

*Peter Auer (Freiburg)*

## Language mixing and language fusion: when bilingual talk becomes monolingual\*

### 1. Introduction

After a time of intense and partly controversial discussion, the topic of “mixed languages” seems to have settled on a definition of **mixed languages as languages with a “split ancestry”** (Mátrás and Bakker 2003: 1; Bakker 1996: 13; Golovko 2003: 191; Thomason 2003: 21). There is no agreement, however, on whether these mixed languages should be seen as strictly delimited from “unmixed” languages (which just show some ‘ordinary’ borrowing) or whether they are merely extreme cases on a continuum whose other extreme is ‘ordinary’ borrowing as it occurs in every language. The first case, advocated most strongly by Bakker (e.g. 2003), stipulates that **mixed languages “must have numerically (roughly) equal and identifiable components from two (or more?) other languages”** (2003: 108–109); only in this case would the above definition hold true in a strict sense. If the number of components from the two languages are not “(roughly) equal”, a mixed variety may still **be said to have language A as its (main) ancestor, even though it contains any number of elements from another language B**. **The continuum view, on the other hand** (propagated by the present author and also by e.g. Thomason 2003: 21; Myers-Scotton 2003: 101), argues that there is no clear boundary between mixed languages and ‘ordinary’ borrowing. Mixed languages simply represent a rare case of massive borrowing from one language to another. This view implies that the split ancestry definition must be gradient: A mixed variety may be related to two ancestor languages, but in most cases, one of these ancestor languages will be the dominant (real?) one.

I will follow the gradient view of mixed languages in this paper and argue that there is nothing special or extraordinary about mixed languages from a grammatical point of view. **The individual linguistic processes involved are those we know from language contact in general**. What is special in the radical cases of mixed languages is their **sociolinguistic status and history, i.e. the circumstances which led to such radical borrowing**, sometimes within a

---

\* I wish to thank Victor Friedman and Carmel O’Shannessy for comments on an earlier version of this paper.

short time-span. Very often (Matras 2009: 304), this happens in situations of language shift and language obsolescence, in which elements of the former group language are maintained by fusing them with the dominant, out-group language. As a rule (but not always) the sociolinguistically dominant language is also the structurally dominant one, i.e. the new variety is genetically speaking a variant of the shifted-to language, rather than a descendent of the lost or endangered language.

GA?

Looking at mixed languages as an extreme case of borrowing implies a shift from the result (a mixed language) to the process of their emergence (extensive borrowing). In order to underline this shift of perspective, I suggest the term *fusion* for the processes of extensive borrowing of elements from another language into the receiving language, and the term *fused lect* to refer to its outcome. While the usual suspects for radically mixed languages are a few, partly insufficiently documented varieties whose status is often disputed due to a simple lack of data, the focus on fusion as a process opens up the possibility of investigating examples of partial fusion as well. In most of these cases, the ancestor language question will not be central; they can clearly be attributed to one ancestor language much more strongly than another one. What is central is the structural conventionalization and indeed grammaticization of bilingual mixing patterns.

The main argument of this paper is that the basis of language fusion is *language mixing* (often also called *code switching*), a phenomenon which we need to separate strictly from “mixed languages” (in the above sense of fused lects in their extreme form), and which has been abundantly documented in bilingual studies. Language mixing is a surface phenomenon which consists of the juxtaposition of two languages. It will be argued that most cases of fusion originate from such mixing (even though other routes cannot be excluded in principle). This leads to the most important question: What kind of surface patterns of bilingual talk become part of the language structure of a fused lect?

The main argument I will put forward here is that language fusion is the result of the conventionalization of the two basic strategies of insertional mixing. These two basic mixing strategies have been amply documented in studies on bilingual talk, where they may occur in isolation or in combination. The first strategy, called *minimal insertion* here, consists of inserting stems (or uninflected words) from one language into the grammatical frame (“matrix”) of the other language without any accompanying grammar. In this case the lexical material from the inserted language is accommodated completely to the grammatical structure of the matrix language. (Phonology and phonetics are not necessarily affected.) The second strategy, called *maxi-*

mal insertion or “embedded island” insertion here, prefers to bring the other-language item(s) into the matrix frame together with the accompanying grammatical markers (affixes or grammatical words) and often extends to the phrase level.

I will argue that these two strategies are the equivalents of the two main patterns of language fusion which can account for the structure of radically fused (“mixed”) languages. Minimal insertion in mixing is the basis of the structure of “symbiotic” fused lects, in which the fusion affects the lexicon of one language and the grammar of another. The grammar is entirely in language B, and lexical items (usually stems) are selectively inserted into this pattern (examples are *Media Lengua*, or *Para-Romani*). Maximal insertion, on the other hand, accounts for the second group of fused lects in which two grammars merge.

The idea that fusion arises out of mixing is not new (see Myers-Scotton 1998, 2000; Auer 1999; Thomason 2001: 131–133). It has sometimes been criticized on the basis that “no documentation of a transitory phase between the supposed CS behavior preceding the mixed languages” existed, as Bakker (2003: 129), one of its strongest opponents, writes. In the meantime, however, there is historical evidence for such a transition (McConvell and Meakins 2003; McConvell 2008; O’Shannessy 2012; Schaengold 2004). In this paper, additional evidence will be given by approaching the issue in a somewhat indirect way by showing that the structural similarities between language mixing and fused varieties are so substantial that it is highly likely that the former provide the basis of the latter. Intermediate stages of fusion will be presented which provide a better understanding of how fused lects emerge.

With regard to the overarching topic of this volume, it should be noted that the notion of split ancestry and hence the concept of language fusion is theoretically speaking orthogonal to that of family resemblance, i.e.: language fusion should occur regardless of whether the two languages in contact are genetically related or come from different language families. However, a look into existing research on “mixed languages” shows that by far the majority of cases discussed so far come from language contact between unrelated or genetically hugely distant languages. This raises the question of whether the same concept of fusion can be applied to closely related varieties as well or whether fusion contradicts family resemblance. I will briefly (and somewhat speculatively) address this question in the last section of this paper.

## 2. Fused lects in their extreme form (“mixed languages”)

An overview of the better-known extreme results of fusion – the so-called mixed languages – is given e.g. in Matras (2009: Ch. 10.3.), Thomason (2001: Ch. 8.), Matras and Bakker (2003) and Thomason and Kaufmann (1988). It will therefore suffice here to summarize some of their structural features. The sociolinguistic motivation in all cases can be said to be identity-related. The fused variety arises out of a need to serve as a symbol of ethnic or group identity under circumstances of heavy pressure from majority groups which force the community to shift language. The fused lect usually, but not always, appears to be a variant of this shifted-to language into which elements of the abandoned minority language are borrowed. Deliberate strategies to create such a linguistic identity symbol through language manipulation seem to have played a role in at least some of these cases. This may explain the rapidity under which some mixed languages have emerged.

Bakker’s (2003) contention that the prototypical mixed language is one in which a split between vocabulary and lexicon occurs (“language intertwining”)<sup>1</sup> has been refuted by Matras already (2009). I suggest distinguishing instead between two basic types of fusion. One of them indeed comes close to Bakker’s “language intertwining”. It consists of the borrowing of lexical material only into the grammatical frames provided by the other language. No grammatical affixes and only some grammatical words are borrowed alongside the lexical items (words or stems). The second type involves the fusion of a part of the grammar of one language with another. As a consequence, a grammatical split may occur, such that, for instance, the verbal system – including stems and morphology – comes from one language, and the nominal system from another, often together with the relevant parts of the vocabulary.

GA?

Good examples of the first type are the following languages or varieties:

- *Media Lengua*, a variety of Quechua, spoken by ethnic Quechua in Quito, Ecuador, in which Quechua grammar combines with 90% Spanish-based vocabulary. The Spanish words are fully integrated into Quechua grammar and even phonology (Muysken 1981). Lexical borrowing includes grammatical words such as pronouns and negators. In addition to this “relaxification” (Muysken), copying of Spanish discourse markers occurs which have no equivalent in Quechua.

<sup>1</sup> Also cf. Winford (2003: 24), who defines mixed languages as languages incorporating “large portions of an external vocabulary into a maintained grammatical frame”.

– *Ma'á* (or *Outer Mbugu*), a register used by a community in Tanzania (Mous 2003). The lect consists of a parallel lexicon of mainly Cushitic words superimposed on the grammar of the Bantu language Mbugu. The difference between *Ma'á* and *Media Lengua* is that in addition to grammatical words, *Ma'á* also has some structural non-Bantu features, such as a voiceless lateral fricative apparently used to overtly distort Bantu words (phonology), and possessive affixes and constructions (morphology), among others. The exact history of *Ma'á* remains unclear due to a lack of data.

– *Anglo-Romani* (and other Para-Romani varieties), a variety of English into which some hundred Romani words have been borrowed. *Anglo-Romani* is used as a secret code and as an in-group way of speaking (Boretzky and Igla 1994). Some grammatical words are also borrowed from Romani. *Anglo-Romani* also shows some deviations from autochthonous English (simplifications due to the omission of grammatical words such as determiners), but these are not due to contact with Romani (cf. Matras 2010).

– *Jenisch*, *Lekoudesch* and other in-group/secret varieties spoken by cattle dealers and itinerant groups in Germany and Austria (Matras 1988, 1991). Words from *Rotwelsch* (the German thieves' cant), from Hebrew and Romani, are inserted into (dialectal) German. Only few grammatical words are affected.

– *Old Helsinki Slang* (Jarva 2008; de Smit 2010), a variety spoken in working-class areas of Helsinki from the end of the nineteenth century until the mid-twentieth century by speakers of Finnish and Swedish as a first language at a time when Helsinki changed from dominantly Swedish-speaking to a dominantly Finnish-speaking city during a process of industrialization and an influx of Finnish speakers from other parts of Finland. *Old Helsinki Slang* seems to have been used as an expression of the new urban identity at the margins of society, i.e. as an anti-language. The structure of *Old Helsinki Slang* is clearly Finnish, but up to 75% of the vocabulary is said to be of Swedish origin. There are some deviations from standard and dialectal Finnish morphology. The phonology may be deliberately distorted by deviations from the patterns of initial consonant clusters from those permissible in Swedish.

– *Bilingual Navajo* (Schaengold 2004), a variety that emerged from language mixing between English and monolingual Navajo through conventionalization. *Bilingual Navajo* is another example of partial relexification, since in this variety, the grammar is Navajo (with some changes), while the lexicon is increasingly English. English nouns are inserted with nativized Navajo phonology; English verbs are integrated via the use of a light verb (the auxiliary *ash'tééh* 'to prepare, do').

