

Rromani and Gazikani: Where is the Boundary?

— A typological approach to the lexical intercourse between a minor tongue and the surrounding linguistic milieu

Marcel Courthiade (INALCO & IRU)

ABSTRACT

In languages in contact, there is no rigid boundary between inherited and loan vocabulary but rather a complex continuum which requires more attention than has been paid so far. The boundary between the two can be made on the basis of various criteria and it will differ according to which criteria priority is given: etymological, functional or other. A ten point scale from 0 (lost original vocabulary) to 9 (items occurring only in code-switching) is proposed in order to highlight the degree of mobility of these words between the two languages in contact. Although all languages in contact provide examples of the phenomenon described by this scale, languages in diaspora exhibit the most complex variations in this respect, due to their successive stages of language contact. Similarities and differences between "migrant" languages and their "local" contact languages in this respect are discussed. Examples are based on the Mećkar Rromani dialect of central Albania with further illustrations from three other Rromani dialects of the Balkan, and on the Romanian-based Bunyash dialect spoken by some groups of unascertained origin in central and southern Serbia. The purpose of the article is to exemplify the complexity of this "border" between languages and get rid of the impression it is a clearly cut line.

The fact that many linguists have had to study individual languages as isolated units for the sake of practicality has caused some scholars to treat languages as completely closed systems. Investigations of languages in contact have shown, however, that such an ideal situation is far from achieved in reality. The degree of openness of the system depends on the language considered and on the kind of intercourse it has with the surrounding world, so it is clear that languages in diaspora and *Wandersprachen* ("wandering languages" – a category proposed by Paul Wexler, Tel Aviv University) are of particular interest when investigating this phenomenon. The Rromani language presents one of the most striking examples of a language contact situation, particularly in Europe.

In this paper, I attempt to describe the lack of a strict lexical border between a given Rromani dialect and the surrounding language, namely that besides a number of lexical items which are definitely Rromani and others which are definitely Gazikani, there are many other items constituting a sort of continuum between the languages concerned. Such a situation by no means impairs the identity of any of the languages, just as the identity of black and white colours is not questioned by the existence of a huge number of shades of grey between them. For the purpose of my analysis, I took primarily the Mećkar dialect of Rromani, as spoken in Tirana (Albania — Mećkar belongs to the first stratum of Rromani in terms of dialectology, *i. e.* it is an **O**-dialect without mutation or **O**ḥ), and this for several reasons:

- Although partly Albanianized in vocabulary, it is very much alive in family and social contacts.
- It is in almost everyday contact with three other Rromani dialects: mainly Kabuži, but also Rupano (both of stratum 1 or **O**ḥ) and Čergar (of stratum 2 or **E**ḥ).
- In addition, Mećkar is closely related to Baćor and Fićir from Greece and Erlides from Sofia.
- The Albanian language is itself a subject of particular interest for sociolinguistic studies, since it is currently going through a complete restructuring on the basis of local dialects, within the framework of common modern usage.
- And finally, the background of Turkish influence on both Mećkar and Albanian is actually common to all the Balkan languages, and consequently to a very large number of Rroms.

The present study deals only with the lexical aspect of the problem, but it is clear, as we shall see later, that it is linked with all other levels of the language. To begin with, we shall examine briefly the relationship of Mećkar to the other Rromani dialects and to Albanian, then we shall attempt a classification of the lexical units that link Mećkar and Albanian. We shall show that this type of classification also applies to Albanian with respect to Turkish and that the loan/inherited polarization of the migrant population is in fact superimposed on background of Turkish/Albanian polarization. Four further examples of a similar continuum between languages in contact are given, three of which are Rromani dialects in contact with other Balkan languages, while the fourth one is Bunyash (also called in other areas Boyash, Beás, Rudar or Karavlox), basically Transdanubian (to the South) strongly Serbianized Rumanian and spoken but some bear-leaders of unascertained origin in Serbia – possibly an autochthonous population eventually dislocated by the arrival of Slavic tribes). Finally, an attempt is made to use the results of this investigation in two practical fields: the enrichment of the emerging modern vocabulary and the transcription of texts with a high percentage of borrowings from the contact language(s), a problem with which field workers are familiar.

Relationship of Mećkar to the Other Rromani Dialects of Albania and to Albanian itself

The distinction between Mećkar and Kabuži, Rupano and/or Čergar is quite obvious for native speakers. Although all three of these are closely related dialects, speakers of each of them can quote perfectly in, or imitate speakers of, other dialects. Furthermore, when asked about a word, they hardly ever hesitate before identifying it as common to all dialects or peculiar to just one or two of them, even for the most obsolete words. On the other hand, it is far harder to define the boundary between that which is felt to be Albanian and what is perceived as Mećkar, since the latter borrows constantly from the former, not only in order to express new concepts, but also in apparently unmotivated doublets with inherited words, e.g. **pika**¹ "knife" from Alb. *thikë* "id." beside Rromani **čhurik** (cf. Hindi *čhuri* "id.") or **gèpe** "leaf" from Alb. *gjethe* "id." beside Rromani **patrn** (cf. Hindi *patt-*, *patr-*). As we indicated above, no borrowing from one Rromani dialects to another occurs except in quotations and imitations.

When speaking of *inherited* and *foreign* lexical items, we must bear in mind that the distinction may be held on two different grounds. The first can be called *subjective*. This criterion relies on the speaker's feeling. In this respect we shall distinguish "Rromani elements" from "borrowed elements" as perceived by the speaker, who at any moment may encounter and actually does encounter the relevant items in the contact language (Albanian). We limit this criterion to the feeling of illiterate people and do not take into consideration, e.g., the opinion of Mećkars, who has learnt some literary Turkish with the aid of books and other didactic material². It is important to point out here that the presence of a Rromani ending is not decisive in this respect, as one can see from the following sentence recently recorded: **urìnàva les isi gązikane, kùrse rromane vakàras amen theràva les xoli but** "I hate him [from Alb. stem *urr-* "hate" + Rromani verbal morpheme *~in-*+ person ending *~av*] is Gązikani, while we say in Rromani *I have him gall* (malice, anger) *much* [even if the meaning is not exactly the same]". The second ground of distinction we call *objective*. This criterion is based on a scholarly knowledge of etymology and distinguishes on the one hand the Asian roots (primarily from Indic, Iranian – including Kurd and Ossetic, Armenian, Georgian and other Caucasus languages, as well as early Greek stems borrowed in Asia Minor) and on the other hand European borrowings of

¹ Except if otherwise stated, all Rromani examples are given in integrated script (so-called Warsaw alphabet); when necessary, specific letters are added to render non-Rromani sounds. In order to render the strong or long Albanian [ř] present in Albanian borrowings, we use here conventionally an acute accent upon the consonant ř, since double **rr** has a different value in Mećkar, usually a flap.

² The first version of this article was written in 1986; in the meantime, a lot of Albanian citizens, including numerous Mećkars, have travelled to Turkey and got in touch with spoken Turkish. Many others watch at the Turkish TV programmes through satellite channels.

any epoch (late Greek, Turkish, Slavic etc... to which Albanian has to be added in the case of Mećkar).

This difference is indeed extremely important, since the knowledge of the source language is not always decisive: a word like **pika** is perceived as a "borrowed" word, whereas the numerals from 7 to 9 are perceived as Rromani words, even by Rroms who can speak Greek. This is also the case for a number of Albanian borrowings in Mećkar, e.g., **blëta** "bee", to quote but one, although all Mećkars can speak Albanian. A Mećkar speaker first identifies this word as Rromani and needs a while of reflection before he realizes that the word is "also" Albanian. Thus the native sense of the opposition Rromani/foreign is not exclusively subordinated to the knowledge of the source language³.

Turkish vocabulary presents a more difficult case. Although there are virtually no more Turkish-speaking people in the country, most Albanians do not hesitate to identify most Turkish loan words in their language and neither do the Mećkars, both in Albanian and in their mother tongue. According to the distinction proposed above, all Turkish items are "borrowings" (*objective* level) but some of them are "Rromani" words (*subjective* level) in so far as they are not perceived as Turkish by the speaker. Otherwise they are regular loan words. Actually, the situation is slightly more complex, since one can come across cases, when hardly literate and likewise non-Turkish speaking Rroms can tell that certain words are not "Rromani" but Turkish although they have not entered Albanian and are known only from Rromani. This is an example of oral information acquired along with the word itself, e.g., **tausàni** "rabbit" (regular Rromani **šoşoj**), **ćicëki** "flower" (Rr. **luludi**), **makàzi** "scissors" (Rr. **kat**), **kehàti** "paper, letter" (Rr. **lil**) in Kabuži — cf. Turkish *tavşan*, *çiçek*, *makas*, *kâfüt*. There are also cases when the Rromani form of a Turkish loan word differs from the Albanian one; the former then usually passes for "genuine Rromani" while the latter is ascribed to the "borrowed" vocabulary: Mećkar **Xelëzi** beside Albanian *Herdelezi* (the Feast of Spring in Islamic countries, on May the 6th — note that Kabuži has also **Herdelëzi** [Turkish *hidrellez* or *hidrellez*]).

