

The phonology of mixed languages*

Rik van Gijn

Max Planck Institute for Psycholinguistics

Mixed languages are said to be the result of a process of intertwining (e.g. Bakker & Muysken 1995, Bakker 1997), a regular process in which the grammar of one language is combined with the lexicon of another. However, the outcome of this process differs from language pair to language pair. As far as morphosyntax is concerned, people have discussed these different outcomes and the reasons for them extensively, e.g. Bakker 1997 for Michif, Mous 2003 for Ma'a, Muysken 1997a for Media Lengua and 1997b for Callahuaya. The issue of phonology, however, has not generated a large debate. This paper compares the phonological systems of mixed languages Media Lengua, Callahuaya, Mednyj Aleut and Michif. It will be argued that the outcome of the process of intertwining as far as phonology is concerned, is at least partly determined by the extent to which unmixed phonological domains exist.

Keywords: bilingual mixed languages, phonology, stratification, prosodic hierarchy, Media Lengua, Callahuaya, Mednyj Aleut, Michif

1. Introduction

Mixed languages (also called bilingual mixed languages or intertwined languages) combine large, unreduced subcomponents of two different parent languages in such a way, that they cannot be classified as genetically descending from a single parent language.

The central question in many studies concerning mixed languages is how the two (largely) unreduced linguistic systems get incorporated into one language. The general answer is that normally the grammar of one language is combined

* This paper is based on Van Gijn (2000), which was written at the University of Amsterdam, the Netherlands, and the University of Aarhus, Denmark. I am grateful to Peter Bakker, Mily Crevels, Pieter Muysken, Norval Smith and three anonymous reviewers for their comments on earlier drafts of this paper. All remaining errors are my responsibility.

with the lexicon of another through a process called language intertwining (Bakker & Muysken 1995:42). However, different structural characteristics of the languages in question may lead to different instantiations of this general principle. In *Media Lengua* for instance, the split in the language is between root words, which come from Spanish, and grammatical, mostly bound, morphemes, which come from Quechua; consider the following example from Muysken (1997a:366), where Spanish items are in italics:

- (1) *yo-ga* *awa-bi* *kay-mu-ni*
 I-TOP water-LOC fall-CIS-1¹
 ‘I come after falling into the water.’

Michif, spoken in Canada and the USA, displays a different picture. Here, the main split is between noun phrases which come from French, and verbs which come from Cree. The French items in the example, taken from Bakker (1997:78) are in italics:²

- (2) e:gwanIgi *li: sava:z* *ki:pa:šamwak* *la vjād*
 they the Indians dry.by.heat the meat
 ‘They, the Indians, dried the meat.’

According to Bakker (1994:25), the fact that the combination of French and Cree morphemes turned out the way it did is due to the fact that Cree verbs ‘cannot easily be divided into a stem and affixes, since the stem itself is already morphologically complex and the affixes have a gliding scale from lexicalized derivational via more productive derivational to clearly inflectional affixes. The Cree verb should as a whole be taken as part of the grammatical system of the language, and therefore the whole verb and not only the morphological markers are from Cree in Michif’.

Quechua, on the other hand, is an agglutinating language. There is generally a one-to-one correspondence between meaning and form, and morpheme boundaries are clear-cut. There is furthermore a clear distinction between roots and grammatical morphemes, this is true for both the nominal and verbal system. Therefore, replacing Quechua verbal stems by Spanish ones does not involve a heavy infringement on the cohesion of the system as in Michif.

1. TOP = topic, LOC = locative, CIS = cislocative.

2. The transcription is taken from Bakker (1997), which stays relatively close to surface realizations. The voiced plosives in the examples taken from the work of Bakker represent surface realizations of underlyingly voiceless plosives. Cree, and the Cree part of Michif does not have phonemic voiced plosives.

In short, the realization of the intertwining of the two languages is determined to a large extent by the structural characteristics of the source languages. Mixed languages look for the most 'natural' way of combining lexicon and grammar from two different source languages, i.e. in such a way, that the subcomponents do not require much adaptation compared to their parent languages.³

It is generally assumed that the phonology of mixed languages is supplied by the grammar language (cf. Bakker 1997:202). However, phonological systems that are the result of language intertwining may show rather different pictures. Some mixed languages (e.g. *Media Lengua*) indeed mainly follow the phonology of the grammar language, other mixed languages (e.g. *Michif*) seem to have a stratified lexicon, each stratum with its own phonemic inventories, phonotactics and morphophonological rules. This paper is intended as a first exploration into underlying reasons for the different outcome. The focus will be on structural, linguistic factors, but at the end some suggestions for other factors at play will be made.

In the next sections I will briefly discuss the phonological systems of four mixed languages: *Media Lengua* (Section 2.1), *Callahuaya* (Section 2.2), *Mednyj Aleut* (Section 2.3), and *Michif* (Section 2.4). In Section 3 I will review the different languages and try to come to an explanation for the differences between the respective phonological systems in terms of Nespors & Vogel's (1986) universal prosodic hierarchy. I will also consider other factors (sociolinguistics and history) that may play a role. In Section 4, finally I will come to a conclusion and make some suggestions for further research.

2. Case studies

2.1 *Media Lengua*

Media Lengua (lit. 'halfway/middle language') is a mixed language that combines the grammar of Ecuadorian Quechua (*QUECHUAN*) with the lexicon of Ecuadorian Spanish (*INDO-EUROPEAN, ROMANCE*). It is spoken in a few communities in central Ecuador by Indians that live on the slopes of the Andes, in between the valleys, where the Spanish-speaking mestizo culture dominates, and Quechua communities in the mountains.

3. The most natural way of combining structures admittedly is an intuitive notion, but it seems to revolve around unit-meaning correspondence, i.e.: if a unit (noun root, verb root, etc.) functions in a relatively independent way in both parent languages in the sense that it is not highly dependent on or requires grammatical information in order to be interpretable, it can more easily be integrated into a foreign grammatical structure.

The phonological system of Media Lengua seems to be identical to that of Quechua. In the discussion of Media Lengua phonology, I rely on Muysken (1997a) and on field notes provided to me by the same author. In the phoneme inventory, all Quechua characteristics absent in the Spanish inventory have been retained: the aspirated stops, palatal fricatives and alveolar fricatives are unaltered. On the other hand, Spanish phonemes not present in the Quechua inventory are often adapted. Spanish /f/ is mostly pronounced [p^h], phonetically the nearest consonant in Quechua:

- | | | | |
|-----|--------|--------------------------|--------------|
| (3) | phalta | 'to be missing, to lack' | < Sp. faltar |
| | philu | 'sharp' | < Sp. filo |
| | phruta | 'fruit' | < Sp. fruta |

Sometimes, however, [f] is retained:

- | | | | |
|-----|----------|----------------|----------------|
| (4) | fishtexa | 'to celebrate' | < Sp. festejar |
| | fri | 'to fry' | < Sp. freír |
| | flor | 'flower' | < Sp. flor |

There are too few examples to draw any firm conclusions on this topic. It might be the case, for instance, that there is variation between speakers or that unadapted items represent newer loans.

Vowels /e/ and /o/ are generally adapted to /i/ and /u/, respectively:

- | | | | | |
|-----|----|---------|--------------------|----------------|
| (5) | a. | intindi | 'hear, understand' | < Sp. entender |
| | | bindi | 'sell' | < Sp. vender |
| | | dixa | 'leave' | < Sp. dejar |
| | b. | kumu | 'how' | < Sp. como |
| | | kwatru | 'four' | < Sp. cuatro |
| | | bula | 'fly' | < Sp. volar |

There are also a number of instances where /e/ and /o/ are retained:

- | | | | |
|-----|---------|-----------|-----------------|
| (6) | komensa | 'begin' | < Sp. comenzar |
| | konsegi | 'succeed' | < Sp. conseguir |
| | molesta | 'molest' | < Sp. molestar |

Muysken (1997a:381) gives a short account of how non-Quechua vowels /e/ and /o/ are treated in Media Lengua: 'we find that *e* and *o* are often, but not always pronounced as *i* and *u*, respectively (with some variation that also occurs in the Quechua pronunciation of Spanish loans). The Spanish vowels [e] and [o] are often retained in names and interjections. In stressed position [e] and [o] are more frequently retained than in unstressed position'. Another tendency, especially valid for /o/, is that retention occurs more often in closed syllables.

