Marcel COURTHIADE

A SUCCINT HISTORY OF THE RROMANI LANGUAGE

The Indian orign of the Rromani language and people is not anymore seriously questionned. Seminal discussions are currently rather held about its current Indian profile and the part of other elements in its constitution.

As highlighted by Sir Ralph Turner, one of the most outstanding indologists of the 20th century, the craddle of the old Rromani language can be located among the languages of the so-called "Central Group", namely in the Aryavarta area, today covered approximately by the raśtra of Uttar Pradesh. This requires some more detailed explanations:

- the arrival of a branch of Indo-European languages in Northern India
- how did Indian languages get constituted ?
- Turner's location of proto-Rromani
- Turner's erroneous datation of the Rroms' exodus
- the development of Sanskrit to Prakrit phonology
- the word ḍomba in this context
- the development of Sanskrit to Prakrit grammar
- the relationship between Indic and Persic fund – Persian among Salçuk Turks

1 - The arrival of a branch of Indo-European languages in Northern India

Although contestations, mainly politically motivated, can still be heard in the field of Indo-European linguistic history, most researchers agree that all languages of this family (namely the Balto-Slavic, Germanic, Italo-Celtic, Romanic, Illyro-Albanian, Anatolian, Achaio-Greek, Arsho-Kuśinean, Persic and Indic linguistic groups) arose from the evolution of a patchwork of vernaculars in use as basilects over a wide area from southern Ukraine to Southern Kazakhstan by various populations, not necessarily (and probably not) related between them by origin ("blood") but with the common name of "Arya". The meaning of this word was "noble, winner, a person able to reign thanks to well performed sacrifices" – therefore without any connexion with the nazi fabrications ascribed to this word. Groups of these populations migrated, mainly to the West but also to the South-East beginning by 2,800 BC. The group which migrated to the South-East probably staid around 2,3000-1,800 in the Bactro-Margian area (mainly today's Afghanistan and Turkmenistan) and encompassed the ancestors of the Iranian and Indo-Aryan peoples. The Iranian peoples dispersed eventually upon a wide territory encompassing in historical Chinese Xīnjiāng to the East and Anatolia to the West, whereas as Indo-Aryan peoples moved toward today's Pakistan, where they arrived around 1,800 BC and further through Northern India, reaching today's Bengal around 600 BC and covering thus all the plains of Indus and Ganges, a huge area they called Ariāvarta, and speaking old Indic vernaculars. They met along road local populations, known as Dasa, and has various kinds of intercourses, sometimes hostile, with them. Those of the Dasa who accepted to be integrated in the new social system became the fourth varṇa, called Śudra.

2 - How did Indian languages get constituted ?

One can read quite often that Rromani developped directly from Sanskrit, a misleadingly inaccurate and simplified statement. As a matter of fact, the history of Indo-Aryan languages is classically divided in three period: old Indic, middle Indic and modern Indic – all of them being actualized firest of all in a series of vernaculars.

A) In early Vedic times (probably 1,800-1,200 BC), a so-called Vedic language was coined out of old Indic basilects by sages as an acrolect for ritual purpose and embeded first in Rig-Veda. This

occured most probably in regions of present day Northern Pakistan (Gandhara). Vedic language was very carefully standardized because it was its creators' creed that the efficiency of rituals to manage the universe depends on the accurateness of the uttered ritual formulæ. All four Veda were composed (and learned by heart) in this carefully made up language. One sould point out that these people were able to write and read but they were reluctanbt to trust the suporeme knowledge of the Veda to perishable material supports. After the Veda, came a series of other erudite compositions in a slightly simplified but still very elaborated language, known as Sanskrit, a word meaning "perfect". True enough Sanskrit is still deemed as logically perfect, as computers have recently confirmed. Due to the fact that old Indic vernacular had developped in the meantime into middle Indic vernaculars, Sanskrit was less and less understood even by Brāhmaṇs and it became necessary to describe it by means of explicit grammatical rules. This was the task of Pāṇini (560-480 BC) and later of his commentator Patañjali (cc. 200 BC).

B) During the same era, some reformers, called śramans, began to protest against the monopolisation of knowledge in the hands (and brain) of the brāhmaṇs' varṇa and to preach in more popular forms of language than Sanskrit, debasilectalizing the vernacular in order to express elaborated philosophical thought. Two of them are still famous: Mahāvīra (599-527 BC) who preached jainism and Siddhārta Gautama (490-410 BC) who preached buddhism. They started a movement of standardization middle Indic vernaculars in order to teach their creed to popular masses, using for philosophical notions a vocabulary often borrowed from Sanskrit. These standardized middle Indic languages are known as Prakrits ("natural") and were geographically distributed all over Ariāvarta and beyond. It was also the time when theater emerged and used also a specific Prakrit form, namely Mahārāṣṭri, for parts played by women, including high varṇa heroines, and petty servants, whereas all main (and male) characters would speak Sanskrit. In the same period, emperor Aśoka (304-232 BC) decided to write down on pillars and huge rocks his "edicts", in dozens of locations almost all over India and translated the texts into the respective local vernaculars for this purpose, giving us the first attestations of these middle Indic languages. The following table gives an idea of the main Prakrits (out of some 20 altogether):

name	area	century	developed into	attestations
Aśokan (304-232)	almost all India	- 3	various vernaculars	edicts on pillars and rocks
Ardhamāgadhi	Kosala (Udh)	- 5	> Māgadhi	Jain scriptures: svetambara
Māgadhi	Bihar, Nepal, part of Bengal	- 4 > - 2	> Bhojpuri, Maithili, Magahī, Bangla, Assamese, Oriya > also Pāli	drama;
Śauraseni	Midland (Eastern U.P.)	+3 > + 10	> Braj Bhasa, Eastern Panjabi	drama; Jain scriptures: digambara [nange]
Mahārāṣṭri	SW	- 5 > + 5	> Marathi, Konkani, Divehi, Siŋhala	drama (the heroine and her ♀ fiends) – Kalidāsa's works; lìrika + "jaina Mahārāṣṭri"
Pāli	Midland	- 2 > today	religious use only	Buddhist Theravada
Nia	Serìnda	+ 2-3	extinct	administration (along with Śaka, Agnean, Kuchean)

One should add Buddhistic simplified Sanskrit (2-3 BC.).

C) Further evolution resulted in the emergence of modern Indic languages, some of them mentionned in the fourth column of the table above. Since proto-Rromani left India at the turning point between middle and modern Indic periods, more details about the last generation of languages would be beyond the scope of this presentation.

