic Hebrew word *schäff-* as a copula but in example (4) also from Leukoudesch *Die goja isch haggel doff, dia kennt-m'r lekächa* (English: The woman is very pretty, one could take her), and among the descriptive list of structures by language, the copula is listed together with other lexical items coming from German. At another point of the paper, the discussion of Romani brings up facts having to do with variation in the language of the copula as one can observe in examples (14) *Dik, savo* ***see****? A gorgio?* (English: Look, who is (that)? A stranger?) and (15) *Coovo Moosh is a gryengro* (English: This man is a horse-dealer). In the first example, the copula appears in Romani and, in the other, English. Matras attempts to account for this discrepancy with the functional turnover hypothesis. The functional turnover hypothesis may be a plausible explanation for language shift in mixed languages, but it is being invoked for the above example in a rather ad hoc way since it implies that language variation can only exist in intermediate stages of language shift and that after the functional turnover has taken place forms become necessarily conventionalized and fixed.

Another important consideration to bear in mind regarding Matras' approach has to do with the four contact mechanisms (lexical reorientation, selective replication, convergence, and fusion) and his claims regarding their cognitive and communicative motivation whereby speakers appear "to model certain functions of language on an alternative linguistic system". The function or motivation for the compartmentalisation of mixed languages is based on a form/function relation which lacks support from real data of language use where meanings and functions are negotiated in the course of verbal interactions (Gumperz, 1982; Auer, 1998). The functional meaning Matras has in mind is static and the contact mechanisms he proposes refer to functional motivations at a given point in time, an issue which is not made sufficiently clear in the discussion. The contact mechanisms proposed describe the historical development of mixed languages but they have little to do with the cognitive processes used by actual speakers for communicating functional meaning. Some suggestions

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

would also be appreciated for how certain uses associated with linguistic structures from the participating languages developed and became conventionalized. So, when Matras talks about the functional motivations for the lexical elements in mixed languages he is distancing himself from the users of mixed languages and their more immediate communicative intentions. A cognitively motivated account of mixed languages should rely on the actual ways people use language to communicate meaning in the course of their interactions rather than on hypothetical abstractions which do not bring the speaker into consideration.

I am in full agreement with Thomason and Kaufman (1988) who argue that mixed languages cannot be accounted for strictly on the basis of linguistic constraints and that the sociolinguistic history and communicative functions of the speakers are important for determining the outcome of language contact. Matras, however, takes his theory of the functional compartmentalisation of structures as the main explanation for the lexical and grammatical make-up of mixed languages. Priority is given to the function of specific words without acknowledging in a systematic way that there are other sorts of constraints which influence the way mixed languages get to be the way they are. First of all, there are the social and historical considerations of the contact situation, and second, one needs to take into account the grammars of the participating languages contributing to the mixed linguistic system.

Matras does provide some general information about the social and historical background of the mixed languages he discusses, but not enough detail is given about the speakers, their language acquisition histories, the socioeconomic pressures influencing their lives, nor the different sorts of contexts and functions of the contact situation which led to the mixed system. For example, what were the reasons and the circumstances which led to the interruption of contact and the direction it took in a mixed language like Ma'a? The shift in orientation by speakers towards the different languages in contact is crucial for the structural outcome of a mixed language. The social and historical circumstances provide indirect evidence that should complement any functional account intending to explain the structural compartmentalisation of mixed languages

A further consideration is the grammatical constraints on mixed languages imposed by the grammars of the participating languages. These constraints are particularly relevant in the mixed languages which Matras claims have undergone convergence, namely, Petjo, Javindo, Copper Island Aleut, Media Lengua, and Michif. Convergence is identified and analysed by surface level phenomena which are mostly lexical. The restrictions on syntactic relations such as structural dependency, agreement features, and locality requirements as well as specific features and syntactic properties associated with verbs and other lexical items (i.e. argument structure, subcategorisation) from different languages are not taken into account. These syntactic relations or others within a non-generative framework which appear to be relevant for constraining the structure of single languages as well as code-mixed sentences are not brought in to explain mixed languages through processes of convergence. In part this is because of the limitation, mentioned earlier, imposed by the single-element form–function relation. In future attempts to explain mixed languages, a deeper understanding could be reached by looking at the ways words from different languages create dependencies which may be explained in relation to the grammars of the languages participating in the contact situation.

References

Auer, Peter (1998) (ed.). *Code-switching in conversation*. London: Routledge.

Gumperz, John (1982). *Discourse strategies*. Cambridge: Cambridge University Press.

Thomason, Sarah G. & Terrence Kaufman (1988). *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

What matters: the out of sight in mixed languages

CAROL MYERS-SCOTTON
Linguistics Program, University of South Carolina, Columbia, SC 29208, USA
E-mail: carolms@sc.edu

Matras deserves admiration for his ambitious attempt to systematize our understanding of mixed languages (hereafter MLs). In addition, that it is innovative also makes his attempt attractive; he departs from the main previous treatments of MLs to argue for a functional basis for their development. Finally, by going beyond the isolation of case study approaches characterizing the bulk of ML studies and taking on the entire ML fabric, Matras opens up a new way to study MLs. In doing so and contrary to many assumptions, he claims that MLs are not cut the same way from the linguistic cloth; and, at least from the surface point of view, he demonstrates he is right. That a fair amount of surface variation does exist is the impetus for him to suggest that several different explanations account for this diversity. Yet, as much as I admire Matras' industry and some of his insights, I see problems in three areas of his analyses. At least the first two stem from his missing what is "out of sight" in mixed languages; that is, he does not look below the surface representations and therefore misses important generalizations that go far in explaining why MLs can have the surface structures they do.

The problems are these. First, his embracing of the power of functional or social factors (e.g. "cultural defiance") to determine linguistic structural outcomes is not convincing. Second, his analyses are essentially ways of showing how four mechanisms at work in MLs produce a taxonomy of MLs. The results are descriptive, but not explanatory. Third, Matras' style of writing makes his arguments hard to follow. There are many abstract terms, many under-defined words and phrases (e.g. *linguistic-mental procedures*, neologisms (e.g. *non-self-contained mixtures*), too few specific examples, and no morph-by-morph glosses. Matras clearly is fairly bristling with ideas – all to the good, but his style does the ideas a dis-service. Space permits my taking up only the first two problems.

If I read Matras correctly, his sense of "functional" is broad and somewhat idiosyncratic. Basically, and as might be expected, "functional" refers to social motivations for structural outcomes; that is, community members develop an ML, often consciously, to achieve social ends. (I have trouble with "consciously", but there is not space to debate that issue.) However, functionalism for Matras also seems to be a psycholinguistic concept. As such, it includes some ideas about the abstract linguistic system in the brain and about how language production works. That is, at various points, Matras refers to "natural compartmentalisation of functions", seeming to mean that different parts of the linguistic system are separate from each other, perhaps accessed differently. He also refers at times to "processing operations". A more specific psycholinguistic reference is to "a cognitive motivation for fusion: it is triggered by cognitive pressure to reduce the mental processing load …". But he does not develop these notions nor, just as important, provide any motivation for their relevance to ML development.

