# Mixed Languages[1]

*Felicity Meakins*
*University of Queensland/University of Manchester*
*f.meakins@uq.edu.au*
*felicity.meakins@manchester.ac.uk*

## 1. Introduction

Mixed languages[2] are the result of the fusion of two identifiable source languages, normally in situations of community bilingualism. As recently as the 1990s, the existence of these languages had often been denied or labelled as cases of code-switching, adstrate influence or borrowing (see e.g. Greenberg 1999). Nonetheless they were brought to the attention of contact linguistics by Thomason and Kaufman (1988) as a legitimate form of contact language. Bakker and Mous' (1994) and Thomason's (1997) edited volumes drew together substantial amounts of data from various languages which have been identified as being 'mixed', and Bakker's (1997) *A Language of Our Own* provided the first detailed account of a mixed language, Michif. These descriptions of mixed languages have allowed cross-linguistic comparisons of the socio-historical origins of these languages and their typological make-up. They have also stimulated debates such as whether a definable subclass of contact languages called 'mixed languages' actually exists and whether mixed languages are the result of other contact phenomena such as code-switching, or whether special contact processes are required. These discussions formed the basis of a volume edited by Matras and Bakker, *The Mixed Language Debate* (2003).

One of the issues with this earlier work was the dearth of data on mixed languages. Many of the languages studied were spoken by few speakers and scanty information about their socio-historical context and development was available. More recently a number of documentation projects have focused specifically on mixed languages, aiming to provide a rich documentation of actively-spoken mixed languages (c.f. Himmelmann 1998). For example Sri Lankan Malay, Gurindji Kriol[3,4] and Angloromani[5] have been sampled across various social contexts and have been placed within their socio-historical context. As a result, grammatical sketches, ethnographic descriptions of the language ecologies and information about the socio-historical origins of these languages have become available[6]. Additionally, corpora of sound-linked annotated transcripts are now accessible. More specific aspects of mixed languages have been the focus of other projects. For example, the central question of the Aboriginal Child Language project[7] in Australia has been the relationship between language acquisition and shift in highly fluid multilingual contexts. The documentation of new Australian varieties such as the mixed languages, Light Warlpiri and Gurindji Kriol, has formed a large part of this project.

These documentation projects provide the field of contact linguistics with a more solid foundation for the discussion of these mixed varieties. This chapter utilises these new contributions to provide an overview of the debates and more recent developments and issues arising in the study of mixed languages. I discuss the classification and origins of mixed languages in relation to their socio-historical context (§3.1) and typological features (§3.2), as well as some more recent issues including how to ascertain whether

mixed languages are autonomous language systems (§4.1), the description and role of variation in mixed language systems (§4.2), whether two phonological systems can be present in the one language (§4.3) and whether code-switching can lead to mixed language genesis (§4.4). I use data from well-known cases of mixed languages such as Angloromani (§2.1), Ma'á (§2.2), Media Lengua (§2.4), Mednyj Aleut (§2.6) and Michif (§2.7), as well as more recently documented examples such as those from Australia - Gurindji Kriol and Light Warlpiri (§2.8), and other mixed varieties such as Bilingual Navajo (§2.3), Old Helsinki Slang (§2.5) and Sri Lankan Malay (§2.9).

## 2. Examples of mixed languages

The following table lists mixed varieties which have been labelled as mixed languages. The status of a number of these languages is questionable. Indeed it is often the case that following closely in pursuit of every claim for the existence of a mixed language lies a counter-claim about its non-existence. For example, the mixed status of Sri Lanka Malay has been challenged by Smith and Paauw (2006) who suggest it is a creole (§2.9). Other identified cases are more doubtful, such as Barranquenho (Clements 2009; Clements, Amaral, and Luís 2008). Barranquenho is Portuguese with some Spanish influence such as clitic placement. Due to its close proximity to the Portuguese/Spanish border, it is not clear how the restructuring found in this variety would differ from that found along a dialect chain. Given some of these arguments over language classification, this table is meant as an overview of the literature rather than a definitive statement on the status of these languages.

*Table 1     Some languages which have been identified as mixed languages*

| LANGUAGE | COUNTRY | SPEAKERS | MIX | SELECTED SOURCES |
|---|---|---|---|---|
| Angloromani | England | Romani | Grammar: English Lexicon: English and Romani | (Smart and Crofton 1875; Hancock 1970, 1976; Kendrick 1979; Boretzky and Igla 1994; Bakker 1998, 2002; Thomason 2001; Matras and Bakker 2003; Matras et al. 2007; Matras 2010) |
| Barranquenho | Border of Spain and Portugal | Barranquenho | Grammar and lexicon: Portuguese with some Spanish influence | (Clements, Amaral, and Luís 2008; Clements 2009) |
| Callahuaya | Bolivia | Callahuaya Travelling Healers | Lexicon: Puquina Grammar: Quechua | (Muysken 1994, 1997) |
| Chindo | Indonesia | Peranakan Chinese | Lexicon: Malay Grammar: Javanese | (Dreyfuss and Oka 1979; Matras and Bakker 2003) |
| Gurindji Kriol | Australia | Gurindji | VP: Kriol[8] NP: Gurindji Lexicon: Gurindji and Kriol | (Dalton et al. 1995; Charola 2002; McConvell and Meakins 2005; McConvell 2008; Meakins 2008a, 2008b, 2009, 2010a, 2010b, to appear-a, to appear-b, to appear-c) |
| Jenisch | Germany | Jenisch traders | Grammar: German | (Matras 2003, 2000, 2009) |

| | | | Lexicon: Rotwelsch[9], Hebrew, Romani, Romance | |
|---|---|---|---|---|
| Lekoudesch | Germany | Jewish cattle traders | Grammar: Judeo-German<br>Lexicon: some Hebrew | (Matras 2003, 2000, 2009) |
| Light Warlpiri | Australia | Warlpiri | VP: Kriol<br>NP: Warlpiri<br>Lexicon: Warlpiri and Kriol | (O'Shannessy 2005, 2006, 2008, forthcoming) |
| Ma'á | Tanzania | Mbugu | Grammar: Bantu Core<br>Lexicon: Cushitic | (Brenzinger 1987; Mous 1994, 2000, 2003b, 2003a; Thomason and Kaufman 1988; Thomason 1997, 1997; Myers-Scotton 2003) |
| Media Lengua | Ecuador | Quechua | Lexicon: 90% Spanish<br>Grammar: Quechua | (Muysken 1981, 1994, 1997; Myers-Scotton 2003; Gómez Rendón 2008) |
| Mednyj Aleut | Bering Strait, Russia | Aleut | VP (finite): Russian<br>NP: Aleut<br>Lexicon: 90% Aleut | (Golovko and Vakhtin 1990; Golovko 1994, 1996; Sekerina 1994; Thomason 1997; Myers-Scotton 2003) |
| Michif | Canada | Metis | VP: Cree<br>NP: French<br>Lexicon: Cree verbs, French nouns | (Bakker 1994, 1997; Bakker and Papen 1997; Papen 1987, 1987, 2003; Rosen 2000; Myers-Scotton 2003) |
| (Bilingual) Navajo | United States | Navajo | Grammar: Navajo<br>Lexicon: Navajo and English | (Schaengold 2003, 2004) |
| New Tiwi | Australia | Tiwi | VP: Tiwi<br>NP: Aboriginal English/Kriol | (Lee 1987; McConvell 2002) |
| Old Helsinki Slang | Finland | Finnish and Swedish gangs | Grammar: Finnish<br>Lexicon: 80% Swedish | (Paunonen 2006; Jarva 2008; de Smit 2010) |
| Shelta | Ireland | Irish Travellers | Grammar: English<br>Lexicon: Irish | (Grant 1994; McCann, Síochain, and Ruane 1994) |
| Sri Lanka Malay | Sri Lanka | Malay | Forms: Contact variety of Malay<br>Grammar: Tamil and Sinhala | (Smith, Paauw, and Hussainmiya 2004; Smith and Paauw 2006; Bakker 2003; Slomanson 2006, 2007; Ansaldo 2008; Nordhoff 2009) |
| Sri Lanka Portuguese | Sri Lanka | Portuguese | Forms: Contact variety of Portuguese<br>Grammar: Tamil and Sinhala | (Smith 1977, 1979a, 1979b, 1984, 2001; Bakker 2003) |
| Wutun | China/Tibet | Tibetan | Lexicon: Northwest Mandarin, Amdo Tibetan, Mongolian<br>Grammar: Amdo Tibetan | (Chen 1986; Lee-Smith and Wurm 1996; Janhunen et al. 2008) |

The languages which are discussed in this paper have been chosen partly because they are representative of the typological composition and socio-historical background of the languages described as 'mixed', and partly because they are the best documented examples of mixed languages. Some of these languages such as Media Lengua (§2.4), Ma'á (§2.2), Michif (§2.7) and Mednyj Aleut (§2.6) have been discussed extensively in the literature and require little introduction. Other languages have been identified more recently, for example Gurindji Kriol and Light Warlpiri (§2.8). Mixed languages such as Angloromani (§2.1) and Sri Lanka Malay (§2.9) have been present in the literature on mixed languages; however more information about their origins and structure is now available as a result of recent documentation projects. These languages will be presented in order of degree of mixing from the languages which exhibit predominantly lexical mixing such as Angloromani, Ma'á, Bilingual Navajo, Media Lengua and Old Helsinki Slang, to languages which source significant amounts of structural resources from two languages, such as Mednyj Aleut, Michif, Gurindji Kriol and Light Warlpiri. This continuum of mixing is discussed in §2.10, with the sections that follow an exploration of the socio-historical and linguistic origins of these outcomes of language contact.

**2.1 Angloromani**

Most mixed languages exhibit a split between the lexicon and grammar. Bakker (2003: 125) calls these G(rammar)-L(exicon) mixed languages and lists 25 in a typological survey. The first example of a G-L mixed language presented here is Angloromani. Angloromani is spoken by some Romanies in Britain. They continue to be a travelling population, many of whom live in caravans. Romanies who now live in permanent accommodation generally do not speak this mixed language (Matras 2010). Matras et al (2007) consider Angloromani to be endangered. Currently it is not the language of conversation but rather it is restricted to individual utterances. These utterances can be characterised as the use of a restricted set of Romani-derived lexicon, which Matras et al (2007) call a lexical reservoir, within an English grammatical frame. This lexical reservoir exists largely in parallel with English lexicon and is drawn on in situations where speakers want to mark a sense of solidarity or group cohesion. In this respect, Matras et al (2007) find that the use of these utterances is prompted by the presence of outsiders or the emotive content of the speech act. An example of the coupling of Romani-derived lexicon with English grammar is given in (1). This sentence contains Romani words inserted into an English frame, for example nouns *mush* (man), a verb *rocker* (talk) and a function word *maw* (NEG). The Romani-derived words are given in italics.

(1)     *Maw* be *rocker*ing in front of the *mush* and *rakli*!
        "Don't be talking in front of the man and [the] girl!" (Matras et al. 2007: 170)

Though the grammatical frame of Angloromani is predominantly English, some structural differences can be observed. For instance, Angloromani also contains some remnants of Romani morphology such as a genitive *-engra* suffix which attaches to lexical roots to create a related word - *masengra* (from *mas* 'meat') (Matras et al. 2007).

Matras et al (2007) also observe that Angloromani speakers do not always use the definite article, aspect and existential auxiliaries and co-referential pronouns in places where they would be expected in English. They argue that these features are not specifically Romani but they indicate that Angloromani has slightly different grammatical rules to English.

The status of Angloromani as a mixed language is somewhat controversial. It was originally described as a dialect of English which contained large amounts of Romani vocabulary (Smart and Crofton 1875). Hancock (1970, 1976) then raised the possibility of it being classified as a pidgin or creole language. Kendrick (1979) suggested it was an English ethnolect. It was only later that it was described as a mixed language (Bakker 1998; Thomason 2001)[10]. Matras et al (2007) argue that these various descriptions of Angloromani have been largely a product of the interests of linguistic research at the time. For example, Hancock's suggestion of labelling of Angloromani as a creole language came about at a time when this was the dominant framework for discussing contact languages. Similarly the classification of Angloromani as a mixed language occurred when mixed languages became a serious object of study with Thomason and Kaufman's (1988) seminal work. Matras et al's (2007) description of Angloromani as making use of a lexical reservoir is the most recent characterisation of this language.

Equally contentious is the origins of Angloromani. Thomason and Kaufman (1988: 103-104) suggest that Angloromani is the result of the wholesale adoption of the English grammatical system coupled with the maintenance of lexical material from Romani. However the case for massive grammatical borrowing is not clear cut, and counter claims which favour relexification of the dominant language using the ancestral language are made for Angloromani. For example Boretzky & Igla (1994: 61) suggest that Angloromani came about after the Romani had already shifted to English. Under this model, the mixed language is the result of a U-turn, that is an attempt by the Romanies to reclaim their heritage language through the use of Romani vocabulary.

Angloromani is not the only mixed language to make use of Romani. Other varieties on mainland Europe include languages which mix Romani with Portuguese, Spanish, Catalan, Basque, French, Low German, Norwegian, Swedish, and Turkish.

## 2.2 Ma'á

Another mixed language, which makes use of a reservoir of lexical material from the ancestral language, is Ma'á. This mixed language is spoken by Mbugu communities in the Usambara mountains in Tanzania. Like Angloromani, Ma'á is spoken alongside one of its source languages, Mbugu (Bantu). The Mbugu were originally a Cushitic-speaking group from Lackipya in Kenya. In order to escape persecution from the Maasai, they moved to the Usambara mountains via the Pare mountains (Mous 2003a). The mixed language, Ma'á is considered to be the result of resisting assimilation with the neighbouring Pare. In this respect it represents the stubborn persistence of an ethnic group (Mous 1994: 175-76). Ma'á combines a Bantu grammar, similar to Pare, with a lexicon composed of Southern Cushitic and Bantu words. In the following example, the non-Bantu elements are italicised.

(2)     hé-*ló*      mw-a*girú* é-*sé*-we         kimwéri    *dilaó* w-a

16-have     1-elder     1-call-PST:PF    Kimweri   king  1-CON
"There was an elder called Kimweri." (Mous 2003a: 9)

Thomason (1997: 481-83) believes that Ma'á is the product of massive grammatical borrowing from Pare, including inflectional categories, for example noun classes. However this view differs sharply from Brenzinger (1987), Sasse (1992) and particularly Mous (1994, 2000, 2003a, 2003b) who believe that Ma'á is a conscious and deliberate result of an attempt to undo a shift to Pare, where speakers tried to relearn their ancestral language. Mous (2003a, 2003b) suggests this happened through a paralexification process where a Bantu lexicon, and a combined Cushitic and Maasai lexicon exist in parallel. Mbugu draws from the Bantu lexicon, whereas speakers use Cushitic and Maasai words in Ma'á. In this sense, he considers Ma'á to be a register of Mbugu (1994: 96-97; 2003a), not entirely unlike, but probably more extreme than, the cases of lexical manipulation found in urban youth languages, slang and taboo codes (Mous 2003b: 217). This notion of paralexification is also similar to lexical reservoir described for Angloromani.

## 2.3 Bilingual Navajo

Both Ma'á and Angloromani can be characterised as deriving their grammar from the introduced language, Bantu and English respectively, and lexical items from the ancestral language, Cushitic and Romani, respectively. Bilingual Navajo demonstrates the reverse, combining the grammar of Navajo, the ancestral language, with some English nouns, adjectives and verbs (Schaengold 2003, 2004).

Bilingual Navajo was developed by Navajo children when they were placed in boarding schools in the early to mid-twentieth century. They were often placed in these schools against their will and were punished for speaking Navajo. As a result, some children ceased to use Navajo, though many more spoke it in secret, using English words where they had not had a chance to learn the Navajo equivalents. This mixed code became an in-group language for many Navajo-English bilinguals who graduated from the boarding school system and moved back to their Navajo communities. Eventually this mixed code stabilised into the mixed language spoken today (Schaengold 2004: 8).

Nowadays few Navajo children learn Navajo as their first language, although they hear it spoken among adults. Indeed most Navajos over 60 years of age are monolingual in Navajo. All children learn English through the school system and require English to function outside of the Navajo Nation. Thus the language ecology of the Navajo Nation now consists of Navajo, Navajo English, and Bilingual Navajo. In addition, bilingual speakers may change from one code to another in accordance with the situation and interlocutor. Depending on the age of a Bilingual Navajo speaker, s/he may be also fluent in either Standard Navajo or English or both. Younger speakers show greater capacity in English and older speakers tend to have full control over Standard Navajo. The situation described for Navajo people and their languages is also the case for Gurindji and Warlpiri people in northern Australia as discussed in §2.8.

Bilingual Navajo is characterised by English noun, adjective and verb insertions in a Navajo morpho-syntactic frame. English insertions are nativised to Navajo phonotactic rules. Although Navajo nouns are relatively uninflected with only some directional and possessive suffixes, English nouns are fully integrated morphologically. English nouns

may also be inflected with Navaho discourse and interrogative clitics. Adverbs are created from English nouns with the use of the Navaho suffix *-go* (Schaengold 2004: 58-60). English verbs are also found but they require the use of a conjugated Navajo verb *áshlééh* 'to prepare/make' as an auxiliary (see also §4.1.2) The following example demonstrates the use of an English noun *face* and a verb *clean* in Bilingual Navajo. Navajo elements are italicised.

(3)    *bi*-face          clean          *doo   bee       áshlééh    da*
       3POSS-face      clean          not    3INST     1SG.make  not
       "I didn't wash his face." (Schaengold 2004: 53)

Navajo word order is also maintained in Bilingual Navajo. Constituents are ordered according to topic and comment rather than by ordering grammatical roles such as subject and object, although the default SOV order is found in sentences with two arguments. More generally, information about grammatical roles is provided by the polysynthetic verbs, therefore a strict word order is not required by the syntax (Schaengold 2004: 44).

## 2.4 Media Lengua

Another mixed language which retains the grammar of the ancestry language is Media Lengua. This mixed language is spoken in Central Ecuador by a Quechuan group known as the Obreros (Muysken 1997: 374). The morpho-syntactic frame of Media Lengua is essentially Quechua (the ancestral language) and therefore agglutinating with around 90% of its stems replaced by Spanish forms (the introduced language). Muysken (1997: 399) also claims that Spanish does contribute some structural features which are not found in Quechua, for example the structure of embedded WH-questions. Other features of the grammar seem to have developed independently. The following example demonstrates the pattern of Spanish stems with Quechuan suffixes (italicised).