3 - Turner's location of proto-Rromani

In his brilliant comparative study "The Position of Romani in Indo-Aryan", published in 1927 in Eddimbourg, Turner pinpoints the position of "ancestral Rromani" in the Central group of modern Indic languages, scrutinizing linguistic data beyond similarities and discrepancies, namely he distinguishes "firstly between conservations and innovations, and of innovations between the earlier and the later" shared by two or more of these languages. He shows that shared innovations are more pregnant than shared conservations to attest kinship between two languages and that in turn shared late innovations are stronger than both series of early features for this purpose.
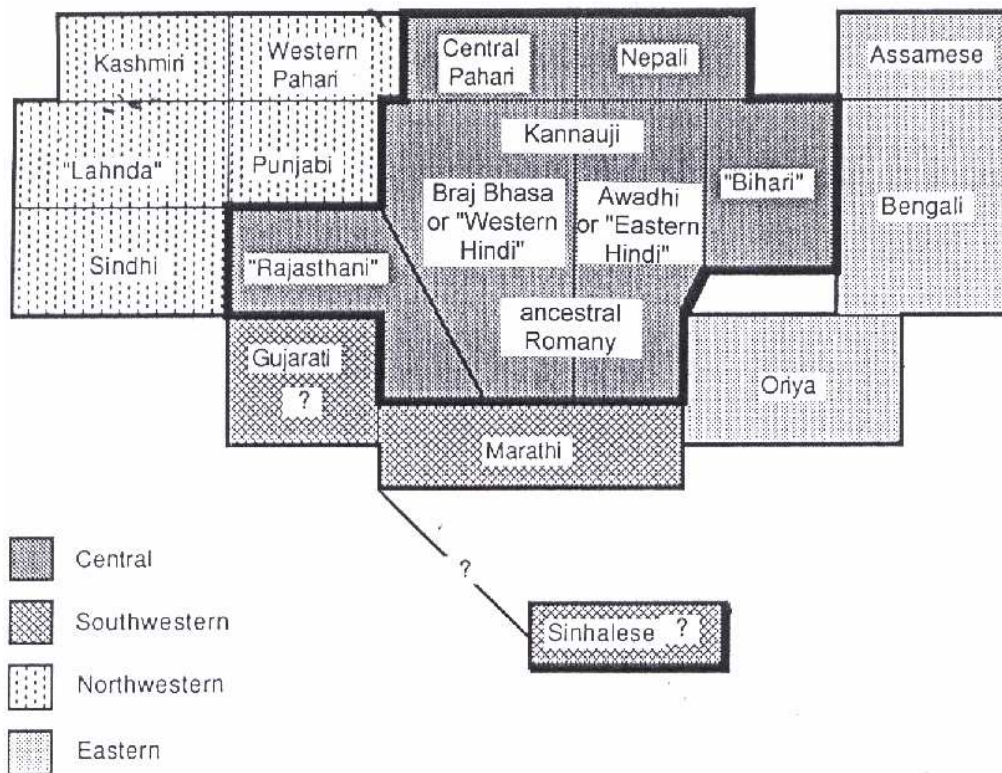
NIA subclassification based on Turner

After Colin P. Masica "The Indo-Aryan Languages" Cambridge, 1991 pp. 453 sq.

4 - Turner's erroneous datation of the Rroms' exodus; a correct datation

While Turner's geographical identification of the Rromani language origins is quite correct when he includes it into the Central group, he made an error in dating the separation of Proto-Rromani from this group. This error originates in a lack of information about linguistic phenomena outside India and the convictions that all features of modern Rromani have to be ascribed to it early period of formation, in India. In this respect, he considers that a specific innovation, the Rromani evolution *Nasal + Voiceless Stop > Nasal + Voiced Stop* (namely *nt > nd*, *nk > ng* and *mp > mb*), is a feature shared with the Dardic group (Kasmiri and surrounding languages) and point at a commonality of fate between this group and Rromani.

In § 43 of his booklet, Turner suggests that this evolution in Rromani took place in a Dardic surrounding, due to the fact that a similar evolution occured in this group in the 3rd century BC. (attested in the Karoshthi documents). He concludes that Proto-Rromani was spoken at the time in the same area as Dardic languages and that the evolution in question was common to them and to Proto-Rromani. In truth Turner overlooked the three following facts:

• first of all, voicing of voiceless stops after a nasal consonant is a quite common-place case of progressive assimilation in homorganic clusters, encountered in most various groups of languages all over the world and that there is no need of language vicinity or contact to make it occur.

• secondly, it developed in Rromani not only in stems — as Turner noticed, but also in the postpositional system which emerged during the second half of the first millennium a. D.: *-n + ke > -nge* (spelled –nqe), *-n + tar > -ndar* (spelled –nθar) etc... namely more than one thousand years after the Dardic evolution in question. This excludes clearly any connection between these therefore independant evolutions.

• this very evolution did occur in popular Greek, a language spoken all over Asia Minor – alongside with Kurdish, Armenian and some other minority languages, now extinct, at the beginning of the second millenium AD. that is to say when the Proto-Rroms reached this area. The evolution nasal sonant + voiceless stop consonant giving nasal sonant + voiced stop is evidenced in all positions in Greek: within a stem (πέντε "five" pronounced ['pende]), with an affix (εμπόριο "trade, business" pr. [em'borio]) and at grammatical junctions (την ταβέρνα "the tavern" pr. [tin da'verna], τον πρότο "the first one" pr. [tom 'broto], την κασέτα "the tape" pr. [tiŋ ga'seta] etc... – just like at the Rromani junction B-form + postposition; as a matter of fact this Rromani evolution seems to be more similar to the Greek one than to the Dardic one). In his "Traité de phonétique" (Paris, 1933 p. 189), Maurice Grammont describes this evolution of homorganic clusters in Albanian and Syriac.

In fact, Turner took on account in his study on strictly etymological, not morphological, data and this is the reason why he put the separation between Rromani and the Central group before this evolution in Dardic was evidence by early inscriptions – what totally mislead further researchers. In fact it is obvious that the voicing of voiceless stops after a nasal in Rromani in stems as well as before a postposition occurred much later, outside India and namely in the mediaeval Greek context, and there is no need of searching the explanation in any analogous evolution which could have occurred on its own among some North-Western Indic languages around 250 BC. This omission is probably the only mistake in all Turner's study but it led to erroneous conclusions in dating the Proto-Rromani exodus.

Other data give better indications as for the Rroms' separation from the Indian soil, namely:

a) the presence of postpositions in Rromani, an innovation shared with other modern Indic languages, which appeared by the end of the first millenium;

b) the reascription to masculine and/or feminine of ancient neuter substantives, when neuter disappeared from Indic languages. Hancock emphisazes that this reascription is almost always the same in Rromani and other modern Indic languages, dating this phenomenon back to the end of the first millenium. Rromani exodus should have occured therefore at the turn of the millenium or at the beginning of the second millenium AC.

5 - The development of Sanskrit to Prakrit phonology

The most striking feature in the development of Sanskrit into Prakrit phonology is the simplification of consonantal clusters into geminates:

pt > MI tt > Rr. t: tapta > MI tatta "heated" > Rr. tato "warm"
                        supta > MI sutta "slept" > Rr. suto "asleep"
                        sapta > MI satta but Rr. efta < Gr. εφτά
kt > MI kk > Rr. k: pakta > MI pakka "ripe, cooked" > Rr. peko (arch. pako)
                        yukta > MI ʒutta "yoked" > Rr. ʒuto "pair"

rakta > MI ratta "red, blood" > rat "blood"
                          muktā > MI (lacking) > Rr. mutǐ "pearl" (Russian dial.)
                          śukta > MI ? > Rr. śut "vinagre"
gn > MI gg > Rr. g   agni > MI aggi > Rr. jag (Baltic Rr. ag)
                          ātman > Aś. ātpa "self" > Rr. [a]pes
rp > MI pp >:   sarpa > MI sappa > Rr. sap "snake"
rṇ > MI ṇṇ > Rr. n: karṇa > MI kaṇṇa > Rr. kan "ear"
The assimilation obeys three main rules:
a) in the case of two stops, the second one prevails
b) stops assimilate liquids (r, l) before or after them
c) nasals are assimilated to a preceeding stop
Other rules are more specific to the various cases.
Examples are numberless. As a rule no consonant cluster in the middle of a word may exceed two
consonants and in such a case, they have to be:
- geminates (tt, pp, kk, kkh etc...)
- homorganic nasals before a stop (mp, mb, nd, etc...)
- aspirated nasals (mh, nh etc.) or lh.