A strict terminology appears to be a necessary condition for a satisfactory analysis of this puzzling situation, especially a distinction between the level of the speaker's feelings (opposition "Rromani/loan word") and the level of the etymological data (opposition Asian/non-Asian — or simply Asian/European⁴). One should also remember that in individual cases the border of both oppositions may be unclear: depending on the speaker's education or background in the first one and on the reliability of etymological data in the second.

Such a distinction does not make sense at the phonological and syntactic levels: as far as phonemes are concerned, one cannot speak of a speaker's feeling of "alien nature" of such and such phoneme

³ We can mention here an anecdote that illustrates the unpredictability of this feeling. While a Rrom from the Netherlands was visiting his relatives in Tarnów (Southern Poland), he was urged to provide them with **xaćikos muxòvos** "fly-hook" (for fishing — Polish *haczyk muchowy*) and everybody was consternated that he couldn't understand this word. Some even doubted the cousin remembered Rromani at all, although they had previously talked several times with him. The embarrassment continued until one of them realized that it was not a genuine Rromani term, and then the tension relaxed with a general burst of laughter. Adam Bartosz, who told me this anecdote, maintains that it is now a part of the local folk stories and that every so often somebody is asked to tell how it was with **xaćikos muxòvos**.

⁴ One should bear in mind that a significant part of the Greek words, early borrowings like **skamind** "chair", **kris** "judgment", **drom** "road" etc... entered Rromani in Asia Minor, since Greek, Kurd and Armenian were the main languages in use in this area when the Rroms reached it (there was then hardly any speaker of Turkish). The lack of an extra vowel ~o or ~i at the end, characteristic of European borrowings in Rromani, indicates that these words belong to the oldest layer of Greek lexical items which entered Rromani. In this respect, this part of the Greek element, which is shared by all Rromani dialects in Europe, can be regarded as Asian — as well as the Armenian, Iranian and Indic stems. Conversely Turkish lexical items have been borrowed by Rromani on the European soil in the various countries of the Ottoman empire. Interesting enough, there are cases of cross borrowing of such words, then perceived as xenisms, and therefore bearing an "illegitimate" final extra vowel, e. g. **krisi**, **drùmo** in some areas of Serbia.

when pertaining to a different language⁵. The introduction of new phonemes into a language follows usually a long process over several generations and most native speakers are deprived of landmarks to differentiate in synchrony which phoneme is alien and which is not. It is the researcher's task to point out the distinction between those phonemes pertaining exclusively to the contact language and occurring in Rromani only in words borrowed from it, on the one hand, and the "inherited" phonemes, which in principle constitute the sound pattern of the "inherited" lexical units, on the other. (As a result, it is clear that a particular spelling convention is necessary for the transcription of borrowed phonemes in dialectal texts – see above note 1.) The difficulty here is obvious, since one may define the "inherited" vocabulary strictly as that which is inherited from Asian languages (including early Greek loan words) or view this inherited vocabulary as the totality of all the lexemes perceived as genuine "Rromani" words. A third possibility consists of also including all the loan words that occur in the borrowing language. Not only do all three approaches have their bases, but, as we shall see below, there is no strict boundary between them. We can note that some phonemes may be encountered exclusively in the oldest element of the language. This is the case of the aspirated voiceless stops in Rromani⁶.

Actually, all strata of Rromani vocabulary can absorb new phonemic realizations taken from the surrounding languages (e. g. the opposition [w]/[ɹ] in some Rromani dialects of Poland, vowel length in Lovàri from Hungary, [y], [ø] in Germany, Sweden, Hungary etc... [θ] and [ð] in Greece and Albania, etc.). In addition new phonemes occur as a result of the integration of non-inherited oppositions within the inherited lexicon (e.g., the reorganization of the Rromani vowel system in dialects from Romania, based on the vocalic sound pattern of Romanian: *a/ă, e/ë*, to use Romanian orthography)⁷.

As for syntax and idioms, we note an almost complete inability on the part of speakers to distinguish inherited patterns from loan-translations (or *calques*) on the contact language — this leads often to the unnoticed emergence of totally opaque structures in local dialects. Such a discrimination would require a high degree of abstraction from the speaker, along with a broad and fluent knowledge of different languages not belonging to the same family or linguistic league — just to be aware of the contrast. The only way to acquire this type of competence is through formal study, and such knowledge would change the structure of the speaker's perceptions about his/her native language, making the research irrelevant.

Along with the Albanian lexical material, a considerable number of expressions, such as idioms, phrases, or proverbs (the Albanian language is famous for its richness in this respect) are calqued or quoted randomly into Mečkar, often marked by phrases like **sar vakăřena o gaze** "as the Albanians (would) say".

Calques are deserving of a special study, but we shall not treat this topic here for two reasons. First of all, a huge number of the idioms and calques in the Balkans are common to all the languages of the area, so the contrast between Rromani and Albanian or Turkish is minimal in this respect. Secondly, calquing goes beyond the two main *foci* of our study, viz. defining the nature of lexical borrowing in its relation to the enrichment of the modern Rromani language and outlining a theoretical approach to the problem of the spelling in linguistically heterogeneous folk texts.

⁵ This assertion has to be restricted to the situation of a linguistic league (or union of languages, Trubetzkoy's *Sprachbund*) with intensive contact between unrelated or distantly related languages spoken on a specific area. When the languages or dialects are closely related, speakers have an acute sense of the "alien" character of some phonemes, e.g. every Albanian recognizes the Geg dialect in a few second upon hearing the nasal vowels, which do not occur in other dialects.

⁶ The only exception in Mečkar is **ather** "then", from Alb. *atëherë* "id." i.e. *atë + herë* lit. "that time"; the aspirated [th] results from the unpronounced *ë*.

⁷ In other words, new phonemes can also arise in both inherited and borrowed vocabulary due to internal sound changes. Such evolutions have been evidenced in Rromani dialects (Courthiade 1995) as well as in languages of various groups (for example in Kajnas, a Slavic language from eastern Albania; Courthiade 1993).

The foregoing discussion leads to the conclusion that the concept of a strict demarcation between both lexicons is an illusion, as also is the myth of "pure dialect". We actually have a continuum that begins with the Indic material, goes through the alien inherited elements and the integrated contact borrowings (the integration of which is showed by the fact that they follow Mečkar morphological paradigms) and ends at the definitely foreign elements which occur in the replacement of inherited Rromani words and in borrowed phrases and idioms (which in contrast are governed by the rules of the contact language's grammar). None of these categories is closed, and we can observe even in synchronic terms a constant fluctuation among them, not only from one generation to another but also in the utterances of a single speaker.

A Classification of the Lexical Units of Mečkar According to Their Status

Here we shall attempt a classification of the different statuses of these lexicological items ("lexicological" is used here instead of "lexical" in order to allow the introduction of phrases and idioms). The scale proposed here ranges from 0 for inherited items that have been totally lost to 9 for borrowed items that only occur in overt code-switching:

<u>0 — totally lost inherited items</u>	replaced by
* šaj "can, may"	>mun (Alb. <i>mund</i>)
* [š]užo "clean, neat, nice, pretty"	>ipàstər (Alb. <i>i/e pastër</i>), Rromani synonyms ⁸
* sa "everything, all"	> tətər (Alb. <i>të terë</i>)
* čhumut, čhon "moon"	>hànza (Alb. dial. <i>hânëz</i>) ⁹
<u>1 — vanishing inherited words</u>	
čekat "forehead"	>bàli (Alb. <i>ball</i>)
čhurik "knife"	>pìka ¹⁰ (Alb. <i>thikë</i>)
phal "board"	>dfàsa (Alb. <i>dërrasë</i>)
kamlöla pe "is needed"	>dùet (Alb. <i>duhet</i>)

These inherited words can be restored at any time by speakers if required by the situation (for example while speaking with a foreign Rrom), although they do not use them in regular conversation. A separate mention should be made of figures, prices and dates, which are very often said in the contact language while speaking Rromani due to the contact function of this terminology with non-Rroms.

A special attention has to be devoted to the ethnonym **Rrom** which has lost its value in Myzeqe Mečkar dialect, where it is replaced by **Mečkàri** "Rrom". In this dialect, **Rrom** means only "husband" (as it is the case in Sinto from Germany for example). Therefore the sentence **Mečkàri isinan tu ?** has to be understood as "Are you a Rrom (not necessarily a Mečkar) ?", while **lačo rrom isi ov** means "he is a good husband", not "he is a good Rrom" which would sound **lačo Mečkàri si ov**.