This entire first group of “mixed languages” may be called “symbiotic” (following Matras 2009), as the fused lects heavily rely on the grammar of one of the two contact languages, in this case on Quechua, Mbugu, English, German, Finnish and Navajo, respectively. In the case of Anglo-Romani, Jenisch, Lekouesch, and perhaps also Old Helsinki Slang, one of the functions of the copied verbal material is to make the majority language incomprehensible to its monolingual speakers (in addition to an identity-related function which is central for all fused lects in this group, as we have seen). In some cases, it can indeed be questioned whether we are dealing with a fused *variety* (“language”) at all; a better term for Anglo-Romani, Jenisch, Lekouesch, and Old Helsinki Slang may be *register*, since the functional range of these ways of speaking is limited. Also, structural conventionalization seems to be only partial, i.e. there is lexical variation (speakers select from a repertoire of languages and varieties: Bakker and Matras [2003: 8] speak of a “style of speech, consisting of occasional lexical insertions”, their number depending on the knowledge of Romani of the speaker). In this sense, Para-Romani, Jenisch, Lekouesch, and Old Helsinki Slang may not be conventionalized mixed languages at all, but rather examples of how speakers *mix* languages for secretive and identity-related purposes. Ma’á and Media Lengua, on the other hand, have reached beyond this status of a register for special purposes, since they are used in a variety of situations by their speakers. In all cases, the symbiotically superimposed variety is the weaker one, since it only provides a fraction of varying size of the vocabulary, while the grammar is clearly that of the other, dominant language. This even holds for Media Lengua: Structurally speaking, it is a variety of Quechua, not of Spanish, since the grammatical structure of Quechua is fully intact and only (re-)lexified by Spanish words.

GA?

Decisive for this first group of radically fused lects is the fact that the borrowed lexical material is fully integrated into the grammatical structure of the receiving language. All grammatical affixes are those of the receiving language, as is their syntactic structure including word order. A typical example from Lekouesch (from Matras 2009: 293) is the following:

- (1) Der *schäff*-t de ganze *Jomm* im *Ushpiss*,  
 he sit-3.SG the whole day in-the pub  
 und *dua*-t immer *harme schasskenn*-a und *melouch*-t *lou*.  
 and do-3.SG always much drinking and work-3.SG not

cf. Alemannic dialect:

*Der sitz-t de ganze Daag im Gaschtbaus und dua-t immer viel drink-a und schaff-t net.*

‘He sits all day in the pub, and drinks a lot, and doesn’t work.’

The Hebrew words (*italics*) are fully integrated into the Alemannic dialectal frame. Both word order and inflectional affixes are exactly those of the dialect. While the nouns *Jomm*, *Uuschpiss* (Hebrew *yōm* ‘day’, *ušpīz* ‘pub’, from Lat. *hospitium* ‘hospitality’) and the adverbs *harne* and *lou* (Hebrew *harbe*, *lō*) are not inflected in Lekoudesch, simply because they would not be in German either, the infinitive *schasskenn-* (Hebrew *šāṭa* ‘drink’) receives the Alemannic *-a* infinitive suffix, and the inflected verbs *melouch-* (Hebrew *mēlaxa* ‘work’) and *schäff-* (Hebrew *šev-* ‘to sit’) receive the third-person singular suffix *-t*. (The full integration of the Hebrew words into the Alemannic frame makes some phonological adaptations necessary, such as the deletion of Hebrew final vowels.)

In a *second group* of mixed languages, grammatical subsystems from two different languages are combined to form one grammar. This can happen with or without *compartmentalization* of the two languages. By *compartmentalization* I mean that grammatical elements are borrowed together with the corresponding lexical items (following Friedman 2008). What results is often (but not always) a grammatical split along the lines of verbal grammar vs. nominal grammar. Examples are the following:

– *Michif* (Bakker 1997), the language of the Métis, a community which owes its existence to the marriages of Cree women with French fur traders in the nineteenth century. The complete verbal lexicon, the verbal grammatical system, and the word order are Cree, while the NPs – including adjectives and determiners – are almost entirely (Métis) French. We therefore find a clear case of compartmentalization (stems from one language combine with its grammar, stems from the second language combine with the other grammar) which even includes phonology. In addition, a split between the verbal and the nominal part of the language has occurred. Nonetheless, since Cree is an agglutinating language with strong head-marking features, there can be no doubt that French and Cree do not contribute equally to Michif: The verbal system is at the heart of Cree and has not been affected by borrowing from French, while the more peripheral NPs have been replaced to a large degree by French ones. The imbalance between Cree and French also shows in the “leakage” of Cree into French NPs, where a Cree demonstrative and obviative suffix are used. Michif is Cree with extensive borrowing from

GA?

**Métis French.** Obviously, the borrowing strategy is very different from the one observed in the first group of “mixed languages” discussed above: **French words are integrated into Cree together with their grammatical marking, resulting in the insertion of French NPs into Cree matrices.** Matras and Bakker (2003) give the following example:

- (2) *La Cendrilleuse* mâka tout kî-piskeyiht-am tout  
 The Cinderella however all PAST-clean-it all  
  
*la maison, le plancher* kî-kisîpêkin-am.  
 the house the floor PAST-wash.by.hand-it  
 ‘Cinderella, however, cleaned everything. She washed the house, the floor.’

On the surface, we find **juxtapositions of French NPs (*La Cendrilleuse, tout, la maison, le plancher*) and Cree VPs (*kî-piskeyiht-am, kî-kisîpêkin-am*).** Note the head-marking structure of the Cree VPs which incorporates an anaphoric element referring to the NPs in the periphery: The VPs alone could form a full-fledged sentence. **The VPs are therefore structurally quite independent from the NPs.** The strategy of fusion underlying Michif is obviously entirely different from the strategy of fusion exemplified in (1) for Lekouesch.

**The basis of a compartmentalized language is the borrowing of lexical elements together with their grammatical entourage – certainly their inflectional morphology, but often also grammatical words (such as functional heads) that form the phrase together with them.** In Michif, it was the structure of Cree and French that additionally led to a split between a verbal and a nominal domain of the language, respectively. Compartmentalization resulted from the combination of a polysynthetic and an inflectional language. Other contact situations in which extensive borrowing occurred and led to fusion had different results, **since the contact languages made different types of grammatical structures fuse.** An example is Kormakiti Arabic, a fusion between Arab and Greek, whose status as a full-fledged “mixed language” is somewhat disputed these days (Thomason and Kaufman 1988: 105–106 count it in, following Newton 1964; also cf. Thomason 2003: 24). But even if Kormakiti Arabic (which was the language of a community of Arab-speaking Maronite Christians who migrated to Cyprus in the twelfth century, and is extinct today) was only partially fused and/or **still in the mixing stage**, its structure is telling. **It is similar to that of Michif without being subject to the same structural constraints.** There is no nominal/verbal split, but nonetheless a **strong compartmentalization:** All Arabic words (from whatever

GA?

word class) retain Arabic morphology, while Greek words exhibit Greek morphology (the phonology is Greek throughout).<sup>2</sup> This shows that the decisive feature of the second group of fused lects is not compartmentalization as such (since this is a by-product of the languages involved), but the tendency to borrow lexical elements together with their grammar, and not as isolated words or stems.

– *Mednyi Aleut* (Copper Island Aleut), like Michif, came into being through intermarriage, in this case between Russian fur traders and Aleut women on Mednyi Island (Vakhtin 1998; Golovko 1994). As in Michif, there is a grammar split with an Aleut nominal part (including lexical stems, inflection and word order) and a Russian verbal part (including verbal inflection and word order). However, the division is much less clear, since function words pattern less predictably (object personal pronouns, demonstratives, indefinite pronouns and most numerals are Aleut, as might be predicted, but subject personal pronouns and possessives are Russian), and there are numerous Aleut verb stems which take on Russian inflection. Matras and Bakker (2003: 4) give the following example:

- (3) *ya bud ivo ha yaa-t' ukushka-â haksii-t'.*  
 I will him ask-INF window-CASE open-INF  
 'I will ask him to open the window'

The pronouns and the future auxiliary *bud* (from Russian *budu*) are Russian. The object NP receives Aleut nominal inflection (-â), and the two infinitives Russian verbal inflection (-t'), as predicted by the nominal/verbal grammar split, but the nominal stem is Russian (from *okoshko* 'small window') and the verbal stems are Aleut. Mednyi Aleut thus combines lexical (stem) insertion (the first pattern) – in most of the VP grammar – with compartmentalization and grammar split (the second pattern) in the NPs. I would follow Vakhtin (1998) in arguing that a fusion pattern like this must have emerged from a Russian basis, i.e. from a situation in which Russian speakers (re-)introduced Aleut elements into their monolingual Russian after the Aleut language had already been lost.

– *Light Warlpiri* (O'Shannessy 2007), a new Australian variety fused from Kriol/Aboriginal English and Warlpiri. Again there is compartmentalization and a split between nominal and verbal grammar, since most verbs and ver-

<sup>2</sup> Borg (1984) thinks that Kormakiti Arabic is not a fused lect, but merely a case of language mixing, which supports my view that on the surface, mixing and fusion are highly similarly patterned.

bal morphology are Kriol, while nouns come from English or Warlpiri, and nominal morphology from Warlpiri. Light Warlpiri has additionally developed some new structural features (the auxiliary system).

– *Gurindji Kriol* (McConvell and Meakins 2005; Meakins 2011, 2012), like Light Warlpiri spoken in the Australian Northern Territory, shows a similar split. As a consequence of the attrition of Gurindji, a fused lect has emerged in which Kriol provides the verbal part of the grammar (particularly the TAM system) and the basic verbs, while Gurindji structures are maintained in the nominal part. Since the nominal part of the lexicon and important sections of the verbal lexicon come from both languages, there is little compartmentalization. (More on Gurindji Kriol below, section 3.)

So far, I have argued that there are only two structural types of radically fused lects (“mixed languages”). In the one case, lexical material (words, or more often stems) is inserted into the receiving language without any accompanying grammar; the other-language material is integrated completely into the grammar of the receiving language. In the second case, the grammatical systems merge. I will argue in the next sections of this paper that the basis of such a merger is the insertion of phrasal constituents from another language into the syntactic frame of the dominant language. Arguably, this is the stronger type of fusion. The insertion of phrasal constituents implies at least an initial stage of compartmentalization, i.e. the morphology is copied together with the lexical stems. This compartmentalization is not always complete, as Mednyi Aleut and Gurindji Kriol show. A split between nominal and verbal grammar may or may not occur, depending on the structure of the languages in contact.

Since the two Australian cases are sociolinguistically and historically well documented, we know that in their case, fusion (a stage which may not have been reached entirely yet) was preceded by a period of language mixing. In addition, there can be no doubt that in these cases, mixing not only preceded fusion in time, but that fusion is the result of the grammaticization of mixing. It is likely that the same applies to Michif and Mednyi Aleut, although detailed historical evidence is lacking in their case. During the time of mixing in a number of the documented cases (surely Light Warlpiri and Gurindji Kriol, perhaps also Mednyi Aleut), a “matrix language turn-over” (Myers-Scotton 1998) must have occurred. Before that time, the dominant language was the ethnic minority language, and since it was in contact with a sociolinguistically more powerful language such as Kriol or Russian, it is unlikely that it was used in a monolingual way; rather, a mixed way of speaking must have existed in which Kriol or Russian elements were inserted into the minority language. However, at some point in time (the point of the matrix language turn-over),

GA?

GA?  
emerg  
new  
ethnic  
lang

the majority language took over in providing the grammatical frame for most sentences into which elements of the old language were then inserted through frequent mixing. The old ethnic language (Aleut, Gurindji, Warlpiri) was factually lost at that time, and language shift occurred, but for the speakers this shift was not dramatic because the extensive admixture of elements of the ethnic minority language into talk in the main language provided the impression that they were still speaking the minority language. The language mixing which emerged after the matrix language turn-over became conventionalized and provided the basis for fusion. Michif differs from this scenario since no matrix language turn-over took place.