In addition, Matras' *Functional Turnover* has a psycholinguistic flavor because it refers to the turn that the language production process can take. This "turnover hypothesis" is clearly reminiscent of my Matrix Language Turnover Hypothesis (Myers-Scotton, 1993a, 1997, 1998 *inter alia*); however, Matras does not even mention this parallel nor make clear the linguistic mechanism that is involved. Matras also refers throughout to "functional categories" (presumably "grammatical categories"), but it seems that he had social motivations in mind, not specific lexical categories, when he named his *Functional Turnover.* Yet, when he presents his taxonomy, which does refer – after all – to types of structural outcome, the term he uses is "four function-based mechanisms of contact". Precisely what this use of "function" means is left open for any or all of the above interpretations, it seems.

Finally, perhaps also under the rubric of functionalism, he comes close to invoking ideas reminiscent of the Whorfian hypothesis. He refers to a seemingly cross-linguistic difference in the "realm of experience and representation of collective knowledge" as motivating the incorporation of elements from an L2 into an L1.

The most important point to make about Matras' use of "functionalism", however, is that he offers very little motivation for any of the claims he makes under these different possible views of functionalism. It would have been most relevant to this journal's readers for him to develop his notions of how the abstract nature of language production affects surface outcomes. Be that as it may, Matras has most to say about functionalism as a social construct, where he is clearest about what he has in mind as "functionalism". For example, when discussing the *Functional Turnover*, he refers to "the nature of social relations between the minority community and the majority society".

Because he has most to say about functionalism in this regard, my comments about his use of this construct refer to its social side. Matras sees community motivations driving the constructing of MLs, but he also clearly implies that these motivations affect *the structural forms* MLs can and do take. I remain unconvinced that the social functions of MLs account for their structural differences. True, Matras convincingly demonstrates that all outcomes are not the same. However, if social factors were the main determinants of ML structure, then how would one explain the many *similarities* in outcomes *when social factors vary*? I agree wholeheartedly that social functions figure in structural outcomes. But *figuring* is not *determining.*

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

I argue that the role of social factors in grammatical structure is constrained to one of *influencing the selection of options* from among a set of possibilities that the abstract blueprint of language (with a big "L") offers humans (cf. Myers-Scotton, 1993b on codeswitching and Myers-Scotton, 2000 on creoles). What this claim means is that MLs cannot be explained by looking only at their surface structure. Rather, what is out of sight – the abstract configuration of elements in linguistic competence that are accessed in language production – is what determines surface possibilities. Seemingly diverse surface outcomes in MLs lose their inconsistencies when viewed in terms of their abstract origins: the underlying principles and processes limiting the surface possibilities are the same. In addition and for the same reason, MLs resemble *other* contact varieties *at the abstract level*. This point will be developed later when I introduce two new sub-models to the Matrix Language Frame model; they explain Matras' ML mechanisms in a way that relates them to other linguistic phenomena.

My overall point of departure is that language competence and production is no different in bilingual speech (including obviously MLs) than in monolingual speech, with only one exception. To put it metaphorically, while it may take two to tango, one partner has to lead. This dictum applies to bilingual speech as well. That is, one variety is structurally dominant in the sense that its grammatical frame prevails. A theoretical construct, the Matrix Language, is a heuristic to label this frame. In some contact varieties (e.g. classic codeswitching; cf. Myers-Scotton, 1993a), the Matrix Language is synonymous with the frame of one of the varieties. However, even then, the Matrix Language is best thought of as an abstract grammatical frame, not as "a specific language". This becomes more important in cases of language shift (or language creation in creole formation) when the grammar of the preferred target is not fully accessible, or when speakers have competing ideas about a target. When this happens, the Matrix Language is not entirely equivalent to the frame of any one language, but rather is a composite of abstract grammatical directions for the two or more participating varieties. However, even in a composite Matrix Language, abstract lexical structure from one variety always seems to dominate.

Matras devotes most attention to establishing a typology of four mechanisms that he says give rise to MLs: lexical re-orientation, selective replication, convergence, and categorial fusion. His typology has a problem that plagues many typologies: three of Matras' mechanisms clearly overlap. For example, he says that lexical re-orientation (né *re-lexification*, at least in my world) often means that nouns, verbs, sentence adverbs, etc. are inserted from one language into the grammatical framework of another. Thus, of the four mechanisms, lexical re-orientation best represents what most previous researchers have considered MLs to be. The functional intention behind lexical re-orientation is to create a language secret from outsiders. What may separate lexical re-orientation from the other mechanisms is that bilingualism in a language unknown to the dominant community is a prerequisite for using this mechanism. In Matras' example, Lekoudesch, speakers of a German dialect who also know another language (Ashkenazic Hebrew) are the type of persons who could produce such an ML. Matras adds to lexical re-orientation a puzzling sub-category, "non-self-contained mixtures". This puzzling name refers to bilingual speech that consists mainly of short formulaic routines that happen to involve more than one linguistic variety. From the examples Matras cites, non-self-contained mixtures seem to be none other than codeswitching, and no more: if it walks like a duck, if it quacks like a duck … The only reason for differentiating it from some other instances of codeswitching is Matras' claim that such lexical insertions are used to "conceal meaning", but some codeswitching has this motivation as well.

Selective replication does not seem to differ measurably from lexical re-orientation. First, the lexical categories of the two mechanisms certainly overlap. Among affected elements in selection replication, are "salient elements of a language" and "lexical vocabulary" (presumably nouns, etc. again). One difference is that now deictic elements and interjections are included. If the structural outcome in selective replication is not much different, neither is the motivation. That is, Matras says that speakers use this mechanism to create a secret language as well. The only difference is that the innovators are mainly members of the younger generation who are already engaged in shifting from their L1, using lexical remnants of the L1 to create a secret language. He does not make it clear if the lexical input under lexical re-orientation is from the L1 or an L2; so the two mechanisms may differ in this regard. His claim that the difference is in "reproducing contextually relevant actions of speech" is only mysterious.