2 — well preserved genuine items

- a) **luv** "hut, cabin"
ušarèla "to praise"
gešt, geštani "brother/sister-in-law"¹¹

⁸ Note that the notion of "clean" is basically split in Rromani into two concepts: "pure" (as such, in itself) expressed by [**ž**]užo, (**v**)užo and "cleaned" (as a result of the active human intervention of cleaning) expressed by **thovdo, šulavdo, khoslo** etc... "washed, swept, wiped"

⁹ Jeta Duka says "I do not know why we lost **čhumut**, in fact we put a ~**za** at the end of Albanian *hënë* and it looks Mečkar..."

¹⁰ An attempt of explanation for the substitution of **čhuri** is proposed by Duka, see Duka 2001:184.

¹¹ As a matter of fact, the Albanian loan **kunàtka** is also in use but it conveys a slight mocking connotation which is definitely lacking in the regular Rromani word **geštani**.

čhumuni "thing"
praxni "tinder, amadou"
kuri "blanket, bedcover"

As might be expected, this class is very large (we chose here as examples items which have vanished in most other dialects).

One can add here also Asian non-Indic words:

b) **xer** "donkey" (Persian)
bov "oven" (Armenian)
camcàle "eye-lashes" (Georgian, as brought to light by V. Friedman)
pètalo "horseshoe" (Greek)

and words of unknown origin:

c) **ràvalo** "(big) sack"
bërzik "span (in fact distance between the tip of the thumb and the tip of the forefinger)"

In some cases, the semantic value of an inherited word has changed or is changing in a given dialect:

gad "undershirt (for women in Mečkar)" ≠ **gad** "shirt (for males)" in all other dialects
phivli (currently) "widow" — this word tends to mean "easy woman" among young people, due to the typical Islamo-sexist connection "widow = unwed = abandoned = repudiated due to her immorality".

3 — loan words from a European language no longer in contact with Mečkar

žila "nerve, vein" (Slavic)
pòlica "shelve" (Slavic)
žòlka "tortoise" (Macedonian)
eksina "sixty" (Greek)
glàva "big and ugly face" (Slavic, maybe through Greek)

Apparently, there is in Mečkar no Turkish borrowing which would not exist in Albanian, while such cases are encountered in Kabuži (cf. **taušàni**, **kehàti** etc... quoted before, as well as numerous verbs **sandäm** "I thought", **čakladäm** "I rinsed" etc... further examples below).

4 — loan words from a language in contact but with various changes

a) phonological change (which often arose at the time of early contact, before the given Rromani system had absorbed elements of the new contact languages sound system)

loxòna "women confined in childbed" (Alb. *lehonë*)
dořonèla "to (ex)change" (Alb. *ndërron* "exchange [3 sg.]" The Albanian cluster [nd] usually gives [n] in Mečkar¹²)
kurzunèla "to spare, to save" (Alb. *kursen* "id." — with [z]; but note Mečkar **kursimtàri** "savior", borrowed later with [s])
sadomos "especially" (Alb. *sidomos*)
züşà "female teacher" (obsolete Alb. address *zonjusha* "Miss")
bërcinèla "to shout, to yell" (Alb. *bërtet* "id.")

¹² One main notice that the same Albanian borrowing developed into a quite different form in the Rromani variety spoken in Mitrovica e Kosovaqi, namely into **undruil** to be split up as *u-ndr-u-i-Ø-l* = prosthetic **u-** + Albanian stem **ndr** + Eastern Romance verbal morpheme **-ui-** (a secondary loan from Slavic verbal **-ov/uj-** into Rumanian) + Ø standing for dropped **-sare-** (typical for E-dialects verbal borrowings) + personal ending **-l** — leaving little chance if any to perceive kinship with **dořonèla**. This confirms the general view that most problems in communication are due to local loan words, even when they originate from the same foreign lexeme.

kištinëla "he locks" (Alb. *kyç* "id.")
prenunëla (past **prenosajlo**) "(the sun) sets" (Alb. *perëndon* "id.")

b) morphological (changes in gender, number, grammatical category, etc.)

fürri masc. "oven" (fem. in Alb.)
fisa fem. "clan, tribe" (masc. in Alb.)
bišta fem. "tail" (masc. in Alb.)
fàja fem. "fault" (masc. in Alb.)
ndërǎ fem. "honor" (Alb. *nderi* masc.)
pëlpa fem. "clove (of garlic) with both phonological and morphological changes (Alb. *thelb* m. — [θelp] in Southern pronunciation)
harrono'man "I forgot" (reflexive in Mečkar, not in Alb.)

c) derivational (closely linked to the preceding; addition of suffixes foreign to the source language)

gühetàrka "she-hunter; boyfriend-hunter" (in Alb. dial. *gjyhetar*, standard *gjahtar* "hunter" + Rromani feminine suffix of Slavic origin *~ka*)
kušerica "female cousin" (in Alb. *kushëri* "id." + Sl. *~ica*)
dřasuno "made of boards" (no Alb. counterpart < Alb. *dërrasa* "board" + Rromani inherited suffix *~uno*)

See also below **xoxamžia**, **škròipa**, **prënopa** etc.

d) semantic (consisting primarily of restriction to a particular meaning already existing in the source language; complete semantic changes are rare)

vešinëla "make hazy with steam" (of windows) and fig. "dim with anger" (also in Alb. but first meaning: "dress")
lainëla "pay (a debt)" (also in Alb. but first meaning: "wash")
štàti "shirt" (maybe from Alb. *shtat* "stature, height")
prehinëla "to rest" (in Alb. *prehet* is the medio-passive of *preh* "to sharpen" — a pure Albanism is also widely used in Mečkar: **člōđinëla** "to rest" from Alb. *çlodhet*, while Kabuži has kept the genuine Rromani word **dinglōla**).
šòipa "fainting fit" in the idiom **pelo laqe šòipa** "she fainted [lit. "it fell to her šòipa]" (Alb. *të shuar* "extinguishing [fire, light]")
künëla "be valid" (from Alb. *quhet* "is called, is worth mentioning")
sionëla "emancipate, empower" (from Alb. *sqaron* "explain, clear up")
atonomëvete "independent, self-assured, self-willed" (determinologization of a political term *autonom më vete* "autonomous by him/herself" — also in informal Albanian).

e) enigmatic (although the derivational patterns points out to a very likely Albanian origin, there is no counterpart in current Albanian):

muzgonëla "to smooth out the soil of a hut with cattle dung"
paparime "slightly corrupted (of meat, fish)"
tarasinëla "to diminish (of rain, when it lets you go through without umbrella)"
ngainëla "to plough"
lōda "butterfly"
dūfa "enthusiasm"
tavalinëla "to roll, to drag in the mud" (a simple Albanian origin is excluded for this verb, since it is encountered also outside Albania and Dardania, even in stratum 3).
ngrov "stand up!" (only imperative; Albanian origin unascertained, no Rromani cognate).

Note that all borrowed verbs automatically fall into Class 4c, since they require a derivational suffix. In Mečkar such suffixes are *~in~*, *~on~* and *~un~* while in other dialects the most widespread are

~in~, ~iz~ and ~isar~. On the other hand adverbs and other invariable parts of speech are usually borrowed unchanged. Substantives stand in between, since feminine substantives in ~a as a rule sound identical to the source language item in the definite non-oblique singular (A form), while most masculine add suffix ~i or ~o, as shown in the following table:

	<i>unchanged</i>	<i>changed</i>
invariable lexical units	usually	very seldom
substantives, feminine	all in -a	seldom
substantives, masculine	not in Mečkar	as a rule + ~i or ~o
verbs	not in Mečkar	as a rule + ~in~ etc...

Unassimilated verbs occur in Kabuži, but these are all borrowed from Turkish (e.g. **săzărăm** "I am wringing out", **kajderăm** "I slip", **bekărăm** "I feel dull, anxious", **bezerăm** "I am ashamed"; occasionally such verbs occur both with Turkish and with regular Rromani inflection: **yazărăm**, **yazărsăn** besides **jazëla**, past **jazdas** "write"). This phenomenon has also been observed in a number of Rromani dialects from Greece (Messing 1988:passim). Unassimilated masculine nouns are common in some Arli dialects from Topàna (Skopje), which indicates an increasing importance of class 5 over class 4. The shift of loan words between classes depends therefore not only on formal criteria, such as gender, part of speech etc., but also on the type of relationship of the dialect under consideration with its linguistic environment. Thus the question of whether or not a lexical item takes on Rromani morphological characteristics is not a decisive criterion for determining whether it is assimilated (see also the comments on **urìnàva** above).

Another remark may be made at this point: It is not rare to find the introduction of derivational elements into the borrowing language:

siavžia "miller" = **siav** "mill" + **ži** < Turkish *ci* (a suffix of profession)

xoxamžëška "(female) liar" = Rr. **xoxav~** "to lie" + **ž~** (as above) + **~ëš~** < Alb. *~eshë* < Latin *~issa* (a feminine suffix) + **~ka** < Slavic *~ka* (a diminutive/feminine suffix)¹³

škròipa "handwriting" = Alb. *shkruan* "write" + Rromani abstract suffix **~pa**

prènoqa "sunset" = Alb. *perëndon* "(the sun) sets" + the same suffix.