In sum, we get the following typology of the varieties discussed in the literature as “mixed languages” (fused lects, in the terminology used here):

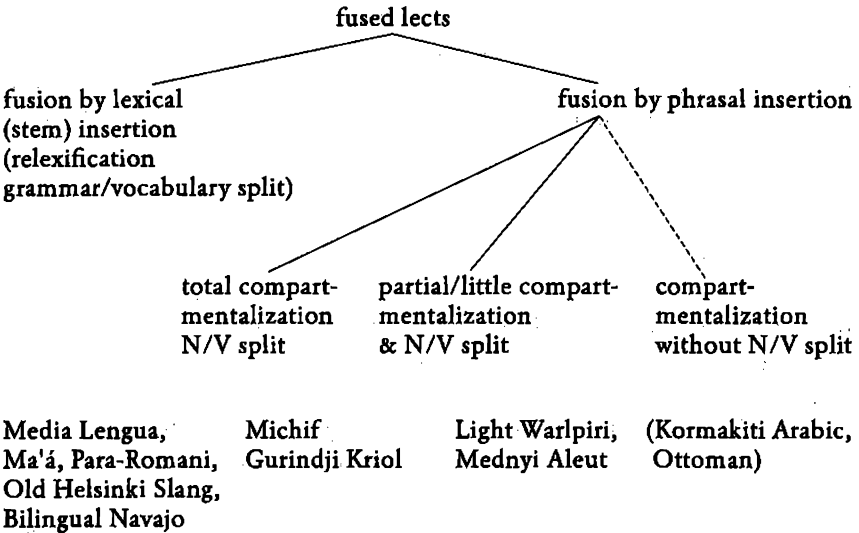


Figure 1: Typology of the varieties classified as fused lects

Kormakiti Arabic and Ottoman are linked by a dashed line in this figure. Whether a language with strong compartmentalization but no split in the grammar (such as Kormakiti Arabic) should be called a “mixed language” is a matter of definition. Strictly speaking, what we are dealing with here is not so much a fusion of two grammatical systems as a partial duplication of the grammatical system (as also observed in Ottoman Turkish and its Arabic-Persian components; cf. Németh 1953).

This grouping is different from the one suggested by Thomason (1995), where the mixed languages resulting from prolonged language contact (“per-

sistent group”: Kormakiti Arabic; Ma'á, and Caló, a Para-Romani variety with Romani lexicon and Spanish grammar) are opposed to those whose emergence was presumably rather abrupt (Michif, Mednyi Aleut, Media Lengua). Thomason's sociolinguistically defined groups are also linguistically coherent as she classifies Ma'á as a Cushitic language with extensive Bantu borrowing in lexicon *and* grammar, and Caló as a Romani variety with Spanish grammatical *and* lexical borrowing. She argues that in these varieties, the dominant language has affected the lexicon and the grammar in substantial ways, whereas in the second group of newly emerging varieties, the effect of the dominant language is less pervasive and restricted to certain domains of the language/grammar. However, this parallelism between social and linguistic facts cannot be upheld if Ma'á and Caló – which includes only slightly more Romani grammar than Anglo-Romani and can therefore also be classified as a Para-Romani variety – are taken to be varieties of Bantu or Spanish, respectively, into which some Cushitic or Romani words are inserted.

### 3. The structure of mixing

In studies of bilingual speech, the term *mixing* refers to a way of speaking in which the two (or more) languages occur within grammatically independent units (“sentences”) or/and conversational turns.<sup>3</sup> For the present discussion, the structure of mixing is important. There is widespread agreement that mixing can be *alternational* or *insertional* (cf. Auer 1999, Muysken 2000).

*Alternational mixing* means that a syntactically independent unit such as a sentence begins in one language and ends in another (multiple mixing is also possible); it is therefore often difficult to define the matrix language of the sentence as a whole. Alternational mixing is a surface phenomenon of bilingual talk. The speaker starts out in one language and changes the language on the fly. The point of alternation is constrained to a certain degree by the grammar of the two languages. Often, their syntactic structures converge in

<sup>3</sup> *Mixing* is often conflated with *code-switching*, but there is an important difference that should be made between discourse functional juxtapositions of the two languages and juxtapositions that are not. I therefore restrict the term *mixing* to the latter case (Auer 1999). Frequent code-switching can result in mixing, as the discourse meaning of the individual switches diminishes. Even though mixing is not discourse-functional in its single occurrences, it is a form of bilingual talk that carries *sociolinguistic* meaning, being as an expression of ethnic or social identity. The social function of mixing is therefore similar to that of fused lects which often serve the identity needs of severely pressured minority groups or new social groups (intermarriage type), as shown in section 2.

this point. In many cases of alternational mixing, the second stretch of talk begins with what looks like an insertion into the matrix of the sentence that has emerged so far. However, the end of the would-be insertion does not coincide with a return to the matrix language. Rather, the language of the would-be insertion is also the language of the remainder of the sentence. Typically, the stretches in language A and B are therefore not coextensive with single constituents.

Now, an important generalization is that alternational mixing never grammaticizes into a fused lect (i.e. the structure of fused lects always reflects an insertional mixing structure). Why is this so? Consider some examples of alternational mixing:

(4) (Twi [italics]/English [capitals] mixing in Ghana, from Flamenbaum in press):

*Ye be tumi a* MAINTAIN, IF THE COUNTRY CAN AFFORD IT,  
3.PLU fut can COND maintain, if the country can afford it,

*eye* CONVENIENT FOR EVERYBODY.  
3SG-be convenient for everybody

'We will be able to maintain [it], if the country can afford it,  
it is convenient for everybody.'

(5) (Russian/German mixing in Germany, data collected by N. Khakimov, ongoing PhD project Freiburg University)

(a) *мы как-то так с девочками всегда были вместе*  
*my kak-to tak s devčonkami vseгда byli vmeste*  
we somehow so with girls always were-PL.PAST together  
von der hauptschule bis komplettem ende  
from the secondary school to ((the)) complete end

'We, the girls, had always kind of stuck together from general education secondary school to the very end.'

(b) *кола без* kohlensäure schmeckt doch  
*kola bez*  
cola without gas tastes-3.SG.PRES particle  
'Coke without gas does taste good.'

In (4), the speaker in both lines starts the utterances with a sequence of Twi grammatical elements which are followed by an English predicate (*maintain, convenient*) in the focus position. This English predicate turns the language of the emerging sentence around (the following conditional clause and prepositional object dependent on the adjectival predicate are also produced in English). Despite the tendency for the alternation point to coincide regularly with the topic/focus boundary in these data (as Flamenbaum shows), the grammatical structure of the stretch of talking up to the point of alternation as well as that following it is highly variable and not predictable.

The same applies to the Russian/German data in (5). In (5a), the speaker adds a German prepositional phrase as a kind of afterthought to her sentence which is continued in this language. In (b), the subject NP is a (presumably) Russian head noun modified by a Russian preposition which governs a German noun (*Kohlensäure, 'gas'*). The use of this German noun arguably triggers the continuation of the sentence in German, with a German predicate and a final emphasizing particle. Once more, the alternation could have occurred at any point in the emerging sentence, provided that the grammars of the two languages in play converge sufficiently to make the transition possible. Such a highly variable pattern cannot grammaticize. As Backus (2003) aptly puts it: **"Alternational CS<sup>4</sup> (...) entails unbridled variation". Grammaticization, like any type of conventionalization, presupposes repetition. If there is no structural similarity between the individual instances of alternational mixing, no conventionalization can take place. Therefore, mixed bilingual talk which is of the insertional *and* alternational type must eliminate the alternational elements on the way to fusion. In the further discussion of how mixing becomes grammaticized in fused lects, we can therefore disregard the structure of alternational mixing.**

Let us now look at *insertional mixing* in more detail. In insertional mixing, there is a matrix language for each syntactically independent unit ("sentence"). Mixing elements of the other language into this matrix language structure follows **two basic strategies (combinations occur, of course, but often one of the two strategies is usually dominant): *single lexeme insertions* and *embedded island insertions*.**

"*Embedded language island*" is the term introduced by Myers-Scotton to refer to "full constituents consisting only of Embedded Language morphemes occurring in a bilingual C[omplementiser]P[hrase] that is otherwise framed by the Matrix Language. An Embedded Language island shows structural dependency relations; minimally it can be two content morphemes

---

<sup>4</sup> CS = code-switching, the term he uses instead of mixing.

(e.g. noun and modifier), or a content morpheme and a non-derivational system morpheme” (202: 139). Numerous examples have been given in the literature (see Myers-Scotton 1997, among others). Typical are Luther’s bilingual (Latin/Early New High German) dinner conversations (cf. Stolt 1965), particularly the German sentences into which Latin material is inserted. As a rule, the **Latin material is structured as islands which consist minimally of a noun with its Latin inflection, but often of a more complex NP.**<sup>5</sup> (Some other inflected word classes such as adjectives, but rarely verbs, are also involved, as well as non-inflected word classes such as adverbs, conjunctions, and discourse markers.)

(6) (Luther’s dinner conversations, from Veit Dietrich’s notes; Latin in italics)

(a) (about the devil) (122)

Wolan, der gifftig      Geist, er thutt uns vil    zu leyd.  
well,    the poisonous spirit, he gives us    a lot of pain

Ich ways, ich will yhn ein mal sehen, *in novissimo*  
I    know, I    want him one time see,    on young-SUP-ABL

*die,*            und seine *ignita*      *tela.*  
day-ABL, and his    blistering projectiles:

‘Well, this poisonous spirit gives us a lot of pain. I know that I want to meet him once, on the Last Day, and his blistering bullets.’

(b) (about the devil) (590)

Judas ist *in vita sua* nit angefochten;

Judas is in life his not tempted;

*Ideo*            da            das stundlin    kam, gieng er *securus*    dahin,  
Therefore when the hour-DIM came, went he assured there,

wuste nit,    wo aus.  
knew not,    where out.

‘Judas had never been tempted in his life; so when the hour came, he went there self-assured and did not know what to do.’

<sup>5</sup> See Stolt (1964) for details. Quoted from the Weimarer Edition of Luther’s Complete Works, edited by Ernst Kroker.

## (c) (about professions) (519)

*Fratris*                      *consilio*                      sol    man    folgen.  
 brother-SG.GEN    advise-SG.DAT    shall one    follow.  
 'You should follow the advice of your brother.'

## (d) (about the Holy Spirit) (521)

*Ideo*                      mus    er    cum    *viperis*                      et    *pharisaeis*  
 Therefore    must    he    with    snakes-PL.DAT    and    pharisees-PL.DAT  
 anderst            reden.  
 differently    talk

'Therefore he has to talk differently with snakes and pharisees.'

(e) *Sed* unse Herr    Got    will    *finem*                      machen  
 but our    Master    God    will    end-SG.ACC    make

*in ultima stultitia.* (522)  
 in last    stupidity

'But our Lord will put an end to this last stupidity.'