The only sounds of Spanish origin that are consistently unadapted are the voiced plosives /b/, /d/ and /g/,⁴ sounds not present in the original Quechua phonemic inventory:

- | | | | | |
|--------|---------|--------------|---|--------------|
| (7) a. | bibi | ‘drink’ | < | Sp. beber |
| | kambia | ‘(ex)change’ | < | Sp. cambiar |
| b. | aprendi | ‘learn’ | < | Sp. aprender |
| | dos | ‘two’ | < | Sp. dos |
| c. | agacha | ‘stoop’ | < | Sp. agachar |
| | grandi | ‘big’ | < | Sp. grande |

This is not a characteristic specific to Media Lengua, however, but rather a trait inherited from Quechua, where voiced stops in Spanish loanwords in Quechua are also retained. Possibly the presence in Ecuadorian Quechua of earlier loans with voiced stops from now extinct indigenous languages played a role in facilitating the retention of these sounds (P. Muysken, p.c.). Whatever the reason, this fact does not set the phoneme inventory of Media Lengua apart from Quechua.

Syllable structure in Media Lengua generally follows Quechua rules. There are two main points of divergence between Spanish and Quechua syllabic structure. Whereas Spanish allows for onsetless syllables both word-initially and word-internally, Quechua syllables must have an onset word-internally (Lombeida Naranjo 1976:89). The other difference is that Spanish allows for complex onsets, whereas Quechua does not.

With regard to the first difference, adjacent vowels of Spanish origin are mostly adapted to Quechua phonology in Media Lengua:

- | | | | | |
|--------|--------|---|-------------|--------|
| (8) a. | [uji] | < | Sp. [uir] | ‘flee’ |
| | [kaji] | < | Sp. [kaer] | ‘fall’ |
| | [liji] | < | Sp. [leer] | ‘read’ |
| b. | [fri] | < | Sp. [freir] | ‘fry’ |

As can be seen in (8), there are two strategies to avoid adjacent vowels. One is to insert a glide (as in 8a), the other — a marginal one — is to delete one of the vowels (as in 8b).

In a few cases, adjacent vowels /i/ and /a/ are retained:

- | | | | | | |
|-----|--------|---------|---|------------|----------|
| (9) | feria | ‘sell’ | < | Sp. feria | ‘market’ |
| | lluvia | ‘rain’ | < | Sp. lluvia | ‘rain’ |
| | pilia | ‘fight’ | < | Sp. pelear | ‘fight’ |

4. It is not known to me whether these phonemes also have respective allophones [β] [ð] and [ɣ] found in Spanish.

Even though Muysken's transcription does not include a semi-vowel /j/, the sequence /i/ + /a/ is very hard to pronounce phonetically without /j/. Therefore we might argue that these adjacent vowels are reanalyzed as sequences of /ja/, in which case they do not violate Quechua syllable structure.

Consonant clusters in the onset, allowed in Spanish when certain cooccurrence restrictions are met (cf. Harris 1983:21), are retained:

- | | | | | | |
|------|----------|------------|---|--------------|------------|
| (10) | mustra | 'show' | < | Sp. mostrar | 'show' |
| | pregunta | 'question' | < | Sp. pregunta | 'question' |
| | simbra | 'sow' | < | Sp. sembrar | 'sow' |

However, Quechua does not adapt these types of consonant clusters either (P. Muysken, p.c.), so this is not a characteristic specific to Media Lengua.

Clusters at word beginnings consisting of a consonant and a glide are sometimes adapted, sometimes not. Clusters involving /j/ are adapted, clusters involving /w/ are often retained:

- | | | | | | | |
|------|----|--------|--------|---|------------|----------|
| (11) | a. | bwinu | 'good' | < | Sp. bueno | [bweno] |
| | | pwirta | 'door' | < | Sp. puerta | [pwerta] |
| | | kwandu | 'when' | < | Sp. cuándo | [kwando] |
| | b. | kin | 'who' | < | Sp. quién | [kjen] |
| | | silu | 'sky' | < | Sp. cielo | [sjelo] |

An exception to the retention of clusters involving /w/ is when there is an intervocalic sequence of /gw/, which is adapted to /w/:

- | | | | | |
|------|--------|-----------|---|--------------|
| (12) | awa | 'water' | < | Sp. agua |
| | awanta | 'sustain' | < | Sp. aguantar |

However, in some Spanish dialects, this happens too, so that it might not be an adaptation.⁵

As far as higher domains (morphophonology, stress) are concerned, Muysken (1997a), in discussing the phonological system of Media Lengua, mentions two traits specific to Spanish-derived items: voicing of /s/ between vowels, and palatalization of /s/ before /t/. However, these 'changes' are no morphophonemic alternations, but rather an integral part of the phonetic representation of Spanish roots. Moreover, voicing of /s/ in intervocalic position also occurs in the local Spanish (Lipski 1994: 248–249). Stress follows the Quechua pattern (Pieter Muysken, p.c.).

There seems to be no need to analyze the phonological system of Media Lengua as being stratified (i.e. consisting of a Quechua stratum and a Spanish stratum,

5. I am thankful to an anonymous reviewer for this observation.

each associated with their own phonological rules). In Muysken's (2004:160) words: "The Spanish stems inserted in Media Lengua retain some of their features on the lower levels, but not on the higher levels of prosodic structure."

2.2 Callahuaya

Callahuaya (Muysken 1997b) is a mixed language with roughly Quechua (QUECHUAN) structure and a lexicon that is predominantly Puquina (UNCLASSIFIED), but contains material from other source languages as well. The group of the Callahuaya consists of approximately 2,000 itinerant healers. The older members of this group may still be able to speak Callahuaya, but the language is probably not used anymore.

The area northwest of Lake Titicaca, where Callahuaya is or was spoken, used to be Puquina speaking, before Aymara- and later Quechua-speaking invaders colonized the area. It is not clear when exactly Callahuaya emerged. It may have happened at any time from the point where Puquina was still spoken widely to somewhere during the shift from Puquina to Quechua, probably in the 18th or early 19th century (cf. Muysken 1997b:429, 442). It is also not clear which group invented Callahuaya: Quechua- or Puquina-predominant speakers.

Structurally, Callahuaya is similar to Media Lengua, combining Quechua grammatical morphemes with content morphemes from another language, in the case of Callahuaya mainly from Puquina. It differs from Media Lengua in that there is much more deviation from the general Quechua pattern in several areas, such as head-modifier order in some constructions, a non-Quechua nominalizing suffix, the distribution of case markers sometimes differs from Quechua, the extension of the second person subject marker to some third person environments, and nominal possession (cf. Muysken 1997b:432–441).

An example of a Callahuaya sentence is the following, with Quechua elements in italics (from Stark 1972:216, glossing by Muysken 1997b:432):

- (13) *laja-kuna atasi-kuna alkalde-tah isna-n-ku*
 man-PL woman-PL mayor-EMP go-3-PL⁶
 'The men, the women, and the mayor went.'

Stark (1972:200–206) gives an overview of the phonology of Callahuaya and describes how it compares to the respective phonological systems of Puquina and Quechua. Even though Bakker (1997:209) mentions Callahuaya as one of the mixed languages that possibly have a stratified phonological system, Stark's description does not seem to validate this claim.