(4)    Unu   fabur-*ta*          pidi-*nga-bu*        bini-*xu-ni*.
       one   favour-ACC        ask-NOM-BEN    come-PROG-1
       "I come to ask a favour." (Muysken 1997: 365)

Muysken suggests that the structure of Media Lengua is the result of a relexification process where Quechuan stems were replaced with Spanish forms on the basis of semantic equivalence. It began developing from around 1967. Around this time many young men started working in the construction industry in a nearby provincial town and learning Spanish. This was the group who created Media Lengua. Muysken (1997: 376) claims that the genesis of this mixed language occurred "because acculturated Indians could not completely identify with the traditional Quechua culture or the urban Spanish culture". Now Media Lengua has native speakers, although the language ability of Media Lengua speakers varies. Older speakers of Media Lengua continue to speak Quechua and no Spanish whereas younger speakers speak a variety of urban Spanish fluently and rarely speak Quechua without some mixing (Muysken 1997: 374). This community profile is not unlike that of Bilingual Navaho (and Gurindji Kriol and Light Warlpiri as

discussed in §2.8) although there is nothing to suggest that it represents a language shift situation where Spanish is becoming increasingly dominant, with the mixed language representing a stage in this shift.

## 2.5 Old Helsinki Slang

Mixes consisting of the grammar from one language and the lexicon from another may also be the result of a compromise between two different groups wishing to mark a new identity. Old Helsinki Slang is one such example. This mixed language was spoken in Helsinki between 1890 and 1950 by *saki* gangs which consisted of both Finnish and Swedish speaking boys and young men (de Smit 2010; Jarva 2008; Paunonen 2006). This mixed language has a similar structure to Media Lengua where Swedish stems are inserted into a Finnish morph-syntactic frame. According to Paunonen (2006: 51) approximately 80% of stems in Old Helsinki Slang are of Swedish origin, although, as Jarva (2008: 66) suggests, this figure is mostly likely the upper limit and a lot of variation probably existed between the use of Swedish and Finnish vocabulary making it more like the lexical reservoir described for Angloromani (§2.1) or the process of paralexification which is used by Ma'á speakers (§2.2). Jarva (2008: 68) admits, however, that 80% of verbs, adjectives and nouns in a 200-word Swadesh list are of non-Finnish origin. Nonetheless function words and closed class lexical items in Old Helsinki Slang are derived from Finnish, for example conjunctions, adpositions, pronouns and numerals (Jarva 2008: 67). An example of Old Helsinki Slang is given below. Finnish elements are represented by italics.

(5)  föra-*kaa*       *nyt*  *toi*  *Väiski*      bastu-*un*  *ja*        tvetta-*kaa*
    take-IMP.2PL   now  that  NAME       sauna-ILL  and        wash-IMP.2PL

    se-*n*        klabbi-*t*
    he-GEN      foot-PL

"Take Väiski to the sauna and wash his feet." (Jarva 2008: 53)

Old Helsinki Slang originated in Helsinki at the end of the nineteenth century as a result the migration of Finnish to Helsinki and the increased bilingualism among Swedish speakers. Helsinki was established in a Swedish-speaking area of Finland. At the time Swedish had a high status and was spoken by the upper-class in Finland. Swedish was also the language of most teachers, university staff, government officials and priests, and Finnish-speaking migrants to Helsinki would often transition to Swedish. This situation changed after the 1870s with increasing numbers of Finnish immigrants to Helsinki and the increasing use of Finnish by the upper-classes. Increasingly Swedish speakers became bilingual in Finnish, though Finnish speakers generally remained monolingual (Jarva 2008: 54-55). Old Helsinki Slang was born in the working class areas in the northern quarters of Helsinki. Two thirds of the population were Finnish speakers, however Swedish was still considered more prestigious. It was common for boys and young men to spend most of their time on the streets due to the lack of compulsory education and over-crowding in apartments. Many of these boys formed gangs called *saki* which

consisted of both Finnish and Swedish speakers. The language mixture which emerged from these gangs was probably the result of a communicative compromise between these different groups of speakers and a way of marking a new in-group identity.

Although Old Helsinki Slang can be characterised as Swedish stems alternating with Finnish suffixes, this mixed language also contains innovative forms. For example, the variety of Finnish used in Old Helsinki Slang differs from both Standard Finnish and Finnish dialects spoken at the time. Verbs (of both Swedish and Finnish origin) are conjugated according to a mix of the first and fourth conjugational classes of a Finnish dialect spoken near Helsinki (Jarva 2008: 73). Additionally many words of both Swedish and Finnish origin are augmented by slang suffixes such as *-ari, -is* and *-tsi*. These suffixes are unique to Old Helsinki Slang, differing from the epenthetic vowels required to borrow Swedish words into standard Finnish (Jarva 2008: 70).

## 2.6 Mednyj Aleut

Although most mixed languages such as Angloromani, Ma'á, Bilingual Navajo, Media Lengua and Old Helsinki Slang exhibit a split between the lexicon and grammar, other languages are more structurally mixed as the following sketches of Mednyj Aleut, Michif, Gurindji Kriol and Light Warlpiri demonstrate. In these languages, both of the source languages contribute to the structure of the resultant mix creating a composite morpho-syntactic frame. I begin with *Mednyj Aleut*.

Mednyj Aleut was spoken on Mednyj Island in the Bering Strait until recently. The island was first settled by Russian fur seal hunters in the early 19th century, and Aleutians were brought to the island soon after. Marriages between Russian men and Aleutian women resulted, and the subsequent population were called "creoles" (Golovko and Vakhtin 1990; Sekerina 1994). Thomason (1997: 462 onwards) suggests that it was the creoles who created Mednyj Aleut. She assumes that they were bilingual in both languages but their half-way position in society led them to mark themselves out as a separate group. Nonetheless Golovko (1994: 117) claims that they considered themselves Aleut, and regarded their language as a variety of Aleut. The use of Mednyj Aleut declined in the 1940s when the Russians introduced Russian education (Thomason 1997). At the last report, only 10-12 Mednyj Aleut speakers remained (Golovko 1994: 113).

The lexicon of Mednyj Aleut is 90% Aleut. The structure consists of many Aleut nominal inflections, including two case distinctions, absolutive and relative. Aleut also provides various derivational suffixes such as agent, instrumental, location, detransitive, inchoative markers and so on. Mednyj Aleut derives much of its finite verbal inflectional morphology from Russian, including portmanteau morphemes which express tense, number, person markers; and a negative verb prefix derived from the Russian negative particle *ne* (Golovko and Vakhtin 1990; Sekerina 1994; Thomason 1997). Some of this structure is demonstrated in (6). All Aleut elements are in italics.

(6)  segodnja  *ta:ŋa-x̂*  bud-ut  *su-la-x̂či*-t'
      today      spirits-SG  FUT- 3PL  take-MULT-CAUS-INF
      "Today they will sell spirits." (Golovko 1996: 67)

This structural outcome does not provide clear clues about the genesis of Mednyj Aleut. Though it is the result of mixed marriages, whether Aleut elements were transferred to Russian or vice versa remains a point of contention. Different linguistic environments at the point of formation have also been proposed. Golovko (1994) suggests that it is the result of creative word play, whereas Thomason (1997) proposes the less consciously manipulative route of code-switching.

## 2.7 Michif

Michif is also the result of mixed marriages, in this case between Plains Cree-speaking women and French Canadian fur traders. Its genesis probably occurred in the early 1800s when the fur trade was strong. Nowadays it is spoken in Métis communities in Canada, northern parts of the United States, and the prairie provinces of Canada. Michif is spoken by the elderly descendants of the fur traders and Cree women. It is an endangered language with fewer than 1000 speakers, mostly elderly, and no children acquiring it as a first language (Bakker 1994, 1997).

Like Mednyj Aleut, Michif shows a great degree of structural mixing. It combines the verbal system of Cree with the nominal system of French, and is classified as a V(erb)-N(oun) mixed language under Bakker's (2003: 122) typology. Bakker observes that the systems from both languages are absorbed in their entirety without simplification. In terms of word classes, Michif is composed of 83-94% French nouns and 88-99% of Cree verbs depending on the speaker. Question words, post-positions, demonstratives and person pronouns are mostly Cree; and prepositions and numerals are almost exclusively French. This NP/VP language divide is also reflected in the grammar - verbal inflections are derived from Cree and the nominal system is dominated by French, as demonstrated in the presence of French plural marking in the article and adjectival agreement. The NP-VP split is clearing demonstrated in (7). Cree elements italicised.

(7)  *êkwa pâstin-am*  sa  bouche  *ôhi*  le  loup  *ê-wî-otin-át*
     and  open-he.it  his.F mouth  this.OBV  the.M wolf  COM-want-take-he.him
     "And when the wolf came to him, he opened his mouth." (Bakker 1997: 5)

Some Cree influence in the NP can also be observed. For instance, the Cree obviative marker is used with French nouns and the locative suffix is calqued (Bakker 1997: 89). However the use of Cree nominal morphology is only found in conjunction with the very few Cree nouns. For example, Michif speakers use both the French preposition and Cree locative suffix, yet this suffix is only found on the very few Cree stems present (Bakker 1997: 110). This lack of mixing between stems and morphology contrasts with Gurindji Kriol and Light Warlpiri where stems and suffixes are commonly derived from different languages in the same word (see §2.8).

As with all mixed languages, the origins of Michif is a matter of speculation based on its resultant structure. Bakker claims that Michif is the result of language intertwining, a process particular to mixed languages (see §3.2.2.1). McConvell (2002 c.f. Bakker 1997; 2008) suggests that the typology of Cree (head-marking) contributed to the maintenance of Cree verbal morphology (see also §3.2.2.3). The conventionalisation of French-Cree code-switching has also been offered as an explanation for its formation

(e.g. Drapeau 1991), though Bakker (2003: 128 onwards) gives arguments against this claim.

## 2.8 Gurindji Kriol and Light Warlpiri

Two more recently identified V(erb)-N(oun) mixed languages come from Australia - Gurindji Kriol and Light Warlpiri. *Gurindji Kriol*[11] is spoken by Gurindji people in northern Australia. Its source languages are apparent from the name of the language - Gurindji, a Pama-Nyungan language, and Kriol, an English-lexified creole language. Gurindji Kriol originated from contact between non-Indigenous settlers and Gurindji people. From 1855 onwards, the traditional lands of the Gurindji and neighbouring groups were seized by colonists who were searching for good cattle pastures. After initial attempts to cull the original inhabitants, cattle stations were set up and the remaining Gurindji people were brought to work on the stations in slave-like conditions with other Aboriginal groups (Hardy 1968; Wavehill 2000; Meakins 2008b).

  The cattle station owners communicated with the Aboriginal workers in a cattle station pidgin which later developed into a creole now referred to as Kriol. The Gurindji added this language to their communicative repertoire and it is likely that code-switching and a certain amount of levelling between Gurindji and the neighbouring dialects provided fertile ground for the formation of this mixed language (McConvell and Meakins 2005). The fact that a mixed language formed amongst Gurindji people is significant given that many Aboriginal groups in northern Australia have shifted almost completely to Kriol. In this respect Gurindji Kriol represents an attempt to maintain an ancestry language under severe cultural incursion and functional pressure from Kriol (Meakins 2008b). Gurindji Kriol continues to be spoken alongside Gurindji and Kriol. It is the first language of all Gurindji people under the age of 35 years. Older Gurindji people speak Gurindji amongst themselves albeit often mixed with Kriol in code-switching. All Gurindji people speak Kriol to varying extents when they visit Kriol-speaking areas to the north, though they do not speak Kriol among themselves. English is the language of the school, media and other services in Kalkaringi but plays little role in people's home lives (Meakins 2008a).

  Structurally, Kriol contributes much of the verbal grammar including tense and mood auxiliaries, and transitive, aspect and derivational morphemes. Gurindji supplies most of the nominal structure including case and derivational morphology. In this respect the structure of Gurindji Kriol is quite similar to the V-N split seen in Michif. However, in Gurindji Kriol nouns and verbs also come from both source languages which contrasts with Michif where the nominal structure and nouns are derived from one language and the verbal structure and verbs are derived from the other. For example, although the case morphology is derived from Gurindji, it is used productively with Kriol nouns, as shown in (8). Gurindji elements are in italics. Note that Gurindji verbs are also in evidence in Gurindji Kriol, although Kriol provides the TMA markers.

(8) dat *karu-ngku*  i=m  luk hol-*ta* *walyak*.
   the child-ERG   3SG=PRS look hole-LOC inside
   "The child looks inside the hole." (Meakins 2007: 364)

Both languages also contribute small amounts of grammar to the systems they do not dominate in Gurindji Kriol. For example, the Gurindji continuative suffix is found in the VP, and Kriol determiners are common in the NP. Kriol also provides Gurindji Kriol with an SVO word order, though the word order is more flexible than Kriol with information structure determining word order to some extent which reflects word order patterns in Gurindji. Complex clauses are constructed using both Gurindji and Kriol strategies, for example coordinating and relative clauses use Kriol conjunctions and relative pronouns, and subordinate clauses are formed using Gurindji-derived case and inchoative marking. In terms of the lexicon, Gurindji Kriol derives its lexicon relatively evenly from both languages. (Charola 2002; Meakins 2007, 2008a, 2009; Dalton et al. 1995).

Although it is easy to identify which language the lexemes and morphemes are derived from, this approach is actually quite deceptive. Gurindji Kriol is not merely the result of a simple replication of features from these languages. Though Gurindji Kriol bears some resemblance to both of its source languages, it uses the forms from these languages to function within a unique language system. For example Gurindji Kriol has adopted ergative marking from Gurindji; however, where ergative marking is obligatory in Gurindji, it is optional in Gurindji Kriol. The ergative marker has become optional with the adoption of word order from Kriol as the main means of marking arguments. As a result, the function of the ergative marker has shifted to marking the prominence and agentivity of a discourse entity (Meakins 2009, to appear-b; Meakins and O'Shannessy 2010). Ergative marking is just one example of the types of changes which language features have undergone in the process of mixed language genesis. The need to describe the structure of mixed languages beyond the mixing of forms is the topic of §5.

*Light Warlpiri* is spoken just 100 km from Gurindji Kriol. Structurally it is very similar to Gurindji Kriol in exhibiting a split in the language dominance of the nominal and verbal system. For example, the nominal structure comes from the heritage language, Warlpiri, and Aboriginal English/Kriol provides the verb structure (O'Shannessy 2005), and is also classified as a V-N mixed language in this respect. Light Warlpiri is also quite mixed lexically; however a crucial difference between Gurindji Kriol and Light Warlpiri is that verbs are almost always derived from Aboriginal English/Kriol in Light Warlpiri. This differs from Gurindji Kriol where many Gurindji Kriol verbs are Gurindji in origin. In the example below, Warlpiri elements are italicised.

(9)     en      *karnta-pawu*      i-m                  kam  geit-*kirra*.
        and    girl-DIM            3SG.S-NFUT      come gate-ALL
        "And the girl came to the gate." (O'Shannessy 2006: 32)

Although Light Warlpiri is structurally similar to Gurindji Kriol, it has a different social function. O'Shannessy (2006) suggests that Light Warlpiri is an expression of the identity of a particular Warlpiri community rather than an attempt to maintain an ancestry language. Unlike Gurindji Kriol speakers, Light Warlpiri speakers also speak their heritage language, Warlpiri. Warlpiri also continues to be acquired by children, though their main language of use is Light Warlpiri. Adults and children alike code-switch between these languages and O'Shannessy (2008) suggests that it is likely that Light Warlpiri also found its origins in such linguistic practices.

The close proximity of Gurindji Kriol and Light Warlpiri begs the question of

whether other such mixed languages exist or have existed in Australia. Indeed a youth variety of Tiwi (spoken in the north on an island off Darwin) has been described (Lee 1987). Additionally Disbray (2008, 2009) and Disbray and Simpson (2005) have described a creolised variety of English spoken 500 kms away called *Wumpurrarni[12] English*. This variety includes some inflectional morphology (possessive and allative case suffixes) from the local Aboriginal language, Warumungu. What is interesting about Wumpurrarni English is that older speakers speak something which is akin in structure to Gurindji Kriol and Light Warlpiri. For instance, their variety includes more Warumungu structural elements, and more rarely ergative case marking, as shown in (10) on the corrected nominal. Warumungu elements are shown in italics. This similarity in structure suggests that a mixed language may have been spoken in this area with the Warumungu structural features gradually eroded by Kriol functional equivalents such as prepositions.

(10)  *nyili*      bin   pok-im   *jina*, no   masbi     *wintirrij-ja*.
      prickle    PST   pierce-TR   foot, no   must.be   stick-ERG
      "A prickle pierced (his) foot, no maybe a stick."[13]

## 2.9 Sri Lanka Malay

The final mixed language to be discussed is Sri Lanka Malay. This language is quite different from other mixed languages. Not only is most of the lexicon derived from Malay, but all of the morphology and function words also find their forms in Malay though the underlying structure is Tamil. In this respect it contrasts with Media Lengua, for example, because the forms of the grammatical elements are still Quechuan though the lexicon is mostly Spanish-derived. The composition of Sri Lanka Malay makes it difficult to classify as either an Austronesian (lexicon) or Dravidian language (grammar).

Sri Lanka Malay is spoken in a number of communities in Sri Lanka by the Malay minority who make up less than 1% of the population of Sri Lanka. Lexically Sri Lanka Malay consists almost entirely of words from a Malay-based trade language called Vehicular or Bazaar Malay (Austronesian) (Smith, Paauw, and Hussainmiya 2004: 200). However its grammar is derived from Tamil (Dravidian) and perhaps Sinhala[14] (Indo-Aryan) with the result that it became unintelligible to other Malay speakers. Sri Lanka Malay seems to be the result of Vehicular Malay developing from an isolating language to an agglutinating language under the influence of Tamil. It has also developed SOV word order, postpositions, and pre-nominal determiners and adjectives due to this contact (Hussainmiya 1987; Nordhoff 2009). This structure is demonstrated in (11).

(11)  *Sir*        *anak-pada-yang*  *ruuma-nang*    *e-luppa*.
      teacher     child-PL-ACC    house-DAT      PST-send
      "The teacher sent the children to school." (Ansaldo 2008: 27)

Sri Lanka Malay is an endangered language, now restricted to home use and generally not being spoken by younger generations, with Sinhala/English bilingualism becoming dominant (Smith and Paauw 2006: 160; Ansaldo 2008; Nordhoff 2009). Traditionally the Malay community have had close ties with the Tamil-speaking Moor community who are also Muslims. The Sri Lanka Malays are descendants of immigrants who were brought to

Sri Lanka at different times by Dutch (1656 onwards) and British colonists (1796 onwards). Although they are called Malays, they came from a number of places including Banda, Balu, Java, with only a Malay trade language in common.

There are different views on how Sri Lanka Malay developed. The first is that it is a pidgin which was creolised by the children of mixed marriages between Malay men who spoke a Vehicular Malay and Tamil-speaking Moor women. Smith and Paauw (2006) suggest that the mothers must have tried to make Malay the language of the home, and the children nativised the Malay pidgin in a process of creolisation. Perhaps an added factor here is that the women learnt Vehicular Malay or the Malay language of their husbands but their version of Malay showed first language interference from Tamil which was the inter-language that the children nativised. Smith and Paauw (2006) propose that this whole process occurred before Malay was reintroduced by the British in the schools between 1802 to 1873 therefore creating a diglossic situation. During this period Sri Lanka Malay had extended contact with the Malay language (Smith, Paauw, and Hussainmiya 2004).