## (f) (about the Turks, he = God) (206)

Er    mus    in            die    *caecos*                      et    *perversos*                      *homines*                      zeichnen.  
 He    must    against    the    blind-PL.ACC    and    rotten-PL.ACC    people-PL.ACC    sign.  
 'He must give signs against the blind and rotten people.'

## (g) (about the devil) (588)

Der    kann    das    *verbum*                      nit    leyden.  
 He    can    the    word-SG.ACC    not    stand  
 'He can't stand the (holy) word.'

In all these examples, the matrix language is German. Yet Latin elements are inserted into the grammatical frame of German. In the examples given here,<sup>6</sup> the insertions are either single non-inflected elements such as conjunctive adverbials (*ideo, sed*), or they are phrases (adjectival, prepositional, or NPs), whose internal structure is entirely Latin. These EL islands are either single stems with a Latin inflectional ending (such as the accusative *fin+em* in [e]), or more complex NPs in which the noun is modified by an adjective (*ultima*

<sup>6</sup> There are also numerous cases of alternational mixing in the data.

*stultitia* [e]), or a genitive NP (*fratris consilio* [c]). If the NP is governed by the German verb it takes on the Latin case ending required by this verb (such as the dative in [c]).<sup>7</sup> Prepositional phrases can be inserted as a Latin EL island together with the preposition (as in [e], *in ultima stultitia*, or in [d], *cum viperis et pharisaeis* or in [b], *in vita sua*); alternatively, the preposition can also be German but the dependent NP Latin (f). In a number of cases, the embedded Latin phrase is additionally marked by German grammatical words such as a definite article ([g]; but cf. [e], where the German definite article is lacking before *finem*). In a few cases, other parts of speech, such as the adjective modifying the subject pronoun *securus* in (b), are treated as an EL island. There is then a strong tendency in Latin/German code-mixing as practiced by the humanists around Luther to avoid full integration of Latin elements into the grammar of New High German. Instead, in insertional mixing, Latin elements almost always carry along their grammatical marking, and often they even occur in larger EL islands.

In other bilingual communities, a different strategy of *single lexeme mixing* is preferred in which uninflected words or stems of the embedded language are inserted into the matrix of the other language. In this case, all grammatical elements come from the matrix language. As the grammatical integration of the borrowed lexical material is nearly total, some researchers have argued that we are not dealing with mixing here at all, but rather with a completely different type of bilingual talk, i.e. ad hoc (“nonce”) borrowing (see Poplack 2004 for a summary of arguments). This is largely a terminological issue, but it shows that the two strategies are very different.

An example of a multilingual community in which minimal lexical insertion and total integration into the matrix language is preferred is the Hungarian-German-Romanian community of Neu-Palota (as described by Szabó 2010), an old German-speaking settlement in the Banat region in present-day Romania.<sup>8</sup> Although EL islands are not absent, the dominant pattern is one of complete integration of single stems. Since case marking on the noun is hardly existent in German, this means, for instance, that Hungarian nouns are inserted in their bare (nominative) form in German matrix sentences, even though Hungarian would require a case ending (Hungarian in italics):

<sup>7</sup> Some exceptions to this rule – i.e. cases in which the Latin ending does not follow the case frame of the German verb, but rather its Latin equivalent – are discussed in Auer and Muhamedova (2005).

<sup>8</sup> I’m glossing over some details here, such as simplifications in the structure of Hungarian verbs when inserted into German. Double marking is not unusual, particularly in the plural. The German dialect found in mixing has converged in several aspects towards Hungarian. See Szabó (2010) for a full description.

(7) (from Szabó 2010)

(a) weil et arweit och dort drauß beim *olasz* (p. 341)  
because she works also there out with-the Italian  
'because she also works out there at the Italian's'

(b) *bát* ich weiß einmal wie of *kezelés* war,  
well I know once when on treatment was-1./3.SG  
demal warscht du och of *kezelés* mit der (.) mit*mama*  
at-that-time were you also on treatment with the with-mother  
net, (p. 338)  
right

'well, I remember, once when I was away for therapy, then you too were away for therapy together with the sister of your brother/sister-in-law'

In (7a), Hungarian would require an adessive case suffix (*-nál*), in (7b) a superessive suffix (*-en*).

If German requires a suffix, such as in the plural, the Hungarian insertion usually receives the German suffix *-e* or *-er*:

(8) awer ich brich vielleicht mei schuh net ausziehe,  
but I need perhaps my shoe not take-off

dass ich soll abzähle wie viel *fagyalt-e*,  
so-that I should count how often ice cream-PL

wie viel (-) eis ich geese han. in zweiunsiebzig jahrt.  
how much ice cream I eaten have. in seventy two years.

weil nur ((claps hands)) zammelese die *garas-er*,  
because only ((claps hands)) collect the mite-PL

da mer kunnen ofbau wat.  
so-that we can build-up something.

'But perhaps I don't have to take off my shoe to be able to count how many ice creams – how much ice cream I have eaten in seventy two years. Because we only had to contribute our mite so that we were able to build up something.'

There are two Hungarian insertions in this extract (from Szabó 2010: 353) – *fagylalt* and *garas* – which are both marked by the German plural suffixes attached to the base form (= nominative singular) of the Hungarian word. The insertion of an EL island is avoided.

Even more remarkable are the following examples in which a derivational affix is attached to a Hungarian adjectival stem that is inserted into German. Since German adjectives are not regularly marked morphologically at all, the affixation of the German adjective-forming *-ig* ending to Hungarian stems is not required by German grammar. Even more so, it is evidence for the speakers' preference to integrate the Hungarian material explicitly into the German sentence frame:

- (9) (a) und de *péterke* is *gyáva-ig*. (p. 360)  
 and the Peter-DIM is coward-ish  
 'and little Peter is a coward'
- (b) na waren son (-) solche *szivacs-ig-e*. (.)  
 then were such-a these sponge-ish-PL  
 fressel so *mackor* welche, und has (p. 376)  
 food-stuff like little-bears some, and rabbit  
 'and there were those (-) these spongy food stuff like  
 some little bears and rabbits'

The Hungarian nouns *gyáva* 'coward' and *szivacs* 'sponge' are turned into adjectives by the German derivational suffix *-ig*; the newly formed adjective *szivacsig* is additionally pluralized by the German plural suffix *-e*. (The projected head noun of the utterance is not produced; rather, the speaker breaks off and continues with a German singular noun, i.e. *fressel* 'food'.)

The mixed style of the Banat Swabian settlers even allows for the German dialectal participle circumfix *g-X-t* to be attached to Hungarian verbs in their base form (Szabó 2010: 299–317; also cf. Földes 2005: 133):

- (10) (a) welchen tag hat K. *g-vállal-t*? (p. 300)  
 which day has K. PART-take-PART  
 'which day did K. choose?'
- (b) na han mer uns immer *g-vitáz-t* (p. 301)  
 then have we us always PART-quarrel-PART  
 'so we always quarreled'

The basic forms of the Hungarian verbs *vállal* and *vitáz* are used here just like German verb stems and receive the German participle-forming circumfix. The result looks like a German verb in all respects.

Note in passing that the lexical elements that are inserted into the German sentence frames are not restricted to those referring to ‘new’ cultural spheres added to the traditional German lexicon in order to cope with modern conditions of life. Some of them clearly affect the core area of the vocabulary (as shown particularly well by the examples in [10]). This is important since some researchers have claimed that language mixing only serves to integrate additional vocabulary (“cultural borrowings”) from the majority language (and majority society) into a language, while in “mixed languages” (fusion) it affects the core vocabulary (Backus 2003). The data from Neu Palota demonstrate that this is not the case.

In this section, I have shown that the two basic strategies of insertional mixing are minimal lexical insertions (“ad hoc borrowing”) with maximal integration into the matrix language, and maximal insertion in the format of EL island insertions into the matrix language. *These strategies correspond with the two basic formats observed in fused lects.* Minimal lexical insertions are the equivalent of fused lects with relexification (fusion by lexical (stem) insertion); maximal insertions in the format of EL islands are the equivalent of grammatical fusion in fused lects. These parallels strongly suggest that the two types of fused lects have their basis in corresponding insertional strategies of mixing. **In fact, the data from Gurindji/Kriol mixing presented by McConvell (2008: 191–194) prove this point; they show that the mixing patterns are reproduced in the newly emerged Gurindji Kriol:**

- (11) Gurindji/Kriol mixing (McConvell 2008: 191; Gurindji in italics)

*kaa-rni-mpal* said orait yutubala kat-im *ngaji-rlang-kulu*  
 east-UP-ACROSS side alright 2DU cut-TRN father-DYAD-ERG  
 ‘You two, father and son, cut it across the east (side of the cow)’

- (12) Gurindji Kriol (McConvell 2008: 193)

an skul-*ta-ma* jei bin hab-im sport *karu-walija-ngku*.  
 And school-LOC-TOP 3PL.S PST have-TRN sport child-PAUC-ERG  
 ‘And the kids had sport at school.’

The two insertions from Gurindji into the Kriol matrix sentence in (11) are EL islands with an internal constituent structure conforming to the grammar of Gurindji: an initial directional expression and a transitive subject NP

in final position. The Gurindji Kriol utterance in (12) is based on the fusion of a Kriol matrix into which the two islands are inserted, a local adverbial phrase, and a final subject NP.

**But what exactly happens on the way from mixing to fusion?** As in all processes of grammaticization, the first change is obviously the *conventionalization* of certain mixing patterns. In a fused lect, the speakers no longer have a choice; they speak one, new language. In mixing, there is always a choice between juxtaposing elements from the two languages or using material from only one language. But conventionalization is not everything; the emergence of a fused lect out of language mixing also implies a considerable amount of *regularization*. This means that some mixing patterns are lost while others become more frequent. We already pointed out in the beginning of this section that perhaps the most regular type of regularization is the loss of alternational mixing patterns; in addition, a mixing style which shows minimal and maximal insertions into the matrix language may prioritize one of them in fusion.

A further difference between mixing and fusion only applies to *EL islands/phrasal insertion*. Here, **compartmentalization** of the language may lead to a split between the nominal and the verbal grammar and lexicon as part of the regularization of the insertional pattern, usually conditioned by the grammatical structure of the language (head marking vs. depending marking, etc.; cf. McConvell 2008). But often, the compartmentalization is not total but allows combinations of nominal lexical material from language B with the nominal grammar of language A, or verbal lexical material from language B with the verbal grammar of language A. An example can be seen in ex. (12), i.e. the locational phrase *skul-ta-ma*. Gurindji Kriol has a nominal/verbal split, and the nominal grammar of the language tends to be from Gurindji. But in *skul-ta-ma*, the nominal stem is Kriol, and only the suffixes are Gurindji, i.e. the nominal/verbal split is **maintained in the grammatical suffixes but not the stem (no compartmentalization)**. Fusions like this cannot be a left-over from mixing according to the minimal insertion strategy (as in ex. [3] from Mednyi Aleut), since the direction of the insertion is reversed: Whereas in Kriol/Gurindji mixing the matrix language was Kriol, the grammatical elements in *skul-ta-ma* come from Gurindji, the **Embedded Language**. A solution to this problem will be presented in the next section (4.2.–4.4.).