Such formations may occur within the borrowing language with elements both pertaining to the surrounding language:

durumžia "a patient person" = Alb. *duron* "be patient" + Alb. *~xhi* < Turkish *ci*¹⁴; fem. **durumžëška**.

This phenomenon raises the question of the study of derivational elements as separate items in the classification of loan words and can be exemplified in Mečkar by the suffix **~ka**, from Slavic directly and/or through Albanian. This suffix, *~kë* in Alb. (especially widespread in the Korça dialect), has a double value: diminutive and hypocoristic, as in *sheqerkë* "candy, caramel", from *sheqer* "sugar", *lulkë* "florete (especially on clothes)" from *lule* "flower" etc¹⁵. The result is a feminine noun, but feminization is not the basic function of this element. On the other hand, its Rromani counterpart **~ka**

¹³ **Xoxamžëška** is probably one of the most intermingled lexemes to be found in any language of the world since it consists in four elements, each of them coming from a different language: Indic, Turkish, Latin, Slavic (but see also **undruil**, note 12).

¹⁴ Such formations are common in language contact, e.g. Yiddish *nudnik* "a pest, a nag, an annoyer" from Polish *nud~* "to bore" + Polish suffix *~nik*, although the compound does not occur in Polish (Pol. *nudziarz*), or Alb. *kalemxhi* "a bad writer or journalist" from Turkish *kalem* "pen" + Turkish suffix *~ci*, but not found in Turkish (cf. also the well-known series Russian *парикмахер* "barber, hairdresser" based on the German (or possibly Yiddish) words meaning "wig maker", but not occurring as "barber, hairdresser" in German, which has *Frizör*, which in turn is based on a deverbal noun from the French *friser* "to curl", while the real French term for "barber, hairdresser" is *coiffeur*).

¹⁵ Due to its high rate of hypocoristic expressions, the Korça dialect is characterized among Albanians by the pleasant phrase: *tepsičkë-revaničkë* both diminutives, for *tepsi* "tray used for baking bread and various dishes" and *revani* "Savoy cake".

has only a feminizing function, even in adjectives: **puntôrka** "(female) worker, hard-worker" (as a noun and as an adjective, for a woman), **kursimtârka** "savior" (but rather "sparing, saving", as an adjective), **seriôska** "serious" (only as an adjective).

5 — borrowings integrated to the Rromani lexicon without any change

blêta "bee"

gôta "glass, tumbler"

This class is very abundant and in everyday extension; it presents also many particles of all kinds:

mun(d) "can, may" (but the negative has keep original Rromani **našti**)

meniêher "at once"

prandaj "hence"

šidi "who knows?" (informal Albanian *shidi* < *kush e di?* — cf. also below in Class 8)

Note the adverb **nèpara** "not especially, in a quite limited way", from Alb. *s'para~* with Alb. negation *s'* probably replaced by Slavic negation *ne*, due to unknown reasons (regular Mečkar negation is **na**).

Among these elements, some may keep contact language's flexional characteristics as in the following examples, taken from Mečkar:

Gender of Albanian adjectives: **ipastër** "clean", fem. **epastër**, pl. **tëpastër** (both genders). Such adjectives are not numerous.

Albanian negation *nuk* with the modal *duhet*, while the Rromani negation **na** is impossible with it: *nuk duhet* (never ***na duhet** — which would mean "we need")

Albanian comparative particle **ma** (Geg *mâ*) is encountered with a few Albanian adverbs, e.g., **von** (Alb. *vonë*) "late", **keq** "badly" besides the alternative Rromani particle **po** (of Slavic origin): **ma von** or **po von** "later" **ma keq** or **po keq** "worse" (with a cumulative and emphatic variant **po mâ keq**).

6 — loan words gradually replacing inherited lexical items (=the borrowed element in Class 1)

These borrowings may be easily replaced by the Rromani inherited word when required by the circumstances (e.g., when speaking with a foreign Rrom who cannot speak Albanian or in order to avoid being understood by non-Rroms); they are the counterpart of class 1.

þika "knife", **bàli** "forehead", **dràsa** "board", **kmîsa** "shirt", **dùet** "is needed" etc...

7 — occasional new borrowings (expressing modern life)

aksiòni "mass work"

dogàna "customs"

penalti "penalty (sport)"

analiza "medical test"

škòla "school"

fletòra "notebook"

These words cannot be replaced by a Rromani equivalent, but only, if necessary, by descriptive or periphrastic expressions (usually words with the participial suffix **~ando** or with the possessive postposition **~qo**, fem. **~qi**, pl. **~qe** "belonging, related to"). Such expressions can be considered as a type of slang.

8 — borrowing of a phrase, idiom or expression (written here purposely in Albanian spelling)

ç'ke ti!? "don't worry" (lit. "what do you have!")

ç'më duhet!? "it's not my concern" (lit. "what is needed to me" — usually with Rromani repetition of the pronoun, thus *çmëduhet manqe*)

hajre qoftë "all the best"

kush e di "who knows" (arch. **šidi** cf.5)

për të mirën e shtëpisë "for the sake of the household"

për nder të familjës [sic] "for the honour of the family"

Such expressions are often completely blended with the Rromani text: **cidinom pikën e zezë leça** "I suffered martyrdom with him" (Alb. *pikën e zezë* "martyrdom", lit. "black point"), **lila amen në qafë** "he caused our ruin" (Alb. *na mori në qafë* "idem", lit. "he took us in the neck"). These expressions are never translated into Rromani while others, without any apparent reason, are commonly reworded in Rromani: **perëna but amare menăthe (akala manuša)** or **perëna but amenqe k-i men** "they disturb us a lot (with various demands)" which corresponds to Albanian *na bien shumë në qafë (ata njerëz)* lit. "they fall much to us at the neck".

9 — items which never occur in the Rromani text (except as code switching in longer quotations (of Class 8 type) or in very specialized meanings)

vëlla "brother"; it can be used in the Class 8 type expression *të paça vëlla!* "may I have you for my brother"

shtëpi "house"; only in expressions like *shtëpia e fëmijëve* "orphanage", *shpia e Partisë* "house of the Party" (iron. about a family involved all the time in secret arguments – note dialectal form of *shpia*) or similar

ujë "water"; it is encountered only in **uj-Glīne** "sparkling water" (from the Glina spring in the Gjirokastër district)

bukë "bread" is used only in the compound Albanian neologism *bukë-peshku* "expanded polystyrene" (lit. "fish-bread"), which can be borrowed without changes into Mećkar.

motër "sister". No occurrence of the Albanian word *motër* in Mećkar, but this situation occurs with its counterpart in Yugoslavia, where Serbo-Croatian *sestra* means "sister" and "nurse (in a hospital)"; as a result **sèstra** is the regular Rromani word for "nurse" in Yugoslavia.

We are excluding the following phenomena, which do not belong to the lexical unit level, but to the text level:

— reported speech: conversations originally held in Albanian are usually reported in the same language, frequently interspersed with **phenom** "I said" and **phena** "he/she said" (**phenëla** "he/she says" has a shortened realization [pła], which punctuates long quotations as a speech reporting particle).

— the echo-effect of repeating a phrase in the other language in order to emphasize or to dignify it.

— Albanian phrases inserted in formal speech to give more solemnity (in feasts, weddings etc...).

The continuum from purely Rromani words to strictly non-Romani items is even more gradual than one could conclude from the preceding outline, since even the boundaries between these ten classes are far from being strict, as evidenced in the following review of the limits between them:

0 to 1: extinction in progress (the eldest recognize the word, but are not always sure of the meaning).

1 to 2: obviously a question of generation and degree

2 to 3: when of unascertained origin

3 to 4: the precise route of borrowing may be unclear, e.g., when in Mećkar it is impossible to state whether a particular Turkism was introduced directly during the period of contact with Turkish or through Albanian (or another Balkan language). The sound structure (phonotactism) of the loan does not always provide sufficient data to determine the borrowing route. It is interesting to cite here another example taken from some Rromani dialects from Yugoslavia. While most dialects use the word **drom** "road", borrowed directly from Greek in Byzantine times, some, e.g., Mohazer, Thare-Gone and related dialects from Southern Serbia, have the form **drūmo**, also from Greek but via Serbo-Croatian *drum*, as evidenced by the paroxytonic stress and the final ~o (see above note 4). Another case of hesitation between classes 3 and 4 is when the loans have not been taken from the current contact language, but from a dialect which is no longer in contact, e.g., there is no current evidence to determine whether **keflëska** "woman

fond of good living" is an old loan from the Korça dialect (Alb. *qejlle*, *qefleshë*; Alb. dial. *qefleshkë* "id." — Class 3) or an internal formation (Class 4).