Bakker (2003: 116) proposes another interpretation of the resultant structure of Sri Lanka Malay. He does not dispute the socio-historical scenario presented by Smith et al, but interprets the structural outcome of Sri Lanka Malay as a subcategory of mixed language which he calls a *converted* language (see §3.2). By a converted language, Bakker is drawing similarities to a process Ross (1996, 2001) calls *metatypy*. Metatypy refers to the maintenance of the surface forms of one language, that is its lexical and morphological material, which are restructured according to the grammar of another language. Bakker suggests that this convergence of Vehicular Malay and Tamil occurred much later than suggested by Smith. Both Smith (2006) and Ansaldo (2008) dispute this claim on historical and linguistic grounds which will not be discussed here.

The third theory about the origins of Sri Lanka Malay is proposed by Ansaldo (2008). Ansaldo supports Bakker's genesis hypothesis; however he believes that it took place over a longer time span. Ansaldo challenges the claim that pervasive intermarriage occurred between Malay immigrant soldiers and Moor women, noting that the Malay soldiers generally brought their wives and children to Sri Lanka, at least during the British colonial phase. This observation has ramifications for a genesis scenario. Tamil and Sinhala would have been present but external to the Malay community and therefore not languages of acquisition. This picture does not support an abrupt nativisation hypothesis but rather change that took place over an extended period of contact. Under Ansaldo's hypothesis, the Malays were more segregated as a community than suggested by Smith, which suggests they must have been highly multilingual in order for such a pervasive restructuring to have take place. How this degree of multilingualism was achieved if they were more segregated is not clear.

## 2.10    Degrees and types of mixing

The languages sketched out in the previous sections demonstrate the degree of typological variation which exists in the category of contact languages labelled 'mixed languages' (see also Matras 2000). One obvious way in which these mixed languages vary is in their degree of grammatical mixing. For most mixed languages, the grammar is predominantly derived from only one of the source languages, for example Angloromani,

Ma'á, Bilingual Navaho, Media Lengua and Old Helsinki Slang. In other languages both source languages contribute significant amounts of grammar, for example Michif, Mednyj Aleut, Light Warlpiri and Gurindji Kriol. This degree of grammatical mixing is represented on the continuum below.
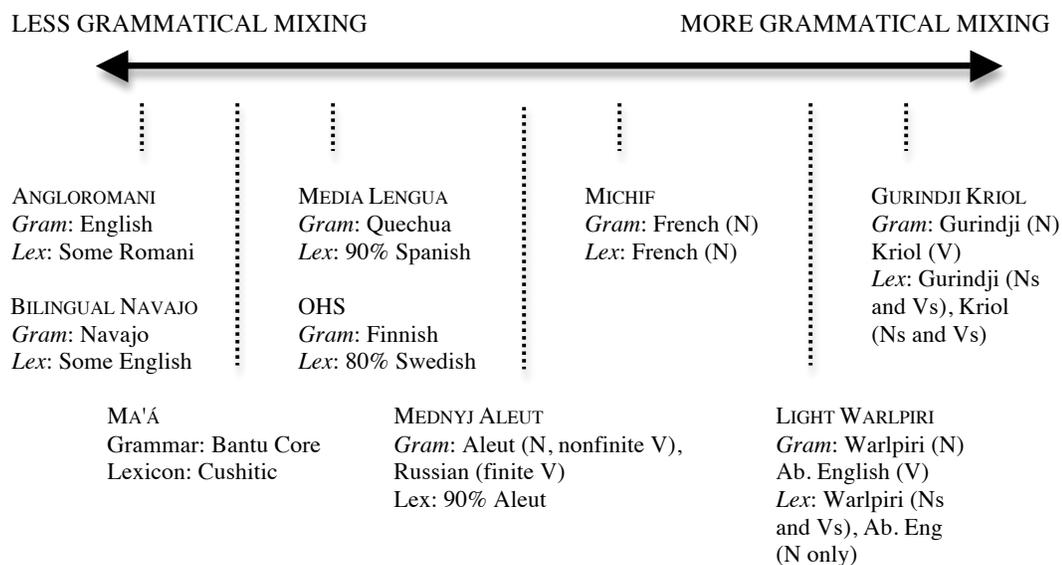
LESS GRAMMATICAL MIXING                    MORE GRAMMATICAL MIXING

ANGLOROMANI
*Gram*: English
*Lex*: Some Romani

BILINGUAL NAVAJO
*Gram*: Navajo
*Lex*: Some English

MEDIA LENGUA
*Gram*: Quechua
*Lex*: 90% Spanish

OHS
*Gram*: Finnish
*Lex*: 80% Swedish

MICHIF
*Gram*: French (N)
*Lex*: French (N)

GURINDJI KRIOL
*Gram*: Gurindji (N)
Kriol (V)
*Lex*: Gurindji (Ns and Vs), Kriol (Ns and Vs)

MA'Á
Grammar: Bantu Core
Lexicon: Cushitic

MEDNYJ ALEUT
*Gram*: Aleut (N, nonfinite V), Russian (finite V)
Lex: 90% Aleut

LIGHT WARLPIRI
*Gram*: Warlpiri (N)
Ab. English (V)
*Lex*: Warlpiri (Ns and Vs), Ab. Eng (N only)

*Figure 1*      *A continuum of grammatical mixing in mixed languages*

Most of these mixed languages can be categorised as either G-L mixed languages or V-N mixed languages (see Bakker 2003 for an overview of ML types). Interestingly these types cannot be not predicted from the sociohistorical background of the languages and their speakers. For example, although Michif and Gurindji Kriol are both V-N mixed languages, Michif is the result of mixed marriages and Gurindji Kriol is the result of language shift by a single cultural group. Similarly, although Angloromani and Media Lengua exhibit a grammar-lexicon split, the grammar of Media Lengua is provided by the ancestral language whereas the grammar of Angloromani comes from the new language. Thus, although the structures of some mixed languages look similar from a synchronic typological perspective, they originated from very different social circumstances. The following sections discuss the socio-historical origins of these mixed languages and the linguistic practices of the speaker communities at the time of genesis. The difficulty of matching social and structural typologies becomes clear in these sections.

## 3.  The origins and features of mixed languages

A number of general definitions have been used to identify mixed languages. The earliest definition was that of *genetic ambiguity*. Thomason and Kaufman (1988) argue that mixed languages can be identified on the basis of their non-classification in historical terms. That is, mixed languages are deemed to have no clear genetic heritage and cannot be classified according to standard historical methods. This definition is still used by Thomason (2001: 198); however this criterion does not distinguish mixed languages from creole languages which are also often claimed to have an unclear genetic heritage (though

see Mufwene (2001) and DeGraff (2005) for arguments against this claim). Another way of framing the notion of genetic ambiguity is to say that mixed languages have two clear parents. This criterion relates to another general characterisation of mixed languages - they are said to be the product of contact between only *two* languages (McWhorter 2005: 253; Bakker 1994: 27). Indeed it seems to be the case that most identified mixed languages have two clear and often equal parents, in contrast, for example, with pidgin and creole languages which usually have one lexifier parent and a number of source languages which contribute to varying extents to the grammar.

Apart from this more general definition, two main approaches have been taken to describing the origins and features of mixed languages - socio-historical approaches (§3.1) and structural approaches (§3.2). This divide is typical of the kind found for other contact languages such as creole languages. Examples of approaches which combine socio-historical and typological features can be found, see for example Matras and Bakker (2003) who posit six types of mixed languages based on the mixed language speakers' knowledge of the source languages, the level of functionality of the mixed language, the typology of its structure, and various social factors.

In general the next sections will demonstrate that the category of contact language contains a diverse range of structural mixes which derive from varied socio-historical backgrounds.

## 3.1 Socio-historical origins and features of mixed languages

Most generally mixed languages contrast with creole languages in that their genesis was a product of *expressive* rather than communicative needs (Golovko 2003: 191; Muysken 1997: 375). This social definition argues that pidgin and creole languages are born out of the need for communication between people of a number of language groups, whereas mixed languages are created in situations where a common language already exists and communication is not at issue. Thus the mixed language serves as an expression of an altered identity, be it new, or differing significantly from an older identity. For example Light Warlpiri is spoken by Warlpiri people who had no need for a new language to communicate with. Warlpiri was already fulfilling this function, and is indeed still spoken. Thus Light Warlpiri marks an identity shift for younger Warlpiri people who simultaneously express their continuing Warlpiri heritage coupled with a more modern Aboriginal identity in the mixed language (O'Shannessy 2008).

More specifically, discussions about the sociolinguistic and historical features of mixed languages have fallen into three categories: the *direction of language shift* which contributed to the formation of a mixed language (§3.1.1), whether speakers of mixed languages constitute a *separate ethnic group* (§3.1.2) and whether the mixed language is used as *the native language of the group* (§3.1.3).

### 3.1.1    The direction of language shift

A number of theories of mixed language genesis are based on the direction of language shift between the ancestry language and the introduced language. Many of these theories are set within borrowing or code-switching theories (see §3.2.1). Most of this work focuses on the grammatical interaction of the source languages. Other work takes a

broader approach looking at the socio-linguistic conditions which influence the direction of the language shift (Croft 2003; Matras 2000; Croft 2000). Croft (2000: 214-221; 2003: 52-60) proposes a social typology of mixed languages which is based on the change in dominance between languages during the process of mixed language genesis. His approach relies on the relationship between the ancestry language and the introduced language, and the direction of shift between the two languages. He suggests this process may take one of three forms: *death by borrowing*, *semi-shift* and *mixed marriages* (identification with a new society). Mixed languages which are the result of mixed marriages represent a fusion of two languages. Semi-shift occurs when speakers of an ancestry language move part-way towards the introduced language but do not complete the shift. Finally death by borrowing involves languages which borrow to such an extent that they replace much of their basic vocabulary, and in more intense cases, grammatical elements. Croft's categories represent two directions of shift - firstly a shift from the ancestral language to the introduced languages which occurs in degrees (lexicon then grammar) and secondly the convergence of two languages. Another possibility is the deliberate 'undoing' of a shift towards an introduced language where a group attempts to reclaim its ancestral language. Boretzky & Igla (1994) call such a shift a 'U-Turn'. These three directions of shift are summarised in Figure 2.

**Mixed marriage**

ANCESTRY LANGUAGE        ➡ ⬅        INTRODUCED LANGUAGE

Fusion of structural and
lexical features

**Shift by degree**

ANCESTRY LANGUAGE        ———➡        INTRODUCED LANGUAGE

Borrowing of lexical and
grammatical material

**Reversal of shift**

ANCESTRY LANGUAGE        ⬅———        INTRODUCED LANGUAGE

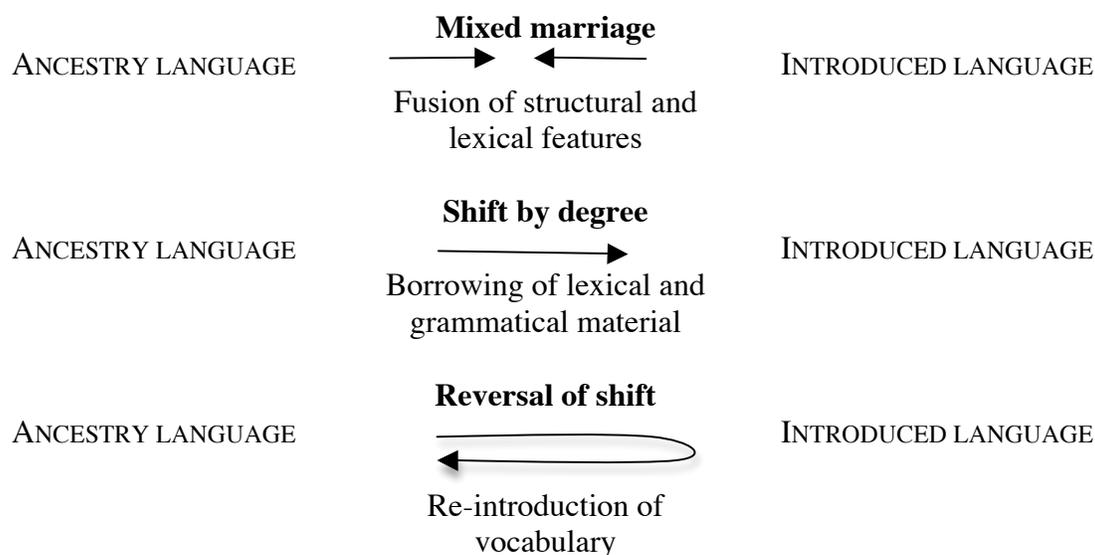Re-introduction of
vocabulary

*Figure 2*        *Direction of shift in mixed language genesis*

The first category of language shift involves *mixed marriages* between men from one society and women from another (Bakker 1997). The children of these mixed marriages are said to form their own distinct cultural identity, with the mixed language as an enactment of this identity. The relative dominance of the languages in this situation is less clear, and it is likely that they converge and fuse (see also §3.2.2). In this respect they do not represent a clear shift in either direction. Michif is the classic example of this type of mixed language genesis. As was shown in §2.7, this Canadian mixed language and its speakers are the product of marriage between French-Canadian fur traders and

Amerindian women. Mednyj Aleut also seems to fit into this category. It emerged from mixed marriages between Aleut women and Russian seal fur traders in the early 1800s.

The second category consists of mixed languages which undergo a *shift by degree* from the ancestral language to the introduced language. A change in the dominance of languages occurs when speakers shift towards the introduced language. This process does not go to completion, and what remains is the mixed language. Thus the mixed language can having varying amounts of material from the introduced language from borrowed nouns right up to morphology. This shift stops part-way because the speakers may not have full access to the introduced language, or because the remaining part of the ancestral language may be a marker of social identity. Media Lengua as an example of a mixed language which resulted from a partial shift. The relexification of Quechua with Spanish was a consequence of Quechuan men becoming more fluent in Spanish. Similarly Bilingual Navajo and Old Helsinki Slang fit well in this category. Gurindji Kriol and Light Warlpiri are examples of languages which have progressed further along the scale, incorporating large amounts of structural material from Kriol/Aboriginal English as they shifted away from their ancestral languages, Gurindji and Warlpiri respectively.

Croft (2003) considers Ma'á and Angloromani cases of languages at the extreme end of this scale of shift. He suggests that they are the mirror-opposite of languages such as Media Lengua, resulting from the extreme absorption of another language's grammatical structure into the ancestral language. In the case of Ma'á, the relevant grammatical system is Bantu (§2.2), and English for Angloromani (§2.1). Croft (2003: 53) suggests that these languages were created within societies under great social pressure. Whilst they attempted to avoid cultural assimilation, they gradually adopted more and more of the introduced language as it became more dominant, leaving only vocabulary from the ancestral language. This view of the formation of Ma'á and Para-Romani varieties follows Thomason's (2001) proposal, however, it is not without controversy. As was discussed in §2.2, Mous (2003a) suggests that Ma'á represents an attempt to undo the shift to Mbugu by maintaining basic Cushitic vocabulary. In this respect, Mous' view of Ma'á fits better with the third category, the reversal of a language shift.

The third category of language shift includes mixed languages which formed as the result of a *reversal of shift*. In this category are mixed languages which are the product of a group shifting almost completely to an introduced language and attempting to reverse this shift by re-introducing material from the ancestral language. In this scenario, the ancestral language is still available, perhaps still spoken by older generations. Generally only lexical material is found from the ancestral language in the mix. Mixed languages such as Angloromani or Ma'á which utilise lexical reservoirs from the ancestral language could be regarded as examples of such a linguistic U-turn.

### 3.1.2   Mixed language speakers as a separate ethnic group

Another sociolinguistic feature, which has been discussed in relation to mixed languages, is whether speakers constitute a separate ethnic group or a continuation of an older identity (Bakker and Mous 1994: 2). Creoles are considered to be markers of a separate or more encompassing group of people. For example, though speakers of the English-based Australian creole language, Kriol, still identify with the smaller substrate languages

such as Marra or Rembarrnga, a larger identity of Aboriginality is marked through the use of Kriol. This identity contrasts again with the group of non-Indigenous Australians which is associated with English. Unlike creole languages, a single coherent story of ethnic identity associated with the genesis and use of mixed languages does not emerge from the data available. Two main processes seem to be apparent. As Thomason (2003: 25) suggests, the new language is either associated with a new ethnic identity, or it is seen as a means of continuing an ancestral group membership.

First, some mixed languages are spoken by new ethnic groups which are generally derived from mixed marriages. For example, Michif speakers call themselves Métis which reflects the mixed identity of their group (Matras and Bakker 2003: 2; Bakker 1994: 14). At the time Mednyj Aleut was spoken, the speakers also formed a separate group, though they considered themselves Aleut (Golovko 1994: 117). The Sri Lankan Malays also constitute a separate ethnic group. Although this language combines a contact variety of Malay with Tamil and Sinhala, the speakers do not identify directly with any of these groups. This is probably because they come from a variety of Austronesian speaking backgrounds and were originally immigrants to Sri Lanka and remain a minority group. However whether this mixed language is the result of mixed marriages is controversial as discussed in §2.9. One mixed language which is marks a new identity without the background of mixed marriage is Old Helsinki Slang. In this case, the new identity was formed between mixed Swedish and Finnish-speaking gangs.

Mixed languages are also spoken by people who *do not* constitute a separate ethnic identity. Speakers of Media Lengua are not separate from Quechua people though they are a sub-group who identify to a certain extent with urban Hispanic society. There are also a number of in-group mixed languages labelled secret languages. Angloromani is spoken by people who simultaneously mark their Romani heritage and maintain their separation from mainstream English society. There are suggestions in earlier literature that Romanies would relexify English more densely with Romani vocabulary to prevent non-Romani people from understanding their conversations (Boretzky and Igla 1994). Speakers of Bilingual Navajo speakers are also not deemed separate from Navajo people in general. This mixed language represents an attempt to maintain Navajo under the continuing colonial pressure of English. Similarly speakers of Gurindji Kriol also do not belong to a separate ethnic group. This mixed language is spoken by Gurindji people whose parents and grandparents were predominantly speakers of Gurindji until at least the late 1940s. Gurindji Kriol speakers continue to identify as Gurindji and also call their mixed language, Gurindji, despite the Kriol content (Meakins to appear-a). The strong maintenance of Gurindji lexicon and noun phrase structure in this mixed language marks the Gurindji as separate from other Kriol speakers, and indeed the encroaching and assimilative non-Indigenous world (Meakins 2008b).

### 3.1.3   Mixed languages used as native languages

Related to the association of mixed languages with a new or continuing identity is their use as native languages by speaker communities. There are a couple of ways that 'native-ness' may be defined. A mixed language may be independent from its input languages, that is, speakers have no knowledge of the input languages. Alternatively a mixed language may be the main language of use within a community of speakers, where the

source languages are still spoken. In this respect it may also be the first language of the community (see also §4.1.1).