Matras discusses the well-known case of Ma'a as an instance of selective replication. If an older generation that is using the L1 still exists, all well and good. But Matras states that under selective replication speakers try to "re-activate impressions of an ancestral language". Matras is not alone in the literature in implying that speakers can conjure up a language no longer spoken and that this is an element of how MLs are created (e.g. cf. Mous, 1994). I do not join them in attributing to humans such powers. Matras recognizes that alternative suggestions have been made for how Ma'a emerged, but he does not mention mine (Myers-Scotton, 1998) even though he uses the idea of a Matrix Language Turnover. What makes his *Functional Turnover* different from my hypothesis is that he uses it without the theoretical motivations that would accomplish the present day state of affairs in Ma'a. These motivations are captured in the Matrix Language Frame Model (Myers-Scotton, 1993a; but especially in Myers-Scotton, 1997) and in the 4–M Model (Myers-Scotton and Jake, 2000). These models were alluded to above when I referred to "out of sight" abstract configurations as explaining surface structures. The explanatory power of both models relies on a distinction between content and system morphemes that is motivated by evidence from other linguistic phenomena (e.g. speech errors, aphasic speech, and even

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

differing frequencies of accuracy of production in interlanguage in SLA (Wei, 2000, *inter alia*). Note that system morphemes are similar to, but not the same as, functional categories or closed class items, an important distinction. The 4-M Model refines the content–system morpheme distinction by classifying system morphemes into three types. The model assumes that conceptually activated morphemes are accessed differently in language production from morphemes that are structurally assigned. There is a distinction between the salience and level of accessibility of conceptually activated morphemes (content morphemes and the "early" system morpheme that they select) in contrast to structurally assigned morphemes (those "late" system morphemes that are not salient until larger constituents are assembled at the level of the "formulator", e.g. morphemes signaling subject–verb agreement). This split in the salience and accessibility of morphemes in production motivates the view that a mixed language, such as Ma'a, is a feasible form of bilingual speech. That is, it is a variety with many content morphemes from one source (a Cushitic variety) but a grammatical frame from another variety (system morphemes from neighboring Bantu languages). The 4-M model also explains splits in which variety supplies different types of morpheme in other language contact phenomena (cf. Bolonyai, 2000 and Schmitt, 2000 on bilingual child language attrition/acquisition, and Gross, 2000 on creole formation).

The third mechanism is convergence. While Matras' discussion of this mechanism is his most insightful, I disagree with his claim that convergence occurs "while at the same time maintaining the material-structural autonomy of the linguistic systems". From my point of view, convergence necessarily involves composite structures. This is where another "out of sight" model, the Abstract Level Model, is useful in explaining outcomes (Myers-Scotton and Jake, 1995, 2000). Under this model, there are three levels of abstract grammatical structure: the levels of lexical-conceptual structure (semantics and pragmatics), predicate argument structure (arguments in relation to thematic role assigners), and morphological realization patterns (surface realizations including morpheme order). While the idea of abstract lexical levels is not new or original, the Abstract Level Model does include an innovation. This is that these levels may be split in bilingual speech so that all, or parts of, one level may come from one variety and other levels from another variety. This view explains what results in convergence: an apparently monolingual surface (monolingual at the level of morphological realization patterns), but with some elements of out-of-sight (abstract) structure from the other levels of abstract lexical structure from another variety.

When Matras gets to fusion, what we have is simply a combination of the other three mechanisms. In his discussion of Javindo, Petjo, Copper Island Aleut, Media Lengua, and Michif, Matras provides a nice list of the linguistic elements that their different contributing varieties bring to the net result. The 4-M Model mentioned above provides an explanation for the divisions. Very briefly, both varieties supply those morphemes that are conceptually activated (content morphemes and those system morphemes that flesh out the meanings of their content morpheme heads – referred to as early system morphemes). Seen in this light, certain puzzles are solved; e.g. the French component of Michif (French nouns with their determiners) is not contradictory. The French nouns are content morphemes and their determiners are early system morphemes. However, generally the two other types of system morpheme that are structurally assigned come *only from one variety*. A cautionary note: this prediction is overturned if evidence points to a language shift as underway. Then the ML is structured by a Composite Matrix Language, meaning some structurally assigned morphemes are taken from both sources.

In closing, let me commend Matras for a far-reaching discussion of MLs that has both prompted us to see MLs in a new way and to re-evaluate our own positions more carefully.

References

Bolonyai, A. (2000). "Elective affinities": language contact in the abstract lexicon and its structural consequences. *International Journal of Bilingualism* 4 (1), 81–106.

Gross, S. (2000). When two become one: creating a composite grammar in creole formation. *International Journal of Bilingualism* 4 (1), 59–80.

Mous, M. (1994). Ma'a or Mbugu. In P. Baker & Mous, M. (eds.), *Mixed languages: 15 case studies in language interwining*, pp. 175–200. Amsterdam: IFOTT.

Myers-Scotton, C. (1993a). *Duelling languages*. Oxford: Clarendon.

Myers-Scotton, C. (1993b). Common and uncommon ground: social and structural factors in codeswitching. *Language in society* 22, 475–503.

Myers-Scotton, C. (1997). Afterword. In Myers-Scotton, *Duelling languages*, pp. 240–259. Oxford: Clarendon.

Myers-Scotton, C. (1998). A way to dusty death: the Matrix Language Turnover Hypothesis. In L. Grenoble and L. Whaley (eds.), *Endangered languages*, pp. 289–316. Cambridge: Cambridge University Press.

Myers-Scotton, C. (2000). A target in creole formation? No, two targets. Unpublished ms.

Myers-Scotton, C. & Jake, J. (1995). Matching lemmas in a bilingual language competence and production model: evidence from intrasentential code switching. *Linguistics* 33, 981–1024.

Myers-Scotton, C. & Jake, J. (2000). Explaining aspects of codeswitching and their implications. In J. Nicol (ed.), *One mind, two languages: bilingual language processing*, pp. 91–125. Oxford: Blackwell.

Schmitt, E. (2000). Overt and covert codeswitching in immigrant Russian children. *International Journal of Bilingualism* 4 (1), 9–28.

Wei, L. (2000). Unequal election of morphemes in adult second language acquisition. *Applied Linguistics* 21, 106–140.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

 Bilingualism: Language and Cognition 3 (2), 2000, 122–123 © 2000 Cambridge University Press

Symbiotic mixed languages: a question of terminology

NORVAL SMITH
Department of Theoretical Linguistics, University of Amsterdam, Spuistraat 210, 1012 VT Amsterdam, The Netherlands
E-mail: nsmith@hum.uva.nl

Matras makes use of a distinction introduced in Smith (1995) between "plain" mixed languages and "symbiotic" mixed languages. He interprets this distinction in the following terms: "*plain* mixed languages ... serve as everyday community languages, and *symbiotic* varieties ... are specialized varieties of a non-mixed language used in the same community, typically secret languages." I will take issue with two points here. Firstly, I do not consider that the distinction Matras makes here corresponds to the distinction I was trying to make in Smith (1995, forthcoming). Secondly, I will in particular take issue with Matras' characterization of secret languages as mere *varieties* of non-mixed languages.