Initial clusters are also reduced:
sth- > th-   sthāna > Pāli thāna "breast" > Rr. than "place"
                          sthūlā "solid, strong" > MI thulla "massive, fat" > Rr. thulo "fat, thick"
sph- > ph-   sphāṭayati > Pāli phāleti, Śaur. phāḍedi > Rr. pharravel, pharavel
gr- > [g]g-   grāma > MI gāma > Rr. gav, cf. Hindi gaõ, gãv
sv- > s-   svapati > MI suvai > Rr. sovel
                          sveda > MI seda > Rr. sed "sweat"
                          śvaśrū > MI sassu > Rr. sasuj "mother in law"
                          śvaśura > MI sasura > Rr. sastro "father in law"
                          svastha > MI ? > Rr. sasto "entire, healthy"
Note the preservation of the clusters in the two last examples in Rromani.

As a rule, initial y- becomes ʒ- in MI (except in Māgadhi) or it may be dropped, whereas it
develops usually in ʒ- in Rromani:
                          yukta > MI ʒutta "yoked" > Rr. ʒuto "pair"
                          yuvatī > Śaur. ʒuvadi, Māgadhi yuvadi > Rr. ʒuvli, ʒuvel "female"
                          yūkā > Pāli ūkā but Rr. ʒuv "louse" (cf. Hindi ʒũ)
but   yāti > MI yāti > Rr. ʒal "goes"
                          yava > Pāli yava > Rr. ʒov "barley" (cf. Hindi ʒau)
                          Pāli hiyyo > Rr. iʒ "yesterday"

Initial bh drops the stop and only the aspiuration remains:
bh- > h-   bhavati "to happen, to be" >MI bhavati "to become" > hoti "to be, to
                          exist" > Rr. ovel and Hindi ho-nā "to be".

Initial ś may turn into ćh, which explains the etymology of ćhavo:
śāba or śāva "young of an animal" > ArdhM. ćhāva, Pāli ćhāpa (but Māgadhi śāvaka) > Rr.
ćhavo

Sanskrit intervocalic simple consonants are usually dropped in MI (through a stage of
sonorisation: -t- > -d- > -Ø-), leading to very surprising forms, especially in Māgadhi:
raʒata > raaa "sovereign"   hṛdaya > hiaa, hia "heart", but Pāli hadaya

viyoga > vioa "dissociation"   avapāta > oāa "elephant trap"

For example, the verbal enfing of 3rd person (present tense) -ati develops in MI as -aï, except in Śauraseni, where it preserves a voiced consonant -adi, -edi. Rromani dropped the final i and developed this consonant further into the lateral [l]: -el.

Other examples: gata > Rr. gelo but MI gaa, gada, gata, yāta
                   gitā, gīti > Rr. gili "song" but MI gīa "sung"
                   mr̥ta > mulo [MI maa]
                   ghr̥ta > Māgadhi ghaa, but Śaur. ghida, hence Rr. khil
                   śata > Māgadhi saa, but Śaur. sada, hence Rr. śel
                   marati > Māgadhi marai, but Śaur. maradi, hence Rr. marel "he beats"

In Pāli devatā "deity", we have to do with a Sanskritism; cf. Rromani devel.

Old Indic r remained r in Rromani through MI:
                   rakta "decorated, red, nice, in ove; blood, safran" > MI ratta "red;
                   blood" > rat "blood"
                   rāʒā "lord" > MI rāʒa "king" > Rr. raj "gentleman"
                   kṣurika/ćhurikā > MI ćhūrikā > Rr. ćhuri[k] "knife"
                   ćaur > MI ćora > Rr. ćor "thief"

Intervocalic simple retroflex ṭ developped usually into regular r in Rromani
ṭ > r   ghuṭ[ī] > MI khura > Rr. khur "[shoe] heel, hoof"
                   ghoṭā "horse" > Aś.ghoḍā > Rr. khuro "foal, colt"
                   akṣoṭa > akkhoḍa > Rr. akhor "walnut"
                   ghaṭa > MI ghaḍa > Rr. khoro "jug"
                   kīṭa "insect, worm" > Rr. kir[i] "ant"
                   ? kāṭava "sharpness" > MI kāṭa "pudendum virile" > kar "id." (one may still
                   encounter the pronunciation [kaɾ] in the rural Mećkar vernacular of Myzeqe
                   in Southern Albania)
                   vāṭa(ka) "enclosure" > MI vāṭa "enclosure" > bar "fence, hedge"
                   kuḍī "hut" > MI ? Rr. kuri "
                   sphāṭayati > Pāli phāleti, Śaur. phāḍedi > Rr. pharravel, pharavel

The evolutions seems to give also r, but sometimes ḍ in MI. In some cases the word is not available.

Intervocalic doubled retroflex ṭṭ developped in Rromani into retroflex ṛ (written rr and pronounced in various manners, according to the dialectal group, *vide infra*) as a first stage and later into various pronunciations, basically differing from regular [r], which is written r.