6. PL = plural, EMP = emphatic.

The consonant inventories of Quechua and Puquina are identical, except for a set of aspirated stops and a set of glottalized stops present in Quechua but not in Puquina. These stops are retained in words of Quechua origin in Callahuaya, and they even occur in words of Puquina origin, an indication that the phoneme system of Callahuaya is treated as a single system. The vowel inventory of Puquina is more elaborate than that of Quechua: in addition to /i/, /u/ and /a/ also present in Quechua, it has mid vowels /e/ and /o/, as well as long counterparts for every vowel. The mid vowels as well as the long vowels are retained in Callahuaya. 'In essence, the system of segmental phonemes in Callahuaya appears to be a merger of Puquina and Quechua' (Stark 1972:201), but Quechua phonology seems to predominate, having affected some words from Puquina origin.

The distribution of sounds differs somewhat in Quechua and Puquina. The latter language allows consonant clusters both at word edges (a stop plus a resonant or lateral word-initially, a nasal plus stop word-finally), as well as word-internally (i.e. CCC clusters). There are furthermore differences in the restrictions on the sounds that are allowed in certain positions. Puquina does not allow /l/, /i/, or the long vowels in word-initial position, nor does it allow long vowels, laterals, and semivowels in word-final position. In intervocalic clusters of two consonants C_1C_2 , C_1 may not be /l/ or /w/, C_2 may not be /w/ or /j/. In Quechua, words cannot begin with /ʃ/, and the sounds /ʃ, m, n, l/ as well as the stops may not occur word-finally. In intervocalic C_1C_2 clusters, stops are not allowed in the position of C_1 . These generalizations, as well as the restrictions in Callahuaya, have been captured in Table 1, where the restrictions stemming from Puquina in Callahuaya are in bold print in the rightmost column.

Stark (1972:205) calls the distribution of the sounds a merger of the Puquina and Quechua distributions, but from Table 1 it appears that Callahuaya follows the distribution of Quechua sounds, the only exceptions being the Puquina constraints involving long vowels, which have been retained in Callahuaya, and the fact that in Quechua, word-initial /ʃ/ is disallowed, which is not the case in Callahuaya. Apparently, preconsonantal plosives and /ʃ/ in words of Puquina origin are not retained; Callahuaya follows Quechua word-final constraints, except for the prohibition of word-final long vowels, which is a characteristic inherited from Puquina. This latter point may be due to the fact that words in Callahuaya generally end in Quechua suffixes, since the grammatical morphemes of Quechua are exclusively suffixing. The fact that the restrictions on long vowels have been retained is due to the fact that Quechua does not have any long vowels, so there are no Quechua restrictions to overrule Puquina rules.

7. Stark also mentions /f/, but this sound is not present in both the Quechua and the Puquina inventory she gives, so it is not clear what the status of this sound is either in Callahuaya.

Table 1. The distribution of sounds in Puquina, Quechua and Callahuaya

	Puquina	Quechua	Callahuaya
word-initial	*/l/	*clusters	*clusters
	*/i/	*/ʃ/	*V:
	*V:		
word-internal consonant clusters	C ₁ : */l/, /w/	C ₁ : *[plos], /ʃ/	C ₁ : *[plos], /ʃ/
	C ₂ : */w/, /j/	*>2 C's	*>2 C's
word-final	*/w/	*[plos]	*[plos]
	*/j/	*/ʃ/	*V:
	*/l/	*/m/	*/ʃ/
	*/ʌ/	*/n/	*/m/
	*V:	*/l/	*/n/
		*clusters	*/l/
		*clusters	

In conclusion we can say that, although at the phonemic level, the sounds of the contributing languages are not adapted, on the higher, more abstract level of phonotactic constraints, Callahuaya largely follows Quechua. Unfortunately, there are no data available for higher phonological domains, such as stress and intonation.

2.3 Mednyj Aleut

Mednyj Aleut (also known as Copper Island Aleut) is spoken by a handful of people on Copper Island, the western-most of the Aleutian Commander Islands. It is a mixture of Russian (INDO-EUROPEAN, SLAVIC) and Attu Aleut (ESKIMO-ALEUT). These languages are intertwined in an unusual way: the lexicon of Mednyj Aleut is a mixture of both Russian and Aleut items. Lexical words come mainly from (Attu) Aleut (especially verbs), grammatical words often come from Russian, as can be seen in the following table, taken from Sekerina (1994:29):

As far as bound morphemes are concerned, derivational morphology for verbs as well as for nouns show an intact Aleut system. Nominal inflection comes from Aleut, but verbal inflection is largely Russian, as is syntax.

This raises the question which language should be seen as the grammar language and which language as the lexicon language. On the one hand, Aleut seems to be the grammar language since it supplies bound derivational morphemes for both verbs and nouns, as well as nominal inflectional forms. On the other hand, Russian seems to be the grammar language, since it supplies verbal inflection, grammatical words and syntactic structure. Looking at the lexicon only, it seems that Aleut is the lexicon language, since it supplies the vast majority of lexical items.

Mednyj is rapidly being replaced by Russian. Thomason (1997a) mentions some differences between the two periods fieldwork was conducted on the language, in

Table 2. Mednyj lexicon (in types)

	total nr	Aleut nr	Aleut %	Russ nr	Russ %
Verbs	180	169	94%	11	6%
Nouns	150	92	61.5%	58	38.5%
Pronouns	18	6	33.5%	12	66.5%
Function words	152	48	31.5%	104	68.5%

the 1960s (Menovščikov 1969) and the 1980s (Golovko 1994, Golovko & Vakhtin 1987, 1990, Vakhtin 1998). For our present purposes, the most important difference between the two periods is represented in Menovščikov's remark, cited in Thomason (1997a:456), that early Russian loanwords (up to the 1940s) in Aleut were nativized into Aleut phonology, while more recent loanwords keep their Russian phonemic and phonotactic structure.

Not much is known about the phonology of Mednyj Aleut, even though Golovko & Vakhtin (1987, 1990) and Sekerina (1994) mention a number of aspects of the phonology of Mednyj, focusing mainly on the Russian phonological traits in the language. The phonology of Mednyj Aleut in many ways seems to be a mixture of the two parent systems of Russian and Attu Aleut, the source dialect of Mednyj. The inventory of Mednyj Aleut consonants contains all the sounds common to both Russian and Attu, and in addition it preserved the aspirated sonorants and uvulars of Attu, not present in Russian. It also contains Russian labial stops and fricatives, both voiceless and voiced (Sekerina 1994:21), not present in Aleut. There are points of convergence as well. Mednyj is losing the distinction between velars and uvulars due to Russian influence (Golovko & Vakhtin 1987:173, 1990:101), but this seems to be a recent development. Furthermore, the Attu Aleut sound /tʰ/ has been lost. According to Bergsland (1994) this is a marked sound also within the Aleut family. In the vowel inventory, Mednyj Aleut has preserved the long vowels of Attu, and in addition it has preserved the mid vowels /o/ and /e/ of Russian, although /o/ is often pronounced [u] (Sekerina 1994:21).⁸

Sekerina (1994:21–22) mentions a number of Russian traits that have been preserved in Mednyj. From the examples she gives, however, it does not become clear whether the characteristics are direct phonetic copies of lexical forms or whether they are incorporated as active rules which only apply to the Russian part of Mednyj. The characteristics and the examples Sekerina gives are the following:⁹

- i. palatalization of consonants (*bol'she* < Ru. *bol'she* 'more')

8. Sekerina also claims that mid vowels /e/ and /o/ have long counterparts. Thomason (1997a:456), however, does not.

9. The examples in Russian are given in their orthographic representations.

- ii. affricates (*bul'niitsaġ* < Ru. *bol'nica* 'hospital')¹⁰
- iii. vowel reduction (*kagda/kada* < Ru. *kogda* 'when')¹¹
- iv. allophone [v] of /g/ (*ivo* < Ru. *ego* 'him')
- v. idiosyncratic pronunciation (*s'as* < coll. Ru. *ščas* 'now')
- vi. voicing assimilation (*safseem* < Ru. *sovsem* 'quite')
- vii. final devoicing of obstruents (*shtop* < Ru. *čtob* 'that')

These examples, however, do not reveal too much about whether Russian phonological rules apply to the (Russian part of the) system of Mednyj. Rather, the examples Sekerina gives suggest that the words have been incorporated into Mednyj in their Russian phonetic realizations. It remains to be seen whether some rules (candidates are i, iii, vi and vii) can be shown to be active rather than fossilized specific realizations.