#### 4. Partial fusion

In order to come to a better understanding of the way from mixing to fusion, it is not sufficient to consider the final outcome of this grammaticization, i.e. the “mixed language” in its final state of development. Rather, it is useful to look at incipient and intermediate stages in the conventionalization of mixing formats as well. Their investigation can provide a glimpse into the way fused lects emerge. In this section, I will look at some examples of this partial fusion. Far from providing a full and coherent picture, the following is only intended as a first unsystematic collection of cases, each of which is very different and deserves further extensive treatment. The varieties discussed show some amount of fusion, but they would not be called “mixed languages” since the fusion only affects a section of the language. We will start with minor and unspectacular kinds of fusion and end with some examples in which the classification of the variety in question as a “mixed language” in the sense of extreme fusion is already possible, although not suggested here.

##### 4.1. Fusion of discourse marker and modal particle systems

The most frequent case of incipient fusion beyond the mere borrowing of lexical material is certainly the borrowing of another language’s system of discourse markers (including tag questions and hesitation markers) and/or modal particles. Conjunctions (particularly coordinating ones) are often included in this list as well, which is plausible as their discourse meaning in spoken interaction often comes close to that of discourse markers. The phenomenon is widespread and has been documented extensively (cf. Matras 1998, 2000, 2007). In order to speak of fusion, a substantial part of the system of discourse markers/particles has to be borrowed, not just a single marker, either replacing the system of the receiving language or adding to it (particularly if the receiving language has few or no particles). The borrowed markers may change their meaning in order to fit into the new fused system. An example is the Sinti variety spoken in Germany which Holzinger (1993) describes. Here, discourse markers, modal particles, and some conjunctive elements (in addition to separable verb prefixes, which present another case of fusion, see 4.2.) are taken from German and solidly fused with the Romani variety:

(13) (from Holzinger 1993: 322; German-origin elements in italics)

*Ach* kai *denn?*  
oh where PART

Kon dšajas *denn* koi *hin*, me *doch* garl  
who wen PART there to, I PART not

*Naja*, *dann* his *noch* mire *gešvistre* ap o vurdi pre.  
well, then was still my siblings on the wagon up

'So where? Who went there indeed, certainly not me! Well, then my siblings were still on the wagon.'

This short passage includes two German-origin discourse markers/interjections (*ach*, *naja*), three modal particles (*denn* in the first and second line, which is used for emphatic questions, and *doch* in the second line, used as an appeal to common knowledge), and one temporal adverbial (*dann*, used primarily to advance the narrative by introducing a next narrative step).

Other examples are the system of French discourse markers borrowed into Shaba Swahili (de Rooij 1996), the system of English discourse markers copied into German varieties spoken in the U.S. (Salmon 1990), or the system of Russian interjections and discourse markers borrowed into Karelian (Sarhimaa 1999: 234–235).

The borrowing of discourse markers, particles and conjunctions (which Matras subsumes under the term "utterance modifiers") also occurs in more heavily fused varieties (such as the ones discussed in section 2), and of course it is also part of language mixing in many bilingual communities. The important step towards fusion is reached when the utterance modifiers can no longer be replaced by the respective elements of the receiving language.

Mixing in discourse markers is different from simple lexical borrowing as it clearly affects another layer of meaning. Discourse markers and modal particles operate on a "metapragmatic" level (as Matras rightly argues), i.e. they direct the hearer towards the intended interpretation of the speaker's utterance. This system of metapragmatics can be said to be part of a discourse grammar of the language.

#### 4.2. Fusion of two systems of derivational morphology

English and German both show traces of fusion as a consequence of extensive lexical borrowing from French, even though this language contact was

too much controlled by norm-setting language authorities to proceed beyond a certain point. (A similar point could be made for Scandinavian languages, particularly Danish, with respect to (Low) German influence.)

As is well known, both English and German derivational morphology show a characteristic compartmentalization. The reason is language contact: From the thirteenth/fourteenth century onwards, but particularly between 1500–1700 (and again in the nineteenth century), both languages borrowed a large amount of vocabulary from French (and Latin), above all learned vocabulary. But the French impact went far beyond the borrowing of words; it also brought into the languages a whole system of new derivational means, which were generalized out of the borrowed complex words to become part of the derivational system of the receiving languages. However, the productive domain of these borrowed derivational affixes is up to the present day limited almost entirely to the ‘non-native’ lexicon: For instance, ‘native’ Germanic words only very rarely combine with non-native Germanic derivational affixes. Examples are the German noun-forming derivational suffixes *-ität* (from French *-ité*), *(at)ion*, *-ant/-ent*, *-eur*, *-ismus*, *-ist*, the adjective-forming derivational suffixes *-abel/-ibel*, *-esk*, *-ös* (fr. *-able/-ible*, *-esque*, *-eux*), and the verb-forming derivational suffixes *-ifizieren*, *-isieren*; they almost exclusively attach to (neo-)Latin, (neo-)Greek or French stems (cf. *Dignität*, *Student*, *Voyeur*, *grotesk*, *melodiös*, *harmonisieren*). A look into the history of this compartmentalization shows that it was not consistently adhered to throughout. For instance, Hilpert (2013) analyses the borrowing of the French derivational suffix *-ment* into English. New words (types) with *-ment* entering the language were particularly frequent up to around 1600, after which their number declined (although the total number of nouns ending in *-ment* continued to increase). Whereas the number of derivations borrowed as such was particularly high until 1500, the number of newly derived forms created in English after the pattern of the French words increased over this time, and, from 1550 onwards, was much higher than that of the direct borrowings. The interesting point for our discussion is that around 1500, there were no restrictions against the use of *-ment* with a Germanic stem (as shown by words like *renewment*, *embitterment*), even though some of these words later disappeared again, cf. *desightment* (‘the act of making unsightly’), *shatterment* (‘the act of shattering’) or *worsenment* (‘the process of getting worse’).<sup>9</sup>

Since the influence of French (and Latin) was largely restricted to the domain of derivation, we would not want to call English or German a fused

---

<sup>9</sup> All examples from OED, as cited by Hilpert (2013).

language. However, the way in which derivational grammar was borrowed into these languages is instructive. It demonstrates how grammatical affixes are borrowed from one language into another. They start out as borrowed words, without a productive grammatical structure that can be used in the receiving language. Once a sufficient number of borrowed words of the same type are available, speakers distill a constructional pattern out of them, which can then become productive. This productivity can be unconstrained when all restrictions on stem–affix compatibility are lifted (a tendency which can be observed for *-ment* in English around 1500), but the more usual way is to limit the productivity of the grammatical affix to stems of a language which is deemed to be compatible with the affix (usually the same language from which the affix was copied). The early phases of *in toto* borrowing initially require bilingual speakers, who insert the words in question into their speech through intentional code-switching and mixing. Once the derivational pattern is part of the receiving language, no bilingualism is required to introduce new words following this pattern.

The fusion of the derivational system of a language with that of another one is a process which is also observed in languages other than English and German, of course. Benítez-Torres (2009) discusses the borrowing of derivational morphology from Berber into Tagdal, a Northern Songhay language of northern and central Niger. Present-day Tagdal has many features of Berber which separate it from the Southern Songhai languages. Among them is phonology, a large part of the lexicon, the lack of tone, as well as many morphological influences. The derivational system has three valence-changing prefixes which are all of Berber origin: the causative, passive, and reciprocal prefix. The important point is that these prefixes only attach to Berber stems, i.e. the fused part of the language is compartmentalized just like in English and German. If, for instance, a Songhay active verb is to be passivized, the Songhay stem has to be replaced by a Berber stem (which in turn would not be used without the derivational affix in Tagdal).

(14)(Tagdal, from Benítez-Torres 2009: 77; Berber origin elements in italics)

(a) active voice

ɣa- b- baay- a  
 1s IMP know 3s  
 'I know it.'

## (b) passive voice

a- b- *tuw-asən*

3s IMP PASS.known

'It is known.'

While the Songhay verb for 'to know' is *bay*, it is replaced in the passive by its Berber equivalent *sən*, so that the passivizing, originally Berber prefix *tuw*- can be attached to it.

The inflectional system of Tagdal is entirely Songhay. The verbal grammar therefore shows fusion between a Berber derivational and a Songhay inflectional system. Benítez-Torres argues that Tagdal is a Songhay language acquired by Berber-speaking conquerors who learned the language of the people whose lands they had occupied. He also argues that the fusion must have been the result of widespread Berber/Songhay bilingualism and code-mixing, in an "abrupt genesis" (80) which did not leave enough time to exchange the derivational system as well. Regardless of whether this reconstruction is correct, Tagdal is a language which clearly shows fusion in its verbal grammar, accompanied by compartmentalization. **Just like in English or German, borrowed derivational affixes only combine with borrowed lexical stems.** Benítez-Torres states that the basis of this fusion was mixing, presumably after a matrix-language turn-over. In the mixing phase, stems and derivational prefixes were inserted together into frames that included the matrix language's inflectional morphology. This is what we find in German and English as well. The mixing strategy was somewhere between the minimal and the maximal strategy outlined above.

## 4.3. Fusion by the borrowing of grammatical words

"Utterance modifiers" are not always easy to separate from grammatical words, and as argued above, some grammatical words (conjunctions and adverbials in particular) tend to change their status and become discourse markers (Auer and Günthner 2005). Yet, while discourse markers may be said to be part of the discourse grammar of a language, grammatical words are part of the sentence grammar which is more tightly structured. Borrowing material from another language into this part of the grammar is therefore a more radical step. Nevertheless, this kind of fusion is quite frequently observed. Let us consider two examples.

The first example comes from the speech of German settlements in Romania (previously Hungary) which we already discussed as an example of

language mixing in section 3. In some cases, the communities have gone beyond mixing and proceeded towards grammatical fusion. A case in point is the system of subordinating irrelevance particles which in most varieties of “Danube Swabian” (*Donauschwäbisch*, including the variety of Neu-Palota) is regularly constructed on the basis of the Hungarian stem *akar-* and a German dialectal interrogative (cf. Szabó 2010: 381):

“Danube Swabian”	std. German	Hungarian	gloss
<i>akár-wer</i>	<i>wer auch immer</i>	<i>akár-ki</i>	‘whoever’
<i>akár-was</i>	<i>was auch immer</i>	<i>akár-mi</i>	‘whatever’
<i>akár-wu</i>	<i>wo auch immer</i>	<i>akár-hol</i>	‘wherever’ (local)
<i>akár-wuhin</i>	<i>wohin auch immer</i>	<i>akár-ova</i>	‘wherever’ (directional)
<i>akár-wann</i>	<i>wann auch immer</i>	<i>akár-mikor</i>	‘whenever’
<i>akár-wie</i>	<i>wie auch immer</i>	<i>akár-bogy</i>	‘however’

Since the new system of irrelevance particles is based on the compounding of a Hungarian and a German element, this is also an example of the borrowing of word-formation patterns. An example of the discourse use of *akár-* is the following:

- (15) (German dialect of Neu-Palota, Romania, from Szabó 2010: 383; Hungarian-origin elements in italics)

weil *ákerwann* i hnausgon of die gass,  
because whenever I out-go on the street

niemal einmal grieben se net ungarisch  
never once greet they not Hungarian

‘because whenever I go out into the street, they don’t even greet (me) in Hungarian.’