Initial ḍ also lead to retroflex initial ṛ, as doubled intervocalic ṭ equally did:
ḍ- > rr-   daṇḍa "rod" > MI ḍaṇḍa > Rr. rran "branch"
                   MI *ḍohi, ḍoha > rroj "spoon" (cf. Hindi ḍo[h]ī "wooden ladle, dipper"
                   - could originate from the retroflex treatment of an initial d-: doha
                   "small wooden vessel used for milking cattle")
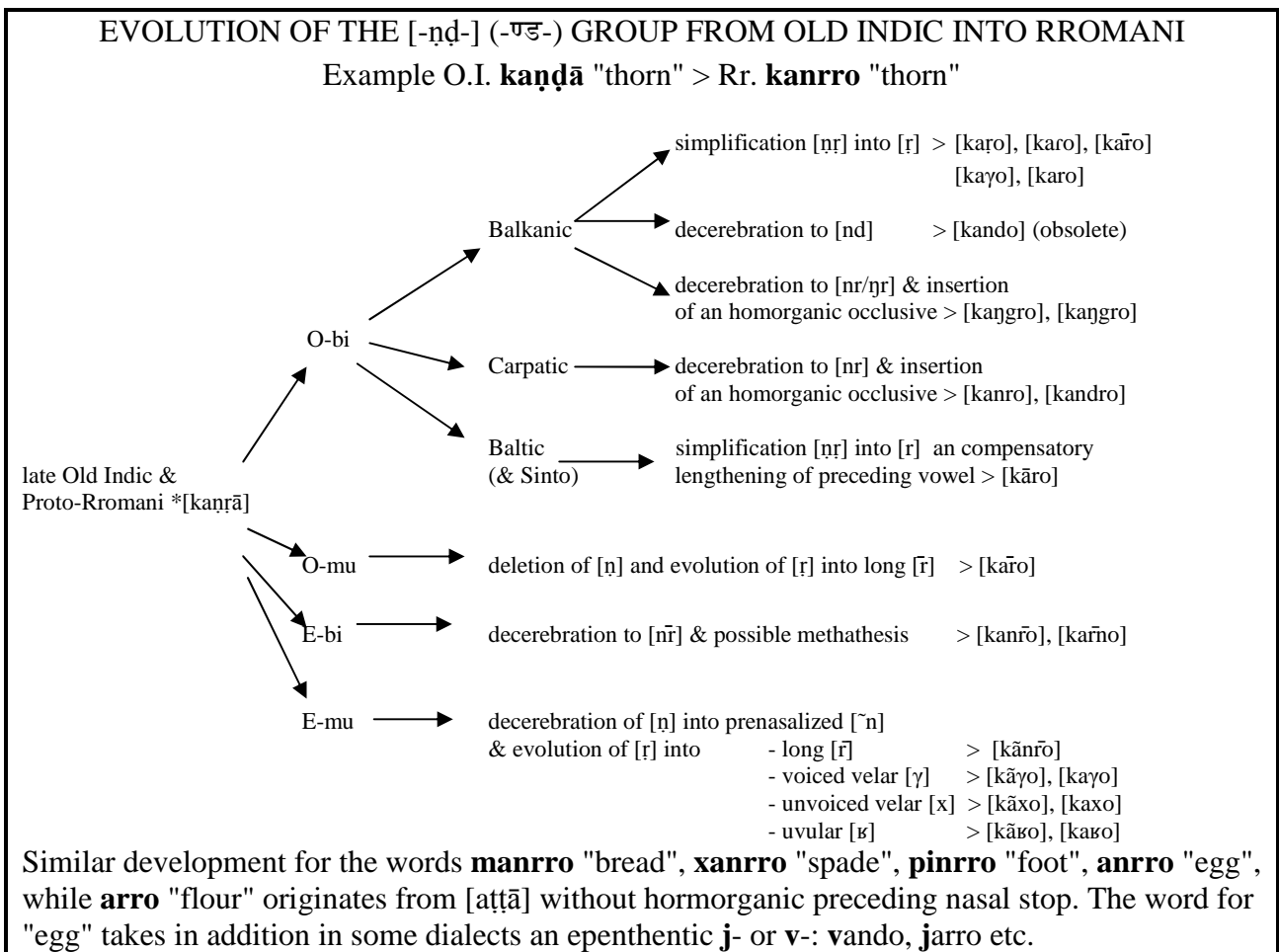ṭṭ > rr   *āṛṭā [reconstructed on the basis of Pers. ārd "flour" and Gr. άρτος]
                   "bread" > MI aṭṭa "meal" > arro, varro "flour"
                   MI peṭṭa > perr "belly" (cf. Hindi peṭ "id.")
                   Māgadhi ćaṭṭei "to eat, to grind" > ćarrel, ćarel "to lick"
                   kuṭṭ[ayati] "to crush, to smash, to strike" > MI kuṭṭeï > Rr. kurrel "to
                   beat, to strike" (in some archaic dialects, whereas it took an offensive
                   meaning elsewhere).

sphuṭati, sphoṭati "to burst, to split open" > MI puḍaï (poḍaï non attested) > Rr. porravel

As one may observe, most of these words are not attested before the MI period, which means that they were introduced into North Indian languages comparatively late (probably from autochthonous languages or of onomatopeic origin). If they already existed in Old Indic times, they had proably a peripheral status and are not mirrored in the texts which passes through up to us. The situation is however different in the case of homorganic retroflex groups (nasal + stop), always intervocalic, which developed differently according to the dialectal group of Rromani:

ṇṭ > nrr   aṇḍa > Pāli aṇḍa > Rr. anrro "egg"

maṇḍa "gruel (among other meanings) or maṇḍaka "a kind of baked flour" > Pāli maṇḍā, maṇḍakān > manrro "bread"

khaḍga > khagga "sword", possibly also khaḍḍa "" > xanrro "sword"

kaṇṭaka > Pāli kaṇṭaka > kanrro "thorn"

piṇḍa "globe, ball; calf of the leg" > piṇḍa "lump, ball of food" [?] > pinrro "leg".

The evolution of this nasal cluster is fully coherent accross the various Rromani vernaculars, as depicted in the following table:

EVOLUTION OF THE [-ṇḍ-] (-ण्ड-) GROUP FROM OLD INDIC INTO RROMANI

Example O.I. **kaṇḍā** "thorn" > Rr. **kanrro** "thorn"

late Old Indic & Proto-Rromani *[kaṇṛā]

O-bi
  → Balkanic
    → simplification [ṇṛ] into [ɽ] > [kaɽo], [kaɾo], [kaῘo] [kaɣo], [karo]
    → decerebration to [nd] > [kando] (obsolete)
    → decerebration to [nr/ŋr] & insertion of an homorganic occlusive > [kaŋgro], [kaŋgro]
  → Carpatic → decerebration to [nr] & insertion of an homorganic occlusive > [kanro], [kandro]
  → Baltic (& Sinto) → simplification [ṇṛ] into [r] an compensatory lengthening of preceding vowel > [kāro]

O-mu → deletion of [ṇ] and evolution of [ṛ] into long [r̄] > [kar̄o]

E-bi → decerebration to [nr̄] & possible methathesis > [kanr̄o], [kar̄no]

E-mu → decerebration of [ṇ] into prenasalized [˜n] & evolution of [ṛ] into
  - long [r̄] > [kãnr̄o]
  - voiced velar [γ] > [kãɣo], [kaɣo]
  - unvoiced velar [x] > [kãxo], [kaxo]
  - uvular [ʁ] > [kãʁo], [kaʁo]

Similar development for the words **manrro** "bread", **xanrro** "spade", **pinrro** "foot", **anrro** "egg", while **arro** "flour" originates from [aṭṭā] without hormorganic preceding nasal stop. The word for "egg" takes in addition in some dialects an epenthentic **j**- or **v**-: **v**ando, **j**arro etc.

The word xarr "pit" developed from Old Indic khaṇḍa according to the same pattern but it survived only in the dialects, where the evolution resulted in [r̄/rr].