In Smith (1995), I say the following on the subject of the various types of mixed language (p. 332):

> **Mixed languages** arise under conditions of bilingualism, when groups attempt to define, redefine, retain or even regain their ethnic status. This often results in a degree of language mixture. The most frequent type encountered is that where the grammar of one of the languages originally spoken in the group in question is combined with the content-words of another language known to the group ... The resultant language replaces the original ethnic language(s), and is in general the only language spoken. This type we will refer to as (plain) mixed languages (M).

and (p. 333):

> An important subtype of mixed language is what we term the **symbiotic mixed language** (MS). This type combines the grammatical structure of one language, and a varying number of lexical items – from hundreds to thousands in number – either from another language (often the original language of the group), or else from a variety of different sources, some words possibly being constructed or deformed deliberately. These languages exist in a symbiotic and dependent relationship with (dominant) unmixed languages with (virtually) the same grammar, and a lexicon from the same source as that grammar. This type of situation must presumably have pertained originally in the case of all mixed languages.
>
> An MS is by definition never the only language of its speakers. Often an MS will have the function of a **secret language**.

Matras draws a clearcut distinction between the two types – the "plain" mixed language which is the sole language of a community versus the "symbiotic" mixed language which is typically a secret language. He talks of systems involving a *lexical reservoir* used to conceal meaning "by manipulating key propositional items in key utterances". However, "the manipulation of key propositional items" does not seem to mean much more than the replacement of lexical stems as against the retention of the grammatical structure of the dominant community language.

Matras mentions the controversial status of symbiotic mixed languages as *registers* (quoting Mous 1994) rather than languages. However, he distances himself from the idea that these are actually languages, claiming in the case of Jenisch that: "*special lexicon* is inserted only sporadically, consciously, and deliberately in order to shift meaning, and so settings, from a serious, straightforward or factual interaction to one that challenges social and conversational taboos, and so might be less committing or even humorous." I would say in answer firstly that the actual domain or functions for which a mixed language is used is irrelevant when we are considering the question of its status as a linguistic code. I will return to the question of the nature of the code shortly. The example he gives of a conversation in Jenisch (11), does not differ essentially in nature from the two "isolated utterances" which he claims "will tempt 'descriptivists' into concluding that non-German [basic] lexicon is combined ... with German (Franconian) grammar".

It is the task of *descriptivists* to *describe* linguistic systems and structures. I refuse to see the term "descriptivist" as referring to some kind of invalid or illegitimate activity. An account of the functions any particular utterance or code is utilized for belongs to the field of *pragmatics*, which is a separate discipline concerned with the use that linguistic structures are put to. However, pragmatics does not concern itself with the identification of different linguistic systems – that is the prerogative of *descriptivists*. These are the linguists whose task it is to *describe* these codes. The use to which codes or particular utterances within them are put is quite a different kettle of fish.

As far as I can see there is little difference between the meaning of the term *register* and what Matras refers to by the terms *lexical reservoir*, *specialized variety of a non-mixed language*, or *special lexicon*. Let us examine what is meant by the term *register* more closely. The standard usage of the term *register* is for a special set of lexical items used by some group *within* a particular linguistic community.

Instances of *registers* are, for example, the technical vocabularies used in the various sciences, like chemistry, astronomy, and linguistics, in various trades, such as plumbing, joinery, or coal-mining, in connection with different hobbies, and so on. They do not involve basic vocabulary at all, but solely specialized vocabulary.

"Well then, that sounds not too different from the lexica of secret languages", one might say. "They are also used by groups in the community, just like scientists, tradesmen, and so on."

There are two essential differences, I will claim. Firstly, these users of symbiotic mixed languages are *external* to

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

their host communities, forming separate dependent communities within the host communities, unlike those groups, like scientists, which use *registers* containing specialized lexical items. These latter groups are *internal* to their communities, not *external* to them. Secondly, the lexical items in question are not the technical vocabularies that registers typically are, they comprise first and foremost a set of *basic* vocabulary items. There may be specialized technical vocabularies in a symbiotic mixed language in addition, just like in any language, but of prime importance is the set of *basic* lexical items. It is the combination of these two factors that qualifies these "languages" for this title: a separate ethnic identity, or sometimes a temporary ethnicity, combined with a different set of basic lexical items. A different cultural identity is combined with a linguistic code which is basically incomprehensible to the host community.

So, I would challenge Matras when he defines symbiotic mixed languages as "specialized varieties of a non-mixed language used in the same community". Here the word "community" is being used ambiguously or imprecisely. In the cases he is referring to here – the secret languages – there are in fact two communities involved: the dominant community which does not use the secret language, and a separate group with a separate ethnic identity, or in the case of peripatetic traders, a temporary, or maybe more correctly, a concealed ethnicity, which uses both the dominant language and their own secret language.

I will now turn to a consideration of two types of symbiotic mixed language. The *closed* symbiotic mixed language, and the *open* symbiotic mixed language. The closed type does not make direct use of the lexical words of the dominant language, and the open type does. The secret languages discussed by Matras are typically of the open type. These have a circumscribed set of lexical items, varying in number from hundreds to thousands. The number appears to correlate with the age at which the symbiotic mixed language is acquired – the earlier the acquisition the larger the number of lexical items. Any meaning element which does not occur in the open mixed language is expressed by utilizing the resources of the dominant language of the host community.

Closed mixed languages do not in principle make use of the content words of the host language at all. Examples of closed mixed systems would be Ma'a (Inner Mbugu (Mous, 1994)) – a mixed language of the East African Mbugu – and the so-called mother-in-law language of the Australian Dyirbal – Dyirbal Dyalŋuy (Dixon, 1972). The functions of these are, in the first case, to mark the Mbugu off ethnically from the neighbouring Pare who share one Bantu language with the Mbugu – Pare and the so-called Normal Mbugu are closely related dialects – and in the second case, an avoidance language used in the presence of particular relations, such as mothers-in-law.

Every utterance in Normal Mbugu has an Inner Mbugu equivalent according to Mous. In the case of Dyirbal Dyalŋuy, the situation is different. Due to its smaller lexicon recourse must be had to circumlocutions and semantically vaguer lexical items (Dixon, 1972).

These closed symbiotic mixed languages form a bridge to the *plain* mixed languages. For these last are closed, must themselves have arisen under conditions of bilingualism, and must have passed through a period of symbiosis with non-mixed languages. However, this is not the place to go into the question of the formation of the various types of mixed language. For this see Matras' article itself, and Smith (forthcoming).

References

Dixon, R. M. W. (1972). *The Dyirbal language of north Queensland* (Cambridge Studies in Linguistics 9). Cambridge: Cambridge University Press.

Mous, M. (1994). Ma'a or Mbugu. In P. Bakker & M. Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, pp. 175–200. Amsterdam: IFOTT.

Smith, N. (1995). An annotated list of creoles, pidgins, and mixed languages. In J. Arends, P. Muysken & N. Smith (eds.), *Pidgins and creoles: an introduction*, pp. 331–374. Amsterdam: Benjamins.