Syllable structure in Mednyj also seems to be rather unadapted in comparison to the source languages. Russian allows for much more complexity in the onset of a syllable than Aleut. Word-initial clusters in Russian can consist of up to three consonants in a great variety of combinations. The first consonant of a word-initial cluster does not seem to be part of the onset but rather connected to the word level (Yearley 1995:547), with the result that word-initial clusters in Russian do not follow the typologically recognized tendency of rising sonority (e.g. *mgla* 'haze', *kto* 'who'). Attu allows for word-initial clusters of only two consonants, the first of these is always /s/; these clusters of two in Aleut only occur with monosyllabic words. In Mednyj, Russian words that violate Aleut constraints are generally not adapted, e.g. *tsvetki* 'flower', *vtorom* 'two', *vnuk* 'grandson', *zvirazavodam* 'fur farm' (lit. wild animal-factory), *vchira* 'yesterday'.

As far as word-internal consonant clusters are concerned, Attu allows clusters of at most two consonants, whereas Russian allows clusters of three consonants. The few examples that I found of CCC word-internal clusters do not show a consistent pattern.

- (14) *Russian Mednyj meaning*
 /kamenu(t)ka/ /kamenu(t)kaχ/ kind of animal
 /prazdnika/ /prazgnika/ holiday
 /poloten(t)e/ /pulutint(a)χ/ towel

Golovko & Vakhtin (1987:173–174) mention the fact that Mednyj has, in comparison to other Aleut dialects, a unique consonant cluster organisation due to

10. The symbol *c* stands for an affricate.

11. With this is meant that vowels are pronounced differently depending on whether they are stressed or not.

historical processes of metathesis, in which it ‘places velars and uvulars before any other consonant’ (*ibid*, p.173). They mention very briefly that this might be due to Russian influence, but this does not seem very likely, as Russian does not have a cluster organisation that requires velars and uvulars to appear before other consonants.

As mentioned, scholars have focused mainly on tracking down Russian elements in the Mednyj phonological system, and not so much on the preservation of Attu features. Nevertheless, Mednyj has preserved some features of Attu that are on a par with what Sekerina mentions as traits preserved from Russian, for instance distinctive vowel length, pre-aspirated nasals, liquids and glides (including the innovation of a pre-aspirated bilabial glide), the velar-uvular distinction, now being lost, and, as Thomason (1997a:456) mentions: ‘the characteristic Aleut pronunciation of /s/ and /z/ is also quite different from Russian’.

In general, Mednyj gives the impression of having developed from rather Aleut-like to rather Russian-like in three stages. As mentioned, Thomason (1997a:456) quotes Menovščikov (1969) who states that older Russian loans were adapted to Aleut phonology, whereas newer loans (after the 1940s) preserved their Russian pronunciation. Golovko & Vakhtin (1987, 1990) furthermore mention the fact that the velar-uvular distinction is now being lost due to Russian influence, indicating a third stage in which there are hardly any speakers of Aleut left among the Mednyj speakers, and that Russian takes over.

Mednyj Aleut phonology is neither like Russian, nor like Aleut. On a par with its morphosyntactic structure, the language has preserved some characteristics of both languages, at least at the phonemic and syllabic level, and possibly also at the level of morphophonology. Unfortunately, we have no data on stress and intonation.

2.4 Michif

Michif (Bakker 1997) has received much attention in the literature on mixed languages. It is spoken by descendants of European fur traders and native American women in scattered communities mainly in the provinces of Saskatchewan and Manitoba in Canada and in North Dakota and Montana in the United States by the Métis (Bakker 1997:3).¹² Roughly speaking, Michif combines Cree (ALGONQUIAN)

12. There is some controversy on the question of the origin of the Métis. Thomason & Kaufman (1988:228) mention that “the extent to which the Métis people arose from actual intermarriage between Indians and whites is a controversial question; it is likely that different Métis communities show different degrees of admixture.” It is clear, however, that both French and Plains Cree were widely spoken in the area.

verbs (including morphology) with French (INDO-EUROPEAN, ROMANCE) noun phrases and French nominal morphology. Bakker has argued in several publications that this pattern is not a deviation from the general distinction between grammar and lexicon in mixed languages, since Cree verbs often consist of all bound elements (cf. Section 1 above). This analysis makes it possible to maintain the general pattern of language intertwining, where grammatical, bound morphemes of language A (in this case Cree) are combined with lexical items of language B (in this case French).

The phoneme inventory of Michif, as far as consonants are concerned, is largely a combination of the phonemic inventories of the parent languages Plains Cree and Métis French, both components retaining their own inventories. Papen (2005:335) mentions that, unlike Métis French, the French part of Michif has affricates /tʃ/ and /dʒ/, which are the result of a historical process changing /t/ and /d/ into palatalized affricates before closed front vowels or /j/. Furthermore the French component has additional sounds /h/ and /ŋ/ — the latter absent in the Cree part — and misses /ɥ/ compared to the Métis French inventory. Apart from that, the consonant inventories of the two parts are identical to the inventories of their parent languages.

As for the vowels, the oral vowels of the Cree part of Michif and Plains Cree are identical, but unlike Plains Cree, the Cree component of Michif has nasalized vowels /ĩ/ and /ũ/. However, the vowels in the Cree component that can be nasalized differ from the vowels in the French component that can be nasalized, which makes French influence less likely. Furthermore, some Plains Cree dialects *do* have nasalized vowels, also in some of the same words as Michif.¹³ The appearance of the nasalized vowels in the Cree component is possibly due to influence from Salteaux (Papen 2005:336).

Constraints and requirements at the level of syllable structure also seem to be different in the two components of the language. Cree does not allow complex onsets, and only the sounds /h/ and /s/ are allowed in coda position word-internally (cf. Wolfart 1973), whereas French allows much more complexity. In the French part of Michif, clusters of the kind VC₁C₂V, where C₁ ≠ h or s are never adapted to Cree phonology. The same holds for word-initial complex onsets (*glace* > /glæs/ 'icecream'; *bras* > /bra/ 'arm').

Syllable structure is not the only part of Michif phonology that displays split systematicity. A number of facts about the realization of phonemes show that the

13. It is not clear which dialect of Plains Cree the Cree component of Michif is closest to. Only the Western dialect has been described in the literature, and so this is the only dialect Bakker reviews, while he admits that, for geographic reasons, the eastern variant would have been more appropriate for a comparison (Bakker 1997:254).

Cree part and the French part of Michif function as different systems. The following facts are taken from Rhodes (1977, 1986), Bakker (1997) and Papen (2003), with additional comments of an anonymous reviewer.

- i. /p/, /t/ and /k/ are optionally pre-aspirated only in the Cree part, not in the French part. In the French part these sounds are optionally post-aspirated.¹⁴
- ii. The voiceless plosives are optionally voiced after nasals in the Cree part, not in the French part.
- iii. The Cree part follows the allophonic variation of Cree between [u] and [ʊ]. In positions where Cree words have *only* phonetic [ʊ], French words (may) have [u] (examples from Rhodes 1986:294):¹⁵

(15) a. *Cree part*

anoš	[anoʃ]	‘today’
minoš	[minʊʃ]	‘cat’
namo	[namʊ]	‘no, not’
wāpámišo	[wapamiʃʊ]	‘look at yourself’

b. *French part*

bouche	[buʃ] ~ [buʃʰ]	‘mouth’
rouge	[ruʒ]	‘red’
mou	[mu]	‘soft’
chou	[ʃu]	‘cabbage’

- iv. /a/ has different allophones in the Cree and French part of Michif. In the Cree part, /a/ is realized as [æ] or [ɛ] in closed final syllables, elsewhere it is realised as the back, unrounded low vowel [a] or as [ʌ]. In the French part on the other hand, the realisation of /a/ varies between [a] and [æ], the latter pronunciation is only encountered in stressed closed syllables.
- v. Michif has developed a system of vowel assimilation that only occurs in the French part, which is illustrated in (16). In the first example the second vowel is articulated somewhat higher, in the second and third example it is the first vowel that is pronounced higher and in the last two examples there is an adjustment in front or back pronunciation (examples from Rhodes 1986:293):

14. The post-aspiration in the French part occurs particularly with speakers of Michif that speak English and not French in addition to Michif. I am grateful to an anonymous reviewer for this observation.