Old Indic had a special r-vowel, written [ṛ] in phonetics and similar to the one heard in Serbo-Croatian, Slovak and Czech. This sound has nothing in common with retroflex [ṛ]. It developed

in MI and Rromani mainly into [u] and [i], but also [a] in MI, according to the specific language; both evolutions are observed in Rromani:

r̥ > u	pr̥ćchati > Śaur. pućchadi > puchel "to ask"
	vr̥kka "kidney" > Pāli vakka > buko "internal organ"
	vr̥tti "activity, work" > butĭ "work, thing"
	mr̥ta > MI muda, mua, maa > Rr. mulo "dead"
r̥ > i	ghr̥ta > Māgadhi ghaa, Śaur. ghida > khil "butter"
	kr̥ś[t]a > Śaur. Māgadhi kisa > Rr. kiślo "thin, meager"
	śr̥ŋga[ka] > MI siŋga, saŋga > śing "horn"
	hr̥d[aya] > MI hida, hiaa > ilo "heart"

The old Indic group kṣ developped in Rromani mainly as kh through MI kh in initial position and kkh in the middle of the word:

kṣ > kh	kṣīṇa "wasted" > khiṇa, ʒhiṇa > khino "tired"
	kṣvelati > Śaur. kheledi > khelel [but Sanskrit has already khel- "shake, play" as early as the Ramayan; these could be two different stems, which became homophone as a result of two different evolutions]
	drākṣā > MI ? > drakh "grape, vine"
	akṣi > MI akkhi/aćchi > jakh "eye" (akh in Baltic dialect)
	makṣ > Pāli makkhi[kā] > makh[i] "fly (insect)"
	vr̥kṣah/rukṣah > Śaur. rukkha [unique form in Śaur. – others have mainly vaćcha] > rukh "tree"
	akṣoṭa > akkhoḍa > Rr. akhor "walnut"
	mrakṣa[ya]ti > Śaur. makkhedi > makhel
kṣ > ćh	kṣipati > Māgadhi khivaï, Śaur. ʒhivadi > ćhivel "to put"
	r̥kṣa > AMāgadhi aćcha, Śaur. rićcha > rićh "bear"

If the specific evolution of the word for "bear" can be possibly explained by the fact that these anials were sold by speakers of montaneous areas, speaking a different MI language, there is no explanaition for the dveloppement into ćh of the kṣ in kṣipati.

Intervocaliv m and p develop into v in Rromani:

-m- > -v-	nāma[n] [neuter] > Śaur. ṇāmaṁ [masc.] > (a)nav "name"
	komala > Pāli komala > kovlo "soft, mellow"
	hima "cold, chill, snow" > MI hima "snow, ice" > iv "snow"
-p- > -v-	svapati > Śaur. sovadi > sovel "to sleep"
	apara > MI avara > aver "other"

Another important feature is the desonorization of initial sounded consonant after the MI stage:

gh- > kh-	ghāsa "food, fodder" (cf. verb ghas "to eat") > MI ghāra > khas "hay"
	ghoṭā "horse" > Aś.ghoḍā > Rr. khuro "foal, colt"
	gharma > MI ghamma "heat, hot sesaon" > kham "sun"
	ghaṭa > MI ghaḍa > Rr. khoro "jug"
bh- > ph-	bhūmī > MI bhūmi > phuv "earth"
	bhagna > MI bhagga > phago "broken"
	bhanati > MI bhaṇadi > phenel "to say"
dh- > th-	dhūma > MI dhúma > thuv "smoke"
	dharati "to hold" > MI dhareti > therel "to possess, to have"
	dharma > MI dhamma > thami "law"

As a rule, an aspiration initially located within the old Indic word migrates to the initial position in Rromani:

kakṣa "hiden place, armpit" > IM kakkha > Rr. khak "armpit"

With concomitant desonorization of the newly formed initial aspirate:

gandha "flagrance" > IM gandha "smell, scent" > *ghand > Rr. khand "stench"
garbhiṇī > IM gabbhinī > *ghabni > Rr. khamni "pregnant"
ʒyotsnā > Pr. ʒoṇha > *ʒhoṇ > Rr. ćhon "moon"
gu[ṁ]phati "to weave" > *ghuvadi > Rr. khuvel
tuććhya > MI ćuććha > Rr. ćhućho "empty"

The group ṣṭh gave śt in Rromani:

-ṣṭh- > -śt-   oṣṭhā > oṭṭha > Rr. uśt "lip"
                      kāṣṭha > kaṭṭha > Rr. kaśt "wood"
                      aŋguṣṭha > aŋguṭṭha > Rr. anguśt "finger (arch. dial.)"

Intervocalic aspirated stops lost their plosive element and the [h] remained alone in Prakrits, developing eventually into [j]:

-kh- > -h-   mukha > muha > Rr. muj "mouth; face"
                      nakha > naha > Rr. naj "nail; finger; spoke (of a wheel)"
                      nidāgha > nidahi > Rr. nilaj "summer"

Other consonantic evolutions are of lesser scope and specific studies should be devoted to each of them. In so far vowels are concerned, two are of doubtless significance, namely the evolution of many Indic -a- into Rromani -e- (a comparatively late evolution):

-a- > -e-   khara > MI khara > Rr. xer "donkey" (but also Persian xar "id.")
                     verb ending (3rd sg. present tense) -ati > Śaur. -adi > Rr. -el
                     taruṇa > MI tarūna > Rr. terno "young" (but Sinto tarno)
                     pakta > MI pakka > peko "ripe; cooked" (but Sinto pako "id.")

and the group -ava- coalesces into -o- (paralleled by other, less frequent, similar shortenings):

-ava- > -o-   lavaṇa > MI lavaṇa, loṇa > Rr. lon "salt"
                     davara > MI *dora > Rr. dori "lace, string"

6 - The word ḍomba in this context and its relationship with "Indian Gypsies" and paṛiya

In this context, it is usefull to investigate the history of the word ḍomba and the etymology of the ethnic name Rrom. As a matter of fact, the filiation of the two words fits with all rules of diachronic phonology, whereas attempts to ascribe to etymons like Rāma (Viśnu's seventh avātar) or to Rūm the origin of Rrom all fail because they disregard the phonologic distinction between two different r's in Rromani, namely [r] and [rr]. One third of the Rroms, mainly under the influence of European languages which do not have this distinction, in fact merge both phonemes into one common sound [r] and they pronounce [rom], in a similar way as some English speakers do not distinguish d [d] from th [ð] and pronounce [de dīp rivə] instead of [ðe dīp rivə] "the deep river"; this doesn't suggest that the opposition between [ð] and [d] has vanished in English. Another factor is however stronger: the first descriptions of Rromani were the work of non-Rromani researchers, who didn't have this distinction in their mother tongue and were not able to percieve it in the Rromani language, even when it was present. Therefore they wrote every where "r" for both [r] and [rr] and it is not random that double rr appeared systematically in the very first dictionary of Rromani written by a Rrom, namely the the "Gyökszótár" by Ferenc Sztojka (Paks, 1890).

The questions related with ḍomba are mainly the following:

a) when did the word appear in Indian documentation ?

b) what was/were its meaning(s)a ?

c) how it developped later ?