Smith, N. (forthcoming). Younger languages: genetically altered. In P. Bakker and F. Field (eds.), *SPRAU: Working Papers of the University of Aarhus (Proceedings of the Workshop on Mixed Languages)*.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

Social context, structural categories and medieval business writing

LAURA WRIGHT
Lucy Cavendish College, University of Cambridge, Cambridge CB2 0BU, UK
E-mail: lcw21@cam.ac.uk

In my response I will consider Matras' observation that "words are not simply words, but are divided into function-orientated categories" from the perspective of Medieval Latin/Middle English and Anglo-Norman/Middle English business texts, as written in Britain from the Norman Conquest to the sixteenth century. I will concentrate on his assumption that the social context of speakers plays a defining role in the resultant language outcome.

Matras presents a discussion of the context in which Lekoudesch and Romani are spoken, detailing the way in which the constraints of the social context (the presence of outsiders) dictate the emergent language mixture. This discussion is given to demonstrate processes of lexical re-orientation, and the function is, apparently, not only to preserve secrecy but also to demarcate group identity.

I am particularly interested in the notion of demarcating group identity. My viewpoint is entirely from that of a written text-type; there is no evidence (either way) that medieval business writing was ever a spoken variety. Medieval Latin and Anglo-Norman became the main languages of governmental administration in England when it became annexed to the Anglo-Norman empire as a result of Duke William of Normandy's succession to the English throne in 1066. His census of his new holdings, the Domesday Book, was written in a Medieval Latin containing some content words in English (presumably as these were technical terms referring to cultural items which it was felt important to retain), although it is not always possible to determine which elements belong to Medieval Latin and which to Old English, as in (1).

(1) in *Eldehā ten̄ Goisfrid*^{9} *de* bech *sub abbe j* hi*d ̄ra ē j car̄ sed deest car̄ ibi ij* co*t̄ silua C porc̄*
in Aldenham holds Geoffrey of Bech under the abbot one hide land for one plough but lacks plough there two cottagers woodland 100 pigs
"In Aldenham Geoffrey of Bech holds 1 hide of land rented from the Abbot; there is land for one plough but the plough is missing; there are two cottagers and woodland and 100 pigs"
(Domesday Book, Hertfordshire, 1086; Wright, 1998: 102)

"Hide" is an Old English word denoting a unit of arable land, usually considered to be the equivalent of 120 acres but subject to wide local variation. The form "co*t̄*" could be "cot", "cottage", "cottager", or "cotarius" – it is simultaneously all of these. The English element in administrative and business writing never went away, and by the fourteenth century had become codified with all linguistic components appearing in either Medieval Latin or Anglo-Norman (it is not always possible to separate the two, as Medieval Latin as written in England became increasingly coloured by Anglo-Norman semantics); and any and all verb roots, nouns, *-ing* forms and certain lexicalised phrases appeared, optionally, in English, as in (2).

(2) in hewyng & apparelyng assheler *pro le* new peere *iux*^{a} *finem borialem ex parte orienti pontis*
in hewing and apparelling ashlar for the new pier next end north of part east of-bridge
"in hewing and dressing ashlar (stone) for the new pier next to the northern end of the eastern part of the bridge"
(1471–72, London Bridge Accounts; Wright, 1995, 368)

There was another, visual, component to this text-type, namely, the visual fusion of morphological information by means of the medieval abbreviation and suspension system.

(3) *Itm̄ solut̄ ꝑ C* pyles *vlni a* Walth^{a}m *vsqʒ* breghous *ꝑ aq*^{a}*m* carian*ꝺ*
item paid for 100 piles elm at Waltham to bridgehouse by water carrying
"And paid for carrying 100 elm piles from Waltham to the Bridgehouse"
(1382–83, London Bridge Accounts; Wright, 1995, 366)

Carianꝺ is both Middle English and Medieval Latin (one of the participial morphemes in Middle English was {*-and(e)*}, which was later superseded by *-ing*). Further morphological information, which would assign it categorically to one or the other language, and which would be expressed orally, is suppressed visually by the abbreviated *ꝺ*. Calques occurred in close proximity as in (4)–(7): there is no suggestion that this was a secret language.

(4) Margarete Spenser Selkwomman *de* So*p*slane Londo*n̄ doit ꝑ . . . dune ℔ de soy crude . . .*
Margaret Spenser silkwoman of Sopers Lane London owes for one pound of silk raw
(1392, Gilbert Maghfield's Accounts; Wright, 1998, 107)

Both *selk* and *soy* occur within a few words of each other; this is a very common phenomenon in this text-type. Similarly, grammatical morphemes could be calqued, as in (5)–(7):

(5) *Itm̄ solut̄ ꝑ st*^{a}*mīe empt̄ ꝑ* dobynge
and paid for straw bought for daubing
(1380–81, London Bridge Accounts; Wright, 1995, 370)

(6) *Itm̄ solut̄ ꝑ stramīe empt̄ ꝑ dobatura*
And paid for straw bought for (?future) daubing
(1380–81, London Bridge Accounts; Wright, 1995, 370)

(7) *Itm̄ solut̄ cuidam dobatori ꝑ molendino apud* Stratford *vocat̄* Spyle*m̂*anesme*ꝺ dobanꝺ*

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

and paid to certain dauber for mill at Stratford called Spylemanesmell daubing
"And paid to a certain dauber for daubing the mill at Stratford called Spylemansmill"
(1382–83, London Bridge Accounts; Wright, 1995, 370)

dobynge is monolingually English, *dobanḍ* is both Medieval Latin and Middle English, and *dobatura* is monolingually Medieval Latin. The presence of calques means that it was not simply a process of lexical borrowing but more a process of codeswitching, although strings cannot always be analysed in terms of which language their elements belong to, due to the fusing effect of the abbreviation and suspension system.

This text-type was mostly used by merchants, accounts-keepers, stewards, salary clerks, customs officers – anybody who dealt in the ingress and egress of commodities for cash. Seen from an original Anglo-Norman empire perspective, such mixing connoted regional locality. Clerks were at first lexically re-orientating away from the dominant romance languages and turning what was at first perhaps a cultural necessity (how do you translate x into romance?) into a variety that was all their own – as distinct from the other Anglo-Norman empire varieties written in France. This subsequently became bureaucratised and fossilised, and remained stable until its disintegration as a linguistic system around 1500.

Medieval business writing behaved like a mixed language in the sense that it didn't just borrow technical terms from English but used basic vocabulary items too, as in the process of lexical re-orientation. But to what extent is such medieval business writing like present-day mixed languages? During its five hundred year-odd life it cannot be regarded as a mixed language in that the optionality rule (i.e. content words had to be in romance or English, but never *all* in romance or *all* in English) meant that the content words never amounted to anywhere near 100 percent English. The exception was at the point of its demise.