4 to 5: Some phonological changes may appear only during inflection, and they can be considerable, as the following case will exemplify. In the non-oblique singular, the Albanian loan **plàčka** "dress, garments" does not differ from the Albanian original; however, the final ~e of the plural feminine palatalizes the dorsal stop into [k'], which sounds so similar to the preceding č [ʧ] that the sequence is treated as a geminate cluster, which is not allowed in the sound system of this dialect. The cluster is therefore simplified, and the sound of the resulting form is intermediate between [plak'e] and [plaf'e], both noticeably different from Alb. *plaçka* (note that this word is virtually pan-Balkan and that its basic meaning is "loot, haul"¹⁶). A similar plural formation is to be noticed in **màčka** "cat", **rìčka** "Albano-Egyptian woman" etc...

A couple of borrowings appear in two forms, one of which is of the same gender as in the source language, e.g., **spitàli** m. or **spitàla** f. "hospital", **kùlmi** m. or **kùlma** "peak (of a roof)", **kazàni** m. or **kazàna** f. "kettle" (all masculine in Albanian). Sometimes, the different gender is ascribed to another dialect: **fàli** masc. "fortune-telling", from Alb. *fall* masc. "id.", besides **fàla** f. "id", perceived as a Kabuzi word¹⁷.

In addition, we should note that, since most of the sound changes occur during the transfer process, due to phonological adaptation to the sound pattern (phonotacticism) of the borrowing language, it is clear that if the loan word does not contain any sound alien to the latter; changes are not likely to occur. The borrowing will thus be similar to the source language word, even if it dates from a very early period. In such cases, the word in question is so deeply embedded in the borrowing language that it is indistinguishable from the few words which are common merely by coincidence (e.g., Albanian and Rromani **pi~** "drink" to mention only one of the most common examples).

5 to 6: No comment required.

6 to 7: Here again there is a great deal of fluctuation, and the problem may be illustrated by the following example. There is in Rromani a large number of phrases to express the notion "I think", the most widespread of which are **dàva man godi** lit. "I give me brain, I remember", **mi godi X-0e** lit. "my brain at X", **isi ma godi so...** lit. "it is [to] me [I have] brain that", **keràva godi** lit. "I am making brain", **maràva godi** lit. "I am beating brain, I am thinking about" etc., but none of them renders exactly the modern meaning "I think", which is semantically weaker, but also broader than all its Rromani counterparts. This verb appears with this breadth of meanings in many European languages, and as a result loan words from contact languages tend to occur in Rromani not only for the broadened European meanings, but also for the inherited meanings. Such borrowings are **menonèla** in Albania (Alb. *mendon*), **mislinèla** in Yugoslavia (S.-C. *misliti*), **gândil** in Romania, **gondolinel** in Hungary and Slovakia (R. *a gândi*, H. *gondolni*), **duminel**, **dumiskirel**, **dumindel** in USSR (Rus. *думать*) etc. This process might be the rise of a new notion or lead loan words to replace inherited phrases.

7 to 8: Not a few Albanian phrases are assimilated by Meçkars as single lexical units: Alb. *burrë e gru(a)* "husband and wife" appears in Meçkar as a regular adverb **buregru**. A further example is the abovementioned expression *kush e di* "who knows", which behaves as an adverb in both Albanian and Meçkar. The old Albanian expression *Allah na rujt!* "God preserve us" is nothing but an indeclinable adjective in both languages: **to çhavo si alanarujt!** "what a roguish lad your son is!"

8 to 9: it is impossible to foresee if and, a fortiori, when a word will shift to another category such as 7 or even 6.

¹⁶ Note that the French word *butin* went through the same semantic evolution in South-West dialects (Mâcon area).

¹⁷ The source is Turkish *fal* "fortune, fortune-telling" from Arabic *fa'l* "omen" — Friedman, pers. comm.

We should also note that some items bridge two or more non-contiguous categories, as do the following:

alanarujt (see above) belongs to class 5 if we consider it as an adjective in Albanian but otherwise to class 8.

çhurik "knife" can belong to class 1 if it has always been present in Mećkar but the ~**k** ending suggests that it belongs rather simultaneously to class 0 and 5: lost inherited word reborrowed from the Kabuzi dialect, which is in constant contact with Mećkar.

dost(i) "comrade, fellow" might be a Persian loan word (class 2b) when used as oxytonic **dost** but rather an Albanian borrowing of Turkish origin when paroxytonic with suffical ~**i**.

oficërka "a female officer", **aktòrka** "actress", **puntòrka** "(lady) worker" are indeed modern words (class 7) but provided with a Rromani feminine suffix of Slavic origin ~**ka** (class 4). These cases can also be treated as separate elements, thus **oficer~**, **aktor~**, **puntor~** pertaining to class 7 and ~**ka** to class 4d (semantic shift). This will not be discussed here since it refers to a far broader debate and has no practical application to our considerations.

A Similar Pattern: The Relationship of Albanian to Turkish and the Turco-Albanian Continuum

In fact the situation in Mećkar is even more complex than drafted above, since the contact language, i.e. Albanian, stands itself in the position of a borrowing language with respect to Turkish (or more precisely Turco-Arabo-Persian vocabulary), in spite of the lack of current linguistic contacts and of the fact that in modern usage Albanian tends to replace Turkisms with its own words. Actually, the Turkisms are not eliminated, but they acquire an archaic or otherwise marked connotation. This is not the case in Mećkar, which has, broadly speaking, only one level of usage: the colloquial register. Consequently, Turkisms have no marked connotation and are regarded as ordinary lexical units in Mećkar. As far as Albanian is concerned, we can illustrate the 0 to 9 scale of Albanian in respect to Turkish with the following selected examples (situation at the beginning of our century):

0 — Alb. *boj* "to banish", not used today, replaced by *degdis* or *syrgjyros* (but also *dëboj* in Dardania-Kosova)

1 — Alb. *ndishk* "tuberculosis" replaced by Turkish *verem* "id." (but lately revived in medical terminology, in competition with international *tuberkulozë*);

2 — most of the Albanian vocabulary;

3 — (this category is relevant essentially for wandering languages);

4 — Alb. *manushaqe* "violet" for Turkish *menekşe*, *dayak* "stick; beating, hiding" from Turkish *dayak* "thrashing, hiding"¹⁸;

5 — Alb. *haman* "Turkish baths" from Turkish *haman* "id."; most of the Turkish borrowings belong to this class;

6 — Alb. *i zbrazët* "empty", formerly threatened by Turkish *boş* "id."; nowadays the tendency is almost completely reversed;

7 — Alb. *defter* "note-book, register"; *pullë* "stamp", *xhami* "mosque", *ramadan* "Islamic fast", *iftar* and *syfy* "night meals during Ramadan", *adet* "costums, habits", *avdez* "ritual washing" and generally all the Islamic vocabulary;

8 — Alb. colloquial *allishverish* "give and take, haggling"¹⁹, *ezmer gjyzel* "dusky is beautiful" (as a proverb), *tashakojnat* "dirty tricks", esp. in Shkodra [from Turkish *taşak oyun*²⁰];

9 — Turkish *ev* "house", *su* "water", *el* "hand" etc... never occurring in Albanian.

¹⁸ The meaning "stick" is archaic in modern Turkish. The word would have been borrowed when the meaning was still usual. Thus this example shows the preservation of a meaning which has subsequently become obsolete in the source language (Friedman, pers. comm.).

¹⁹ *Alışveriş* is already a lexicalized phrase in Turkish.

²⁰ Turkish: *taşak* "testicle" + *oyun* "game, joke" (Friedman, pers. comm.).

In principle, the Mećkar lexical system could be described as a sort of combination of two 0 to 9 scales (Mećkar/Albanian crossed by Albanian/Turkish). However it can quite properly be treated as a one-dimensional scale. All the same one can see the continuity of the lexical material between two languages in contact, especially in a diglossic and diaspora situation. Only for the sake of simplification may one study a dialect as a closed system.

Similar Examples in Other Dialects

The editor of texts, especially ethnotextes, and interviews is constantly confronted with situations to which the scheme of classification given above can be applied. Four pairs of languages in contact have been chosen to provide concrete examples. (When no other comment is required, just the numbers of the classes are given.)

A) RROMANI/GREEK

The relationship between Greek Rromani (Baćòri dialect, str. 1) and surrounding Greek may be exemplified as follows:

0 — **čhumuni, khanć, vareso** replaced by **tìpotes** "something"

urāl lost and replaced by **petazèla**, past tense **petaksàlo** "to fly" (as Fićiri)

patrn lost and replaced by **filos** pl. **fila** "leaf" (as Fićiri)

iv, jevend lost and replaced respectively by **χjòja** "snow" and **χimònas** "winter" (while Fićiri has resp; **viv** and **ven** cf. common Rromani **iv** and **jevend**); similarly **nilaj**, well attested in Fićiri, is replaced by **kalokèri** in Baćòri – as a result, the genuine Rromani words are at least passively understood by everybody

1 — **tatipe** often replaced by **kàpsa** "heat" (while Fićiri prefers **tatipe**)

kir most commonly replaced by **mirnìngi** [sic] "ant"

plaja replaced by **vuna** "mountains", **gurmuso** by **pondiki** "rat" or **amborome** occasionally replaced by **ksaplosàno** "sick, ill" – actually with a slight shade of meaning, the latter expressing rather the idea of "laid up, infirm" (an active inquiry would be necessary to determine if these items definitely belong to class 0 or 1, since it is not obvious from the interview)

2 — the overwhelming majority of the vocabulary in this dialect (more than 90% in a corpus of everyday speech, some 85% in taped ethnotexts) including very rare items or structures: **jidèla** "to press", **dudèla** "hit" (cf. Mećkar **durdèla** "id."), **ovov** "he", **dèla spili** "to push"

Some unexpected changes may occur in this vocabulary: **nakh** "nose" is feminine in Baćòri, whereas it is masculine in all other Rromani dialects. On the phonological plan, Baćòri **jang** "fire" corresponds to common Rromani **jag** "id.". The following idiom is worth mentioning: **dukhàva tuqe** "I love you".