First, independence can be defined by speaker use. In these cases speakers do not use or understand the input languages and therefore the mixed language is spoken in isolation from these languages. This criterion is used to distinguish creole languages from pidgins. Whilst pidgins are usually associated with a particular domain such as trade and therefore are not a first language for speakers, creole languages are usually the first language of a community and tend to exist in a greater degree of isolation from their source languages, though this is not the case for many of the Pacific creoles. In fact Michif is the only mixed language which is spoken today largely independent of its source languages. Most Métis no longer speak either of the contributing languages, French or Cree. English is spoken and has now become the principal language in the communities (Matras and Bakker 2003: 3), with only older people using Michif, Amerindian languages or French.

The majority of mixed languages are the native language of a community, however they are spoken alongside one or more of their source languages. Smith (2000) calls these languages *symbiotic* mixed languages. For example Mednyj Aleut was spoken concurrently with a number of Aleut dialects and Russian, though it is not clear whether Mednyj Aleut speakers had control of one or more of its input languages (Golovko 1994: 114). According to Muysken (1994: 210), Media Lengua is learnt either as a first or second language. Middle-aged speakers of this mixed language also may have access to both input languages. Younger speakers tend to speak Spanish better and older speakers, Quechua. All Bilingual Navajo speakers also have either Navajo or English or both in their linguistic repertoire. In the case of Angloromani, English which provides the grammatical base, is the main language for speakers of this mixed language. Similarly Gurindji Kriol and Light Warlpiri are the first languages of their speaker communities, though Gurindji, Warlpiri and Kriol are still used to varying extents, as will be discussed in more detail in §4.1.1.

## 3.2 Structural origins and features of mixed languages

As §3.1 shows, mixed languages are a product of the social circumstances of their speakers. There is a general consensus that severe social upheaval is a major factor which contributes to the formation of mixed languages, whether the speaker community is marking a new identity or maintaining an old one. What distinguishes mixed languages from other contact varieties is that they emerge as expressions of identity rather than as a result of a communicative need. There is much less agreement on the kinds of grammatical processes which are required for a relatively stable language to emerge. Generally theories of genesis are divided over whether mixed languages are the extraordinary results of ordinary contact processes (Thomason 1995), or whether special processes are required. Ordinary contact processes include borrowing (§3.2.1.1) or code-switching (§3.2.1.2) whereas Bakker's (1997) language intertwining theory (§3.2.2.1) and Muysken's relexification theory (§3.2.1.3) are an examples of special processes.

Theories of mixed language formation take on two main forms - (i) *unidirectional* approaches (§3.2.1), for example borrowing (Matras 2003; Thomason 2001), code-switching (Auer 1999; Myers-Scotton 2002, 2003), relexification (Muysken 1981),

paralexification (Mous 2003a), or Matras et al's (2008) language repertoire approach; and (ii) *fusion* approaches (§3.2.2), such as Bakker's language intertwining theory, language competition approaches (Ansaldo 2008; Meakins 2007) (§3.2.2.2) and McConvell's (2002, 2008) centre of gravity hypothesis (§3.2.2.3). In general, these diachronic accounts of mixed language genesis are based on synchronic descriptions of the languages (though see Matras 2010 for an account the genesis of Angloromani based on an evaluation of historical sources). Further, the diversity of genesis theories is largely the result of extending an explanation for the formation of one mixed language to other varieties. In this respect, no one theory provides a unified account of mixed language genesis which accounts for all languages which have been identified as mixed.

Finally, with the exception of Myers-Scotton and the language competition and relexification approaches, one shortcoming of most of these theories is their reliance on *form* (i.e. the phonological shape of a lexeme or morpheme) as the basis of mixing. Often mixing is more intricate involving different layers of the lexeme or morpheme, such as combining the form from one source language with its function from another language (see §5).

### 3.2.1   Unidirectional approaches to mixed language genesis

Unidirectional approaches to mixed language genesis assume a one-way shift from a source language to a target language. They consist of borrowing or switching scenarios where lexical and morphological material is replicated from the source language into the target language. Mixed languages are considered to be a half-way house for this language shift, no longer classifiable as a variety of one language or the other. Note that the separation of borrowing and code-switching approaches (particularly in relation to insertional code-switching) in the literature is somewhat artificial and largely the product of a separation between historical linguistics, which is focussed on language-level change, and contact linguistics, which tends to have the bilingual individual in mind.

### 3.2.1.1 Borrowing approaches

Thomason and Kaufman (1988; 2001) base a theory of mixed language genesis on their model of borrowing. This model combines language features with social factors, proposing that features can be borrowed regardless of the typological distance between the affected languages (Thomason and Kaufman 1988: 53). Social factors play a fundamental and determinate role in the linguistic outcome of language contact and, given the right level of social disruption, substantial structural borrowing including that of inflectional morphology is possible (Thomason and Kaufman 1988: 37). Under Thomason and Kaufman's (1988: 47) model, two social features are necessary for extensive borrowing - time and a level of bilingualism. Extensive and prolonged community bilingualism is considered a necessary condition for borrowing structural elements of a language, such as inflectional morphology. The end result is a borrowing scale which is reminiscent of Whitney (1881), Haugen (1950), Moravscik's (1978) and Appel and Muysken's (1987) earlier scales. See also Matras and Sakel (2007), Matras (2009) and Aikhenvald and Dixon (2007).

| Degree of Contact | Borrowing Type | Features Borrowed |
|---|---|---|
| 1. Casual contact | lexical | non-basic vocabulary before basic |
| 2. Slightly more intense contact | lexical | functional vocabulary e.g. conjunctions and adverbs |
|  | syntactic | only new functions borrowed |
| 3. More intense contact | lexical | pre/postpositions, derivational affixes, inflect. affixes (attached to stem), pronouns, low numerals |
|  | syntactic | change in word order, borrowing postpos. in a prepos. language |
| 4. Strong cultural pressure | syntactic | extensive word order change, inflectional affixes (e.g. case) |
| 5. Very strong cultural pressure | syntactic | typological disruption, changes in word structure (e.g. adding prefixes in suffixing language), change from flexional to agglutinative morphology |

*Figure 3        Thomason and Kaufman's borrowing scale*
*(based on Thomason and Kaufman 1988, p. 74-75; Thomason 2001)*

Language shift may stabilise somewhere along this scale of borrowing. For example mixed language formation may halt at the least disruptive end of the scale and exhibit only lexical borrowings. Bakker (2003: 109) identifies this pattern as a class of mixed language: *L-G languages*. They comprise by far the largest category, and are characterised by a clear division between the lexicon and the grammar where these systems are dominated by a different source language. Examples include Angloromani (English grammar, Romani words), Ma'á (Bantu grammar, Cushitic core vocabulary), Bilingual Navajo (Navajo grammar, English vocabulary), Media Lengua (Quechua grammar, Spanish lexicon) and Old Helsinki Slang (Finnish grammar, Swedish lexicon). Note that this category includes mixed languages where the language contributing the grammar is the ancestral language, for example Media Lengua and Bilingual Navajo, and languages where the introduced language provides the grammar, for example Angloromani and Ma'á.

At the other end of the scale are fossilised mixes which contain inflectional morphology from both languages. At this stage other borrowings such as lexical and more minor structural borrowings are assumed. A number of mixed languages exhibit this grammar mixture, including Michif, Mednyj Aleut, Gurindji Kriol and Light Warlpiri. For example, inflectional morphology from both French and Cree is present in Michif. Verbal inflections are derived from Cree and in the NP, Michif preserves both French plural marking and adjectival agreement and some case-marking from Cree. For instance, the Cree obviative marker and locative suffix mark French nouns in Michif, albeit in a somewhat reduced distribution (Bakker 1997: 89). Another language, which has also retained inflectional morphology from both source languages is Mednyj Aleut. The structure of Mednyj Aleut includes both Aleut nominal inflections such as two case distinctions (absolutive and relative) and Russian finite verbal inflectional morphology, including portmanteau morphemes which express tense, number and person (Thomason 1997: 457-59). Finally, Gurindji Kriol and Light Warlpiri combine Kriol/Aboriginal

English, the language of the verbal inflectional categories (tense and mood markers), with Gurindji and Warlpiri, respectively. Gurindji and Kriol contribute nominal inflections in the form of case marking, both syntactic (ergative, dative, possessive) and semantic (locative, allative, ablative) (Meakins 2009, to appear-b; Meakins and O'Shannessy 2005, 2010; O'Shannessy to appear).

The situation described for these four languages is exceptional given the fragility of inflectional morphology in other language contact situations. For example, inflectional morphology is rarely borrowed (Thomason and Kaufman 1988; Heath 1978; Weinreich 1974 [1953]; Aikhenvald and Dixon 2007; Gardani 2008; Matras 2007; Meakins to appear-b), mostly derived from the more dominant language in code-switching (Myers-Scotton 2002; Muysken 2000), is rarely found in pidgin and creole languages (Plag 2003, 2003; McWhorter 1998), and is one of the first systems to be reduced or lost in situations of language obsolescence (Maher 1991; Sasse 1992). In this respect, the disappearance or maintenance of inflectional morphology is indicative of the relative strength of the interacting languages. The loss of inflectional morphology is one of the first signs of language death, thereby demonstrating the weakening of the morpho-syntactic frame of the language. In the cases of borrowing and code-switching one language is more dominant, as defined by the presence of inflectional morphology. On the other hand, the maintenance of inflectional morphology from both languages in mixed languages would suggest a relatively equal weighting given to both languages, with neither language definitively stronger. Inflectional morphology in this level of language contact neither disappears nor is selected for by only one language. Therefore no one language can be identified as the grammar language, but rather the morpho-syntactic frame represents a composite of the languages.

Indeed Matras (2003: 158) suggests that a particular feature of mixed languages is the seemingly unconstrained borrowing of grammatical elements, which in the past have been labelled as 'loan proof'. Included in the list of loan proof items is inflectional morphology. Matras (2003: 171) goes on to suggest that this violation of borrowing tendencies is what characterises a mixed language. He includes in his assessment of the presence of these rarely attested borrowings, not only inflectional morphology such as case affixes, but also in/definite articles, bound pronouns and TAM markers, possessive markers, sentential negation, personal pronouns, demonstratives, existentials (copula), place deictics, the basic interrogatives *what* and *who*, numerals under 5, and adpositions which express basic local relations (*in, at, out of*) (2003: 158-159).

**3.2.1.2 Code-switching approaches**

Related to borrowing theories is the idea that mixed languages are derived from code-switching. Auer (1998, 1999) and Myers-Scotton (1993, 2003) suggest that mixed languages are the result of the gradual fossilisation of code-switching. The empirical evidence for code-switching origins is discussed in §4.4. A distinction which is relevant is Muysken's (2000) typological differentiation of *insertional* and *alternational* code-switching. Generally speaking, these two types of code-switching can be distinguished by the level of involvement of the grammars of the interacting languages. Alternational code-switching involves the alternation of structures from different languages. On the other hand, the grammar of one language is more dominant in insertional code-switching,

with elements from another language inserting into the more dominant language's structure (Muysken 2000: 3).

The first account of the transition from code-switching to a mixed language comes from Auer (1999: 309-10; 1998). He presents a model of the fossilisation of code-switching (Stage 1) into a mixed language (Stage 3) via language mixing (Stage 2). A 'cline' from pragmatics to grammar can be observed between these three stages of mixing, where code-switching loses its pragmatic function over time and the shape of the mixing is increasingly determined by grammatical constraints (1998: 16). These three forms of bilingual speech are differentiated by the type of mixing, and their degree of variation and social markedness. First *code-switching* is the most variable and socially-marked form. By "socially marked", Auer (1999: 310) is referring to the social weight carried by each language and the associated social meaning of switching between languages. Typologically, code-switching at this stage can be alternational and/or insertional (1999: 313-314). Stage 2 *language mixing* also exhibits patterns of code-switching, but the social meaning associated with the switches is lost. Instead syntactic factors in the form of structural constraints play a role in the language switches. Both alternational and insertional code-switching may be present in the language mixing stage. However Auer (1999: 315) believes that these patterns converge making it difficult to distinguish them. Finally *mixed languages* differ from language mixing in a number of ways. Auer suggests that they lose any hint of alternational code-switching, looking entirely like insertional structures. They contain much less syntactic variation than language mixing; functionally equivalent structures from both languages may develop more specialised uses in the mixed language; and mixed language speakers do not need to be speakers of either of the contributing languages (1999: 321).

Myers-Scotton (2002, 2003) provides the second major approach which theorises the move from *insertional* code-switching to a mixed language. She frames this transition within her Matrix Language Frame (MLF) model. She labels this transition the Matrix Language Turnover Hypothesis. The relevant concepts in this hypothesis are the Matrix Language (ML) which is the language which provides the grammatical frame for switching or mixing, and the Embedded Language (EL) which is the weaker language which contributes mostly lexical material to the mix. The Matrix Language Turnover Hypothesis is concerned with the change in dominance of the participating languages. Myers-Scotton proposes that mixed languages arise when there is a turnover under way which does not go to completion. That is, the source languages do not entirely change in dominance but stabilise someway through this process resulting in a language which is a combination of the source languages.

Like Auer, Myers Scotton also proposes three steps from insertional code-switching to composite code-switching (convergence) to the mixed language itself. The first step in the Matrix Language Turnover is *classic* code-switching which is basically insertional code-switching, with alternational code-switching incorporated in the form of EL Islands. The more dominant language takes the role of the Matrix Language, with the less dominant language inserting or embedding morphemes within this grammatical frame (2002: 110). More lengthy switches to the less dominant language create EL Islands.  The second stage, *composite* code-switching, occurs when the participating languages begin to converge, such that one of the participating languages loses its undisputed role as the Matrix Language. In this respect, the weaker or embedded

language gains strength. The Matrix Language splits and recombines to form a composite structure consisting of abstract material from both languages. The convergence of the EL and the ML represents a change in the morpho-syntactic frame. This convergence precedes the third mixed language stage. Myers-Scotton suggests that mixed languages "represent turnovers that do not go to completion, but 'stop along the way' " (2002: 249). Mixed languages may stop at different places along the scale, which explains why they surface in different forms and with the split in different places.

Myers-Scotton calls this process the Matrix Language Turnover hypothesis. The outcome is a language which contains structural material such as inflectional morphology from the weaker language, that is the language which was the embedded language in the code-switching (2003: 92). This analysis is similar to Thomason and Kaufman and Matras' view of mixed language genesis which is based on borrowing models. Indeed, as §3.2.1.1 shows, mixed languages with structural material from the weaker language do exist. Michif, Mednyj Aleut, Gurindji Kriol and Light Warlpiri all exhibit inflectional morphology from both of their source languages. Like the borrowing approaches, the gradual conventionalisation of code-switching is unidirectional where material from one language is subsumed into another language. Although some problems exist for both Auer and Myers-Scotton's models, such as the need to lose alternational structures (see §4.4.2), empirical evidence for code-switching preceding mixed language genesis exists (§4.4.1), though it must be noted that empirical evidence for each of the three stages that Auer and Myers-Scotton propose is not available.

**3.2.1.3 Relexification**

The process of relexification is more familiar as an account of the origin of creole languages (see e.g. Lefebvre 1998, 2004), nonetheless relexification has also been offered as an explanation for the formation of one mixed language, Media Lengua (Muysken 1981, 1994, 1997). Relexification differs from borrowing in that it is only the phonological form which is borrowed from another language rather than the whole structure of the lexical and morphological item. This form is then mapped onto the recipient language's own structure. For example, while the form of a verb may be borrowed, it is mapped onto the predicate argument structure of the recipient language.

In the case of Media Lengua, Muysken (1981) argues Spanish forms have been borrowed into Quechua but in the process have adopted the structure and meaning of the equivalent Quechuan forms. For example Muysken (1981: 57) observes that Spanish pronominal forms have been mapped onto Quechuan pronominal paradigm. In the case of third person singular pronouns, Quechua does not distinguish masculine/feminine in the third person whereas Spanish does. Media Lengua also does not make this distinction but uses the form *el* as a general third person pronoun which is a phonological compromise between the Spanish *él* 'he' and *ella* 'she'.

**3.2.1.4 Paralexification and language repertoires**

The final set of unidirectional theories are Mous' (2003b) *paralexification* and Matras et al's (2008) notion of *language repertoires*. These two theories give explanations for the development of Ma'á and Angloromani, respectively. They provide an understanding of

mixed languages which have tried to undo a complete shift to a regionally dominant language. These theories also bear some resemblance to the borrowing and code-switching models of mixed language genesis. However, as will be shown, the major difference is in choice - the presence of alternative means, usually lexical, of expressing a notion. Both of these models challenge the notion of a mixed language as a closed system, rather posturing a fluidity of language choice reliant on social context.

First, paralexification involves the existence of two lexicons in parallel. One language involved in the mix exists in its entirety with its grammar and lexicon intact. The other language merely provides a substantial reservoir of lexical material for inserting into the more dominant language's grammatical frame. Lexical insertion may be a strategy for maintaining a waning ancestry language. For example, as discussed in §2.2, Mous (2003a: 89) believes that Ma'á is a deliberate attempt to undo a shift to Pare (Bantu), which occurred through a paralexification process. In this situation the Bantu lexicon, and a Cushitic and Maasai lexicon exist in parallel. People speak both Mbugu, which draws from the Bantu lexicon, and Ma'á, which replaces many Bantu words with Cushitic and Maasai equivalents.

Matras et al (2008) characterise mixed languages such as Ma'á in a similar way. They consider these mixed varieties as more fluid and less contained entities which do not have clearly definable boundaries. Matras et al suggest that secret languages such as Angloromani which consist largely of lexical insertions with little grammatical alterations are not separate languages but involve lexical reservoirs which may be drawn on to express particular social functions. The broadest social function this kind of mixed language might have is marking a sense of solidarity to express speech acts such as warnings. In this respect, Matras paints a speaker of mixed language (and indeed any bilingual speaker) as someone who uses their language resources consciously and creatively.

Indeed secret languages are not the main language of communication for the community. Rather, they are a marked form of speech used in restricted functional domains. Other mixed languages exhibit less optionality. These tend to be the mixed languages which are the unmarked language of communication for a speech community and are used in a broad range of functional domains. For example, Gurindji Kriol is the language of general interaction for younger Gurindji people. Switches to lesser-used Gurindji or Kriol elements are marked and perform particular social functions. For example some variation is present in the lexicon. 37% of the lexicon contains semantically equivalent forms from Gurindji and Kriol where language choice expresses social meaning (Meakins and O'Shannessy 2005). This is similar to Ma'á's parallel lexicon or Angloromani's lexical reservoir. In the case of Gurindji Kriol, the choice often depends on whether the speaker is talking to an older or younger person. In general, however, most of the structure and lexicon of Gurindji Kriol and other mixed languages such as Michif[15] and Light Warlpiri are quite invariable and cannot be said to find their origins in the social markedness of linguistic choice.