(8) *It᷑ p^{a}e a* laborer^{a}s$\jmath$ *p^{r}* serchyng$_{\varrho}$ off grondys & cowchyng$_{\varrho}$ vp of sto$\hat{n}$ off the same grounde
"And paid to labourers for searching of grounds and couching up of stones off the same ground"
(1432, Grocers' Company Records; Wright, 1998, 108)

This is beginning to look more like a mixed language, but it was highly ephemeral, symptomatic of language death, and the grammatical structure (i.e. word order) became anglicised too. Further, in medieval business writing there never was any vocabulary replacement – the use of calques demonstrates that this text-type was not the result of vocabulary loss (either of English, or, when the Anglo-Norman administration ceased, of Anglo-Norman and Medieval Latin). It shares with mixed languages its property of stability (around five hundred years) and the fact that it had a tightly defined social function.

Matras considers whether there is any functional reason for the structural compartmentalisation of mixed languages and those varieties that share some of their characteristics, such as Lekoudesch and medieval business writing. In terms of motivation, medieval business writing would indeed seem to suggest that there was a pragmatically motivated, social reason for including English in one's Latin and French. It demarcated a particular group of writers from a particular part of the territory. In terms of structure, medieval business writing contains English content words (Matras' category of lexical re-orientation: whilst not wholesale, any content word was eligible for an English realisation), a romance grammar, somewhat reduced and fossilised over time (selective replication), and considerable hybridisation (convergence). However, Matras' category of fusion cannot be identified because it is not possible to say whether any elements were conscious or deliberate, and this forms part of his distinction between lexical re-orientation and fusion. One could, I suppose, guess that particles like "item" and "per" were indeed below the threshold of consciousness, as they are still with us in this register (e.g. when talking about the cost per item); and thus contrast with a word like "pyles" in (3) above, where "piles" was also written in Latin on the same folio and thus bespeaks a conscious choice. But Matras' comments with regard to a hierarchy of likelihood of structural categories cannot be properly addressed from the viewpoint of this text-type, as it was not, so far as we know, a spoken form and does not contain many of the discourse-regulating constituents he lists, like discourse markers, hesitation markers and interjections. Nonetheless, in terms of seeking an underlying mechanism to the multifarious forms of mixed and near-mixed languages, Matras has provided us with an extremely interesting hierarchy of probability based on structural categories, and the fact that medieval business writing cannot bear them all out to the letter does not, I think, invalidate his suggestions. If, as he claims, "we can explain why lexical re-orientation will target items that express meaning, and leave the grammar component of the language largely intact", then the application of structural categories is a powerful tool.

References

Wright, L. (1995). Middle English *-ende* and *-ing*: a possible route to grammaticalization. In J. Fisiak (ed.), *Linguistic change under contact conditions*, pp. 365–382. Berlin: Mouton de Gruyter.

Wright, L. (1998). Mixed-language business writing: five hundred years of codeswitching. In E. H. Jahr (ed.), *Language change: advances in historical sociolinguistics*, pp. 99–118. Berlin: Mouton de Gruyter.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

 Bilingualism: Language and Cognition 3 (2), 2000, 126–129 © 2000 Cambridge University Press

AUTHOR'S RESPONSE

Back to motivations: between discourse and grammar

YARON MATRAS
Department of Linguistics, University of Manchester, Manchester M13 9PL, UK
E-mail: *y.matras@man.ac.uk*

Three principal issues are raised in the commentaries to which I would like to respond. The first involves the role of the study of codeswitching in explaining stable Mixed Languages, the second pertains to the multiplicity of the linguistic processes through which MLs arise, and the third concerns functional methodology.

Cases of stable MLs differ from cases of codeswitching mainly in that the former can be regarded as conventionalised systems, where language alternation is no longer a synchronic option but is merely reflected in the etymology of the various components. Naturally, the conditions for the emergence of an ML depend on the ability of speakers to draw on different languages, at least to some degree and at an initial stage. This, however, does not in itself help explain the different profiles of MLs, and so it does not, as Gardner-Chloros implies, make redundant the effort to define the linguistic mechanisms that participate in the creation of MLs. On the other hand, what Auer has referred to as sedimentation of language mixing need not result in what the mixed language literature generally views as an ML. One of Auer's (1999) examples for a fused lect is Sinte Romani, which, though extreme in its extent of grammatical and lexical borrowing from German, remains indisputably Indo-Aryan in its bound morphology, core vocabulary, and representation of salient grammatical categories such as deictics, personal pronouns, existential and possessive expressions, and definite and indefinite articles, and so it does not qualify as an ML. Admittedly, definitions of MLs in the relevant literature have so far been largely intuitive. We still do not know just how much, and what kind of, mixture a language must contain in order to qualify as Mixed or as genetically non-classifiable, and we still have no catalogue of structural mixing patterns that are "permissible" for a non-mixed language (both issues are currently being investigated as part of the Manchester research project on "Structural and Functional Constraints in Cases of Stable Mixed Languages").[1]

The present discussion has nonetheless brought up a connection with codeswitching surrounding the role of individual mixed utterances in symbiotic mixed languages that are not self-contained (see in particular the comments by Auer, Moyer, Blommaert, and Myers-Scotton; Backus addresses structural constraints, but the issues he raises belong here too). The switch at the discourse level here is between a system that is not mixed and one that is mixed. And while I do not believe that it is helpful to flag conversation analysis as an alternative to the idea of form–function correlations in grammar, as Moyer wishes to do, a conversation analytical approach is certainly called for when attempting to explain the discourse position of such mixed utterances in what is otherwise a non-mixed discourse interaction. (Mous's remarks show how difficult it is to choose proper examples given the fragmented nature of the documentation of MLs.) The fundamental distinction between non-self-contained mixed languages and those that are self-contained is indeed a principal issue in my paper. The question arises of why the former should be discussed at all in connection with stable, all-purpose and self-contained MLs.

The answer lies partly in the fact that the mixed utterances encountered in non-self-contained MLs are not, like classic codeswitches, activations of L2 elements, but activations of L2-origin elements that have already undergone shifts through a process of mixing and have since become conventionalised. (This is the case despite the tendency in secret argots and trade jargons towards flexibility, adaptability, and on-the-spot lexical creativity.) Thus it is the nature of the mixed register itself, which provides the target for switching at the discourse level, that is of interest. Moyer's suggestion that Caló *mansa camelo tuque* "I love you" might be stereotypically reproduced from inflected Romani must be rejected on the grounds that Caló possesses only petrified Romani inflection forms: the original sociative for the 1sg pronoun *mansa*, the 3sg inflected form of the verb *-el* to which Spanish concord is added, and the original dative for the 2sg pronoun *tuque*. It is clearly not the phrase itself that is lifted from Romani (cf. Romani *me kam-av tu-t*), but the individual items, which form part of a fixed reservoir and can be composed to create new phrases. In Lekoudesch, the expression *Schochamajim* for "coffee" (Hebrew *šaxor* "black", *mayim* "water") is not a spontaneous composition, but a fixed part of the lexical reservoir, attested in numerous sources covering two centuries. Yet it is not lifted from Hebrew, except in its individual components; rather, its origin will have been in a playful or deliberate attempt to disguise meaning. To sum, then, mixed utterances in non-self-contained MLs are not codeswitches directly containing L2 material, but switches into a mixed code to which this L2 has contributed.