15. According to an anonymous reviewer, the lax variant [ʊ] in the French part only occurs only in closed syllables with a non-lengthening consonant (/v, z, ʒ, h/) in the coda. In stressed open syllables only the tense variant /u/ can occur. Even though this goes counter some of the observations made by Rhodes (1986:294–295), it does not change the fact that the two variants have different conditions in the Cree and French part of Michif.

(16)	<i>Standard Fr.</i>	<i>Michif</i>	<i>Fr. orthography</i>	<i>English</i>
	/pujø/	/pujy/, /pyjy/	pouilleux	'flea-ridden, dirty'
	/məzyr/	/myzyr/	mesure	'measure'
	/ʃevrøej/	/ʃuvry/	chevreuil	'roe'
	/fyzij/	/fizi/	fusil	'gun'
	/myzik/	/myzyk/	musique	'music'

- vi. There is a difference between Cree and French in the way that adjacent vowels are treated. In Plains Cree, as well as in the Cree part of Michif, there is a productive rule that merges two adjacent vowels if one of them is short, or that inserts a glide /j/ between them if both vowels are long. French has liaison-consonants, which are pronounced between two vowels, as in *lezenfants* vs. *les tables*. There are traces of this liaison-rule, but Bakker (1997:82) considers them no longer productive.¹⁶
- vii. Underlying schwa, pronounced [ɨ] or [ɪ] in Michif (Papen 2003:54), is deleted in the French part only if it follows a single pronounced consonant.

All of the above arguments suggest a stratified phonological system for Michif, which not only incorporates foreign sounds, but also much of the behaviour of these sounds in different environments. Nevertheless, Rosen (2000) argues for a non-stratified account of Michif phonology, claiming that the rules that apply to the different components are not synchronically active. Papen (2003), in an answer to Rosen's (2000) paper and focussing on the French component of the language, shows that there are indeed synchronically active rules that apply to the French component of Michif only (e.g. liaison consonants which Papen considers to be synchronically active, schwa deletion), and that the phonology of Michif should be considered as having different strata depending on the source language for a particular item.

3. Explaining the differences

Bakker & Muysken (1995:50) describe language intertwining as a relatively regular process of combining the grammar of one language with the lexicon of another: 'the way in which intertwined languages are formed is highly uniform.'

16. Bakker draws this conclusion among other things from the fact that some nouns do not have the predicted liaison consonant. However, these occurrences only constitute 10% of all cases (Papen 2003:52), and some of the vowel-initial French nouns do not have the liaison consonant in certain compound constructions. Papen (2003:54) therefore claims that 'liaison is (still) a powerful cue for the child learner to assume that the lexicon is stratified'. A psycholinguistic experiment is needed to give a definite answer to this question. I thank an anonymous reviewer for this comment.

In the view of Bakker and Muysken (1995) as well as Bakker (1997), phonology is part of the grammar and thus supplied by the grammar language. As we have seen, the situation seems to be more complex than this. Mixed languages can incorporate all kinds of elements of the phonology of the lexicon language, from individual phonemes up to rules of allophonic variation (in the case of Michif and possibly Mednyj Aleut). In this section, I will argue for a position that assumes that the phonology of a mixed language prototypically consists of two strata, one from each parent language. This has, in my view, to do with the very type of contact language a mixed language is. I will compare mixed languages briefly with pidgins and creoles in Section 3.1 to argue for this position. Deviations from this general pattern can be explained in different ways. The focus in this section will be on linguistic explanations (cf. Section 3.2), but I will also mention sociolinguistic and historical issues that might be involved in 3.3.

3.1 Contact languages, and the prototypical mixed language phonology

In this paper, we have seen four mixed languages, with different patterns in terms of their phonology. On the one end, we find *Media Lengua* and *Callahuaya*, which have a phonological system that is almost identical to the phonological system of their grammar language, on the other end we find *Michif*, where phonology is stratified, at least up to the level of morphophonology. Finally, we find *Mednyj Aleut*, that seems to be in between these two extremes. The question now is, what is the prototypical state of the phonology of mixed languages? Is it the *Media Lengua* type, as has often been assumed in the literature on mixed languages, or is it rather the *Michif* type? To give a tentative answer to this question, it is useful to compare mixed languages to two other types of contact languages: pidgins and creoles.

Mixed languages, pidgins and creoles have in common that they are not the result of normal, full language transmission, i.e. no single parent language can be identified for these contact languages. Another, related, characteristic that they share is that these languages arise in situations of contact.

In some cases, it may be difficult to assign a particular contact language to one of these types, but the prototypes are well distinguishable. First of all, mixed languages differ from both pidgins and creoles in that they prototypically arise in a two-language contact situation, where there is extensive bilingualism. Pidgins and creoles, on the other hand, prototypically arise in contact situations of more than two languages, without extensive bilingualism (Thomason 1997b:80). A further distinction between pidgins and mixed languages is that the latter are structurally much more stable normally, i.e. there is less interpersonal and intrapersonal variation.

Structurally, mixed languages differ from the other two types of contact language, in that both components (lexicon and grammar), although stemming from a different source, can easily be traced back to their parent languages. The grammar, and sometimes also lexicon of pidgin and creole languages is far less homogeneous (cf. Grant 2001:97, Thomason 1997b:76,78). The lexicon of creoles, and of some pidgins, when there is a socially dominant group, is often dominated by a single lexifier language, but the grammar of these languages is usually not identifiable with one of its source languages. Rather, the structure of pidgins and creoles is a compromise between the different source languages (Thomason 1997b:76–80).

This latter opposition — compromising pidgins and creoles versus uncompromising mixed languages — in my view suggests that the normal state for the phonology of a mixed language would be more like Michif than Media Lengua in the sense that the non-compromising strategy is carried out in the domain of phonology as well. This means that languages like Media Lengua and Callahuaya require additional explaining rather than Michif.

3.2 Linguistic factors: Domains

The explanations alluded to in the previous section can be of different types. In this section, I will consider the linguistic constraints by comparing the two opposites: Media Lengua and Michif. Whereas the phonology of Media Lengua does not seem to differ at all from unmixed Ecuadorian Quechua phonology, the phonology of Michif differs considerably from the phonology of both component languages. This difference might be connected to the way in which the division between grammar and lexicon in the two languages is realized. 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.¹⁷ In other words, Michif has larger unmixed domains (i.e. levels at which certain phonological processes may be active), where French or Cree phonological rules can apply. This observation can be connected to Nespor & Vogel's (1986) widely used universal prosodic hierarchy:

If we compare Media Lengua and Michif on the scale of this hierarchy, the difference becomes clear:¹⁸

17. Mixed forms do exist in Michif, both verbs and nouns. But the number of unmixed forms is probably far greater than in Media Lengua.

18. Q = Quechua, S = Spanish, C = Cree, F = French; the symbol & stands for mixed material in a certain domain (i.e. consisting of material from both parent languages), the symbol ' stands for the existence of unmixed domains consisting of material of either of the two parent languages.