3 — due to the specific position of Baćòri, at the threshold of Europe, no loan words from previous contact languages may occur

4a — **skèbzava** "I think" (Gr. *σκέφτομαι*)

b — **kòkalos** "bone" (Gr. *κόκαλο*)

drom "road" feminine in Baćòri, whereas it is masculine in Greek (*ο δρόμος*)

c — **zimusarèla** "to knead" (Gr. *ζυμώνω*)

stenevzòl (var. **stinindòl**) "to get narrower" (*στενεύεται*)

jirisàlo "he returned" (*γύρισε*) is adapted from surrounding Greek but it belongs to the common Rromani vocabulary all over Europe; similarly **lulùdi** "flower", masculine in Baćòri, is feminine in all other Rromani dialects

mèxum (beside **mèxri**, closer to Greek) "until" (cf. Greek *μέχρι*)

d — **agòra** is often used for "town, city" whereas the basic Greek meaning is "market"; however this confusion is encountered in rural Greek and various other Balkan languages **pazàri, çarsìa, diz** etc... (note that this lexem migrated to Central Europe as **ahor** *fem.* – meaning "value")

sùsta "cart, wagon" (cf. Greek *σούστα* "spring (of steel), press-button")

e) **međixo** "without" standing for **bi ...~qo**

5 — **mòno** "only"

pàdote, pàndote "always"

pràsinos "green"

ekaton eksinda "one hundred and sixty"

jatro pl. **jatri** "physician" (more rarely **hekimi**)

pràma "thing"

potìri "glass (for drinking)" etc...

6 — see 0 and 1

also: **mirisàjlo** "he smelled" (esp. traditional) instead of **khandilàs**

saniða "board" instead of **phal**

Note that in all these cases, the borrowing is motivated by slight semantic divergence (often figurative use) between the inherited Rromani word and the loan.

7 — **sinavlia** "concert"

film "movies"

kasèta tape"

zaxàri "sugar"

iðisis "TV news"

8 — **mufènete** "apparently" *μου φαίνεται*

ðemules "by the way, tell me" *δε μου λες*

aman aman "(have) mercy" (Turkish)

9 — all words such as those meaning "happy", "work", "sing" etc...

B) RROMANI/MACEDONIAN

The examples are taken from an interview with the famous Rromani singer Esmà Redžepova "Queen of the Rromani Music", from Skopje, published in *Etudes Tsiganes* 1988/3.

0 — **čhumuni, khanć, vareso** replaced by **nèšto** "something"

lošavel lost and replaced by **raduinel** "to rejoice" (for)

živèl lost and replaced by **živinel** "to live"

sarkon, sarkoj lost and replaced by **svàko** "every, each" (Serbo-Croatian)

1 — **kana** often replaced by **kad** "when [as a conjunction]" (Serbo-Croatian; cognate forms **ked, kèda** "when" also exist in the area)

pinzarel replaced by **upoznainel** "to make someone's acquaintance" – actually with a slight semantic difference, which explains the borrowing (an active inquiry would be necessary to determine if these items definitely belong to class 0 or 1, since it is not obvious from the interview)

2 — most of the vocabulary of the interview (appr. 80% in frequency)

3 — **pàna** "still" (Greek *πάντα*; Rr. **vàze** in her dialect – but **pànda** is common in most Balkan dialects of str. 1); however the loan may have occurred in Asia Minor and not be viewed as "European" word

4a — **poćminel** "to begin"

b — **valànel** "must" (var. **valàni** also exists)

c — no occurrence in the interview

d — no occurrence in the interview

e — no occurrence in the interview

Class 4 represents appr. 6% of the text of the interview

5 — **sàmo** "only"

život "life"

izglèda "seemingly"

grćke "Greek" (**grćke gilă** "Greek songs")

Class 5 represents appr. 8% of the text

6 — see 0 and 1

also: **nòsnia** "dress, costume" (esp. traditional) instead of **gada, šeja**
istùrito, pràvo "true, genuine" instead of **ćáčo, ćácutno** "true, genuine"

bòrba "combat, struggle" instead of **maripe** "fight"

Note that in all these cases, the borrowing is motivated by slight semantic divergence (often figurative use) between the inherited Rromani word and the loan.

Class 6 represents appr. 3% of the text.

7 — **mùzika** "music"

kultùra "culture"

knjìga "book"

zàsluga "merit, desert"

tòno "record"

Class 7 represents appr. 3% of the text of the interview

8 — *i za rados i za taga* "and for joy and for sadness" (Macedonian)

jedan put za uvek "once for ever" (Serbo-Croatian)

najpoveće bratstvo "greatest brotherhood" (Macedonian)

za jedno dobro na sledeće generacije "for the good of following generations" (Mac. dial.)

9 — all words such as those meaning "happy", "work", "sing" etc...

It is worth noting that Esma's Macedonian is strongly influenced by Serbo-Croatian, which results a matrix of the type Rromani\Macedonian crossed by Macedonian\Serbo-Croatian.

C) RROMANI/SERBO-CROATIAN

The examples are taken from the poetry of Rasim Sejdić in a forthcoming book dedicated to the life of this Rromani poet from Bosnia, who died a few years ago in Italy.

0 — **sung** "fragrance" unknown, replaced by **mirìso** (S.-C. *miris* "id.")

1 — **kamel** "to love" most often replaced by **voli(sarel)** "id." (S.-C. *voleti* "id.")

2 — most of Sejdić's vocabulary, but less than in Esma's interview, in spite of the archaic character of poetic texts.

3 — **kuvèti** "strength, power" (Turkish; actually not unknown in Bosnian dialects of Serbo-Croatian, but quite archaic and missing even in Vuk's dictionary)

4a — **kavćin** "drop" (S.-C. *kapa* "id.")

brìsto "limpid" (S.-C. *bistar* "id.")

b — **svàko** "every" (unchanged also with feminine: **svàko mrva** for S.-C. *svaka mrva* "every crumb, smallest particle")

jedino "sole, unique" (same remark as above: **jedino ljubav** for S.-C. *jedina ljubav* "unique, sole love")

prtljaga "luggage" (*fem.* in Sejdić's Rromani but *masc.* in S.-C. — *fem.* in Slovenian)

d — **oblàko** "sky, heaven" from S.-C. *oblak* "cloud" (actually **oblàko** belongs to an earlier stratum than the other Slavic loans; however, in terms of synchrony its place is in Class 4d)

5 — **zagrljaj** "embrace" (Rrom. **angali** from Greek)

pleténica "plait, twist" (Rrom. **ćurni**)

6 — **istina** in competition with **ćácipe** "truth"

7 — **horizònto** "horizon, sky-line"

8 — [**vazdel**] *ćáša nazdravi* "[raises] a glass for a toast"

D) ROMANIAN/SERBO-CROATIAN IN BUNYASH

Bunyash is spoken by several thousands of persons of unascertained origin (see above) and their native Romanian dialect has been for centuries in contact with surrounding Serbo-Croatian – at last in post-Yugoslavian countries. Most of them are coppersmiths or bear-leaders; they are called Karavlox in Bosnia. The examples below were gathered while shooting a documentary film in Autumn '87 in

Kruševac and can be found in the interviews taped for the film (Bunyash quoted in Romanian spelling):

0 — **împreuna** replaced by *zaiedno* "together"

1 — **trebuie** replaced among the young by *mora* "must"

2 — most words of everyday life, e.g., **ursu** "bear", **lucru** "work", **am** "harness"

3 — **bâlgi** "fair, market" (also in Rom. *bâlgi* from Hung.)

4a — **boaliță** "hospital" (S.-C. *bolnica*)

b — **jivesc** "to live"

se isplatește "it is worth"

lec "medicine", pl. **lecuri** (Rum. *leac*)

d — **țiganu** "husband" (S.-C. **ciganin** "Gypsy"; however, this borrowing with semantic shift is probably a result of the assimilation of Romanian *țigan* with both meanings of Rromani **rrom** "1. Gypsy. 2. husband" and should be ascribed to the earlier stage of Rromani/Romanian interferences South of the Danube; note that the Romania dialect spoken by Beás in Southern Hungary uses **țigan** in the same function: **incá n-ám țigán** lit. "I have not yet a husband" (note the pronunciation of *i/î* in both cases : [i] under Serbian influence, [î] under rural Rumanian influence).