A further consideration is the point on the mixing continuum at which an ML emerges. Auer's (1999) model foresees conventionalised, self-contained fused lects as the final point on this continuum, with stylistic mixing at a point preceding "fusion" (in his sense of the term, which differs considerably from mine). The stylistic distribution of mixed utterances in some non-self-contained MLs gives room to consider whether a heavily mixed variety might in

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

fact be reduced functionally to a discourse-level register of the Ma'a type, and later to the occasional mixed insertion at the utterance level, of the Para-Romani type (cf. Smith's new distinction between "closed" and "open" symbiotic MLs), while the motivation to maintain a mixed register is reinforced by the wish to preserve a means of in-group communication. Myers-Scotton dismisses both the suggestion that speakers may associate distinct extralinguistic contexts and values with the choice of items from a particular language (calling this a "Whorfian" view), and the idea that speakers might show a social motivation to keep alive impressions of an ancestral language. But there is plenty of evidence that peoples' feelings towards languages may lead them to make semi-deliberate efforts to manipulate language use, and that these efforts may result in stable mixtures. Insights into the ongoing process have recently been provided at the idiolect level by Maschler (1998) and others (see also the comments on this issue by Mous and by Gardner-Chloros). At any rate, the crucial distinction between non-self-contained MLs and code-switching is that the former maintain a reservoir for mixed insertions despite the fact that speakers have no access to an autonomous L2.

I now turn to the issue of the processes involved in MLs. The literature on MLs shows at least two attempts to account for the (etymologically) mixed structural profile of MLs through a single, wholesale process: Bakker's (1997) Intertwining model, and Myers-Scotton's (1998) Matrix Language Turnover hypothesis. The two approaches agree on a natural, cognitive compartmentalisation of linguistic structures by which lexicon is distinguished from grammar. Bakker even explains the fact that women in mixed communities supply the grammar as a reflection of their (by implication, cognitive) influence in bringing up children, although in his present commentary he rejects the notion of a native versus an intruding language and argues for hierarchically equivalent "mother tongues". The two approaches differ however on the specific breakdown of grammatical categories. Bakker regards function words as a separate though wholesale category, and the recruitment of certain function words from one or another source language into a given stable ML as arbitrary. Myers-Scotton's model in its traditional form (prior to the "4-M model") doesn't recognise an internal division within system morphemes, and classifies function words that pattern with the lexifier language as content morphs. Both views may be interpreted as attempts to smooth out the grammar–lexicon dichotomy.

My point of departure was to consider whether there are any relevant correlations between functional categories in MLs, and my argument was that the division of categories by source language is not entirely accidental. This has in the meantime been confirmed through further work on a sample of MLs: mixed languages have a clear and straightforward preference for one source language when recruiting finite verb inflection or INFL, though INFL need not necessarily coincide with the source language for nominal inflection (as in Michif and Mednyj Aleut). I am in full agreement with Myers-Scotton that this provides sufficient evidence for there being a base language in each and every ML, though my argument pertains to the crucial status of the predication in the utterance; it is not entirely clear to me whether Myers-Scotton's notion of Matrix similarly emphasises the predication. Symbiotic argots (or secret languages) that have not had an ancestral community language involved in their formation show no L2-origin pronouns or deictics, but they may show distinct negators and copulas, while in-group symbiotic languages that do involve keeping alive impressions of an ancestral language tend to show pronominals and deictics. (This answers Auer's question, and the confusion addressed by Mous, concerning the difference between Jenisch and Para-Romani in regard to deixis.) While all three groups of categories – INFL, negators and copula, pronouns and deictics – are not usually obvious candidates for grammatical borrowing in more conventional situations of contact, clause connecting elements and utterance modifiers (such as focus particles or phasal adverbs) are generally prone to borrowing and their L2-origin is not at all unique to stable MLs.

Targeting the lexicon is indeed a prominent feature of MLs. The ability to do so definitely has to do with the relative ease of cognitive retrievability as well as structural detachability of lexical material. There are two principal differences between what I call "lexical re-orientation" and what has been referred to in the literature as "re-lexification'. The first is an important formality: re-lexification is used to refer to the substitution of a morph through an equivalent morph, irrespective of the system position or function of this morph, hence Muysken's (1997) view that Spanish discourse markers in Media Lengua form exceptions to the re-lexification model since they do not correspond to Quechua markers (which are often bound, and are not necessarily replaced at all). Lexical re-orientation, however, pertains only to those content elements that convey meaning (Bühler's "symbolic field"). The second difference lies in my understanding of the motivation for the process. In fact, only a more general sociolinguistic motivation, if any, is associated with re-lexification: to create or preserve a mixed language. My definition of lexical re-orientation is more specific: the targeting of meaning in language mixing, metaphorically speaking the structural devolution of the forms that convey meaning, allows us to associate distinct values with mixed segments of speech, be they the copying of a prestigious culture or the concealment of meaning (intentions, attitudes, and information) from outsiders.

If we include symbiotic MLs, it is clear that the wholesale adoption of a foreign lexicon does not automatically entail mixing around portions of grammar. Once again, Myers-Scotton claims to be able to predict this through the MLF model, while Bakker leaves the class of function words prone to coincidence. I argue that the fact that deictics and pronouns only appear in MLs whose speakers

[1] In collaboration with Anthony P. Grant, and supported by a research grant from the British Academy's Arts and Humanities Research Board.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

have had access to a full-fledged ancestral language calls for a separate process that is beyond pure lexical re-orientation (even though it may entail the former, or combine with it), and which I call selective replication.

Both Bakker and Myers-Scotton take issue with my notion of fusion. I argue that fusion is a distinct process, because it follows its own regularities, and because it is, once again, not implied automatically by other forms of mixing such as replacement of the lexicon, replication of deictics, or convergent re-arrangement of word order rules, but may either occur independently of those, or accompany them. The fact that Michif adopts an English form for "but", but not for "or" or "and", is in line with my hierarchical understanding of fusion (but > or > and, for coordinating conjunctions), although this admittedly does not help explain the process of emergence of Michif. Exceptions to the hierarchy, as provided by the Cree-derived items, are not in principle inconceivable, though, and they could be connected to the structural nature of the resources involved. For instance, a conjunction with an underlying deictic meaning may be resistant to replacement, a prediction which is entirely consistent with the principles of fusion outlined in my 1998 paper. I remain unconvinced that the fusion of discourse operators can or should be accounted for by re-defining the items in question as content morphs (when switched, and otherwise as system morphs), as Myers-Scotton proposes. As I described in my 1998 paper, fusion is a gradual and gradational process. Why should prolonged contact turn "but" from a system to a content morph? And why should "but" need less time to become a content morph than "or"? However useful Myers-Scotton's content–system split may be in other respects, the insistence on just two classificatory categories (prior to the latest modification of the model in the form of the "4-M model") obliterates the relation between the individual system morphs in question and the distinct language processing functions for which they are each responsible.