### 3.2.2   Fusion approaches to mixed language genesis

Fusion approaches differ from unidirectional approaches in the direction of shift. Unidirectional approaches are one-way, describing language shift which progresses from

the ancestral language to an introduced language or from the introduced language back to the ancestral language. It both cases the shift stops along the way to form the mixed language. On the other hand, fusion approaches consider mixed languages to be the result of two languages merging. Bakker's (1997) *language intertwining* approach was the first of these theories. More recently *language competition* models based on Croft (2000) and Mufwene's (2001) language evolution theories have been proposed (Ansaldo 2008; Meakins 2007).

## 3.2.2.1 Language intertwining

Bakker's (1997) theory of language intertwining was one of the first explanations for the formation of mixed languages. It proposes a process specific to mixed language genesis rather than relying on other language contact processes such as borrowing or code-switching. Bakker (1997: 210) takes *combination* rather than *replacement* to be the central process. In this respect no language is shifting to another, rather a new language is created in the intertwining of structures and vocabulary. Specifically, intertwining involves mixing aspects of the two languages based on two morphological distinctions: lexical/grammatical, free/bound. Generally the free lexical morphemes come from one language and the bound grammatical morphemes from the other. Free grammatical morphemes can come from either language. The grammatical system comes from the language speakers are most familiar with, or the language that the speakers wish the mixed language to 'sound' like (1997: 211). The last specification is needed to account for mixed languages such as Angloromani. Bakker (1997: 206) suggests that in fact the English grammatical system was adopted so that Angloromani would sound more like English, and English speakers would not notice when Romanies were speaking their mixed language around them.

Bakker's theory, along with adjustments for Michif, provides an explanation for how L(exicon)-G(rammar) mixed languages, which constitute most mixed languages, are created. Language competition approaches look beyond this mix, as is shown in the next section. What is also missing in Bakker's theory is an idea of how the speakers used the two source languages at the point of creation of the mixed language, that is their linguistic practices. Code-switching is one way that both languages are used with more or less emphasis. The role of code-switching in mixed language genesis is discussed in §4.4.

## 3.2.2.2 Language competition and evolution

A more recent fusional approach to mixed language genesis utilises concepts from evolutionary theory to provide an explanation for the formation of these languages. This approach sits within general theories of language change which use concepts of variation and competition to account for both internally and externally-motivated language change (Mufwene 2001; Croft 2000; Labov 1994). Explanations for the formation of mixed languages can be positioned within these approaches. What is relevant here is not the direction of language shift but the competition between features in the source languages. Note that the importance of variation for the origin of mixed languages will be discussed further in §4.2.

Croft's model of language change is based on the evolutionary view that repeated copies of objects change their properties over time (imperfect replication). Croft (2003: 41) utilises this view to suggest that languages evolve through the replication and selection of particular linguistic structures in normal language use. In biological evolution the basic unit of replication is the gene. Replication involves copying genes in a process of reproduction. Replicated genes can be identical (normal replication) or altered (altered replication) which may occur via mutation or recombination from parents in sexual reproduction (Croft 2000: 23). Croft suggests that this process can also be applied to language, where features of language structure, called *linguemes* are taken as the basic unit of replication.

> Replication also occurs in the production of utterances. Every time I produce an utterance, I replicate linguistic structures from prior utterances that I have been exposed to and have internalised in some way. These linguistic structures, name *linguemes* of (at least) three types familiar to historical linguists: sounds, words/morphemes and constructions. In any utterance, I replicate a large number of linguemes, and these linguemes may be replicated from a multiplicity of prior utterances. (Croft 2003: 42)

Replication may follow similar patterns to biological evolution in that the replicated structure may be more or less identical to that of the original structure (normal replication). In this situation a speaker is conforming to the linguistic conventions of the speech community. On the other hand, the structure may change in replication (altered replication), which is where linguistic innovation finds its origins. Linguistic innovation can only occur where more than one lingueme or variant is available to mark a particular function. This variation is an essential part of language evolution. Mufwene (2001) calls this a feature pool, that is the set of variables available to speakers from within a language system or as a result of external pressures. It is here that the process of competition, replication and change takes place. External pressures occur when more than one language is a part of the pool. Language mixing results when links are made between linguemes or variants between two languages. In these situations speakers identify forms in one language as performing the same function as a form in another language. Once this identification is made, "the way is open to interference", that is the replication of a lingueme from one language into another (Croft 2003, p. 46). Forms may also be transferred when there are no equivalents in the recipient language, for example when the speakers want to express a new concept which is not encoded in their language. The linguemes or pieces of linguistic structure are either replicated in a way that does not differ from their source, or they undergo altered replication to form unique structures. In general, a high level of bilingualism and fluidity between the languages is a precondition of the language ecology for language competition to occur (Mufwene 2001). Pervasive code-switching as the unmarked linguistic practice is the most obvious way to produce such an intense level of language contact (Meakins 2007: 321).

This model of language change can be applied to mixed languages. Rather than providing an explanation for the overall shape of the mixed language (e.g. grammar-lexicon mix or N-V mix), this model can be used to demonstrate how subsets of the language's structure eventuated. For example, Meakins (2007: 265) describes the

competition between Kriol-derived locative preposition and the Gurindji-derived allative case marker in Gurindji Kriol. These forms are used to mark inanimate goals in their respective languages. In the process of mixed language genesis they came into contact and competition. The end result was the use of the Gurindji-derived *locative* case suffix to mark goals where it is not used in this function in Gurindji. This functional shift occurred as a result of the form from Gurindji being mapped onto a Kriol pattern (locations and goals are not distinguished in Kriol). Competition between linguemes have had other results in the formation of mixed languages. For example, in Sri Lanka Malay, Malay-derived TAM markers have shifted their alignment from a Malay mood and aspect system to a tense system based on Tamil categories (Bakker 2000a, 2000b; Smith and Paauw 2006). The TAM markers have retained the form of Vehicular Malay whilst mapping onto the Tamil tense system. For example the progressive marker has become a present tense marker (Smith and Paauw 2006: 161). More examples of this type of change as the result of contact between functionally equivalent elements is discussed in §4.1.2 and §5. In these examples, descriptions of mixing go beyond the level of form to looking at other features of the morpheme or lexeme including its semantics, function and distribution.

### 3.2.2.3 Centre of gravity hypothesis

Another final fusion theory is McConvell's (2002, 2008) *centre of gravity* hypothesis which relies on the distinction between head and dependent marking languages (c.f. Nichols 1986). This hypothesis differs from most other theories of mixed language genesis because McConvell is mostly concerned with the more intricate mixes which involve NP-VP splits rather than grammar-lexicon mixes. McConvell suggests that the typology of the ancestry language may have something to do with the nature of the split, and the way in which this ancestry language shifts to an introduced language via an intermediary period of code-switching and perhaps more structural language mixing. McConvell observes that head-marking and dependent-marking languages have different centres of gravity. The typology of the ancestry language affects the path of the 'turnover' to the new language by producing different results in the intermediate language stages. He suggests that head-marking languages have a tendency to retain the verbal grammar after the nominal grammar has 'turned over' to the introduced language. For example, Michif has retained the Cree verbal grammar whilst adopting French nominal grammar. Conversely dependent-marking languages keep the nominal structure of the ancestry language whilst turning over to the VP structure of the new language (McConvell 2002: 345). McConvell suggests this is the case with Gurindji Kriol. The turnover to Kriol has frozen midway maintaining the nominal grammar of Gurindji.

Although McConvell's model is heavily based on typology, it focuses on the typology of the ancestry language. McConvell does not discuss the typology of the introduced language which may be relevant the resultant structural mix. By comparison the language competition approach presented in §3.2.2.2 takes into consideration the typology of both languages. It is also incorrect to assume that Cree was the ancestry language in the case of Michif, because Michif is the result of mixed marriages between Cree-speaking women and French-speaking men (Bakker 1997: 208). In this respect neither language can be assumed to be the ancestry or introduced language. McConvell's

analysis works better for Gurindji Kriol, where the original language of the population was clearly Gurindji. Thus it is more reasonable to suggest that Gurindji was the ancestry language and speakers stopped midway in a shift towards Kriol. However the typological analysis of the ancestry language, Gurindji, is problematic given that it is neither strongly head- nor dependent-marking. Gurindji contains both case-marked nouns and cross-referencing bound pronouns which, though not marked on directly on the verb, are a part of the verb complex.

## 4. Current issues for the study of mixed languages

### 4.1 Are mixed languages really autonomous language systems?

One of the criticisms which is often levelled at descriptions of mixed languages is the autonomy of the language variety presented (Bakker 2003; Meakins to appear-a). The term 'autonomous language system' refers to the ability of the language to function as a stand-alone linguistic entity with only minimal continuing input from its source languages. Following Saussure ([1916] 1983: 86), the parts of a language must be "synchronically interdependent". Thus changes in the source languages do not feed into the mixed language and visa versa. Whether such a level of autonomy is possible for a mixed language is indeed questionable given that most are spoken alongside one or both of their source languages (§3.1.3). Further there is often a close synchronic and diachronic relationship between other mixing practices and mixed languages. For example, the composition of mixing in many cases resembles patterns which may also be found in code-switching. Indeed Bakker (2003: 129) observes that the majority of mixed languages, which exhibit grammar-lexical mixed structures such as that found in Media Lengua, reflect the patterns found in insertional code-switching. Historically, code-switching also most likely preceded the formation of many mixed languages, and synchronically the mixed language and code-switching may continue to co-exist within the same speaker population. This connection has been demonstrated for at least one mixed language, Gurindji Kriol (McConvell and Meakins 2005) and has been proposed for many other cases (Myers-Scotton 2003; O'Shannessy 2005) (see §4.4). These kinds of similarities cast some doubt on the 'language-ness' of these mixed languages. Yet, despite the symbiotic nature of these types of language mixing, they can also be distinguished from each other.

The purpose of this section is not to be drawn into debates about what is a 'language' but to demonstrate that mixed languages have as much right to the pre-theoretical label 'language' as other linguistic objects which have been called languages. A number of criteria have been used to support the claim of language autonomy in mixed languages: (i) the stability of the language §4.1.1, (ii) the independent development of the source or mixed language §4.1.2, and (iii) the presence of structural features from both languages in a clause §4.1.3. A detailed discussion of the application of these criteria to Gurindji Kriol can be found in Meakins (to appear-a).

### 4.1.1 Language stability

The term 'stable' is used in conjunction with mixed languages by a number of writers (see e.g. Thomason 2003). The notion of 'stability' is meant in a relative sense, i.e. just how established and predictable patterns are in a mixed system in relation to so-called normal languages. Actually defining and measuring stability is highly problematic. Part of the problem relates to the degree of variation found in the language, an issue which will be discussed more fully in §4.2. The other issue is that stability is not something which is generally in question for normal languages. It is acknowledged that variation exists and is meaningful and quantifiable, and that language change occurs. However there are no benchmarks or measures of stability to gauge a mixed language against. This section offers some social and grammatical indicators which may be used to judge the stability of a mixed language: (i) whether the mixed language is spoken outside of its source language context, (ii) the degree of consistency between speakers, and (iii) whether it is being acquired as an L1.

First, the clearest demonstration of language stability occurs when a mixed language is *spoken outside of the bilingual context in which it arose*, that is speakers are no longer fluent in the source languages (Thomason 2003: 24). Michif is one such example. Although Michif is derived from French and Cree, most of its speakers are not fluent in either language. Indeed nowadays most Michif speakers are elderly and the main language of Michif communities has become English (Bakker 1997). Similarly Mednyj Aleut was spoken long after the Russians had withdrawn from Copper Island and speakers of Mednyj Aleut were no longer fluent in Aleut or Russian (Golovko 1994). In other situations, mixed language speakers may not be fluent in the source languages but they may have a good passive knowledge of one or both of them. This situation may occur where the mixed language is an attempt to maintain an ancestral language which is endangered. For example, Gurindji Kriol speakers have a limited command of Gurindji; however they hear Gurindji being spoken by older people and have little difficulty understanding them. Indeed Gurindji Kriol speakers are still often addressed in Gurindji by their parents and it is common to hear conversations where the older person speaks in Gurindji and the younger person replies in Gurindji Kriol (Meakins 2008a: 296). In these situations, the case for language autonomy becomes quite blurred because, though speakers are not fluent in the source languages, there is still potential for influence or input from the source languages making the language less stable.

Cases of mixed languages which exist in total isolation from their source languages are actually quite rare. Most mixed languages are spoken alongside their source languages, and most mixed language speakers are fluent in one or both of the source languages. Smith (2000) calls these *symbiotic* mixed languages, as discussed in §3.1.3. Language stability is harder to achieve in such situations because of the continued link with the language's origins, and as a result, the claim of an autonomous language system is more difficult to establish in these situations. Mixed languages can resemble code-switching or perhaps a sociolect or variety of one of the source languages. For example, Ma'a (Inner Mbugu) speakers are also fluent in (Outer) Mbugu and even code-switch between these languages (Mous 2003a: 86 onwards). Similarly Light Warlpiri speakers are fluent Warlpiri speakers and are in constant contact with other Warlpiri speakers from neighbouring communities who are not speakers of Light Warlpiri. Moreover they have a reasonable command of Kriol and Aboriginal English from their contact with other Aboriginal people. Light Warlpiri speakers also code-switch between

the mixed language and Warlpiri and a contact variety of English (O'Shannessy 2006). Being able to demonstrate that Ma'á or Light Warlpiri operates as a closed linguistic system which is only minimally influenced by its continued contact with the source languages is difficult in such a language ecology.

These three situations - complete separation from the source languages, contact but non-fluency in the sources, and ongoing bilingualism - illustrate three points along a scale of mixed language autonomy, both in the ecological and structural sense. Generally speaking, increased distance from the source languages promotes stability in the mixed system. At the extreme end of the scale where speakers are not only in contact with the source languages but are fluent in one or both, other indicators for stability must be sought.

Another measure of stability in a mixed language is the *degree of consistency* both between and within speakers in their use of lexicon and grammar. This notion relates closely to variation which is discussed in §4.2. For example the choice of lexical items and syntactic constructions is very consistent across speakers in Gurindji Kriol. As a result, Gurindji Kriol speakers use virtually identical constructions to express the same event. This point can be demonstrated looking at a small subset of data consisting of 18 tokens of the sentence: "the dog bit the man on the hand" from 18 different speakers (Meakins to appear-a). Of these 18 sentences, the Gurindji word *warlaku* (the dog), *marluka* (old man) and *wartan* (hand) is used in all 18 sentences, with the Kriol *baitim* (bite) used in 89% of sentences in variation with the Gurindji equivalent *katurl*. Syntactically all pronouns present are Kriol-derived free forms, and similarly any verbal inflection found is of Kriol origin. The Gurindji-derived ergative marker *-ngku* is used in 61% of the sentences[16], and the locative marker *-ta* is found 83.5% of the time, with the Kriol preposition *la* used in the remaining sentences (Meakins to appear-a).

(12)  det   *warlaku-ngku*    i    bin   bait-im   *marluka*   *wartan-ta*.
      the   dog-ERG          3SG.S PST  bite-TR   old.man    hand-LOC

              ↑      ↑          ↑     ↑     ↑          ↑         ↑     ↑

           100%   61%       100%  100%   89%        100%      100%  83.5%

The level of uniformity in lexical and syntactic choices shown by Gurindji Kriol speakers supports its status as a language independent of its sources. For example, though speakers regularly hear Gurindji verbal morphology from older speakers, they consistently use Kriol tense and aspect markers, e.g. *bin* (non-future) in (12). This degree of consistency and stability in Gurindji Kriol developed as particular lexical choices and structures in the language became dominant and impervious to further outside influence and change. Unfortunately consistency is difficult to demonstrate for all mixed languages because a large corpus is required to characterise variation and consistency within the mixed language. Moreover the corpus needs to include examples of equivalent sentences by different and same speakers. Mixed languages, such as Michif, which have elderly speaker populations or Angloromani which is not the main everyday language of the speech community do not lend themselves well to this type of study.

A final indicator of language stability is whether or not children are learning the language as a first language. Many mixed languages such as Michif and Mednyj Aleut are now highly endangered with only elderly speakers and no child learners. The process of language acquisition is a matter of speculation in these cases. Other mixed languages such as Light Warlpiri, Gurindji Kriol and Bilingual Navajo have child language learners and more observations about the context and process of acquisition can be made. For example, O'Shannessy (2006) recorded 5 Warlpiri children over three years. The recordings were made in naturalistic situations where the children interacted with their caregivers who included parents, grandparents and older children. O'Shannessy found that children are spoken to in both Warlpiri and Light Warlpiri, and the language ecology that children grow up in also includes mixing between these languages and Aboriginal English/Kriol. She observed that, though there is quite a lot of overlap between the Warlpiri and Light Warlpiri, children acquire the languages separately. They begin speaking Light Warlpiri first but from the age of around five, they include Warlpiri in their linguistic repertoire though they remain almost exclusively Light Warlpiri speakers. Observations of children acquiring a mixed language such as Light Warlpiri as a separate code support the case for language stability. Clearly a separate linguistic entity is being identified and singled out for specific attention by children.

These three factors - independence from source languages, speaker consistency and language acquisition - contribute to the stability of a mixed language. Although it is difficult to compare stability across contact and non-contact languages, it is clear that there is a correlation between stability and language autonomy.

### 4.1.2 Independent development of the mixed language and source languages

Language autonomy may also be demonstrated by the independent development of the mixed language and its sources. In this scenario a change in the source languages does not necessarily imply a change in the mixed language and vice versa. Bakker (2003: 126) gives examples of cases where changes in the source languages are not reflected in the mixed language. Here I examine independent developments which have occurred in mixed languages, firstly where unique forms have been created in the mixed language system and secondly where forms from the source languages have developed new functions in the mixed language.

First mixed language developments are not always reflected in the source languages. In the case of Light Warlpiri, O'Shannessy (2005: 39) identifies a unique auxiliary paradigm in Light Warlpiri which consists of a pronominal proclitic and a tense-aspect element.

| AUXILIARY FORM | 1SG.S | 1PL.S | 2SG.S | 2PL.S | 3SG.S | 3PL.S |
|---|---|---|---|---|---|---|
| NON-FUTURE | a=m | **wi=m** | **yu=m** | - | **i=m** | de=m |
| FUTURE | a=rra | wi=rra | yu=rra | **yumob=rra** | i=rra | de=rra |

*Figure 4        Light Warlpiri auxiliary paradigm (adapted from O'Shannessy 2006: 39)*

This system is based on Kriol morphemes but has a Warlpiri flavour to the structure since Warlpiri also has a single auxiliary structure which combines these elements, although in the reverse order. Although part of this paradigm is present in Kriol it is not a complete paradigm. The bolded elements in Figure 4 show the gaps in the Kriol paradigm. In these gaps, Kriol uses full forms rather than clitics - *bin* rather than *=m* and *garra* rather than *=rra*. Thus the Kriol forms have developed further in Light Warlpiri, and there is no evidence that this development is feeding back into Kriol. Further it is not the case that this system is replacing the auxiliary system in Warlpiri. O'Shannessy (2005: 52) also observes speakers do not use this auxiliary when code-switching between Warlpiri and Kriol. She presents this auxiliary system as a diagnostic for Light Warlpiri utterances in mixed conversations. Thus Light Warlpiri has a unique system which has developed independently of its source languages which suggests that Light Warlpiri operates as an autonomous language.