One has to notice first that the word ḍomba appears late and is comparatively rare. Its first occurence dates back to the 6th century CE in a treaty of astrology which mentions a "danger for a king at the sight of celestian musicians [gandharva] accompanied by some ḍomba". The second one is to be found in a collection of 25 tales by Somadeva (7th or 8th century CE) and it means simply "drummer", being probably coined on an onomatopoeia (like English "drum"). The next occurence of this stem occurs in the Daśarūpaka, a treaty of dramaturgy composed around 980 CE by Dhanañjya, who describes a specific minor kind of theater, called ḍombika, which can be translated roughly by "body language" in the kind of the later "commedia del'arte" in Europe. Jeanne Gamonet (to whom we owe these references) suggests, on the basis of Indian testimonies, that ḍombika refers to the main traditional occupation of the ḍomba, namely performing great Indian epic deeds in a colourful and visual way, accessible for simple people, usually around temples and other holy places. Then there is no more mention of ḍomba until 1148 CE, when brāhmaṇ Kalhaṇa composes his famous Rājataraŋgiṇī "River of the Kings", in which the Kashmiri sovereign Jājāpiḍa falls in love with a beautiful ḍombī, the singer and dancer Haṁsī, Raŋga's daughter and marries her in a gorgeous ceremony – in spite of his ministers' jealousy. In Prakrit literature, we can find only one mention of the word, again a love affair between a ḍombī and king, resulting in a curse against the whole kingdom, due to this adharmic union, and the abdication of the king, who prefers to leave with his beloved ḍombī in the woods and spend there the end of his life.

As far as the meaning is concerned, we have to do with a development from percussioners to musicians widely speaking and all kinds of artists of the stage: actors and dancers. More divergent meanings surface eventually in the modern period and it is true that, as opposed to the Sanskrit and Prakrit period, most of them are derogative. The moderm form of the word is ḍom, fem. ḍomni, and it was very fashionable in British publications in the 19th century, which commented the out-law character of the ḍom tribes, classified by the administration as "schedule tribes" or "criminal tribes". A strong connexion was repeatedly made with British Gypsies and kinship was very affirmed between them, leading quite often to the labelling of such tribes as "Gypsies" in India. The word is still in use, yet with low case "g". To our surprise, we could't find any remnant of this word among Indians, be it on the spot, be it outside India, whereas "gypsy" is quite widespread and percieved as the worst possible offence. Despite the popular (and maybe adminitrative) amalgamate between [Indian] "gypsies", [British] "Gypsies" and ḍom, some scholars remained careful and for example Grierson wrote in his monumental Linguistic Survey of India (1904-1928) as early as 1926: "Migratory tribes are found all over India, and are of different kinds. Some of them are descended from adventurers and individuals belonging to various castes and trades; others are occupational units, who wander all over the country in pursuance of their trade; others again are much of the same kind as the Gipsies in Europe, tumblers, jugglers, acrobats, or thieves and robbers, who have come under the Criminal Tribes Act. It has become customary to call these tribes Gipsies, but this designation does not imply any connexion between them and the Gipsies of Europe." In spite of this wise warning by the most outstanding researcher in Indian dialectology so far, it is commonplace to find even in recent publications not only the blind identification of these groups with Rroms, or with their ancestors, but also attempts of seemingly scholarly justification of such an identification, in spite of the blatant lack of rationale for this.

These facts lead to the following suggestion: the word ḍomba entered Indo-Aryan languages at the beginning of the common era, or shortly before it, as an onomatopoiea based borrowing from autochthonous languages, to refer to percussioners and other musicians of the temples. The word

took a derogative meaning when extended to delinquent groups maybe during the Moghol occupation, when all social relationships became more tense, and for sure under the British Raj, even if it was comparatively rare. Discovered by British civil servants, the groups referred to by this odd word were treated according to the same hostility as Rroms (Gypsies) in British Island and gave material for all kinds of speculations in the United Kingdom on both the criminal character of these Gypsies and their link with Rroms, called *Rromani ćhal*s in England. During the 19th century as well, the Dravidian word paṛiya, used in the South of the country, changed its meaning in the speech of British housewives in India: the epithet paṛiya overextended from percussioners of various varṇa (its genuine meaning) to all local employees, a good number of whom were ćaṇḍala (or dalits, intouchables). As a result, paṛiya became popular in the derogative but also compassionate meaning of intouchable, as it is still functionning currently. A further clue to the "filthy" position of musicians was their alleged impurity due to the contact of their hands with animal skin on the drums when playing, an affirmation which is invalidated but the fact that deities also would play drums, as well as brāhmaṇs, as mentionned in literature. In the mind of British civil servants and their wives, the amalgamate was done between Southern Dravidian paṛiya and the Northern form ḍomba, leading to the erroneous identification of musicians to outcasts. Their insolent inscription to Criminal Tribes confirmed this stigma, which encompassed soon European "Gypsies" in the same reluctance for these "thieves and robbers".

To conclude with, it is necessary now to reveal and explain the semantic evolution of these designations from the social point of view and resassign them to their respective genuine meaning. This is a very much needed task because some Indian activists have discovered in Europe the legends conveyed by Britons in the 19th century (and mainly unknown in India) and, unaware of their scientific vacuity, they are attempting to reappropriate them, bringing an indescriptible confusion into a chapter where approximations and fabrications have already proved very harmful.

7 - The development of Sanskrit to Prakrit and Rromani morphology
A) in the nominal group
Sanskrit is a highly inflected language. This means that the various relationships between the elements of the simple sentence (or *functionnal cases* – which can be virtually thousands) are expressed by a set of specific endings of the substantive; such endings are the morphological cases or simply cases. This is not the choice of all languages: functionnal cases can be expressed as well by means of prepositions, as in English and most modern European languages or postpositions, as in modern Indo-Aryan languages but also in Turkish, Hungarian, Finnish, Japanese etc. There exist among the languages of the world more sophisticated means to express sentence relationships (or functionnal cases) but their description would lay beyond the scope of this paper.

Vedic expressed all functionnal cases by means of seven morphological cases (+ vocative), grouping therefore the various meanings of sentence relationships into seven basic cases. Yet progressively, adverbial elements were used to explicit more clearly the meaning of the sentence and they turned into prepositions (a dozen). Concurrently the same functional case began to be expressed by means of two or three different morphological cases, bringing various new semantic nuances. However, at the basilect level, this created a kind of confusion which weared up the ancient system. Dative declined and genitive began to be use in a dative meaning in common speech, an evolution mirrored in Prakrits. Later on accusative merged to nominative, whereas the genitive-dative of late Prakrit turned in Rromani into an oblique case as mirrored in the following table:

Sanskrit (also Mahārāṣṭri)	Śauraseni	Late Prakrit	Rromani	Other NIA languages
Nominative	Nom.-Acc.	Direct case (Nom.)	A-case	Direct case (A)
Accusative				
Instrumental				
Genitiv	Genit.-Dat.	=> Oblique case	B-case	Oblique case (B)
Dativ				
Locative	-**e**		-**e** (adverbial remnant)	-**e** (adverbial remnant)
Ablativ	-**ado** (Śaur.), -**ao** (other MI)		-**al** (adverbial remnant)	
[*Vocative*]				

In addition, the old ablative and locative disappear, leaving their ending (resp. -**e** and -**al** in rromani) only in a handfull of words (like khere "at home", kheral "from home", dromal "on one's way") actually used in adverbial function. The instrumental case began early to disappear progressively and it is absent alltogether in modern languages, where it is replaced by the B-case followed by a postposition (-**ka sath** in formal Hindi, **ko səŋ** in Kannauji, **sa** in Braj Bhasa and **ça** in Rromani).