This brings me to the final and perhaps most difficult area of controversy, that of the functional approach taken in the paper. Mixed languages arise in different socio-linguistic constellations, they may serve different social and communicative functions, and their structural profiles or "compartmentalisations" – the fact that some structures derive from one language, others from another – are not identical. I did not pretend to be able to illuminate new historical facts about the relevant communities, nor did I pretend to present an integrative theory of the genesis of MLs (since I do not believe that they arise through a single process). My aim was certainly not to distance myself from descriptive accounts of MLs or to discourage further descriptive work; on the contrary, given the fragmented documentation of MLs more detailed descriptions are vital for any theoretical discussion of the problems that MLs raise. Rather, the question I tried to pursue was, whether one might be able to make any sense of the structural profiles by taking into consideration overall functions and contexts of emergence, as described for the sample MLs in the literature. Admittedly, this is an ambitious endeavour and the issue cannot be resolved by drawing on just a small sample of languages and within a single paper, if it can be at all.

"Function" in the context of my paper is the notion that the distribution of a linguistic structure or form, and therefore also its behaviour in contact situations, is a product of what speakers universally wish to achieve by using this form, in terms of information processing strategies. This is what is meant by a natural or cognitively motivated compartmentalisation of linguistic structures. I take Bühler and the Bühlerians as an example of an approach that is consistent in associating forms with functions in this sense. Auer is right in pointing out that the Bühlerians of the Ehlich/Rehbein school have never taken the time to display their model in full, yet I find some of their ideas nonetheless inspiring. Irrespective of the limitations of the Bühlerian-type approach, I maintain that the constraints which operate on content-lexical material differ from those that operate on deictics, which in turn differ from those on existentials and negators, which differ from those on discourse markers, which differ from those on finite verb inflection, and so on. This is not accidental, but has to do with the functions that these elements serve in communication, whether monolingual or bilingual, and so with their position in the linguistic system.

On this basis we might be able to make predictions such as the following. If a community develops a mixed register for the purposes of in-group identity flagging, for secret communication, or for trade, and if this community has not at a recent stage shared a distinct all-purpose community language, then the likely process by which the mixed code will develop is "plain" lexical re-orientation, and the outcome is likely to be a reservoir of content-lexical items, perhaps also negators and existentials. This serves a purpose: to mark out content-meaning, and nothing else. If a community's in-group symbiotic mixed code is based on an ancestral all-purpose language, which has lost its role as an everyday language through functional turnover, then we are likely to find additional categories, most notably deictics and pronominals. This can be attributed to the residual presence of impressions of active, situation-bound speech which speakers try to replicate, and to the prominence of deixis and anaphora in situative communication. (Unlike Myers-Scotton's Matrix Language Turnover hypothesis, my notion of a functional turnover pertains specifically to such instances; the two notions partly overlap in the case of Ma'a, which is Myers-Scotton's case study. I do, in fact, refer to her work in the beginning of my paper.)

These two processes are perhaps the most radical, and the least likely to be encountered in more conventional cases of contact, in that they involve a wholesale replacement of the lexicon. If a situation of active bilingualism prevails, then individual occurrences of morphosyntactic convergence may occur, which are indeed common in other situations of bilingualism as well, and which may further complicate the structural profile of the ML. If, in addition, contact prevails with a dominant pressure-exerting language (Spanish in the Media Lengua case, Russian in the Mednyj Aleut case, Mbugu/Pare in the Ma'a case), then

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press

fusion is likely to set off in the predictable hierarchy, affecting clause-combining and utterance-modifying expressions, as in other cases of contact with a dominant language. If, finally, the ancestral language is beginning to lose ground, then fusion may take an even more extreme appearance, affecting even INFL, as in Mednyj Aleut. The essence of the functional argument, then, is that, at least to a considerable extent, the structural profiles of MLs are not entirely coincidential.

There remain of course numerous issues to be resolved. One point, mentioned by Bakker and by Moyer, is the exact distribution of existential expressions. (Bakker's observation on a split between the locative and existential copula might apply to Lekoudesch, but not to Jenisch or Caló.) Another which could be reconsidered is the position of quantifiers, which I have tentatively attributed to situative speech, but where different outcomes might be expected in some contexts, such as trade jargons. In addition there may be semantic or structural constraints that cut through functional categories, such as a split between singular and plural pronouns, or, as in Mednyj Aleut, between emphatic and clitic pronouns (the first pattern with deictics, and second with INFL). This is to say that the functional compartmentalisation idea and the four mechanisms outlined in the paper by no means pretend to account for each and every sub-component of an ML's grammar, though the intention was to provide a blueprint to explain the principles of the respective structural profiles.

Mixed languages may be said to have become one of the most controversial issues in the fields of contact linguistics and historical linguistics (see for example Greenberg's (1999) recent contribution, which challenges the existence of MLs altogether). While it is possible to survey data already discussed in the literature, it remains an extremely difficult and risky task to interpret those data: the situations are highly diverse, they involve many combinations of languages and social and historical facts about communities that are often difficult to access. Though the task might become somewhat easier should empirical investigators of individual MLs follow Bakker's (1997) example and present us with detailed monographs on the languages, the debate is likely to remain challenging for quite some time. I wish to conclude by thanking my critics for the attention which they have awarded my ideas.

References

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

Bakker, P. (1997). *A language of our own: the genesis of Michif – the mixed Cree–French language of the Canadian Métis*. New York: Oxford University Press.

Greenberg, J. (1999). Are there mixed languages? In L. Fleishman et al. (eds.), *Essays in poetics, literary history and linguistics*, pp. 626–633. Moscow: OGI.

Maschler, Y. (1998). On the transition from code-switching to a mixed code. In P. Auer (ed.), *Code-switching in conversation: languages, interaction and identity*, pp. 125–149. London: Routledge.

Matras, Y. (1998). Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36, 281–331.

Muysken P. (1997). Media Lengua. In S. G. Thomason (ed.), *Contact languages*, pp. 365–426. Amsterdam: John Benjamins.

Myers-Scotton, C. (1998). A way to dusty death: the Matrix Language turnover hypothesis. In L. A. Grenoble & L. J. Whaley (eds.), *Endangered languages: language loss and community response*, pp. 289–316. Cambridge: Cambridge University Press.

https://doi.org/10.1017/S1366728900270211 Published online by Cambridge University Press