țigancă "wife" (same remark as above; also in Hungarian Beás: **incá n-ám țigancă** "I have not yet a wife"; **am lot-o de țigancă** "I married her".)

Note that **voinic**, meaning "son" [but S.-C. *vojnik* "soldier"], does not belong to this class, since it already has the same meaning ("son") in Romanian.

5 — **geubre** "rubbish" (S.-C. *đubre* < Turkish *gübre* < Persian *gobar* "dung")

iednostavno "simply"

6 — **brigă** replacing **grija** "care, anxiety" (itself of Slavic origin: S.-C. *griža* "diarrhea" < *gristi* "to bite")

mora see above in Class 1

7 — **zadruǵă** "cooperative"

staniță "station"

8 — **osam razred de școli** "eight classes of schools" (S.-C. *osam razreda*)

nepismen ceovec "illiterate"²¹

It is interesting to note that the proportion of items belonging to the different classes differs from one *parler* (linguistic variety) to another. This could be used as an index to express the degree of creolisation, the permeability, the vividness and many other features of a given *parler*: in some of them most borrowings enter the vocabulary through Class 4 (*parlers* in sudden contact, from which pidgins usually develop), while in some others they enter rather through Class 5 (like Sejdić's dialect). Also the weight of classes 1 and 6 is important for estimating the rate of change.

An Attempt to use the Above Classification as a Model in the Process of Enrichment of the Emerging Modern Rromani Common Language

As has already been indicated, the scheme given above fits all Rromani dialects and most minority languages in natural field conditions. However, Rromani is undergoing at present significant changes due to the efforts of the Rromani intelligentsia. A similar pattern can be used to investigate the structure of the emerging common Rromani language, the vocabulary of which is being expanded by means of several devices: recirculating forgotten, dialectal or marginal words, borrowing or adapting words from Indian languages or from the international modern vocabulary, creating neologisms, etc.

²¹ The echo-effect referred to above can be exemplified in Bunyash with the following sentence of an interview: **cu ursu nu se zaradește ništa, jedino nimic** "with the bear, you don't earn anything, simply NOTHING" (**ništa** "nothing" is S.-C. *ništa*, while **nimic** "id." is Romanian).

In order to classify these new items, it could be convenient to establish a further scale from 10 to N, which would be more or less similar to the first scale, but developing further, as one can see from the following rough outline (note that it is impossible to construct such a scale abstractly for the standard language without reference to a concrete dialect; thus the sample given below refers to the standard language as developed since 1990, with respect to the Mećkar dialect):

10 — words lost but borrowed from another dialect that kept them in use

čhurik "knife" (actually re-introduced by popular channel in Mećkar) also in most dialects
phanri "silk"
dud "light"
heroj "thigh"
mesto "free"
zervo "left (side)" (from Sinto, a word of Greek origin)

10' — derivations from dialectal Rromani stems

dudărel "to light, to enlight" < **dud** "light" (in fact, **dudărel** exists in genuine dialects outside Albania)
mestin, mestipen "freedom" < **mesto** "free"
akharin "appeal" < **akharel** "to call, to invite"
baxtalin "a song broadcasted by a local radio to greet a happy event in a family" < **baxtalo** "happy" (loan translation of Macedonian *čecmumka*)
rangvalo or **butrangvalo** "multicolour, colourful"

11 — vanishing words revived with a new (more precise) meaning

drabarel "to use magic" and, in some dialects "to tell fortunes"²², now "to read"
mastin "colour of a horse", now "colour" (the fact that the word is encountered in Prizren (Kosovia) seems to be a convincing argument against the hypothesis of a Russian origin for it.); in fact **rang** is more commonly used for "colour"
them "country", now usually "State": **Thema-Somdasne** "Member-States"
lupunz "hobble, fetter", now often "brake"
thami "law"
berxum "acorn" (from Spanish Caló)

revived doublets with difference of meaning

halărel "to understand"	vs.	xatărel "to feel"
užo "pure"	vs.	šužo "pretty"
zakărel "wait"	vs.	azukarel "hope" (beside elpidel "id.")

11' — derivation from common Rromani stems

ćirlatuno "ancient" from **ćirla** "a long time ago (in stratum 1)" (also **ćirlatunipen** "antiquity"), after **dumutano** from **dumut** "id. (in strata 2 and 3)" and **xargatuno** from **xarga** "id. (in stratum 1, Baltic)".
agòrdel "to finish" < **agor** "end" + **del** "to give" (**agordisarel** is also used, esp. in Yugoslavia)
lovikano "financial" < **love** "money" + suffix ~**ikano**
lovitro "treasurer" (same etymology)
godisarel or **godărel** "to think" < **godi** "brain, thought" (and after **găndisarel** "id." from Romanian *a gândi* "id.")

²² As a matter of fact, **dorăkerel** is the genuine Rromani verb for "to tell fortune"; Finnish Rromani distinguishes **drabarel** "to tell fortune" and **drabavel** "to read"

xasin "loss" < **xasarel** "to loose" (in fact the word **xasin** does exist in the Rromani dialect of northern Russia); note **xasain** "desert" in Gina Ranjičić's poems (19th century)

dikham or **dikhan** "point of view" (also **dikhipnasqo viram**)

We can give for example a series of neologisms coined on the basis of **som** "along (with)":

somthan "condition" < **som** + **than** "place"

somogi "agreement, accord" from **som** + **ogi** "spirit, soul" (**and-o somogi e thamăça** "in accordance with the law")

somgodi "consciousness, awareness" < **som** + **godi** "brain"

somvakărno "interlocutor" from **som** + **vakărel** "to speak" + suffix of agent **~no**

sombešipen "session, council" from **som** + **bešel** "to sit" + abstract suffix **~pen** etc... Also short form **sombeš** for institutions: **komunaqoombeš** "town council", **Evropaqo Sombeš** "Council of Europe"

Note the following, from a postpositional form:

Kupatni "city hall" < **kupaθe** "together" + empty feminising suffix **~ni** (coined in Macedonia at the end of the 20th century)

raipen "government" < **raj** "gentleman"; **raipnikano** (var. **raipno**) "governmental" and **na-raipnikano** "non-governmental"

Also back formations

gind "number" deverbative from **ginel** "count", **gindo** "counted".

var "time, turn" from **duvar** "twice"

čhand "way, manner" from **averchane** "otherwise".

12 — loan words from other Indic languages

lekhel or **lekhavel** "to write"

pusti(k) "book"

viram "point"

vizaj "victorious"

jažur "sacrifice"

Most of them underwent phonological or morphological changes; **pusti(k)** "book", which appeared first in Skopje, seems to have been introduced as **pustak** fem. [cf. Hindi *pustak*], but it was soon reinterpreted as a fem. in final **~i** (or **~ik** in some dialects, like **kusti(k)** "belt", **kangli(k)** "comb" etc...)

12' — derivation from Indic stems

lehipen "writing, spelling"

lekhavel "to write (a spontaneous Gurbet variant of **lekhel** "id.")

vizavel "to conquer" also **vizavno** "conqueror", **vizaipen** "victory"

viramlin "office, branch, outlet"

jažurno "sacrificator"

13 — bridge words

a) Arabo-Persian vocabulary occurring in both Indic and Balkan (through Turkish) languages are bridge-words (Class 13) for the varieties of Rromani spoken in the Balkan

ževap "answer" (in fact Rromani **anglidipen** is more common)

dünia "world" (in fact genuine Rromani is **sundal**, beside Romanian borrowing **lùmia**)

barabar "equally"

tasvir "picture" (already widely used in Yugoslavia)

hakaj "right" (Rrom. dialects **hak**, **hako**, **xakaj**)

manaj "reason, cause"

majmùni "ape, monkey" (widely known in Balkan languages; also *maimuță* in Romanian and *majom* in Hungarian; possibly from Persian through Arabic and Turkish). Actually these are Class 12 for the varieties of Rromani spoken outside the Balkan²³.

b) English

šťartorro "note-book" (maybe from English "start", but maybe as well a genuine Rromani word as a translation of South Slavic *tetradka* from Greek with *τετρα*~ reinterpreted as **šťar** "four")

duti, dutĩ "duty"

puni "punishment" (in fact from Spanish Caló, readapted after English stem *puni*~)

abstràkta "summary"

13' — derivation from bridge words" stems

biharabaripen "inequality"

bihakaipen "total lack of rights"

manaipen "meaning, significance"

startinel "to start" (**širdel** "to begin" is more common, but some still use the internationalism **startinel** or **startisarel**, due to its particular connotation)

dutisarel "to compel"

14 — borrowings from languages in contact (these are traditionally integrated)

žal "regret" (Slavic)

grìza "worry" (Slavic and Romanian)