The parallel of formes between Rromani and other New Indo-Aryan languages is conspicuous in the following examples:

	Rromani	Hindi	Kannauji (villages)
SINGULAR			
Masc. Direct case	jekh tikno raklo	ek ćhoṭā ləɽkā	ek tikino larika/lariko
Oblique case	jekhe tikne rakles	ek ćhoṭe ləɽke	ek tikine larike
Fem. Direct case.	jekh tikni rakli	ek ćhoṭī ləɽkī	ek tikinī larikini
Oblique case	jekhe tikne raklia	ek ćhoṭī ləɽkī	ek tikinī larikini
PLURAL			
Masc. Direct case	panʒ tikne rakle	pañć ćhoṭe ləɽke	pañć tikine larike
Oblique case	panʒe tikne raklen	pañć ćhoṭe ləɽkõ	pañć tikine larikõ
Fém. Direct case	panʒ tikne raklia	pañć ćhoṭī ləɽkiyã	pañć tikinī larikiyã
Oblique case	panʒe tikne raklien	pañć ćhoṭī ləɽkiyõ	pañć tikinī larikiyõ

The collapse of the ancient Indic case system brought in compensation the elaboration of a new system of postpositions and prepositions, in addition to the remnants of locative and ablative.

During a first stage, half a dozen of postpositions were matched to the B-case to express:
- spatial relationships, namely the place where the subject is/goes: -**θe** and the place whither he comes: -**θar**.
- relationship of instrument or company, as in English "with": -**çar**.
- relationship of aim, objective of beneficiairy, as in English "to, for": -**qe**.
- relationship of possession, expressed by a postposition inflected according to the number, gender and case of the possessor: -**qo**, -**qi**, -**qe**, exactly as in modern Indo-Aryan languages of India. Alongside with this short form of the possessive postposition, the most archaic dialect have kept a long form, unique or emphatic if used alternatively to the short form, and featuring as follows: -**qero**, -**qeri**, -**qere**, -**qoro**, -**qiri**, -**qro**, -**qri**, -**qre** etc. Similar form have been evidenced in literary Kannauji as well as in the rural variants around this city – compare:

Rromani: **lesqo** "his" (short form), **lesqero, lesqoro, lesqro** (long form)
Standard Hindi: **uskā** "id."
Kannauji: **uhkero** "id." (village, but **ūhko** in the city).

The genesis of the Rromani and Modern Indic possessive postposition is worth a bit of explanation. Yamuna Kachru traces it back to Middle Indic kera (< Sanskrt kṛta) "done". She relies on Beames' hypothesis that a syntagm of the type X-kṛta B (fem. X-kṛti B) "X-done B, B done by X" could acquire a meaning of possession "X's B". According to her, kera (fem. keri) was a common form in what she calls "Old Hindi" and it would therefore be a simple archaism in the Kannauj area.

One should still add one postposition, conveying a meaning of corporal position or of time:
beśel koćienqo "he is kneeling"
bićhal les manqe kurkesqo "send him to me on Sunday"
kidisàjlam Herdelezesqo/Neve berśesqo "we gathered on the 6th of May/for the New Year eve"
Combined with the privative preposition bi, it expresses the meaning of "without" – compare:
bi kheresqo bi limoresqo (bi kheresqoro bi limoresqoro in Bulgaria) "without home without grave". A similar structure is encountered in standard Hindi.

However, due to the fact that the spatial indications conveyed by simple postpositions were not sufficient, speakers used to specify the position complementary words which developped into prepositions in the case of Rromani and into compound postpositions in India – compare:
Rromani: amare purane kheresθe "at our old house" (general position, but ≠ khere "at home")
                    paś amare purane kheresθe "near our old house"
Hindi: hamare purane ghar mẽ "at our old house" (general position)
                    hamare purane ghar ke pas "near our old house" (with pas functionning as a noun, litt. "near of our old house").
Later, the new element developped into a real preposition and the postposition was dropped, leading to the structure paś amaro purano kher, namely preposition + nominal group in A-case. It is not excluded that the Rroms departed the Indian plains at the time of the creation of postposition, since the latter in Rromani are only partly identic to those of modern Indo-Aryan languages and before the creation of compound postpositions in India, since there is no trace of such structures in Rromani.
Other categories of the nominal declension are the gender, with neuter disappearing during the MI period and concomitant reascription of former neuter nouns to the two remaining genders: masc. and fem. Ian Hancock explored the fate of these former neuter nouns and noticed that their reascription is almost always the same in Rromani and in the other languages of the same family, what means that the process occured when they were still in contact and gives an indication dating the exodus at the beginning of the second millenium CE. As far as the number is concerned, one may consider that globally dual didn't exist any more in Prakrits.

B) in the verbal group
Unlike the nominal group, which is very similar between Rromani and all other modern Indo-Aryan languages, the verbal group differs drastically between them. Even at its late stages, Prakrits keep a very sophisticated system of tenses and moods, which developed in a very original system in the modern languages, while Rromani lost almost all of the MI system and buil up a totally new paradigm, probably after the exodus, during the contacts with languages of Asia Minor (12th-14th centuries).

Only some endings of the present tense are preserved from Indic paradigm (note that only so-called thematic verbs [√bhŗ] are at stake in this development):

	Sanskrit	*Māgadhi*	*Śaur.*	*Rromani*
sg. 1	pŗććhāmi	puććhāmi	puććhāmi	pućhav
sg. 2	pŗććhaṣi	puććhasi	puććhasi	pućhes
sg. 3	pŗććhati	puććhaï	puććhadi	pućhel
pl. 1	pŗććhāmas	puććhāmo	puććhāmo	pućhas
pl. 2	pŗććhatha	puććhaha	puććhadha	pućhen
pl. 3	pŗććhanti	puććhanti	puććhanti	pućhen

While the length of the ā in the first person sg. and pl. of the present tense explaine the existence of an a in the Rromani ending, there is no clue to account for the evolution of the first person pl. -mas into Rromani -s. The extention of the third person pl. in -anti > -en to the second person may explain the substition of this ending to the original -atha, -adha, -aha forms.
All other evolutions, namely -m- to -v, -t- to -d- > -l and -nti to -n, are perfectly regular.
The other tenses and moods of the verbal paradigm will be explored later in the chapter devoted to the Asian Minor period.
The question of the copula is puzzling, since the Rromani forms are quite specific and don't match with their Indic counterpart, which follow the general verbal paradigm (the symbol Σ in Rromani stands for the various dialectal form of the initial groups of the copula: s-, h-, sin-, isin- etc.):

	Sanskrit	Prakrit	Rromani	Standard Hindi	Awadhi
sg. 1	asmi/bhavāmi	homi	Σ + **om/em**	hũ	hõ
sg. 2	asi/bhavasi	hosi	Σ + **an**	hɛ	hɛ
sg. 3	asti/bhavati	hoti	(**i**)**si/hi[n]**	hɛ	hɛ
pl. 1	smas/bhavāmas	homa	Σ + **am**	hɛ̃	han
pl. 2	stha/bhavatha	hotha	Σ + **en**	ho	ho
pl. 3	santi/bhavanti	honti	(**i**)**si/hi[n]**	hɛ̃	hɛ̃