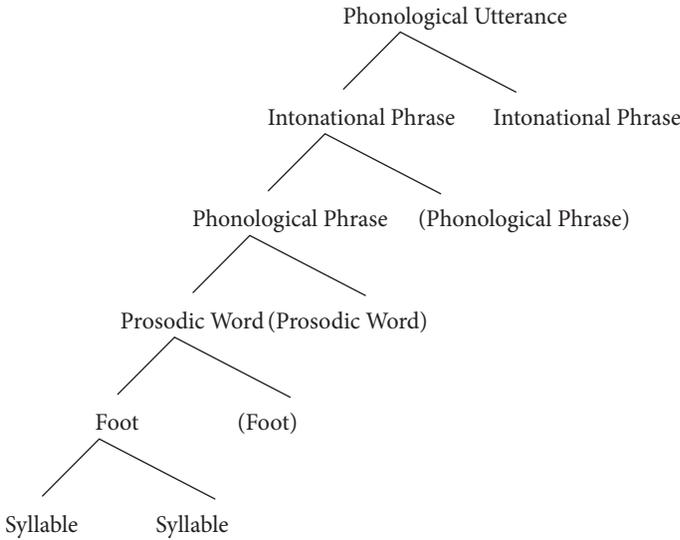


Figure 1. The universal prosodic hierarchy

Table 3. Source material contained in Prosodic domains of Media Lengua and Michif

	Media Lengua	Michif
Phonological Utterance	Q & S	C & F
Intonational Phrase	Q & S	C & F
Phonological Phrase	Q & S	C & F / C, F
Prosodic Word	Q & S	C, F
Syllable ^a	Q & S / Q, S	C, F

^a Nespov and Vogel argue for an intermediate level, the foot, but I have no evidence for either language that the foot level is relevant for phonological rules.

This table should be read as follows: the highest two categories, the phonological utterance and the intonational phrase, are likely to contain material (in the form of morphemes or even entire words) from both languages. This can also be said for the phonological phrase and the prosodic word in Media Lengua (where almost every word contains material from both languages). On these levels, however, there is a difference with Michif. The phonological phrase, which can be described as the head of a constituent and its specifiers, can be either mixed or unmixed in Michif. Michif has French nouns and Cree verbs, but it also has French articles, prepositions, numerals, possessive pronouns and some adjectives. This means that there will be phonological phrases that are purely French. On the other hand, apart from Cree verbs, there are also adverbial particles, personal pronouns and some negation, which means that phonological phrases with a verb as their

head can be purely Cree. There will also be mixed phonological phrases, as demonstratives and postpositions are from Cree, and French is the main supplier for negation as well as some adverbial elements. At the level of the prosodic word, constituents in Michif are mainly unmixed as the split in the language is between verbs and nouns rather than between grammar and lexicon within each word, as in *Media Lengua*. Syllables, finally, will normally be unmixed in Michif, as words are unmixed, in *Media Lengua*, this will also be true for most syllables, but there may be mixed syllables, as is the case for instance when a Spanish verbal root is marked with the Quechua infinitive marker *-y*.

Some of the rules that have been proposed to consider Michif as having two separate phonological systems seem to function at the level of the phoneme (e.g. optional aspiration of stops, allophone [ʊ] of /u/), others function at higher levels such as the syllable (allophones of /a/), the prosodic word (vowel harmony), or even the phonological phrase (adjacent vowels).¹⁹

Nevertheless, the morphophonological operations that are stratified in Michif seem to be restricted to operations that stand in a direct relation to the lexicon. Stress assignment in Michif does not display a stratified system. Though it differs from both French and Cree stress assignment, there seems to be no reason to posit two stress-assigning systems (cf. Rosen 2006). This may be due to the fact that stress assignment in Michif is carried out relatively independent from the lexicon. It seems to be the case that only rules and principles as instantiated in lexical items can be incorporated from the lexicon language. This (i.e. the extent to which principles are tied to the lexicon) may be a further constraining element in the outcome of the phonology of mixed languages which needs more research.

The Creole language Saramaccan provides an example of a language with split stress assignment (cf. Good 2004):²⁰ it combines the stress-accent system of the European lexifier languages English, Portuguese and Dutch, with a tonal system. The two principal African source languages of Saramaccan, Fon (NIGER-CONGO, GBE) and Kikongo (NIGER-CONGO, BANTU) have tone systems. In words of European origin the syllable that bears stress in the origin language in Saramaccan has a lexical high tone. An important observation with respect to the prosodic system of Saramaccan is that the stress-accent system of the lexifier languages is connected directly to the lexicon.²¹

19. It seems to be the case, however, that the liaison consonants are lexicalized and cannot be seen as the result of an active phonological rule synchronically.

20. I am grateful to Norval Smith for drawing my attention to this publication.

21. The loss of vowel length in Saramaccan renders the English stress system completely opaque. In addition the Portuguese vocabulary employs a different stress system (Norval Smith p.c.), it

Mednyj Aleut seems to be somewhere in between *Media Lengua* and *Michif* as far as the mixed nature of the phonological system is concerned. It has incorporated Russian phonemes, possibly including some rules active at higher levels, e.g. voice assimilation and word-final devoicing of stops. On the other hand, it has retained many of the Aleutian phonological characteristics, such as distinctive vowel length, aspirated consonants, a velar-uvular distinction (only now being lost).

There are a few interesting aspects concerning Mednyj Aleut. First of all, the nature of the split in Mednyj is rather peculiar. It is not really possible to tell which of the two languages is the grammar language and which is the lexicon language. Both languages contribute to the grammar as well as the lexicon. A much more detailed analysis of the phonology of Mednyj is necessary to determine the consequences for phonology of this peculiar mixture. For instance, there are probably several words of Aleut origin without Russian material, as the derivational and inflectional morphology of nouns is mainly Aleut, and also a number of words that are unmixed Russian (e.g. inflected verbs without derivational morphemes). It may be that this fact facilitates the preservation of specific phonological rules of the parent languages, as there will be unmixed phonological domains, bearing in mind Nespor & Vogel's prosodic hierarchy.

3.3 The role of sociolinguistics and sociohistory

The focus of this paper has been to search for a structural, linguistic explanation for the different outcomes of language intertwining with respect to phonology. However, other factors might additionally be at play, notably the sociolinguistic situation of the speakers of the mixed language at the time of creation and at later moments.

Sociolinguistic and sociohistorical factors have played a big role in the discussion of contact languages in general and mixed languages in particular. In fact, Thomason (1995:15) claims that "the causes of language mixture are social, not linguistic. Linguistic factors, as far as I can tell, play virtually no role at all." In the same paper, Thomason sets mixed languages apart from other contact languages, pidgin and creole languages, on the basis of differences in sociolinguistic history. She claims that, unlike pidgin and creole languages, mixed languages arise "under conditions of full, or at least extensive, bilingualism" (ibid.:16). It is clear, then, that sociolinguistic factors play a major role in giving rise to, and structuring mixed languages, although the specifics of the interaction between sociolinguistic factors and linguistic structure needs much more research.

is clear that we are dealing with lexically determined placement of tones.

This section is not meant to give an answer to the question what role these other factors play in structuring the phonology of a mixed language, but rather to consider them briefly, and to indicate where possible how they might influence the phonology of mixed languages. Further research is required to get a deeper understanding of these factors in determining the phonology of mixed languages.