ùlica (var. **vùlica**) "street" (encountered even in Spanish Caló as "**ulich**")

trùpo "body" (Slavic and Balkan; its Caló counterpart **drupo** means also "skin")

šarkàni "dragon" (Hungarian)

kùjbo "nest" (Romanian); also "web-site" in IT

15 — borrowings from international terminology

— sometimes in the forms occurring in local languages:

aktìvo "active"

integràcia "integration"

diàspora "diaspora"

paradìgma "paradigm"

— with morphological change (gender):

dialèkta or simply **lèkta** "dialect"

konsonànta "consonant"

kontàkta "contact"

— derivational

istorìsto "historian"

pikturìsto "painter" (in fact this form is quite usual in Mećkar)

artistipen "art" (Djurić; cf. another proposal: **dombipen** "id." < **domb** "artist, creator")

diasporisarel "to scatter into a diaspora"

aktivipen "activité"

This lexical system, in rapid and continuous evolution, is naturally much more complex than what is presented here and it deserves a special study. In this respect, a clear synoptic representation of the

²³ **Dünia**, **zevap**, **hak** and **manaj** are all originally from Arabic, while **barabar** and probably **majmùni** are Persian pers. com. by V. Friedman). These words occur from Indonesia to Morocco and from Central Asia to Tanzania as the result of Islamic conquest. They spread to the Balkan via Turkish (see Friedman 1989 and 1990).

different ways in which the Rromani lexicon has grown over time would be highly desirable. Such a study would be of valuable help to standardisers in their efforts to endow the Rromani nation with a modern and efficient means of modern literary communication. So-called "weak links" in the lexical structure could be pinpointed, and undesirable processes could be altered. At the same time this system helps the linguist to follow as closely as possible the path of natural development used spontaneously by Rromani speakers so far, which is always the safest one. It can also give a higher degree of stability to lexical items introduced for the purpose of enlarging vocabulary for modern use. In this respect, the 0 to 9 scale is of interest not only in sociolinguistic descriptive terms but also for establishing a further 10 to N scale, which in its turn can function as a theoretical framework for the elaboration of the Rromani standard vocabulary and language, one of the most significant means of emancipation of the Rroms as a nation.

The use of the 0 to 9 scale in publishing ethnotexts

Another practical application of this classification is to be found in the publication of ethnotexts, which contain as a rule a significant foreign lexical element. The problem is that the basis for the choice how to spell foreign words is not always clear. While there is no difficulty for items that are clearly Rromani or clearly Gazikani, each class being spelled in its own script, it is more difficult for the intermediate categories. Should the border run between Class 4 and 5, within Class 4, between 5 and 6 or in another place? Formal criteria, seemingly practical, are actually of little help, as seen above, since they are not relevant to the perceived status of the lexical items. An argued choice has to be done initially as to the discrimination between inherited and borrowed element, according to the specific situation of the given language, and it is possible to visualise this differentiation for example by writing the Rromani element in roman characters and Rromani spelling, and the foreign element italicised and in the spelling of the original language. Here is an example from another area – beginning of a folk song from Vojvodina:

*Svanisarda o četvertko rano,*²⁴

maladile o duj phralorra,

maladile o Ivo th-o Zlato,

maladile and-o tang *sokàko*,

thaj pučhen pen pire sastimasθar.

Thursday morning dawned,

both brothers met,

Ivo and Zlato met,

they met in the narrow back street

and asked each other about their health.

Compounds of combined origin are most commonly perceived as Rromani and fall then into the first category. Nevertheless – as shown by the preceding examples, this is indeed a difficult issue, both from the theoretical and practical standpoints, with which field workers are faced every time they publish not only folklore texts, but also interviews, conversations, theater etc... in any minor language. This paper might serve as a framework within which a more rigorous and practical approach to this problem could be found.

Brozović, Dalibor & Dunja • Sociolingvistički aspekti standardizacije romskog jezika. In: *Jezik i kultura Roma*. Sarajevo, 1989. Pp 223-237.

Courthiade, Marcel • Problèmes actuels de lexicologie rromani. In: *Lingua posnaniensis*. Poznań, 1987. Pp. 47-62.

"~" • La standardisation du rromani: aspects sociolinguistiques et méthodologiques. In: *Tsiganes — identité, évolution*. Paris, 1989. Pp. 375-384.

"~" • La langue romani: évolution, standardisation, unification, réforme. In: *Spracheform — Geschichte und Zukunft*. Hagège & Fodor ed. Köln, 1989. Pp. 87-110.

²⁴ The first verse would be in common Rromani: **disili e zòjaqi teharin** (with **disili** < Indian, **zòja** < Rumanian but now widely used all over Europe, **teharin** < Greek; **teharin** belongs to E-dialects, whereas its O-dialect equivalents would be **tàša**, **tajsa**, **tasàra** - same etymology as **tehàra**, **teharin**).

- "~" • Проблеми прилагођавања терминологије неромског порекла ромском језику In: *Развитак рома у Југославији - проблеми и тенденције*. Belgrade, 1992. Pp. 211-224.
- "~" • Un parler slave original: le *kajnas* d'Albanie orientale. In: *Balkanica posnaniensis VI*. Poznań, 1993. Pp. 237-325.
- Duka, Jeta • Pratique interdialectale en situation de diaspora : le lexique – témoignage en synchronie. In: *Faits de langues – langues de diaspora, langues de contact* Anaïd Donabédian éditeur, Paris, 2001. Pp. 181-189.
- Friedman, Victor • Problems in the Codification of a Standard Rromani Literary Language. In: *Proceedings of the Fourth and Fifth Annual Meetings: Gypsy Lore Society*. New-York, 1985. Pp. 56-75.
- "~" • Toward Defining the Position of Turkisms in Romani. In: *Jezik i Kultura Roma*. Sarajevo, 1989. Pp. 251-267.
- "~" • On the Turkish Lexical Component in Romani Dialects and their Relationship to Romani Language Planning. In: *100 Years of Gypsy Studies*. Cheverly, 1990. Pp. 137-144.
- Hancock, Ian • Patterns of English Lexical Adaptation in a Texan Dialect of Romani, ms. 19 pp. 1974.
- Osmani, Lumia • E lavenqe save nane an-i rromani čhib — Sur les mots qui manquent en rromani. In: *Tsiganes — identité, évolution*. Paris, 1989. Pp. 385-394.
- Pathania, J.S. • Четири породице речи-мостова између ромског језика и језика Индије. In: *Развитак рома у Југославији - проблеми и тенденције*. Belgrade, 1992. Pp. 211-224.
- Sharp, Margaret • The Relationship between Romani and English as Spoken by Portland Gypsies. Portland State University, M.A. thesis. 1983.

ABSTRAKTA

Rromani thaj gazikani: kaj si i sindra (limita) maškar lenθe?
Tipològiko dikhipen upr-o stàtus e leksikalone jekhimatenqoro
an-i čhib rromani pal-o avutnipen lenqo.

I analiza sikavel so an-e čhibă so si an-i kontăkta, nanaj zorali limita maškar-e "phurikane" lava ta e lava lile avere čhibēθar. Čacimasθe isi jekh komplėkso kontinuum maškar-e duj agora e lavenqe dorăqere. Maškar e maškarutne elemėnte, šaj te keras jekh klasifikăcia pala jekh skăla kategorienqi kotar-o 0 ži k-o 9.

And-o vaxt so sa e čhibă, save savaxt si an-i kontăkta jekh-avraça, theren akaja polarizăcia, e "phirutne čhibă" si len jekh maj komplėkso struktura, odolesqe so sine len an-i història buteder kontăkte bute thanutne čhibēnça, maškar save phirde. Palem šaj vi on te giren an-o generălo modėli. Avere rigaθe, ma te bistras so jekh "thanutni čhib (biphirutni)" si vi oj jekh čhib an-i kontăkta averença., thaj akalaθar nakhel vi prdal laθe i polarizăcia maškar-e lava "phurikane" ta e lava "avere-čhibăqe". Opr-akaja polarizăcia, avel jekh aver polarizăcia, maškar laθe thaj i čhib e phirutne populaciaqiri.

Te zal pe dureder akala butăça, o artıklo del p-o agor jekh skăla te klasifisaven e statută e neve lavenqere k-o biandöpen an-i khetani čhib rromani. Thovdöl emfăza vi p-e pašimăta, vi p-e ulaimăta (diferençe) maškar-i situăcia jekhe lektaqi ta i situăcia e khetane čhibăqi. Misală si dine šeral kotar-e Mečkarenqi lėkta kotar-i maškarutni Albănia, avere misalēnça kotar aver lėkte andar-e Evröpa, thaj vi kotar-o vakăripen e Serbiaqe Bunăšenqo/Rudarenqo, save si manuša bipinzarde originaqere (šaj čirlatune thanutne), so živen an-i maškarutni Sěrbia thaj vakăren jekh čhib but serbisarde rumanikane leksikaça.