Another example of the development of a unique system comes from Sri Lanka Malay. Smith, Paauw and Hussainmiya (2004: 2004) observe that Malay-derived prepositions have become post-positional case-markers under the influence of Tamil. What is particularly interesting is that syncretism occurs between dative and accusative marking (only distinguished by optionality - accusative marking is optional where dative marking is not)[17]. This syncretism is not observed in any of its source languages. Tamil and Sinhala distinguish these categories with separate markers and Vehicular Malay marks the dative but not the accusative with a preposition. Further, although the forms of the case suffixes come from Vehicular Malay, the origin of the accusative/dative case suffix *-n(y)a(ng)* is obscure. Thus in both distribution and form, the accusative/dative marker is a unique form in Sri Lanka Malay and there is no evidence that this marker has fed back into its source languages.

Forms in a mixed language may also develop different functions from the source language it is derived from. For example, in Bilingual Navajo some constituents function in unique ways which contrast from their language source. In particular, the verb *áshlééh* 'to prepare/make' is used as a light verb in Bilingual Navajo in order to integrate English verbs into the Navajo morpho-syntactic frame. This verb is not used in this manner in Navajo, moreover it is not used in this manner by English-Navajo code-switchers who do not speak the mixed language.

(13)  bi-'éé'          **change**    íí-**lééh**
      3POSS-clothes    change        1DU-make
      "The two of us are changing her clothes." (Schaengold 2004: 68)

(14)  bi-'éé'          lahgo      át'éhí      bii        ndeezhteeh
      3POSS-clothes    different   one         3.into     1DU.move.animate
      "The two of us are changing her clothes." (Schaengold 2004: 68)

These examples demonstrate one of the ways in which elements from Bilingual Navajo's source languages operate in different ways in the mixed language. Through these differences, Bilingual Navajo can be shown to be an autonomous language rather than a mixture of two separate language systems. The development of Navajo elements in the mixed language has not extended to Navajo itself. Moreover the use of Navajo

constituents in code-switching generally does not differ radically from monolingual clauses. Navajo systems, which have been altered in the mixed language, continue to be used unchanged in the source languages which demonstrates that these developments are a feature of the mixed language.

The distribution of forms may also change in the mixed language where they remain unchanged in the source languages. For example, as was discussed above, Sri Lanka Malay case markers consist of a Vehicular Malay-derived phonological form which has become postposed under the influence of Tamil case-marking. Sri Lanka Malay now contains the same case-marking categories as Tamil. Nonetheless some distributional differences occur. For example, the Sri Lanka Malay locative marker *-ka* is derived from the Vehicular Malay dative preposition. It is used to mark both locations and goals in Sri Lanka Malay where Tamil (and Sinhala) uses dative to mark goal and locative to mark location (Smith, Paauw, and Hussainmiya 2004: 206). Although the case markers generally function as they do in Tamil, this is one example where the form has been extended to new functions. Although Sri Lanka Malay is still in contact with Tamil, this change in distribution is not reflected in Tamil. This is evidence that Sri Lanka Malay is operating as an autonomous language.

### 4.1.3    Presence of structural elements from both languages

Another argument for treating a mixed language as a coherent linguistic entity is the presence of structural features from both languages. As was discussed in §3.2.1.1, this is considered to be one of the defining features of a mixed language because it defies trends seen in other language contact scenarios such as borrowing or creole languages where generally one language dominates the morpho-syntactic frame of the mix. Such patterns tend to exist in N(oun)-V(erb) splits such as Michif, Mednyj Aleut and Gurindji Kriol rather than L(exicon)-(G)rammar mixes such as Angloromani and Media Lengua.

The structure of *Michif* is shared by French and Cree, as was introduced in §2.7. French and Cree share the load between the verb and noun systems. In this respect Cree provides the verbal morpho-syntactic frame, as seen in tense marking and cross-referencing pronouns and the French, the NP frame. In terms of the NP, Michif preserves French determiners, plural marking and adjectival agreement in Michif. Another language which distributes the structural load over both source languages is *Mednyj Aleut*. Again this split largely follows an N-V split with Aleut nominal structure including the Aleut case frame and other suffixes such as agent, instrumental and location markers. On the other hand Mednyj Aleut derives much of its finite verbal inflectional morphology from Russian, such as person and tense marking (Thomason 1997: 457-59).

Both source languages also contribute to the structure of *Gurindji Kriol* and *Light Warlpiri* (Meakins to appear-a). Again the different source languages dominate different parts of the grammar. In the nominal structure, inflectional morphology of Gurindji and Warlpiri origin is clearly demonstrated in case marking. Gurindji/Warlpiri-derived ergative, dative, locative, allative and ablative case markers are all present in the mixed language. Kriol is responsible for the verbal frame. In both Gurindji Kriol and Light Warlpiri, Kriol contributes two verbal suffixes, the continuative and transitive markers. Kriol-derived tense, aspect and mood markers are also present, for example *bin* (past), *garra* (potential), *yusta* (habitual). Additionally, TAM clitics, such as *=m* (non-future)

and *=rra* (potential) from Kriol, are also present. These are based on reduced forms of Kriol tense and aspect auxiliary forms. Finally Kriol-derived clitic pronouns cross-reference optional nominals. Thus, like Michif and Mednyj Aleut, the grammatical frame of Gurindji Kriol and Light Warlpiri represents a composite structure, with both source languages contributing significant structural elements to the mix. The rarity of this kind of composite structure in other language contact varieties (see §3.2.1.1) speaks to the uniqueness of this phenomena. It suggests that these languages are not merely cases of code-switching or borrowing but language systems in their own right.

## 4.2 Characterising variation in mixed languages

Related to the issue of autonomy and stability is variation. The presence of variation in mixed languages is seldom discussed and sometimes even played down in order to avoid questions of autonomy (Matras and Bakker 2003: 7-8; Mous 2003a: 7; but see Bakker 1997: 159; Smith and Paauw 2006: 165, for observations of variation in Michif, Angloromani, Ma'á and Sri Lanka Malay, respectively). Nonetheless it plays an important role in the formation of mixed languages and continues to affect their evolution, as was introduced in §3.2.2.2. Variation is, at once, one of the key ingredients of language change and one of the results of this change, and in this respect it is indicative of perpetual change. There are a number of sources of this variation including continuing external input from the source languages, as well as internal variation from shared forms which are unique to the mixed language (see also §4.1.2) and idiolectal varieties. This variation is part of an autonomous language system, which can be mapped using variationist methodology. This section considers the external and internal sources of variation in relation to mixed languages, and demonstrates how variation may be characterised in a mixed language.

 *External* variants exist due to the symbiotic nature of many mixed languages, in that they continue to be spoken alongside one or more of their source languages (see §4.1.1). Though they may not actively speak the languages themselves, speakers still have access to the source language lexicon and grammar, and these may continue to affect the structures used in the mixed language. For example Bakker (1997: 159) admits that in Michif, though many speakers have no knowledge of the source languages, French or Cree, knowledge of these languages creates some variation among speakers. Specifically, he observes that speakers who know French tend to use more French elements, particularly function words. Bakker (1997: 160) also observes that speakers vary in their use of French or Cree elements in Michif depending on their interlocutor.

 The other source of variation is language *internal* where a language may contain a choice of structures or vocabulary which can mark a particular function. Language internal variation may be quite complex, involving forms from the source languages which both appear in the mixed language, as well as unique forms which have formed within the mixed language system (see §4.1.2), and finally idiolectal variation. It is often the situation that more variants exist than are contributed by the source languages. These new variants represent amalgams of the source languages. For example in Gurindji Kriol speakers use the Gurindji-derived locative case marker *-ta* 87% of the time to mark static location. However, younger speakers occasionally use Kriol-derived *langa* preposition (9.5%) or a double-marked construction (3.5%) which contains both the case marker and

preposition. The double-marked constructions are specific to Gurindji Kriol and represent a compromise strategy (Meakins 2007: 242).

The presence of variation does not undermine the notion of an autonomous language system. Variation has long been recognised as a normal and integral part of all language systems. Indeed the program of variationists from Labov (1972) onwards has been to challenge the assumption of linguistic uniformity which is a characteristic of formal grammatical theory. Variation within a language system is not necessarily, and indeed very rarely, random, but is meaningfully distributed. Variationists argue that the use of particular variants in language can be predicted by the presence or absence of other linguistic and sociolinguistic factors. This theoretical underpinning is based on Weinreich, Labov and Herzog's (1968: 100) notion of "ordered heterogeneity". Labov (1969) claims that speakers make choices when they use language and these choices are systematic and can be predicted through statistical modelling. In this respect, the grammar of a language can be characterised as probabilistic, rather than categorical. Thus the data can be described using a set of variable rules which constitute the grammar of a language. As Labov (1969: 759) concludes "we are dealing with a set of quantitative relations which are the form of the grammar itself". Poplack (1993: 253) also observes that treating the patterns of variation as the grammar of a language bases the endeavour of grammatical description in empiricism, rather than unobservable underlying structures and rules.

Within the context of mixed languages, variationist methodology offers a way of characterising the use of functionally equivalent forms, such as the Gurindji locative case marker or Kriol preposition as discussed above for Gurindji Kriol. The presence of variation is not indicative of instability or a lack of language autonomy, but can be treated as meaningful within a contained linguistic and social system. The presence of variation does not undermine the notion of a coherent mixed system because it is motivated by other factors or variables with the system. Variables may be non-linguistic and linguistic. For example, some variationist studies measure the effect of *social* factors such as age, gender and socio-economic classes, and stylistic factors such as the formality of a communicative context on the choice of one linguistic variant over another (see Chambers 1995 for a good overview). Other studies have broadened the pool of factors to include *linguistic* factors. For example, in a study of the constraints on the use of passive vs active sentences in English, Weiner and Labov (1983) found that the choice of syntactic structure carried neither stylistic nor social significance but the choice was constrained almost entirely by syntactic factors[18]. In some cases, both linguistic and social factors may play a role in the realisation of one variant over another. For example age is a factor in Gurindji Kriol. The younger the speakers, the more likely they will use a Kriol preposition or a double-marked construction. Linguistic context also plays a role. Again the preposition is more likely to be used if the following NP includes a determiner (Meakins 2007: 242).

The use of a variable in relation to other variables is measured using statistical modelling, most commonly regression models which measure the likelihood of a linguistic variable such as a phoneme, morpheme or syntactic structure occurring with respect to other social and linguistic variables, of the type described above. Regression models, can be applied to mixed languages to see whether a variable can be characterised as a part of the mixed language system, or whether it is external to the system. Another

example from Gurindji Kriol: the Gurindji-derived ergative marker is only used optionally, which contrasts with the categorical use of the ergative marker in Gurindji. This optionality seems to be the result of the ergative marker coming into contact with the equivalent Kriol system of argument marking - word order. One interpretation of this variation may be that continuing interplay between Gurindji and Kriol argument marking systems has made this Gurindji Kriol system unstable. However Meakins (2009) shows that the variable application of the ergative marker acts in a coherent manner. In a study of the motivations for the appearance of the ergative marker, Meakins (2009: 71 onwards) tests the use of the ergative marker against ten variables: two sociolinguistic variables - age of speaker and the formality of context; a lexical variable - the language of the stem; a number of grammatical and semantic variables relating to the degree of transitivity of the clause: potentiality, actualisation of the event indicated by the verb, agent[19] animacy, object animacy, and whether the object is overt; and finally two variables which relate to the clause structure: the position of the agent in relation to the verb, and the presence of a co-referential pronoun. She finds that the ergative marker is more likely to appear if the agent is inanimate, found post-verbally and in conjunction with a co-referential pronoun. The use of the ergative marker decreases when the verb is marked with continuous aspect and the event denoted by the verb has not come to completion (Meakins 2009: 79). Meakins interprets these results as an indication that the main function of the ergative marker in Gurindji Kriol is discourse related, specifically its presence highlights the agentivity of a subject nominal (see also Meakins and O'Shannessy 2010). Similar analyses exist for languages with optional ergative marking (see McGregor 2010 for an overview). Studies of this kind demonstrate that variation in a mixed language system is not a sign of the fragility of the language but rather is a part of a systematic grammar. Moreover variation can be quantified in such a way that its meaning can be deciphered.

## 4.3 Can two phonological systems exist in the one language?

The fusion of lexical and morphological material from different languages begs the question of what happens at the phonological level. Do the words and affixes maintain their original phonological forms thereby producing a stratified system or are they integrated into one phonological system? The debate about phonological stratification in mixed languages has centred around the phonology of Michif, though more recently van Gijn (2009) has compared the phonological systems of four mixed languages.

Papen (1987) and Bakker and Papen (1997) claim that two separate phonological systems function in Michif, with separate phoneme inventories for both the French and Cree components and separate sets of phonological rules and processes. Implicit in this claim is that each lexical item in Michif must be marked for its language of origin (Bakker and Papen 1997: 312). However Rosen (2000: 6) suggests that this assumption presents 'learnability' problems as there are insufficient phonological cues for learners to determine whether a lexical item is from either French or Cree. She suggests that any differences in the French and Cree components are historical artefacts and are no longer functional in Michif. For example, Rosen (2000: 16) follows Bakker (1997) that the French rule of liaison is no longer productive in the French component of Michif. Liaison applies to the final consonant of a determiner or adjective. In certain syntactic

environments it is pronounced as the onset of the following noun. Instead Rosen claims that all vowel-initial words in Michif have been reanalysed as /n/, /l/ or /z/ initial. Indeed this claim was originally presented by Bakker and Papen (1997: 309). Papen (2003) revisits this problem finding synchronic evidence for a stratified phonology. He (2003: 5) notes that most of the Michif dictionary entries of French vowel initial words include the liaison consonant. Nonetheless Papen suggests that "the only valid argument in favour of considering liaison to no longer be functional in Michif would be to show that the 'wrong' (e.g. unexpected) consonant occurs … or that there is no consonant at all". Papen finds that only 10% of cases contain unexpected liaison consonants, which he posits as evidence for the continuing existence of the liaison rule. However it seems reasonable to suggest correct use of the liaison consonants does not rule out the possibility that these now consonant-initial words are now lexicalised as the result the liaison process which occurred historically. And indeed Papen does not present any arguments which would suggest that Michif speakers analyse these words as vowel initial.

Although it can be argued that Michif has a stratified phonology, this generalisation does not necessarily hold for all mixed languages. Van Gijn (2009) suggests that the typology of the mixed language determines whether one language is subsumed into another's phonology, or whether two phonological systems are maintained. He compares Michif with Media Lengua. Media Lengua differs greatly from Michif in terms of typology, being a grammar-lexicon mixed language rather than exhibiting an N-V split. Van Gijn (2009: 96) notes that, unlike Michif, Media Lengua has only one phonological system, that of Quechua, which may be the product of the level of mixing at the prosodic word level.

> In Media Lengua, almost all words have both Spanish (stems) and Quechua (affixes) elements, in Michif, on the other hand, there are more unmixed words: verbs are generally Cree (both stems and affixes), noun phrases, or at least nouns, are to a large extent unmixed French. In other words, Michif has larger unmixed domains … where French or Cree words can apply. (van Gijn 2009: 109)

He suggests that the mixed nature of Media Lengua words makes it difficult to maintain two phonological systems, whereas language mixing in Michif occurs at a higher point in the prosodic hierarchy making a stratified phonology more viable.

## 4.4 Can code-switching lead to the formation of a mixed language?

A further issue for mixed languages is a characterisation of the language practices of speakers at the time of genesis. This section looks specifically at code-switching and its role in the formation of mixed languages. The extent to which code-switching is a factor in the formation and resulting structure of mixed languages is debated extensively. Bakker (2003: 129) is the strongest critic of code-switching approaches suggesting that they play no role in mixed language genesis, and that typological resemblances between mixed languages and code-switching are the product of selective comparison on the part of researchers. However, McConvell and Meakins (2005), present empirical evidence demonstrating that code-switching contributed to the formation of at least one mixed

language, Gurindji Kriol (§4.4.1). A growing body of work supports the contribution of code-switching to mixed languages (Gardner-Chloros 2000; Auer 1999; Backus 2003; Myers-Scotton 1993). Within this literature, there are two main approaches. The first considers the different structures of mixed languages and compares them directly to different types of code-switching, such as insertional and alternational code-switching (Backus 2003) (§4.4.2). The second approach is more explanatory, proposing a transitory stage between code-switching and a mixed language. This approach utilises structural constraint theories of code-switching to better understand the resultant character of mixed languages (Auer 1999; Myers-Scotton 2003). This approach was discussed in §3.2.1.2.

### 4.4.1 Empirical evidence for the link between code-switching and mixed languages

One of the main disclaimers in the debate about a link between code-switching and mixed languages concerns the lack of empirical evidence to support a claim for the link between code-switching and the formation of mixed languages. While Bakker states that "we have no documentation of a transitory phase between the supposed code-switching behaviour preceding the genesis of the mixed language" (2003: 129), Auer suggests that claims about the transition from code-switching to mixed languages are "plausible guesses rather than empirically based" (1999: 324). Myers-Scotton (2002: 249) believes that the next step in her Matrix Language Turnover theory (see §3.2.1.2) which outlines the progress from code-switching to a mixed language is to demonstrate the process using actual data. Finally, Backus says that all of these claims "call for evidence which, to the best of my knowledge, has not been brought forward" (2000: 104), and later he suggests that it is doubtful whether this sort of evidence would ever be available (2003: 241).

Indeed many of the mixed languages which have been documented are generally 100 or more years old. Evidence, then, is a matter for the historical record in these cases, and few of these languages have a significant body of written work associated with them. Nonetheless empirical evidence for code-switching preceding a mixed language now exists for one mixed language, Gurindji Kriol. McConvell and Meakins (2005) show that code-switching not only preceded the formation of Gurindji Kriol, but that a number of structures found in the mixed language correspond with the pattern of Gurindji-Kriol code-switching.

McConvell (1988) suggests that multilingualism was a traditional social practice of the Gurindji before Kriol was added to this repertoire after colonisation (see §2). By the mid 1970s, it appeared that inter-sentential and intra-sentential code-switching between Gurindji and Kriol/English was a very common style of communication (McConvell 1985: 96). McConvell and Meakins (2005) demonstrate that resonances of the mixed language, Gurindji Kriol, can be found in the patterns of code-switching from this time. In a study of a conversation between six Gurindji stockmen who were butchering a cow in a bush paddock near Kalkaringi, they find that 73% of the mixed utterances used a Kriol verbal structure as become the norm for Gurindji Kriol two decades later. It appears that during this period the Kriol verbal structure was already becoming dominant. Indeed now Kriol forms the basis of the VP structure of Gurindji Kriol with Gurindji verb morphology never found. Despite the predominance of Kriol in the VP of the code-switched utterances, Gurindji morphology, including case and derivational morphemes, was also present in the structure of the noun phrases in code-switched utterances. Code-

switched utterances from the 1970s bear a strong resemblance to the mixed language spoken today. Compare (15) with (16). Gurindji elements are italicised.