It is obvious that this German dialect has restructured the entire system of German irrelevance particles (which in std. German is only moderately grammaticized, as the lack of phonological univerbation indicates). The language-internal reason is perhaps the simplification resulting from such a fusion and the fact that the standard German irrelevance particles are rather alien to the traditional dialects. The newly created irrelevance particles can also be used as indefinite pronouns (16) and adverbials (17), which is not possible in standard German, where the indefinite pronoun is *irgendjemand*, and the indefinite adverbial *irgendwohin*:

- (16) (German dialect of Neu-Palota, Romania, from Szabó 2010: 387; Hungarian-origin elements in italics)

*bát* kannst du net *ákerwer* in de haus tien  
 but<sup>10</sup> can-2SG you not somebody in the house do  
 ‘but you can’t let anybody into the house’

- (17) (German dialect of Neu-Palota, Romania, from Szabó 2010: 387; Hungarian in italics)

*hiába* geht der P. *bácsi* *áker*wohin, dat  
 in-vain goes the P. uncle.SG.NOM wherever, that  
 musse auszahle  
 must-he pay.INF

‘wherever uncle P. goes, he has to pay for it’

*Áker* may have replaced *jeder-* first and now be generalized to the use of an irrelevance conjunction. One would hypothesize that the first step was the insertion of Hungarian indefinite pronouns and indefinite adverbials into the German matrix as a whole.

Indeed, such a process of mixing is documented in my second example, Chiac, the variety of French spoken in Moncton, Canada. In this variety, Perrot (1995, 1998, 2001, 2003, 2005) observes the copying of a number of English grammatical elements (in addition to utterance modifiers) and their fusion with the grammar of monolingual Canadian French among young speakers. The variety described by Perrot has replaced the entire system of French indefinites (determiners, pronouns, adverbials, and irrelevance pronouns) by their English counterparts:

- (18) (youth variety of Chiac in Moncton, Canada; from Perrot 2003: 273; English-origin items italicized)

- (a) *Whoever* qui travaille à MacDonald’s là, *everybody* haït ça.  
 ‘Whoever who works at MacDonalds there, everybody hates it.’  
 (b) *Nobody* se moque de moi.  
 ‘Nobody makes fun of me.’

---

<sup>10</sup> “Danube Swabian” has also partly borrowed the Hungarian system of utterance modifiers.

Fusion in the Chiac variety investigated by Perrot also includes intensifiers (such as *right*, as in *j'aime right ça* 'I like just that'; *that* as in *c'est pas that bon* 'it's not that good', Perrot 2003: 274; *quite a*, as in: *c'était quite a différence* 'it was quite a difference', *on a eu quite a batailles* 'we had quite some fights'; cf. Perrot 2003: 275), possessives (*ils font encore leur own affaire* 'they still do their own thing', instead of post-positioned French *ils font encore leur affaire à eux*; cf. Young 2002: 119) and prepositions (such as *about* as to mark prepositional objects, as in *parler about* 'think about' instead of *de*, and *penser about* 'think about' instead of *à*; or local prepositions as in *i allont over un cliff* 'they go over a cliff'; Perrot 1998: 221; Young 2002: 129–131).

Perrot takes pains to underline that we are not dealing with language mixing here, since the code has stabilized (i.e. it has become conventionalized). This stabilization has led to some adaptations of English to the French matrix, as the example *on a eu quite a batailles* shows: Here, the singular indefinite article of English does seem to have lost its grammatical meaning, and *quite a* has become a single word that can modify the noun *batailles*, which is plural.

Intensifiers need a head to be modified, and prepositions govern an argument. In both cases, the borrowed grammatical word cannot form a phrase on its own. Interestingly, there is variation in Chiac between examples in which the grammatical English word combines with a French element to form a phrase, and others in which the whole phrase is in English: cf. for the latter case *ça m'excite pas that much* 'that doesn't excite me that much', *j'aime pretty much tout except la country* 'I like pretty much everything except country (music)', *c'était quite a trip* 'it was quite a trip', *le TV était improv-é quite a bit* 'the TV has improved quite a bit' (Perrot 2003: 274–275); *i avient fait un movie about it* 'they have made a movie about it' (Perrot 1995: 84). This suggests that the basis of the fusion of grammatical words from English into French grammar may have been mixing of the maximizing, phrase-related type. The variety of Chiac analysed by Perrot is a youth variety. In it, the widespread format of maximizing, phrase-related mixing, which is likely to occur in other, less stabilized varieties of Chiac as well, has developed into a much more marked kind of fusion according to which single grammatical words are inserted into French. Arguably, this variety has an aspect of playful distortion of French to it, a partly conscious, identity-related effort to create something very specific which only young bilingual speakers can produce. Many researchers have suggested that this may also have been one of the mechanisms by which quickly-evolving "mixed languages" came into being. Moncton Chiac youth language may be an example which is happening before our eyes of how this works.

These obviously are only some few examples of a beginning grammatical fusion on the level of grammatical words which is far from complete, and in

which we surely would not want to speak of a fused lect yet. Many more examples are known, among them famous controversial cases such as Chamoru (Chamorro), the Malayo-Polynesian language of the Mariana Islands which used to be in heavy contact with Spanish, and as a consequence has borrowed grammatical words such as indefinite article (*un*); the definite article (*i < el*), the demonstrative (*este*); prepositions (*sin*) or the comparative construction (*mas ... ki < Sp. más ... que*) (Pagel 2010). These grammatical forms are an obligatory and non-variable part of Modern Chamoru grammar, while other borrowings are variable and alternate with Chamoru equivalents. It seems a futile discussion whether Chamoru is a “mixed language” (fused lect) (cf. Stolz 2003); there is no categorical distinction between mixed and unmixed languages, but there are good reasons to believe in gradience in fusion with Chamoru being a good and uncontroversial case of fusional processes at work.

#### 4.4. Fusion of two systems of inflectional morphology

The borrowing of derivational affixes is in the middle range of the borrowability hierarchy; in contrast, the borrowing of inflectional morphology is much rarer. However, there are also cases of partial fusion affecting the inflection of a language. We will look at an example of fusion in the inflectional grammar with a high degree of compartmentalization, and another example in which no such compartmentalization is observed.

An example of inflectional fusion with compartmentalization is Komotini Romani, a Romani variety spoken in Greek Thrace today (Adamou 2010). Nowadays, the Komotini Roma are trilingual, speaking Greek, Turkish, and Romani. The speakers distinguish between “pure Romani” and their own Romani variety, called *xoraxane romane* ‘Turkish-Romani’. Like numerous other Romani dialects of the North and South Balkan as well as of the South Vlax group (cf. Friedman 2008), this variety of Romani borrowed extensively from Turkish due to long-term dialect contact during the Ottoman Empire. Among these borrowings are conjunctions (as expected; cf. 4.1.) such as adversative (*j)ama* ‘but’, the negative answer particle *ayir* (turk. *hayır*), time adverbials and grammatical words such as indefinites (*er*, from turk. *her* ‘all’, ‘every’), the focus marker/coordinator *-da* and the obligation marker *lazım*, in addition to a large number of lexical borrowings. But most remarkably, Komotini Roma has borrowed the time, aspect, mode, number, and person system of Turkish which is used with all Turkish verbal stems. (Borrowed nouns by and large are integrated into Roma morphology.) The use of these morphological forms is obligatory. The borrowed verbal stems are numer-

ous and include a large part of the basic vocabulary. Romani stems never take on Turkish endings. Cf. the following examples:

- (19) (Komotini Romani, from Adamou 2010; Turkish origin parts in italics)

kon        *ama alna-ma-dzak*                      leski kor    ka    tʃindol  
 who.NOM but understand-NEG-FUT.3SG his    throat will cut.3SG  
 ‘but the one who will not understand, I will cut his throat’

The Turkish-origin negative suffix *-ma* and future marker *-dzak* are attached to the Turkish-origin stem *alna-* ‘to understand’ (from Turkish *anla-*). The first clause (*kon ama alna-ma-dzak*) could indeed be classified as Turkish if it did not follow the Romani (Indo-European) pattern of relative clause formation unknown in Turkish. Note that the Romani verb in the second clause has no Turkish, but a Romani person suffix. This is also the case in the following example. (It additionally includes a Greek and a Romanian element which, however, must be considered examples of mixing, not fusion.)

- (20) (Komotini Romani, from Adamou 2010; Turkish origin parts in italics, Greek underlined, Romanian underlined and in italics)

psixoloyos/psixoloyos    *jaz-mi-jor*                      tuke    ap-ora  
 psychologist psychologist write-NEG-PROG.3SG you.DAT pills-PL

kantʃik in    del    tut  
 nothing NEG gives you.ACC

vo mono beʃel        *konuf-ur*                      tusa  
 he just sit.PRS.3SG speak-PRS.3SG you.INSTR

‘a psychologist, a psychologist, doesn’t prescribe pills to you. He gives you nothing. He just sits, talks to you.’

Here, the Turkish verbs stems *jaz-* (‘write’) and *konuf-* (‘speak’) are inserted together with their inflectional endings for the present continuous (*-jor-*), the negation (*-mi*), and the present tense (*-ur*) into a Romani sentence. Romani provides the word order, which clearly makes it the matrix language. Romani verbs are inflected in the Romani way (*del*, *beʃel*).

As mentioned above, the compartmentalized fusion of borrowed (Turkish) and traditional inflectional morphology is typical of many Balkan varieties of Romani. These varieties therefore provide a good picture of the process of fusion. Friedman (2008: 133) suggests the following hierarchy which portrays this process:

preterite < present (general and/or progressive) < clitic *idi* (past) < optative  
< future and negative < infinitive

Komotini Romani has borrowed all elements excluding the infinitive.

Let us now consider an example of non-compartmentalized fusion. I suggest that non-compartmentalized forms of fusion have gone through a similar phase of compartmentalized fusion, but have lifted the structural co-occurrence constraints typical of compartmentalized grammars. For instance, Pakendorf (2009) describes inflectional borrowing in the Northern Tungusic language Even in Yakutia from Sakha (Yakut), the dominant language in which all speakers are fluent. Despite a long history of bilingualism and contact, the process of fusion seems to be an ongoing one, i.e. the patterns are not yet sedimented entirely. Evidence for this ongoing process from mixing to fusion is that the speakers are still aware of the use of the two languages and can identify its components, and that the patterns are still variably used.