Mednyj Aleut shows the importance of the languages spoken by the speaker community, aside of the mixed language. Most scholars (cf. e.g. Golovko 1994, Thomason 1997, but see Vakhtin 1998) claim that the people who created Mednyj Aleut were, to a high degree, bilingual Aleut and Russian speakers, but that Aleut was the language they felt most familiar with. Thomason (1997:450), citing Levin and Potapov (1956), mentions that in the late 1920s, almost all Aleuts that lived on the Commander Islands spoke Aleut (332 out of 345). At the time of Menovščikov's fieldwork in the 1960s this number had dropped to 20 or 30 elderly people speaking Aleut on a population of 300. Golovko (1994:113) reports that at the time of his fieldwork, there were not more than 10 to 12 active speakers of Mednyj Aleut left, but the context in which they live is now almost completely Russian, and they are also fluent in Russian. This means that there has been a complete and rapid change of the situation surrounding Mednyj from predominantly Aleut to almost completely Russian. This turnover is reflected in the phonology of the language. According to Menovščikov (1969, cited in Thomason 1997:450), Russian loans from before the 1940s were adapted to Aleut phonology; after that time they kept their Russian pronunciation, resulting in the preservation of typically Russian features as noted by Sekerina (1994). In the current situation, Russian goes one step further in influencing Mednyj: the Aleut distinction between velars and uvulars is being lost (Golovko & Vakhtin 1987). Taking this into account, it seems reasonable to consider the additional unmixed languages spoken by the speakers of the mixed language to be a highly important factor for determining the outcome of the phonology.

This observation is corroborated by the comparison of Media Lengua with Callahuaya. The structure of the mixture of Media Lengua and Callahuaya is more or less the same: they both combine an intact Quechua grammar with lexical stems from (an)other language(s). Still, there are differences in the phonologies of these mixed languages. At the phoneme level, both languages retain sounds from both parent languages, even though in Media Lengua adaptations of Spanish sounds also takes place. At the higher levels of the syllable and the phonological word, Callahuaya follows the grammar language Quechua more so than Media Lengua. This may have to do with the fact that Callahuaya is what Smith (1995:333) calls a symbiotic mixed language: "this type combines the grammatical structure of one language, and a varying number of lexical items — from hundreds to thousands in number — either from another language (often the original language of the

group), or else from a variety of different sources.” An important characteristic of these languages is that they exist in a symbiotic (hence the term) relationship with an unmixed variety of the grammar language. Callahuaya fits this definition, speakers of Callahuaya are also speakers of Quechua, they take their vocabulary from a number of different sources, but mainly Puquina, which was probably their original language (although we do not know to what extent this was the case at the time of creation of Callahuaya). These languages can be expected to have their phonology dominated by the grammar language, since that is their current mother tongue. Influence from the lexicon language is small, firstly because there is, in the case of Callahuaya, no single lexicon language, but also because borrowings will undergo inference from their unmixed mother tongue Quechua.

Another such language which is often cited in the literature is Ma’a or Inner Mbugu (Mous 2003). Ma’a speakers combine the grammatical structure of unmixed (Normal) Mbugu (BANTU) with a lexicon drawn from a variety of languages, mainly Cushitic. It is generally agreed upon that originally, the speakers of Ma’a were speakers of a Cushitic language (Mous 2003:37–41 argues for Old Kenyan (Eastern) Cushitic), which they currently do not speak anymore. The phonological influence of Cushitic on the phonology of Ma’a is negligible, and restricted to a small number of phonemes, and some lexical tonal patterns (cf. Mous 2003 for details).

The phonological systems of Callahuaya and especially Ma’a in comparison with Mednyj Aleut show that time depth of a certain sociolinguistic situation is of importance. The longer a symbiotic relation with the unmixed form of only one of the parent languages exists, the more influential it is for the phonology of the mixed language, with the effect that the phonology of the mixed language is less likely to be stratified, and more likely to follow the phonological patterns of the parent language still actively spoken.

4. Conclusion and further research

In this paper I have tried to look at linguistic explanations for the differences in the phonology of mixed languages. I have argued that mixed languages tend to incorporate phonological items and principles from both parent languages, but that this is constrained by several factors. The constraint in focus in this paper is structural in nature, and says that a stratified phonology is possible only when there are unmixed domains in a language, since this allows for phonological rules that apply at levels higher than the phoneme or syllable (e.g. the phonological word or the phonological phrase) to take effect.

This would mean that N-V mixed languages (i.e. where the split is between nouns and verbs) are more likely to have separate phonologies for the different components than mixed languages that have a split between roots and bound grammatical morphemes, like *Media Lengua*. Another example (besides Michif and to a lesser extent *Mednyj Aleut*) of such an N-V mixed language is *Gurindji Kriol*, which appears to have a stratified phonological system as well (Felicity Meakins, p.c.).

Nevertheless, as will be clear to the reader, many issues remain for future research. The first type of research needed to see if the model presented in this paper can account for the phonological outcome of language intertwining is, of course, a more thorough description of the phonologies of mixed languages (quality). Information on phoneme inventories is usually well represented, but information on higher prosodic domains such as stress and intonation is scarce. Another issue for further research is, of course, to test the model presented here against more mixed languages.

Secondly, as mentioned in 3.3 above, the role of sociolinguistic and sociohistorical facts needs to be studied more thoroughly. For instance, in many cases, speakers of the mixed language are native speakers of the grammar language and not of the lexicon language (like *Callahuaya*, *Ma'a*). This causes the grammar language to have a big effect on the phonological system, whereas the lexicon language suffers from interference from the unmixed language. This is particularly evident in the case of *Mednyj Aleut*, where there seems to have been a reversal of the dominantly spoken unmixed language from Aleut to Russian, a fact which has, judging from Menovščikov's (1969) remark, a great influence on the phonology of *Mednyj*. Although today *Media Lengua* speakers are shifting towards Spanish,²² at the time of creation, Quechua was their mother tongue, whereas Spanish was their second language. It will be interesting to see whether there is a difference in pronunciation of Spanish-derived items between older and younger speakers, the latter group presumably being more fluent in Spanish. Michif speakers currently do not speak French or Cree anymore, but rather English. According to Bakker (1997:280) the creators of Michif spoke both French and Cree as their mother tongue,²³ which might explain the more equal contribution of both parent languages to the mixed

22. In fact, *Media Lengua* seems to be more stable than Quechua (cf. Muysken 1994:210).

23. Bakker (1997:280) does, however, assume that the first generation Michif speakers were more familiar with Cree than French, but that they can be considered bilingual nonetheless. A further issue is that we do not know anything about the increase of French derived items in Michif over the years of its existence. It might be that speakers of Michif more fluent in French than Cree added French material at different stages after the creation of Michif. I thank an anonymous reviewer for this comment.

variety. Time depth may also be of importance. Speakers of Ma'a and Callahuaya have not spoken Cushitic and Puquina, respectively, for quite some years. Speakers of Mednyj, on the other hand, have only recently given up Aleut, therefore traces of Aleut phonology are more likely to have remained more or less stable, even though currently we see these Aleutian features crumbling down.

A comparison between pidgins, creoles and mixed languages might shed light not only on how these contact languages may be alike or different with respect to phonology, but also on what the prototypical mixed language may look like in terms of its phonology.

Even though all these — and no doubt much more — questions remain, I hope the hypothesis presented here can serve as a starting point for further research.

Submitted: 09/07/06

Revised and resubmitted: 04/23/07

Accepted: 05/22/07

References

- Arends, Jacques, Pieter Muysken & Norval Smith (eds.) (1995) *Pidgins and Creoles: an introduction*. Amsterdam/Philadelphia: John Benjamins
- Bakker, Peter (1994) 'Michif, the Cree-French mixed language of the Métis buffalo hunters in Canada'. In: Bakker & Mous (eds.) (1994), 13–33.
- Bakker, Peter (1997) *A language of our own: the genesis of Michif, the mixed Cree-French language of the Canadian Métis*. Oxford: Oxford University Press.
- Bakker, Peter & Maarten Mous, eds. (1994) *Mixed languages: 15 case studies in language intertwining*. Amsterdam: IFOTT.
- Bakker, Peter & Pieter Muysken (1995) Mixed languages and language intertwining. In: Arends, Muysken & Smith (eds.) (1995), 41–52.
- Bergsland, Knud (1994) *Aleut dictionary/Unangam tunudgusii: an unabridged lexicon of the Aleutian, Pribilof, and Commander Islands Aleut language*. Fairbanks: Alaska Native Language Center, University of Alaska.
- Golovko, Evgenij V. (1994) Mednyj Aleut or Copper Island Aleut: an Aleut-Russian mixed language. In: Bakker & Mous (eds.) (1994), 113–121.
- Golovko, Evgenij V. & Nikolai B. Vakhtin (1987) The convergence of contacting phonological systems: the Aleut dialect of Copper Island. *Proceedings of the eleventh international congress of phonetic sciences 6, Tallin, Estonia*. 172–174.
- Golovko, Evgenij V. & Nikolai B. Vakhtin (1990) Aleut in contact: the CIA enigma. *Acta Linguistica Hafnensia* 22. 97–125.
- Good, Jeff (2004) Tone and accent in Saramaccan: charting a deep split in the phonology of a language. *Lingua* 114. 11–30.
- Grant, Anthony P. (2001) Language intertwining: its depiction in recent literature and its implications for theories of creolisation. *Creolization and contact*, ed. by Norval Smith and Tonjes Veenstra. Amsterdam/Philadelphia: John Benjamins.