(15)  *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 & son, cut it across the east (side of the cow)."
      (1970s Gurindji-Kriol code-switching)

(16)  an    skul-*ta-ma*        jei   bin   hab-im       sport *karu-walija-ngku*.
      and   school-LOC-DIS   3PL.S NF    have-TRN    sport child-PAUC-ERG
      "And the kids had sport at school." (2006 GK mixed language)

Through these general observations of Gurindji-Kriol code-switching in the 1970s, McConvell and Meakins (2005) provide the missing empirical link for arguments about the transition between code-switching and mixed languages for at least this mixed language. Of course this evidence does not exclude the possibility that other mixed languages have derived from different bilingual practices at their point of genesis.

**4.4.2   Typological similarities between code-switching and mixed languages**

A number of studies have examined the typological similarities between code-switching and the structure of mixed languages (Bakker 2003; Backus 2003; Mous 2003a). Muysken's (2000) distinction between *insertional* and *alternational* code-switching is relevant for this work (see §3.2.1.2). Insertional code-switching is generally considered to have the greatest influence on mixed language genesis in these comparisons, because the mixed languages studied generally do not resemble a fossilised form of alternational code-switching, whereas they look remarkably like insertional code-switching. For example, Michif has been compared with various synchronic descriptions of insertional code-switching between pairs of English or French, and an Algonquian language, for example, Plains Cree-English code-switching (Bakker 1997: 181-182) and Montagnais-French (Drapeau 1991, cited in Bakker 1997: 184-86). Insertional code-switching can be characterised by a nested ABA pattern (Muysken 2000: 63). This means that the segments on either side of the inserted constituent are grammatically related and are derived from the same language. Bakker finds this to be a common pattern of switching in all of these cases which involve noun phrase and prepositional phrase insertions from French or English. This insertion pattern reflects the NP-VP split found in Michif, where French dominates the NP structure. Bakker presents a number of objections to the link between this type of code-switching and Michif[20]. Nonetheless he suggests that insertional code-switching and mixed languages in general do show striking typological similarities.

A similar comparison can be made between Gurindji-Kriol code-switching and Gurindji Kriol, the mixed language. This comparison is more diachronic because the language varieties are linked historically and the speakers are socially connected, i.e. the speakers of the mixed language are the children and grandchildren of the code-switchers from the 1970s. Many aspects of the structure of Gurindji Kriol bear a striking resemblance to insertional patterns which are also found in the code-switching from the

1970s (Meakins 2007: 175 onwards). For example, in both cases Gurindji insertions of single content words such as direct objects, intransitive subjects and verbs occur often in both the code-switching and the mixed language. Nonetheless alternational patterns are also present in both forms of language mixing. Alternational switches can be characterised by *peripherality* (Muysken 2000; Treffers-Daller 1994). Peripherality refers to how marginal the switched elements are to the argument structure of the clause and also to the structural position of switches. Constituents which are adjuncts or are switched on the physical edge of a clause are prototypical cases of alternational code-switching. This is a common pattern in Gurindji-Kriol code-switching (Meakins 2007: 183 onwards). This feature often goes hand-in-hand with the relationship of the switched element to the predicate argument structure of the matrix clause. Constituents which are peripheral to the argument structure of the clause often occur at clause boundaries in this data. These kinds of switches are also found in the mixed language. Common alternational switches which are found in both forms of mixing include Gurindji locational adjuncts switched on the right and left periphery of the clause.

The continued presence of alternational patterns in Gurindji Kriol is interesting given that, as discussed above, most comparisons between mixed languages and code-switching find the insertional pattern to be dominant. This observation is probably just a sampling issue in that most identified mixed languages are of the grammar-lexicon kind. In these cases, one language is stronger providing the grammatical frame while the other language contributes only lexicon. This mixing pattern has strongly supported the idea that mixed languages are the result of insertional code-switching. Nonetheless Backus (2003) raises another issue with the notion that alternational code-switching can contribute to the structure of mixed languages - the *predictability* of switches. The key to forming a mixed language lies in the ability of code-switching patterns to conventionalise. Conventionalising requires switching between languages to be predictable, thereby narrowing the degree of structural variation (see §3.2.1.2). Alternational code-switching is considered to be a socially-marked form of language mixing which is unconstrained by the structures of the interacting languages, and is therefore unlikely to grammaticalise into a stable autonomous language. On the other hand, insertional code-switching is assumed to be more rule-governed and predictable, properties which lend themselves to grammaticalisation. Backus suggests that one of the characteristics of alternational code-switching is that it is impossible to predict when a speaker will use language A or language B. In this respect, alternational code-switching "entails unbridled variation" (2003: 240). Backus suggests that the unpredictability of alternational code-switching is the result of discourse functions which he believes are marked by this type of switching. These functions include alignments with another culture and topic changes which he observes in Turkish-Dutch mixed lects (2003: 248). Backus concludes that it is simply not in the nature of alternational code-switching to conventionalise because this would mean losing its inherent communicative function. The choice of language, rather than the lexical item itself, is always a communicative choice.

**5.  Conclusion: Beyond characterising mixed languages as mixed *forms***

There are a variety of languages which have been classified as mixed languages. Lexically they range from languages which derive an extraordinary amount of vocabulary from one language and their grammar from another (Media Lengua, Old Helsinki Slang) to languages which selectively replace lexical items according to specific communicative contexts (Angloromani, Ma'á). Mixed languages also differ according to structure. In some mixed languages the structure is clearly derived from one language (Media Lengua, Angloromani, Ma'á, Bilingual Navajo, Old Helsinki Slang). In other cases, the two source languages contribute relatively equal amounts of structure to the mix in a way that contrasts dramatically with other language contact scenarios (Michif, Mednyj Aleut, Gurindji Kriol, Light Warlpiri). Yet other examples have restructured one language on the basis of another such that, on the surface, the language 'looks' like one of its sources, however it has mapped these forms onto the other language's grammar (Sri Lanka Malay).

Despite these differences, what is common to all of these languages is that they have emerged in situations of bilingualism where a common language is already present. In this respect, they do not serve a communicative function, but rather they are markers of an in-group identity, whether it be a new identity created through mixed marriages (Michif) or the maintenance of an old identity which is under threat (Angloromani, Gurindji Kriol, Light Warlpiri, Ma'á, Bilingual Navajo). Despite this common socio-historical cradle, little more can be predicted. It appears from the variation in structural outcomes that different contact situations can result in similar mixed languages, and different mixed languages may arise from similar contact situations.

One theme which emerged strongly from a discussion of these languages has been the inadequacy of *form* as a basic descriptor of the structure of a mixed language. What is meant by 'form' is the phonological shape of a morpheme or lexeme, that is - what is said. Many mixed languages are described in this manner, for example, as mixes of the stems from one language and the affixes from other, or as splits between nouns and verbs and their accompanying morphology. Yet these descriptions do not adequately characterise mixed languages. Sri Lanka Malay demonstrates this problem most clearly. This language consists almost entirely of the lexical and morphological material from a contact variety of Malay. Yet Sri Lanka Malay is unintelligible to a speaker of any Malay variety. Whilst the forms are Malay-derived, their syntactic position, distribution and function is based on Tamil and Sinhala. Mixing occurs below the level of the form. Gurindji Kriol also demonstrates this kind of mixing. For example, as has been shown, although this mixed language maintains Gurindji-derived case-marking, these suffixes have shifted in their function and distribution under the influence of Kriol-derived prepositions and word order. One needs to look below the surface of this language because the mixing occurs within the lexical entry. It is here that the features of the lexeme or morpheme including its semantics, word class, syntactic distribution, pragmatic force and function may be split between the source languages. In this respect the surface form masks the true extent of language fusion or mixing.

Although some of these types of observations were made in earlier descriptions of mixed languages, analyses such as these have emerged from the more recent large scale documentation of some mixed languages (for example, Sri Lanka Malay, Gurindji Kriol, Light Warlpiri). Indeed other studies of language contact and contact varieties have utilised this notion of layering in the lexical entry as a basis for explaining language

change, for instance studies of first language interference in second language learning, descriptions of processes such as calquing and convergence, composite code-switching (Myers-Scotton 2003) and the Relexification Hypothesis in creole studies (Lefebvre 1998). It is likely that this approach to describing language mixing will influence analyses of mixed languages in further studies.

In general, mixed languages provide a unique opportunity to study the often observable birth, life and death of languages both in terms of their socio-historical context and structural changes. These debates continually benefit from the identification of new mixed languages, some of which reinforce current views, while others challenge us with a new range of structural outcomes that result from the intense interaction between the grammars of two languages.

## 6. References

Aikhenvald, A. Y., and R.M.W. Dixon      2007    *Grammars in Contact: A Cross-Linguistic Typology*. Oxford: Oxford University Press.

Ansaldo, Umberto    2008    Sri Lanka Malay revisited: Genesis and classification. In: A. Dwyer, D. Harrison and D. Rood (eds.), *A World of Many Voices: Lessons from Documented Endangered Languages*, 13-42. Amsterdam: Benjamins.

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

Auer, Peter    1998    Introduction: Bilingual conversation revisited. In: Peter Auer (ed.), *Code-switching in Conversation: Language, Interaction and Identity*, 1-24. London: Routledge.

———      1999    From codeswitching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *Journal of Bilingualism* 3 (4):309-332.

Backus, Ad    2000    Insertional codeswitching in an immigrant language: "just" borrowing or lexical re-orientation. *Bilingualism: Language and Cognition* 3 (2):103-105.

———      2003    Can a mixed language be conventionalized alternational codeswitching? In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate: Theoretical and Empirical Advances*, 237-270. Berlin: Mouton de Gruyter.

Bakker, Peter  1994    Michif, the Cree-French mixed language of the Metis buffalo hunters in Canada. In: Peter Bakker and Maarten Mous (eds.), *Mixed Languages: 15 Case Studies in Language Intertwining*, 13-33. Amsterdam: Uitgave IFOTT.

———1997    *A Language of Our Own: The Genesis of Michif, the Mixed Cree-French language of the Canadian Métis*. New York: Oxford University Press.

———1998    Para-Romani language versus secret languages: Differences in origin, structure and use. In: Yaron Matras (ed.), *The Romani Element in Non-Standard Speech*, 69-96. Wiesbaden: Harrassowitz.

―――2000a   Convergence intertwining: An alternative way towards the genesis of mixed languages. In: D. Gilbers, J. Nerbonne and J. Schaeken (eds.), *Languages in Contact*, 29-35. Amsterdam: Rodopi.

―――2000b   Rapid language change: Creolization, intertwining, convergence. In: C. Renfrew, A. McMahon and R. L. Trask (eds.), *Time Depth in Historical Linguistics*, 585-620. Cambridge: McDonald Institute for Archaeological Research.

―――2002    An early vocabulary of British Romani (1616): A linguistic analysis. *Romani Studies Fifth Series* 12(2):75-101.

―――2003    Mixed languages as autonomous systems. In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate: Theoretical and Empirical Advances*, 107-50. Berlin: Mouton de Gruyter.

Bakker, Peter, and Maarten Mous    1994   Introduction. In: Peter Bakker and Maarten Mous (eds.), *Mixed Languages: 15 Case studies in Language Intertwining,*  1-11. Amsterdam: Uitgave IFOTT.

―――            1994    *Mixed Languages: 15 Case studies in Language Intertwining*. Amsterdam: Institute for Functional Research into Language and Language Use (IFOTT).

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

Bickerton, Derek       1981    *Roots of Language*. Ann Arbor: Karoma.

Boretzky, Norbert, and Birgit Igla    1994   Romani mixed dialects. In: Peter Bakker and Maarten Mous (eds.), *Mixed Languages: 15 Case studies in Language Intertwining,*  35-68. Amsterdam: Uitgave IFOTT.

Brenzinger, Mathias   1987    *Die Sprachliche und Kulturelle Stellung der Mbugu (Ma'a)*. MA, Cologne.

Chambers, J. K       1995    *Sociolinguistic Theory*. Cambridge, Mass.: Blackwell.

Charola, Erika 2002    The verb phrase structure of Gurindji Kriol. Honours dissertation, Dept. of Linguistics and Applied Linguistics, Melbourne University.

Chen, Naixiong       1986    Guanyu Wutun Hua (An outline of the Wutun linguistic structure). *Journal of Asian and African Studies* 31:33-52.

Clements, J. Clements 2009    *Linguistic Legacy of Spanish and Portuguese*. Cambridge: Cambridge University Press.

Clements, J. Clements, Patrícia Amaral, and Ana Luís       2008    Cultural identity and the structure of a mixed language: The case of Barranquenho. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 34S

Croft, William 2000   *Explaining Language Change: An Evolutionary Approach*. Harlow, England: Longman.

———2003   Mixed languages and acts of identity: An evolutionary approach. In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate: Theoretical and Empirical Advances*, 41-72. Berlin: Mouton de Gruyter.

Dalton, Lorraine, Sandra Edwards, Rosaleen Farquarson, Sarah Oscar, and Patrick McConvell   1995   Gurindji children's language and language maintenance. *International Journal of the Sociology of Language* 113: 83-98.

de Smit, Merlijn   2010   Modelling mixed languages: Some remarks on the case of Old Helsinki Slang. *Journal of Language Contact* 3:1-19.

DeGraff, Michel   2005   Linguists' most dangerous myth: The fallacy of Creole Exceptionalism. *Language in Society* 34:533-91.

Disbray, Samantha   2008   More than one way to catch a frog: Children's discourse in a language contact setting. PhD dissertation, Dept. of Linguistics and Applied Linguistics, University of Melbourne.

———2009   More than one way to catch a frog: Children's discourse in a language contact setting. PhD Dissertation, Linguistics and Applied Linguistics, University of Melbourne, Melbourne.

Disbray, Samantha, and Jane Simpson   2005   The expression of possessive in Wumpurrarni English, Tennant Creek. *Monash University Linguistics Papers* 4 (2): 65-85.

Drapeau, Lynn 1991   Michif replicated: The emergence of a mixed language in northern Quebec. Paper read at Tenth International Conference on Historical Linguistics, at Amsterdam.

Dreyfuss, G. R., and D. Oka   1979   Chinese Indonesian: A new kind of language hybrid? *Papers in Pidgin and Creole Linguistics (Pacific Linguistics)* A-57.

Gardani, Francesco   2008   *Borrowing of Inflectional Morphemes in Language Contact*. Frankfurt: Peter Lang.

Gardner-Chloros, Penelope   2000   The tortoise and the hare: Distinguishing processes and end-products in language contact. *Bilingualism: Language and Cognition* 3 (2): 112-114.

Givón, Talmy 1979   *On Understanding Grammar*. New York: Academic Press.

Golovko, Evgeniy   1994   Mednyj Aleut or Copper Island Aleut: An Aleut-Russian mixed language. In: Peter Bakker and Maarten Mous (eds.), *Mixed Languages: 15 Case studies in Language Intertwining*, 113-121. Amsterdam: Uitgave IFOTT.

———1996   A case of nongenetic development in the Arctic area: The contribution of Aleut and Russian to the formation of Copper Island Aleut. In: *Language Contact in the Arctic. Northern Pidgins and Contact Languages*, 63-77. E. Jahr and I. Broch (eds.),

Berlin: Mouton de Gruyter.

———2003   Language contact and group identity: The role of 'folk' linguistic engineering. In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate: Theoretical and Empirical Advances*, 177-208. Berlin: Mouton de Gruyter.

Golovko, Evgeniy and Nikolai Vakhtin      1990    Aleut in contact: The CIA enigma. *Acta Linguistica Hafniensia* 72: 97-125.

Gómez Rendón, Jorge 2008   *Typological and Social Constraints on Language Contact. Amerindian languages in contact with Spanish*. Utrecht: LOT Dissertation Series.

Grant, Anthony       1994    Shelta: The secret language of Irish travellers viewed as a mixed language. In: Peter Bakker and Maarten Mous (eds.), *Mixed Languages: 15 Case studies in Language Intertwining*, 123-150. Amsterdam: Uitgave IFOTT.

Greenberg, Joseph     1999   Are there mixed languages? In: A.A Visagin, R. Vroon, M. L. Gasparov (eds.), *Essays in Poetics, Literary History and Linguistics Presented to Viacheslav Vsevolodovich Ivanov on the Occasion of his Seventieth Birthday*, 626-33. Moscow: OGI.

Hancock, I. F. 1970    Is Anglo-Romanes a creole? *Journal of the Gypsy Lore Society* 49: 41-44.

———. 1976. The pidginization of Angloromani. In: *New Directions in Creole Studies*, G. Cave. 1-23. Georgetown: University of Guyana.

Hardy, Frank   1968    *The Unlucky Australians*. Melbourne: Nelson.

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

Heath, Jeffrey 1978    *Linguistic Diffusion in Arnhem Land*. Canberra: AIAS.

Heine, Bernd, Ulrike Claudi, and Friederieke Hünnemeyer  1991    *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.

Heine, Bernd, and Mechthild Reh      1984    *Grammaticalisation and Reanalysis in African Languages*. Hamburg: Helmut Buske.

Himmelmann, Nikolaus        1998    Documentary and descriptive linguistics. *Linguistics* 36: 161-95.

Hopper, Paul, and Elizabeth Traugott 1993    *Grammaticalization*. Cambridge: Cambridge University Press.

Hussainmiya, B. A     1987    Melayu Bahasa: Some preliminary observations on the Malay creole of Sri Lanka. , B. A. Hussainmiya (ed.), In: *Lost Cousins: The Malays of Sri Lanka*, 153-172. Malaysia: Universiti Kebangsaan.

Janhunen, Juha, Marja Peltomaa, Erika Sandman, and Xiawu Dongzhou    2008    *Wutun*. München: Lincom Europa.

Jarva, Vesa     2008     Old Helsinki Slang and language mixing. *Journal of Language Contact* 1:52-80.

Kendrick, Donald     1979     Romani English. *International Journal of the Sociology of Language* 19:111-120.

Labov, William     1969     Contraction, deletion, and inherent variability of the English copula. *Language* 45 (4): 715-762.

———1972     *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

———1994     *Principles of Linguistic Change: Volumes 1 and 2*. Oxford: Blackwell.

Lavandera, Beatriz     1996     Where does the sociolinguistic variable stop? R. Singh (ed.), In: *Towards a Critical Sociolinguistics*,  17-30. Amsterdam: John Benjamins.

Lee, Jennifer   1987     *Tiwi Today: A Study of Language Change in a Contact Situation*. Canberra: Pacific Linguistics.

Lee-Smith, Mei W., and Stephen Wurm     1996     The Wutun language. In: S. Wurm, P. Mühlhäusler and D. Tryon (eds.), *Atlas of Languages of Intercultural Communication in the Pacific, Asia and the Americas Vol. 2*, 883-97. Berlin: Mouton.