Even has borrowed the mood system from Sakha which is not known in the non-fused Tungusic languages but is highly frequent and highly salient as an expressive marker in Sakha. This mood system includes (among other suffixes such as the negative present participle and the necessitative) the suffix *-TAX* which Pakendorf (2009) calls “assertive-presumptive” and which is frequently used in Sakha to refer to past events in narratives. I suggest that this copying is a consequence of language mixing in which the whole predicate was inserted from Sakha into Even. Evidence for this interpretation comes from the fact that the assertive-presumptive suffix is always copied together with the person suffix:

- (21) (Sebjan-Küöl Even, from Pakendorf 2009: 94; Sakha-origin elements in italics)

<i>ejm-u</i>	<i>koke-ri-ke-nun-ni</i>	<i>tar</i>
mother-POSS.1SG	die-IMPF.PTC-DIM-COM-POSS.3SG	DIST other
<i>honte ahj-w</i>	<i>ga-j-dag-a=di</i>	
woman-ACC	take-CONN-ASS-POSS.3SG-EMPH	

‘as soon as my mother died he took another wife’

In this example, the assertive-presumptive suffix plus the person suffix attach to an Even stem. But there are also examples in which the entire VP is Sakha. To begin with, there is an analytic form of the assertive-presumptive which is based on the Sakha auxiliary *buol* and which is inserted together with the suffix *-TAX* and the person suffix into an Even sentence:

- (22) (Sebjan-Küöl Èven, from Pakendorf 2009: 97; Sakha origin parts in italics)

*nebi:le* ilan-nūn kila:h-a-lkan *buol-lag-im=di:*  
 barely, three-COM grade<sup>11</sup>-EP-PROP AUX-ASS-POSS.1SG=EMPH  
 'I have barely done three grades (at school)'

But even more, whole VPs with a lexical stem from Sakha are also attested:

- (23) (Sebjan-Küöl Èven, from Pakendorf 2009: 101; Sakha in italics)

tarit *buolla*, unta-w-u *buolla kĩajan tik-pet*  
 then DP<sup>12</sup> fur.boots-ACC-POSS-1SG DP be.able sew-NEG  
*buol-lag-im=di:*  
 AUX-ASS-POSS.1SG=EMPH

'and then, well I can't sew fur boots, right'

These phrasal constituents are classified as "switches" by Pakendorf, i.e. language mixing, which she believes to have been the origin of present-day Èven with its partially fused inflectional system. Mixing following the pattern in (23), then, was the entry point for copying the Sakha inflectional system into Èven. A plausible path in the grammaticization of the present-day system is that Sakha verbs were first inserted as a part of VP-islands following the maximal insertion strategy into Èven matrix clauses, together with their modal and person affixes. However, Èven did not compartmentalize, i.e. Sakha suffixes could not only be used with Sakha verb stems, but also with Èven verbs.

The fact that only some inflectional endings (those of the modal system) were fused into the monolingual grammar of Èven is best explained by the fact that they were highly salient in Sakha to bilingual Sakha/Èven speakers, but lacking in their own language. As the process of conventionalization is not finished yet, Èven is indeed

a real-time example of how such forms can enter a language, starting with initial code-switching by some speakers, developing into established codes in their linguistic repertoire, and over time and with constant repetition by this group of innovators finally entering the language of those speakers of Sebjan-Küöl Èven who are not radically opposed to Sakha copies, but who view them as salient features of their dialect (Pakendorf 2009: 106).

<sup>11</sup> Russian loanword.

<sup>12</sup> Discourse markers are also frequently borrowed from Sakha.

## 5. Fusion within a language family?

Due to a curious division of labor, linguists who work on bilingualism are rarely interested in the long-term consequences bilingualism can have on the language systems of the two languages involved. On the opposite side, linguists working on language contact almost exclusively deal with structural outcomes, but show little interest in the question of how these structural changes have originated in bilingual talk. Yet, it is difficult to imagine how structural language contact could take place without prior manifestations in bilingual talk. The aim of this paper was to look into the conventionalization of bilingual speech and into its long-term grammaticization in the structure of a language. The example discussed in some detail here was how language fusion can emerge from language mixing. There seems to be growing agreement today that even extreme results of language fusion – so-called mixed languages – result from discourse-based mixing through regularization and conventionalization. Such a development is also compatible with the sociolinguistic embedding of language mixing at the discourse level and that of radically fused lects: In both cases, matters of group identity seem to be of utmost importance. Both mixing and fused lects are products of identity construction – they do not emerge out of the need to understand each other, as do pidgins, and they are not a mere matter of interference and substrate influence after language acquisition or shift.

The main argument I have put forward here is that language fusion is the result of the conventionalization of the two basic strategies of insertional mixing (alternational mixing was shown to be unsuitable for conventionalization). These two basic mixing strategies have been amply documented in studies on bilingual talk, where they may occur in isolation or in combination. The first strategy, called minimal insertion here, consists of inserting stems (or uninflected words) from one language into the grammatical frame (“matrix”) of the other language without any accompanying grammar. In this case the lexical material from the inserted language is accommodated completely to the grammatical structure of the matrix language. (Phonology and phonetics are not necessarily affected.) The second strategy, called maximal insertion or “embedded island” insertion here, prefers to bring the other-language item(s) into the matrix frame together with the accompanying grammatical markers (affixes or grammatical words) and often extends to the phrase level.

I have argued that these two strategies are the equivalents of the two main patterns of language fusion which can account for the structure of radically fused (“mixed”) languages. Minimal insertion in mixing is the basis of the

structure of “symbiotic” fused lects, in which the fusion affects the lexicon of one language and the grammar of another. The grammar is entirely in language B, and lexical items (usually stems) are selectively inserted into this pattern (examples are *Media Lengua*, or *Para-Romani*). Maximal insertion, on the other hand, accounts for the second group of fused lects in which two grammars merge. While symbiotic fused lects can occur without any traces of maximal insertion, the group of grammatically fused lects usually also shows minimal insertion as well. This group is therefore more heterogeneous.

Fused grammatical systems that most directly reflect the stage of maximal insertional mixing are those in which a strong compartmentalization of grammar and lexicon occurs, which may result in a split between verbal and nominal grammar (as in *Michif* or *Gurindji Kriol*). But such a split is not a necessary correlate of compartmentalization; rather, it is dependent on the grammatical structure of the two languages in contact. Non-split compartmentalized fusion occurs as well (e.g. in radically fused languages such as *Ottoman Turkish* and *Kormakiti Arabic*). Weak or one-way compartmentalization may be due to various reasons, most obviously (if the non-compartmentalized part allows for combinations of *Matrix Language* grammar with *Embedded Language* stems) to the conventionalization of a co-occurrence of minimal and maximal insertion strategies at the mixing stage (*Mednyi Aleut*). On the contrary, if the non-compartmentalized part allows for the combination of *Embedded Language* grammar with *Matrix Language* stems, a generalization of the grammatical structure of the *Embedded Language* islands to *Matrix Language* stems is likely to have occurred (*Even*, *Chiac*).

In section 4 of the paper I have presented varieties that were or are in a process of grammatical fusion but have not reached (and will perhaps never reach) the extreme forms of fusion known from “mixed languages”. The weakest forms are cases in which the system of “utterance modifiers” (*Matras*), most notably discourse markers and conjunctions, are borrowed and made an obligatory part of the receiving varieties. In this case, the distinction between maximal and minimal insertion plays no role. The next step is the copying of derivational grammar (word formation) as in *German*, *English*, or *Tagdal* (in the latter case already deeply linked to syntax). The underlying mixing strategy is still minimal here, but fusion implies the grammatical analysis and productive use of the borrowed grammatical elements. In the case of fusion of language B derivational grammar into language A grammar, it is possible to distinguish between compartmentalized and non-compartmentalized results (in the latter case, the borrowed derivational elements of grammar freely combine with stems of the receiving language in addition to those of the source language). A next step consists in the integration of grammatical words, which

may or may not be phrase-building, into the borrowing languages (Chiac). The basis of this kind of fusion is Embedded Language mixing. Patterns corresponding exactly to the maximal insertion strategy in mixing are observed alongside with the generalized use of the borrowed grammatical words with content words of the receiving variety. Finally, fusion in the domain of inflectional affixes was discussed (Balkan Romani, Èven). Again, the compartmentalized version (corresponding to maximal insertion of Embedded Language islands) and its generalization to all stems was observed.

Considering partial grammatical fusion in addition to radical fusion (“mixed languages”) shows that the latter are less exotic with respect to the patterns of borrowing than one might think. As Thomason (1995) puts it: “It is only the results that are extraordinary. The actual processes that lead to these extreme outcomes are quite ordinary”.

Keeping these results in mind, we can now turn again to the question of whether fusion is restricted to unrelated languages, or requires a certain structural distance between the languages involved. It is true that most examples for “mixed languages” (radical fusion) come from situations in which unrelated languages are in contact. (Exceptions are fusions of the “symbiotic” type, such as Anglo-Romani or Lekoudesh.) However, once we include partial fusion in the picture, and look at fusion as a process rather than a state, the number of examples increases in which e.g. two Indo-European languages are involved. This suggests that the scarcity of examples of radical fusion between two languages from the same language family is probably not due to structural factors but rather a result of the social conditions under which such extreme cases arise. Family resemblance therefore does not exclude fusion.

Now we can take the argument one step further and ask how closely related two varieties can be in order for grammatical fusion to occur. Can fusion even take place between two varieties of a language (two dialects, or a dialect and the standard variety), or between mutually intelligible languages such as Swedish and Norwegian or Belarusian and Russian (cf. Hentschel 2008 and in this volume)? In the spirit of the hypothesis of this paper (i.e. that fusion is grammaticization of mixing), the question cannot be answered but a more precise prediction can be derived. If it is true that fusion is a sedimentation of patterns of mixing on the discourse level, then the absence of such mixing would make fusion impossible. More precisely, if mixing between two varieties follows neither the “maximal insertion” nor the “minimal insertion” strategy, no fusion will emerge – at least not the kind of fusion known from so-called “mixed languages”. This prediction needs to be tested empirically. There are cases of contact between very closely related languages

or varieties which show precisely the patterns of mixing that are the basis of fusion, and nothing should prevent such mixing patterns from becoming sedimentated. However, in structurally very close varieties, other forms of mixing are also observed, and the reasons for these types of mixing – untypical for contact between structurally distant languages – must surely be located on the structural level (cf. Muysken's "congruent lexicalization", Muysken 2000, as well as Berruto 2005 who discusses many pertinent examples from contact between standard Italian and dialects). In these cases, different mechanisms are at work and mixing need not follow the constraints observed between more distant languages. The resulting structures, if grammaticized, would not fall under the term fusion. In sum: family resemblance is unlikely to be a good predictor for the grammaticization from mixing to fusion; however, extremely close resemblance within a language family (such as that between immediate siblings) may trigger additional processes not known from contact between more distant languages.

## References

- Adamou, Evangelia 2010 Bilingual speech and language ecology in Greek Thrace: Romani and Pomak in contact with Turkish. *Language in Society* 39 (2): 147–171.
- Auer, Peter 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.
- Auer, Peter and Susanne Günthner 2005 Die Entstehung von Diskursmarkern im Deutschen – ein Fall von Grammatikalisierung? In: Torsten Leuschner, Tanja Mortelmans and Sarah de Groot (eds.), *Grammatikalisierung im Deutschen*, 335–362. (Linguistik – Impulse und Tendenzen 9.) Berlin, New York: de Gruyter.
- Auer, Peter and Raihan Mohamedova 2005 'Embedded language' and 'matrix language' in insertional language mixing: some problematic cases. *Italian Journal of Linguistics / Rivista di linguistica* 17 (1) (special issue "Commutazione di codice e teoria linguistica", edited by Gaetano Berruto): 35–54.
- Backus, Ad 2003 Can a mixed language be conventionalized alternation code-switching? In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate*, 237–270. Berlin, New York: de Gruyter.
- Bakker, Peter 1996 Language intertwining and convergence: typological aspects of the genesis of mixed languages. *Sprachtypologie und Universalienforschung* 49 (1): 9–20.
- Bakker, Peter 2003 Mixed languages as autonomous systems. In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate*, 107–150. Berlin, New York: de Gruyter.
- Benitez-Torres, Carlos M. 2009 Inflectional vs. derivational morphology in Tagdal: a mixed language. In: Masangu Matondo, Fiona McLaughlin and Eric Potsdam (eds.), *Selected Proceedings of the 38th Annual Conference on African Linguistics: Linguistic Theory and African Language Documentation*, 69–83. Somerville, MA: Cascadia Proceedings Project.