- Harris, James (1983) *Syllable structure and stress in Spanish*. Cambridge: MIT Press.
- Levin, Maxim G. & Leonid B. Potapov (eds.) (1956) *Narody Sibiri*. Moscow: Izdatel'stvo Akademii Nauk SSSR.
- Lipski, John M. (1994) *Latin American Spanish*. London: Longman.
- Lombeida-Naranjo, Ernesto Baldomero (1976) *Ecuadorian highland Quechua phonology*. Dissertation University of Texas at Austin. Ann Arbor, Michigan: UMI University Microfilms International.
- Menovščikov, Georgij Aleksejevič (1969) O nekotoryx social'nyx aspektax èvoljucii jazyka. *Voprozy social'noj lingvistiki*. 110–134.
- Mous, Maarten (2003) *The making of Ma'a*. Amsterdam/Philadelphia: John Benjamins.
- Muysken, Pieter (1994) Media Lengua. In: Bakker & Mous (eds.) (1994), 206–211.
- Muysken, Pieter (1997a) Media Lengua. In: Thomason (ed.) (1997), 365–425.
- Muysken, Pieter (1997b) Callahuaya. In: Thomason (ed.) (1997), 427–448.
- Muysken, Pieter (2004) Two linguistic systems in contact: grammar, phonology and lexicon. *The Handbook of bilingualism*, ed. by Tej K. Bhatia and William Ritchie, 147–168. Malden/Oxford: Blackwell.
- Nespor, Marina & Irene Vogel (1986) *Prosodic phonology*. Dordrecht: Foris Publications.
- Papen, Robert A. (2003) Michif: one phonology or two? *Proceedings of the 8th workshop on structure and constituency in languages of the Americas.*, ed. by Y. Chong, C. Gillon and P. Wojdak, 47–58. University of British Columbia working papers in linguistics 12.
- Papen, Robert A. (2005) Le mitchif: langue franco-crie des Plaines. *Le Français en Amérique du Nord: état présent*, ed. by Albert Valdman, Julie Auger & Deborah Piston-Hatlen, 327–347. Québec: Presses de l'université Laval.
- Rhodes, Richard (1977) French Cree — A case of borrowing. In William Cowan (ed.): *Actes du huitième congrès des Algonquinistes*, ed. by William Cowan, 6–25. Ottawa: Carleton University Press.
- Rhodes, Richard (1986) Métchif — a second look. *Actes du dix-septième congrès des Algonquinistes*, ed. by William Cowan, 287–296. Ottawa: Carleton University Press.
- Rosen, Nicole (2000) Non-stratification in Michif. Paper available at http://www.chass.utoronto.ca/~nrosen/michif_stratification.pdf
- Rosen, Nicole (2006) Language contact and Michif stress assignment. *Sprachtypologie und Universalienforschung* 59 (2). 170–190.
- Sekerina, Irina A. (1994) Copper island (Mednyj) Aleut (CIA): a mixed language. *Languages of the World* 8. 14–31.
- Smith, Norval (1995) An annotated list of creoles, pidgins, and mixed languages. In: Arends, Muysken & Smith (eds.) (1995), 331–374.
- Stark, Louisa R. (1972) Machaj-Juyai: secret language of the Callahuayas. *Papers in Andean Linguistics* 1(2). 199–228.
- Thomason, Sarah Grey (1995) Language mixture: ordinary processes, extraordinary results. *Spanish in four continents: Studies in language contact and bilingualism*, ed. by Carmen Silva-Corvalan. Washington: Georgetown University Press, 15–33
- Thomason, Sarah Grey (ed.) (1997) *Contact languages: a wider perspective*. Amsterdam/Philadelphia: John Benjamins.
- Thomason, Sarah Grey (1997a) Mednyj Aleut. In: Thomason (ed.) (1997), 449–468.
- Thomason, Sarah Grey and Terrence Kaufman (1988) *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

- Van Gijn, Rik (2000) The syllable in mixed languages. M.A. thesis University of Amsterdam, ms.
- Vakhtin, Nikolai B. (1998) Copper Island Aleut: a case of language “resurrection”. *Endangered languages: current issues and future prospects*, ed. by Lenore A. Grenoble & Lindsay J. Whaley, 317–327. Cambridge: Cambridge University Press.
- Wolfart, H. Christoph (1973) *Plains Cree: a grammatical study*. Philadelphia: American philological society.
- Yearley, Jennifer (1995) Jer vowels in Russian. *Papers in Optimality Theory*, ed. by Jill N. Beckman, Laura Walsh Dickey & Suzanne Urbanczyk, 533–571. *University of Massachusetts Occasional Papers 18*. Amherst: Graduate Linguistic Student Association.

Appendix: Phoneme inventories

Media Lengua (based on Muysken 1997a), sounds from Spanish which do not occur in the native *Media Lengua* inventory are in bold and italic:

Consonants					Vowels	
p	t	ts	tʃ	k	i	u
p ^h	t ^h		tʃ ^h	k ^h	e	o
b	d	dz		g		a
f	s		ʃ	x		
	z		ʒ			
m	n		ɲ			
	l		ʎ			
	r					
w			j			

Callahuaya, from Stark (1972:200–201). Sounds from Puquina are in bold. The consonant inventory of the local Quechua and Puquina is identical, except for the series of aspirated and glottalized stops, which all come from Quechua. In the vowel inventory, sounds specific to Puquina are bold and italicized:

Consonants					Vowels	
p	t	tʃ	k	q	i iː	u uː
p ^h	t ^h	tʃ ^h	k ^h	q ^h	e eː	o oː
p ^ʔ	t ^ʔ	tʃ ^ʔ	k ^ʔ	q ^ʔ		a aː
	s	ʃ		h		
m	n	ɲ				
	r					
	l	ʎ				
w		j				

Mednyj Aleut, based on Golovko & Vakhtin (1990), Russian sounds that do not occur in Attu Aleut native phonology are in bold and italics.

Consonants						Vowels		
<i>p</i>	t	<i>ts</i>	tʃ	k	q	i i:		u u:
<i>b</i>	<i>d</i>			<i>g</i>		<i>e</i>		<i>o</i>
<i>f</i>	s		ʃ	x	χ		a a:	
v	<i>z</i>		ʒ	ɣ	ʁ			
m	n			ŋ				
m ^h	n ^h							
	l							
w	r		j					
			j ^h					

Michif (Bakker 1997:80):

Consonants										
French part						Cree part				
p	t		tʃ	k		p	t		tʃ	k
b	d		dʒ	g						
f	s	ʃ			h			ʃ		h
v	z	ʒ								
m	n	ɲ		ŋ		m	n			
	r									
	l									
w			j			w			j	

Vowels: oral

French part			Cree part					
i		y		u		i i:		u u:
	ɪ							
ɛ		œ		ɔ		e:		
æ		a		a			a a:	

Vowels: nasal

French part			Cree part					
				ũ		ĩ		ũ
		õ						
æ̃		ã				æ̃		