Lefebvre, Claire     1998     *Creole Genesis and the Acquisition of Grammar: The Case of Haitian Creole*. Cambridge: Cambridge University Press.

———2004     *Issues in the Study of Pidgin and Creole Languages*. Amsterdam: John Benjamins.

Maher, Julianne     1991     A crosslinguistic study of language contact and language attrition. In: H. Seliger and R. Vago (eds.), *First Language Attrition*,  67-84. Cambridge: Cambridge University Press.

Matras, Yaron 2000    Mixed languages: A functional-communicative approach. *Bilingualism: Language and Cognition* 3 (2): 79-99.

———2003     Mixed languages: Re-examining the structural prototype. In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate: Theoretical and Empirical Advances*, 151-176. Berlin: Mouton de Gruyter.

———2007     The borrowability of structural categories. In: Yaron Matras and Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 31-73. Berlin: Mouton de Gruyter.

———2009     *Language Contact*. Cambridge: Cambridge University Press.

———2010     *Romani in Britain. The Afterlife of a Language*. Edinburgh: Edinburgh University Press.

Matras, Yaron, and Peter Bakker     2003     The study of mixed languages. In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate: Theoretical and Empirical*

*Advances*, 1-20. Berlin: Mouton de Gruyter.

———, eds   2003   *The Mixed Language Debate: Theoretical and Empirical Advances*
Berlin: Mouton de Gruyter.

Matras, Yaron, Hazel Gardner, Charlotte Jones, and Veronica  Schulmann 2007
Angloromani: A different kind of language? *Anthropological Linguistics* 49 (2):142-184.

Matras, Yaron, and Jeanette Sakel   2007   *Grammatical Borrowing in Cross-Linguistic
Perspective*. Berlin: Mouton de Gruyter.

McCann, Mary, Séamas Síochain, and Joseph Ruane (eds.) 1994   *Irish Travellers:
Culture and Ethnicity*. Belfast: Institute of Irish Studies, The Queen's University of
Belfast and the Anthropological Association of Ireland.

McConvell, Patrick   1985   Domains and codeswitching among bilingual Aborigines.
In: Michael Clyne (ed.), *Australia, Meeting Place of Languages*, 95-125. Canberra:
Pacific Linguistics.

———1988   Mix-im-up: Aboriginal codeswitching old and new. In: Monica Heller
(ed.), *Codeswitching: Anthropological and Sociolinguistic Perspectives*, Berlin: Mouton
de Gruyter.

———2002   Mix-im-up speech and emergent mixed languages in Indigenous Australia.
*Texas Linguistic Forum (Proceedings from the 9th Annual Symposium about Language
and Society)* 44 (2): 328-349.

———2008   Mixed Languages as outcomes of code-switching: Recent examples from
Australia and their implications. *Journal of Language Contact* 2:187-212.

McConvell, Patrick, and Felicity Meakins   2005   Gurindji Kriol: A mixed language
emerges from code-switching. *Australian Journal of Linguistics* 25 (1): 9-30.

McWhorter, John    1998   Identifying the creole prototype: Vindicating a typological
claim. *Language* 74: 788-818.

———2005   *Defining Creole*. Oxford: Oxford University Press.

Meakins, Felicity    2007   Case marking in contact: The development and function of
case morphology in Gurindji Kriol, an Australian mixed language. PhD dissertation,
Dept. of Linguistics and Applied Linguistics, University of Melbourne.

———2008a  Unravelling languages: Multilingualism and language contact in
Kalkaringi. In: Jane Simpson and Gillian Wigglesworth (eds.), *Children's Language and
Multilingualism: Indigenous Language Use at Home and School*, 247-264. New York:
Continuum.

———2008b  Land, language and identity: The socio-political origins of Gurindji Kriol.
In: Miriam Meyerhoff and Naomi Nagy (eds.), *Social Lives in Language*, 69-94.
Amsterdam: John Benjamins.

———2009   The case of the shifty ergative marker: A pragmatic shift in the ergative
marker in one Australian mixed language. In: Johanna Barddal and Shobana Chelliah

(eds.), 59-91. *The Role of Semantics and Pragmatics in the Development of Case*, Amsterdam: John Benjamins.

———2010a  The development of asymmetrical serial verb constructions in an Australian mixed language. *Linguistic Typology* 14 (1):1-38.

———2010b  The Importance of understanding language ecologies for revitalisation In: John Hobson and Susan Poetsch (eds.), *Reawakening languages: Theory and practice in the revitalisation of Australia's Indigenous languages*, 225-39. Sydney: Sydney University Press.

———to appear-a.    Which Mix? - Code-switching or a mixed language - Gurindji Kriol. *Journal of Pidgin and Creole Languages*.

———to appear-b.    Borrowing contextual inflection: Evidence from northern Australia. *Morphology*.

———to appear-c    Spaced out: Inter-generational changes in the expression of spatial relations by Gurindji people. *Australian Journal of Linguistics*.

Meakins, Felicity, and Carmel O'Shannessy  2005    Possessing variation: Age and inalienability related variables in the possessive constructions of two Australian mixed languages. *Monash University Linguistics Papers* 4 (2): 43-63.

———2010    Ordering arguments about: Word order and discourse motivations in the development and use of the ergative marker in two Australian mixed languages. *Lingua* 120 (7):1693–1713.

Moravcsik, Edith       1978    Universals of language contact. In: Joseph Greenberg (ed.), *Universals of Human Language: Vol. 1 Method and Theory*, 95-122. Stanford CA: Stanford University Press.

Mous, Maarten       1994    Ma'á or Mbugu. In: Peter Bakker and Maarten Mous (eds.), *Mixed Languages: 15 Case studies in Language Intertwining*, 175-200. Amsterdam: Uitgave IFOTT.

———2000    Selective replacement is extreme lexical reorientation. *Bilingualism: Language and Cognition* 3 (2): 115-116.

———2003a  *The Making of a Mixed Language: The Case of Ma'á/Mbugu*. Amsterdam: John Benjamins.

———2003b  The linguistic properties of lexical manipulation and its relevance or Ma'á. In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate: Theoretical and Empirical Advances*, 209-236. Berlin: Mouton de Gruyter.

Mufwene, Saliko. 1996. The Founder Principle in creole genesis. *Diachronica* 13: 83-134.

———2001    *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.

Muysken, Pieter    1981    Halfway between Quechua and Spanish: The case for relexification. In: A. Highfield and A. Valdman (eds.), *Historicity and Variation in Creole Studies*, 52-78. Ann Arbor: Karoma.

———1994    Callahuaya. In: Peter Bakker and Maarten Mous (eds.), *Mixed Languages: 15 Case studies in Language Intertwining*, 207-211. Amsterdam: Uitgave IFOTT.

———1994    Media Lengua. In: Peter Bakker and Maarten Mous (eds.), *Mixed Languages: 15 Case studies in Language Intertwining*, 201-205. Amsterdam: Uitgave IFOTT.

———1997    Callahuaya. In: Sarah Grey Thomason (ed.), *Contact Languages: A Wider Perspective*, 427-447. Amsterdam: John Benjamins.

———1997    Media Lengua. In: Sarah Grey Thomason (ed.), *Contact Languages: A Wider Perspective*, 365-426. Amsterdam: John Benjamins.

———2000    *Bilingual Speech: A Typology of Code-Mixing*. Cambridge: Cambridge University Press.

Myers-Scotton, Carol 1993    *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: Clarendon Press.

———2002    *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press.

———. 2003. What lies beneath: Split (mixed) languages as contact phenomena. In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate: Theoretical and Empirical Advances*, 73-106. Berlin: Mouton de Gruyter.

Nichols, Johanna    1986    Head-marking and dependent-marking grammar. *Language* 62 (1):56-119.

Nordhoff, Sebastian    2009    *A Grammar of Upcountry Sri Lanka Malay*. Utrecht: LOT Dissertation Series.

O'Shannessy, Carmel  2005    Light Warlpiri: A new language. *Australian Journal of Linguistics* 25 (1): 31-57.

———2006    Language contact and children's bilingual language acquisition: Learning a mixed language and Warlpiri in northern Australia. PhD dissertation, Department of Linguistics, University of Sydney.

———2008    Children's production of their heritage language and a new mixed language. In: Jane Simpson and Gillian Wigglesworth (eds.), *Children's Language and Multilingualism: Indigenous Language Use at Home and School*, 261-282. New York: Continuum.

———2009    Language variation and change in a north Australian Indigenous community. In: D. Preston and J. Stanford (eds.), *Variationist Approaches to Indigenous Minority Languages*, 419-439. Amsterdam: John Benjamins.

———to appear    Competition between word order and case-marking in interpreting grammatical relations: A case study in multilingual acquisition. *Journal of Child Language*.

Papen, Robert A    1987    Can two distinct grammars coexist in a single language? The case of Metif. *Proceedings of the 10th Annual Conference of the APLA*.

———1987    Linguistic variation in the French component of Metif grammar. *Papers of the 18th Algonquian Conference*.

———2003    Michif: One phonology or two. *University of British Columbia Working Papers in Linguistics* 12: 47-58.

Paunonen, Heikki    2006    Vähemmistökielestä varioivaksi valtakieleksi. In: K. Juusela and K. Nisula (eds.), *Helsinki Kieliyhteisönä*, 142-61. Helsinki: Helsingin yliopiston suomen kielen ja kirjallisuuden laitos.

Plag, Ingo    2003    *Phonology and Morphology of Creole Languages*. Tübingen: Max Niemeyer Verlag.

———2003    *Yearbook of Morphology 2002*. Great Britain: Kluwer.

Poplack, Shana    1993    Variation theory and language contact. In: D. Preston (ed.), *American Dialect Research*, 251-283. Amsterdam: John Benjamins.

Rosen, Nicole 2000    Non-stratification in Michif. PhD dissertation, Department of Linguistics, University of Toronto.

Ross, Malcolm    1996    Contact-induced change and the comparative method: Cases from Papua New Guinea. In: Mark Durie and Malcolm Ross (eds.), *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*, 180-217. New York: Oxford University Press.

———2001    Contact-induced change in Oceanic languages in north-west Melanesia. In: A. Y. Aikhenvald and R. M. W. Dixon (eds.), *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, 134-166. Oxford: Oxford University Press.

Sasse, Hans Jurgen    1992    Theory of language death. In: M. Brenzinger (ed.), *Language Death: Factual and Theoretical Explorations with Special References to East Africa*, 7-30. Berlin: Mouton de Gruyter.

Saussure, Ferdinand de    [1916] 1983    *Course in General Linguistics*. Translated by R. Harris. London: Duckworth.

Schaengold, Charlotte. 2003. The Emergence of Bilingual Navajo: English and Navajo Languages in Contact Regardless of Everyone's Best Intentions. In: B. Joseph, J. DeStefano, N. G. Jacobs and I. Lehiste (eds.), *When Languages Collide*, 235-54. Columbus, Ohio: Ohio State University Press.

———2004    Bilingual Navajo: Mixed Codes, Bilingualism and Language Maintenance,

Department of Linguistics, Ohio State University, Columbus.

Sebba, Mark   1998   A congruence approach to the syntax of code-switching. *International Journal of Bilingualism* 2: 1-20.

Slomanson, Peter     2006   Sri Lankan Malay morphosyntax: Lankan or Malay. In: A. Deumert and S. Durrleman (eds.), *Structure and Variation in Language Contact*, 135-158. Amsterdam: John Benjamins.

———2007   On the areal character of Sri Lanka Negation. Paper read at *Society for Pidgin and Creole Linguistics Conference*. Amsterdam.

Smart, H., and B. Crofton     1875   *The Dialect of the English Gypsies*. London: Asher.

Smith, Ian     1977   Sri Lanka Creole Portuguese Phonology. *International Journal of Dravidian Linguistics* 7:248-406.

———1979a  Convergence in South Asia: A creole example. *Lingua* 48:193-222.

———1979b  Substrata versus universalsin the formation of Sri Lanka Portuguese. In *Papers in Pidgin and Creole Linguistics 2*, edited by P. Mühlhäusler. Canberra: Pacific Linguistics.

———1984   The development of morphosyntax in Sri Lanka Creole Portuguese. In: Mark Sebba and Loreto Todd (eds.), *Papers from The Creole Conference, York University*, 291-301. York University: York University.

———2001   Creolization and convergence in morphosyntax: Sri Lanka Portuguese ad Sourashtra nominal marking typology. In: P. Bhaskararao and K. Subbarao (eds.), *The Yearbook of South Asian Languages and Linguistics, 2001: Tokyo Symposium on South Asian Languages : Contact, Convergence and Typology*, 391-409. New Delhi: Sage.

Smith, Ian, and Scott Paauw   2006   Sri Lanka Malay: Creole or convert. In: A. Deumert and S. Durrleman (eds.), *Structure and Variation in Language Contact*, 159-182. Amsterdam: John Benjamins.

Smith, Ian, S. Paauw, and B. A. Hussainmiya       2004   Sri Lankan Malay: The state of the art. In: R. Singh (ed.), *Yearbook of Sourth Asian Languages 2004*, 197-215. Berlin: Mouton de Gruyter.

Smith, Norval 2000   Symbiotic mixed languages: A question of terminology. *Bilingualism: Language and Cognition* 3 (2): 122-123.

Thomason, Sarah Grey       1995   Language mixture: ordinary processes, extraordinary results. In: Carmen Silva-Corvalan (ed.), *Spanish in Four Continents: Studies in Language Contact and Bilingualism*, 15-35. Washington: Georgetown University Press.

———1997   Ma'a (Mbugu). In: Sarah Grey Thomason (ed.), *Contact Languages: A Wider Perspective*, 467-487. Amsterdam: John Benjamins.

———1997    Mednyj Aleut. In: Sarah Grey Thomason (ed.), *Contact Languages: A Wider Perspective,* 449-468. Amsterdam: John Benjamins.

———2001    *Language Contact: An Introduction*. Edinburgh and Washington, DC: Edinburgh University Press and Georgetown University Press.

———2003    Social factors and linguistic processes in the emergence of stable mixed languages. In: Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate: Theoretical and Empirical Advances*, 21-40. Berlin: Mouton de Gruyter.

———, ed.    1997    *Contact Languages: A Wider Perspective*. Amsterdam: John Benjamins.

Thomason, Sarah Grey, and Terrence Kaufman    1988    *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

Traugott, Elizabeth, and Ekkehard König    1991    The semantics-pragmatics of grammaticalisation revisited. In: Elizabeth Traugott and Bernd Heine (eds.), *Approaches to Grammaticalisation*, Amsterdam: John Benjamins.

Treffers-Daller, Jeanine    1994    *Mixing Two Languages: French-Dutch Contact in a Comparative Perspective*. Berlin: Mouton de Gruyter.

van Gijn, Rik  2009   The phonology of mixed languages. *Journal of Pidgin and Creole Languages* 24 (1): 93-119.

Wavehill, Ronnie Jangala    2000    *Nyawuyinangkulu Larrpa Kujilirli Yuwanani Ngumpit-ma Kartiyarlu*. Translated by E. Charola. Katherine: Diwurruwurru-jaru Aboriginal Corporation.

Weiner, Judith, and William Labov   1983    Constraints on the agentless passive. *Journal of Linguistics* 19: 29-58.

Weinreich, Uriel    1974 [1953]    *Languages in Contact: Findings and Problems*. The Hague: Mouton.

Weinreich, Uriel, William Labov, and Marvin Herzog    1968    Empirical foundations for a theory of language change. In: W. Lehmann and Y. Malkeil (eds.), *Directions for Historical Linguistics*, 95-188. Austin, Texas: University of Texas Press.

Whitney, William Dwight    1881    On mixture in language. *TAPA* 12: 5-26.

---

[1] Thanks to Peter Bakker, Yaron Matras and Eva Schultze-Berndt for comments on an earlier draft of this chapter and more generally for conversations in and around the topic of these very strange and fascinating languages.

[2] Also called 'split languages' by Myers-Scotton and 'fused lects' by Auer. For consistency I use the term 'mixed language' which is the most widely-used term.

[3] The Volkswagen Foundation in Germany funds some endangered language documentation projects, see www.mpi.nl/dobes. Two relevant projects are the Sri Lankan Malay project (Umberto Ansaldo, Universiteit van Amsterdam) and the Jaminjungan/Ngumpin project which includes work on Gurindji Kriol (Eva Schultze-Berndt, Patrick McConvell and Felicity Meakins, University of Manchester).

[4] The Endangered Languages Documentation Program (Hans Rausing Foundation, UK, see http://www.hrelp.org/) is currently funding the documentation of Gurindji Kriol (Felicity Meakins, University of Manchester)

[5] The Romani project is based at the University of Manchester (Chief Investigator: Yaron Matras), see http://romani.humanities.manchester.ac.uk.

[6] Angloromani material is held in an archive at the University of Manchester. Sri Lanka Malay and Gurindji Kriol recordings and transcripts are digitally available through the DOBES archive at the Max Planck Institute in Nijmegen (The Netherlands). Light Warlpiri material is also held at the Max Planck Institute in Nijmegen (The Netherlands).

[7] The ACLA project is headed by Gillian Wigglesworth, Jane Simpson and Patrick McConvell at the University of Melbourne, see http://www.linguistics.unimelb.edu.au/research/projects/ACLA/index.html.

[8] Kriol is an English-based creole language spoken across northern Australia. The name 'Gurindji Kriol' is somewhat deceptive as it seems to suggest that Gurindji Kriol is a variety of this creole, however it is a clear and balanced bilingual mix of Gurindji and Kriol.

[9] Rotwelsch is camouflaged German and not an independent language

[10] For a historical overview of the various treatments of Angloromani, see Matras et al (2008).

[11] For a discussion of the origin of the name Gurindji Kriol, see Meakins (to appear-a)

[12] The word *wumpurrarni* means 'black' in the local Aboriginal language Warumungu.

[13] This example comes from Samantha Disbray's corpus of Wumpurrarni English. The speaker is aged in her late 30s.

[14] Smith and Paauw (2006: 160) suggest that the Sri Lanka Malays were bilingual in Tamil and Malay with a knowledge of Sinhala less widespread. Ansaldo (2008) argues for a stronger Sinhala influence though he notes that it is often difficult to separate the Tamil and Sinhala influence because much structural convergence has already taken place between the languages.

[15] There is some variation present in Michif but it is community dependent.

[16] Which is consistent with the pattern of optional ergativity seen in Gurindji Kriol, see §4.2.

[17] Although Ansaldo (2008) observes that these case forms are distinguished in the Kirinda variety of Sri Lanka Malay.

[18] Questions about the legitimacy of the comparison of active and passive sentences given the dubious equivalence of their semantics have been raised in reference to this study and others (see for example Lavandera 1996 )

[19] Where agent is taken to mean the subject of a transitive clause.

[20] Thomason (2003) provides some good counter-arguments to Bakker's objections about a causal link between French-Cree code-switching and Michif.