- Berruto, Gaetano 2005 Dialect/standard convergence, mixing, and models of language contact: the case of Italy. In: Peter Auer, Frans Hinskens and Paul Kerswill (eds.), *Dialect Contact*, 81–95. Cambridge: Cambridge University Press.
- Boretzky, Norbert and Birgit Iгла 1994 *Wörterbuch Romani – Deutsch – Englisch für den südosteuropäischen Raum. Mit einer Grammatik der Dialektvarianten*. Wiesbaden: Harrassowitz.
- Borg, Alexander 1984 Kormakiti Arabic: Phonological notes and texts. *Journal of Arabic Linguistics* 13, 68–85.
- de Rooij, Vincent A. 1996 *Cohesion through Contrast: Discourse Structure in Shaba Swahili/French Conversations*. Amsterdam: IFOTT.
- de Smit, Merlijn 2010 Modelling mixed languages: some remarks on the case of Old Helsinki Slang. *Journal of Language Contact VARIA* 3: 1–19.
- Flamenbaum, Rachel in press The pragmatics of codeswitching on Ghanaian talk-radio. *International Journal of Bilingualism*.
- Földes, Csaba 2005 *Kontaktdeutsch*. Tübingen: Narr.
- Friedman, Victor A. 2008 Codeswitching and code integration in Romani. In: Max Bane, Juan José Bueno Holle, Thomas Grano, April Lynn Grotberg and Yaron McNabb (eds.), *Proceedings from the Annual Meeting of the Chicago Linguistic Society* 44 (2), 123–138. Chicago: Chicago Linguistic Society.
- Golovko, Evgenyi 1994 Mednij Aleut or Copper Island Aleut: an Aleut-Russian mixed language. In: Peter Bakker and Maarten Mous (eds.), *Mixed Languages: 15 case studies in languages intertwining*, 113–121. Amsterdam: IFOTT.
- Golovko, Evgenyi 2003 Language contact and group identity: The Role of ‘Folk’ Linguistic Engineering. In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate*, 177–209. Berlin, New York: de Gruyter.
- Hentschel, Gerd 2008 Zur weißrussisch-russischen Hybridität in der weißrussischen “Trasjanka”. In: Peter Kosta and Daniel Weiss (eds.), *Slavistische Linguistik 2006/2007*, 169–219. München: Otto Sagner.
- Hilpert, Martin 2013 *Constructional Change in English*. Cambridge: Cambridge University Press.
- Holzinger, Daniel 1993 *Das Rómanes: Grammatik und Diskursanalyse der Sprache der Sinte*. Innsbruck: Institut für Sprachwissenschaft.
- Jarva, Vesa 2008 Old Helsinki Slang and language mixing. *Journal of Language Contact VARIA* 1: 52–80.
- Matras, Yaron 1998 Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36: 281–331.
- Matras, Yaron 2000 Fusion and the cognitive basis for bilingual discourse markers. *International Journal of Bilingualism* 4: 505–528.
- Matras, Yaron 2002 *Romani: A Linguistic Introduction*. Cambridge: Cambridge University Press.
- Matras, Yaron 2003 Mixed languages: re-examining the structural prototype. In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate*, 151–176. Berlin, New York: de Gruyter.
- Matras, Yaron 2007 The borrowability of grammatical categories. In: Yaron Matras and Jeanett Sakel (eds.), *Grammatical Borrowing in Cross-Linguistic Perspective*, 31–74. Berlin, New York: de Gruyter.
- Matras, Yaron 2009 *Language Contact*. Cambridge: Cambridge University Press.
- Matras, Yaron and Peter Bakker 2003 The study of mixed languages. In: Yaron Mat-

- ras and Peter Bakker (eds.), *The Mixed Language Debate*, 1–20. Berlin, New York: de Gruyter.
- Matras, Yaron and Peter Bakker (eds.) 2003 *The Mixed Language Debate*. Berlin, New York: de Gruyter.
- Matras, Yaron and Jeanett Sakel (eds.) 2007 *Grammatical Borrowing in Cross-Linguistic Perspective*. Berlin, New York: de Gruyter.
- McConvell, Patrick 2008 Mixed languages as outcomes of code-switching: recent examples from Australia and their implications. *Journal of Language Contact* 2: 186–212.
- McConvell, Patrick and Felicity Meakins 2005 Gurindji Kriol: A mixed language emerges from codeswitching. *Australian Journal of Linguistics* 25: 9–30.
- Meakins, Felicity 2011 *Case-Marking in Contact. The Development and Function of Case Morphology in Gurindji Kriol*. Amsterdam: Benjamins.
- Meakins, Felicity 2012 Which mix – code-switching or a mixed language? – Gurindji Kriol. *Journal of Pidgin and Creole Languages* 27 (1): 105–140.
- Mous, Maarten 2003 The linguistic properties of lexical manipulation and its relevance for Ma'á. In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate*, 209–236. Berlin, New York: de Gruyter.
- Muysken, Pieter 1981 Halfway between Spanish and Quechua: The case for relexification. In: Arnold Highfield and Albert Valdman (eds.), *Historicity and Variation in Creole Studies*, 57–78. Ann Arbor: Karoma Press.
- Muysken, Pieter 2000 *Bilingual Speech: A Typology of Code-mixing*. Cambridge: Cambridge University Press.
- Myers-Scotton, Carol 1997 *Duelling Languages*. Oxford: Clarendon Press.
- Myers-Scotton, Carol 1998 A way to dusty death: the Matrix Language turnover hypothesis. In: Lenore A. Grenoble and Lindsay J. Whaley (eds.), *Endangered Languages: Language Loss and Community Response*, 289–316. Cambridge: Cambridge University Press.
- Myers-Scotton, Carol 2000 What matters: The out of sight in mixed languages. *Bilingualism: Language and Cognition* 3 (2): 119–121.
- Myers-Scotton, Carol 2003 What lies beneath: Split (mixed) languages as contact phenomena. In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate*, 73–106. Berlin, New York: de Gruyter.
- Németh, József 1953 Zur Kenntnis der Mischsprachen. (Das doppelte Sprachsystem des Osmanischen.) *Acta Linguistica Academiae Scientiarum Hungaricae* 3: 159–199.
- Newton, Brian 1964 An Arabic-Greek dialect. *Word* 20 (3): 43–52.
- Nichols, Johanna 1987 Head-marking and dependent-marking grammar. *Language* 62 (1): 56–119.
- O'Shannessy, Carmel 2004 Light Warlpiri: a new language. *Australian Journal of Linguistics* 25 (1): 31–57.
- O'Shannessy, Carmel 2012 The role of code-switched input to children in the origin of a new mixed language. *Linguistics* 50 (2): 305–340.
- Pagel, Steve 2010 *Spanisch in Asien und Ozeanien*. Frankfurt/M.: Lang.
- Pakendorf, Brigitte 2009 Intensive contact and the copying of paradigms: An Even dialect in contact with Sakha (Yakut). *Journal of Language Contact* VARIA 2: 85–110.
- Perrot, Marie-Ève 1995 “Tu worries about ça, toi?” Métissage et restructurations dans le chiac de Moncton. *Situations du français* 33 (2): 79–85.

- Perrot, Marie-Ève 1998 Les modalités du contact français/anglais dans un corpus chiac: Métissage et alternance codique. *Le français en Afrique: Francophonies, Recueil d'études offert en hommage à Suzanne Lafage* 12: 219–336.
- Perrot, Marie-Ève 2001 Bilinguisme en situation minoritaire et contact de langues: l'exemple du chiac. *Faits de Langues* 18: 129–137.
- Perrot, Marie-Ève 2003 Le français acadien en contact avec l'anglais: analyse de situations distinctes. In: André Magord (ed.), *L'Acadie plurielle: dynamiques identitaires collectives et développement au sein des réalités acadiennes*, 267–279. IEAC (Université de Poitiers) et CEA (Université de Moncton).
- Perrot, Marie-Ève 2005 Le chiac de Moncton: description synchronique et tendances évolutives. In: Albert Valdman, Julie Auger and Deborah Piston-Hatlen (eds.), *Le français en Amérique du Nord*, 307–326. Québec: Presses de l'Université Laval.
- Poplack, Shana 2004 Code-Switching. In: Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier and Peter Trudgill (eds.), *Sociolinguistics. An International Handbook of the Science of Language and Society*, 589–596. Berlin: Walter de Gruyter. 2nd edition.
- Salmons, Joe 1990 Bilingual discourse marking: code switching, borrowing, and convergence in some German-American dialects. *Linguistics* 28: 453–480.
- Sarhimaa, Anneli 1999 *Syntactic Transfer, Contact-Induced Change, and the Evolution of Bilingual Mixed Codes*. (Studia Fennica Linguistica 9.) Helsinki: Finnish Literary Society.
- Schaengold, Charlotte C. 2004 Bilingual Navajo: Mixed codes, bilingualism, and language maintenance. Ph.D. dissertation, Ohio State University.
- Stolt, Birgit 1965 *Die Sprachmischung in Luthers Tischreden. Studien zum Problem der Zweisprachigkeit*. Stockholm: Almqvist & Wiksell.
- Stolz, Thomas 2003 Not quite the right mixture: Chamorro and Malti as candidates for the status of mixed language. In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate*, 271–315. Berlin, New York: de Gruyter.
- Szabó, Csilla Anna 2010 *Language shift und Code-mixing. Deutsch-ungarisch-rumänischer Sprachkontakt in einer dörflichen Gemeinde in Nordwestrumänien*. Frankfurt/M: Lang.
- Thomason, Sarah G. 1995 Language mixture: ordinary processes, extraordinary results. In: Carmen Silva-Corvalán (ed.), *Spanish in Four Continents*, 15–33. Washington D.C.: Georgetown University Press.
- Thomason, Sarah G. 2001 *Language Contact – An Introduction*. Washington D.C.: Georgetown University Press.
- Thomason, Sarah G. 2003 Social factors and linguistic processes in the emergence of stable mixed languages. In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate*, 21–40. Berlin, New York: de Gruyter.
- Thomason, Sarah G. and Terrence Kaufman 1988 *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Vakhtin, Nikolay 1998 Copper Island Aleut: A case of language 'resurrection'. In: Lenore A. Grenoble and Lindsay J. Whaley (eds.), *Endangered Languages: Language Loss and Community Response*, 317–327. Amsterdam: Benjamins.
- Winford, Donald 2003 *An Introduction to Contact Linguistics*. Oxford: Blackwell.
- Young, Hilary A. N. 2002 'C'est either que tu parles français, c'est either que tu parles anglais'. A cognitive approach to Chiac as a contact language. unpubl. Ph.D., Rice University.