The semantic difference between the two Sanskrit forms of the copula has been a subject of discussion and most probably they were overlapping each other, asmi meaning rather "to be" (and "to stay" at the third person) and bhavāmi rather "to be(come)". The important element is that both conjugations are based on the general active type and that the Rromani paradigm is almost totally innovative in respect to these ancient forms, while all MI and modern forms originate from bhavāmi, in the sense of "to be" (Hindi infinitive honā), following the aforementioned rules of evolution: initial bh- becomes h- and -ava- coalesces into -o-. The same stem, equally through regular evolution, developed in Rromani into o-, which gave, with the intrusive consonant -v-, the verb ov- (past u- or ondil-), meaning "to become" in most cases but also simply "to be" as suppletive to the copula in tenses and moods where it lacks (future, conjunctive, imperative).
The other Rromani copula, widespread in the Balkan, namely tano, fem. tani, pl. tane (with variants stalo, lo etc...), originates possibly from the Sanskrit stem sthā "to stay".
The Rromani verb is inflected in two diathesis (voices), namely active and medio-passive. However the Rromani medio-passive is not inherited from old Indic but it was probably built up much later than the exodus, already in Asia Minor (*v. infra*). Three other verbal forms are inherited from Indic, namely:
a) the causative -av-, from MI -ape (<-apaya): daral "to be afraid" > daravel "to frighten", ʒanel "to know" > ʒanavel "to inform (let/make know)"
b) the causative -akăr-/-aker-: kalăkărel "to blacken" (< kalo "black"), siklăkărel "to teach" (siklo "well-read, educated") etc... It is encountered in archaic varieties of Rromani, mainly in the Balkan and in Russia and Baltic countries, including Poland. It was formed from the verb karoti

itself "to do", which is attested in such idioms as moghī karoti "to make useless" (< moghī "useless") or viṣami karoti "to make dissimilar" (< viṣami "different, dissimilar").
c) the frequentative -ker-, quite parallel to Indic -kar- and widespread in Central Europe – compare Hindi pūćhkarnā "to bombard with questions" (< pūćhnā "to ask") and Rromani respectively pućhkerel (pućhel).

8 - Common features between Rromani and other Indian languages
Languages don't develop only on the basis of internal changes but also create so called "linguistic unions" while sharing features beyond borders between them and even between linguistic families spoken on neighbouring soils. The most famous linguistic union is the so-called Balkanic one (Balkanische Sprachbund) but such unions are common all over the world. Indianists have defined a dozen of common features linking Indian languages of all origins into one such union. Rromani complies widely to these features:

a) Retroflexion: it originated probably from the Dravidian family and difused into old Indic, since other Indo-European languages don't have this feature, its was more and more frequent in words as time elapsed in Indic (except Vedic retroflex ḷ, which disappeared eventually) and it still characterizes Dravidian languages. As stated above, old Indic retroflexes developed into [r] in Rromani, except when initial, geminated or preceded by an homorganic vowel – then they gave a retroflex [ṛṛ], while intervocalic ṇ developed into [n], [rr].
b) The SOV sentence order is not limited to the Indian area but characterizes also practically all other languages, on the Rroms' way to Europe, namely Persian and Turkish. In spite of this, the order is quite free in Rromani and the SOV order is percieved, at least in the Balkan, not as an Indian but as a Turkish structure (albeit Rumelian Turkish has developed a SVO order under local influence !).
c) Postpositions are also a common feature of Rromani with Indian languages (not only Indo-Aryan) but it affects as well Turkish. The fact that the possessive postposition is inflected in gender, number and cases, is more specific to all modern Indo-Aryan langues, including Rromani. In addition, the existence of a long form of this postpositions (possibly an archaism) links Rromani to the languages of the central group and more specifically the area of Kannauj.
d) Indian languages are allegedly rich in onomatopoiea and so is also Rromani, but many other languages as well.
e) Reduplication is much more common in India languages than in Rromani and the specific intensive adjectival construction implying the repetition of an adjective in diminutive form after its simple form (dilo dilorro "quite crazy") is rather a Serbo-Croatian influence in Rromani than an Indian heritage. In addition, reduplication is so common in Greek (where it is probably even more productive than in India), and to some extend also in Turkish, that it would be pointless to consider Rromani reduplication as an Indian rather than Greek feature.
f) Causatives coined by means of a derivational suffix are quite frequent in Rromani, as in India (v. supra) – whereas the very notion of causative construction is quite alien to many European languages. However, unlike other Indo-Aryan languages, Rromani has no second degree causative ("to have somebody make somebody else do something").
g) Rromani has globally less idioms involving an auxiliary verb than standard Hindi (which uses to the excess the construction noun or adjective + karnā) but more verbs may act in this function. The most productive is del "to give" and one can distinguish cases in which it is fastened to the meaning bearing element (ex.: spìdel "to push", kìdel "to gather" – spi and kid cannot appear alone) and cases in which both elements are loose (ex.: del godi "to think"),while some can be encountered loose in archaic dialects (ex.: del rod "to look for", del phurd "to blow") and attached in others (ròdel, phùrdel). Other auxiliary verbs are kerel "to do", lel "to take", xal "to eat", marel "to beat".

h) Instead of Direct Object marking (by means of a special case in Indo-Aryan and a specific suffix in Dravidian languages), one should rather, in the case of Indo-Aryan and Rromani, speak of two basic morphologic cases, direct and oblique, also called A and B in Rromani grammar. As a matter of fact, there are further parallels in all this group:
- widely similar endings
- use of B-case as a basis for postpositions
- use of B-case in the function of Direct Object for nouns referring to human beings and of A-case for nouns referring to inanimate objects.
- fluctuant border between human beings and inanimate objets, with fluctuant repartition of nouns referring to animals according to their perceived status.
i) The absolutive is a major verbal mood in old and MI languages, which didn't have any system of phrase subordination and used absolutive (to be compared with a kind of past gerundive) to express cause, temporality, concession etc. Absolutive apparently disappeared from modern vernaculars but was reintroduced by scholars into standard languages. As a verbal mood, it dissapeared also in Rromani but it remained as a form expressing the abstract noun derived from the given verb (usually ending in -pen, -ben, -mos < -tvan[na]).
j) The specific construction of non agentive predicates to express physiological, emotional and similar processes, common in India, is encountered as well in Rromani, but it is perceived rather as a Slavic influence than an Indian heritage: dar manqe "I am afraid" (fright to me), avel manqe ʒung "I am disgusted" (comes to me [dative] repugnance), avel man zor "I feel embarrassed" (comes me [B case] strength) etc... In terms of physiological, emotional and similar processes, Rromani presents a specificity, since as a rule the verbs expressing them belong to a specific conjugation, noticeably different of the general pattern. This group of verbs also express often physical movements, there number is limited to some 30 and the paradigm is not any more productive.
k) The lack of verb expressing possession "to have", mentionned by Indianists as an Indic feature, is so widespread all over the world that it can hardly be considered as such. In everyday speach, possession is rendered by the construction: copula + possessor in B case + possessed object in A case:
isi man jekh sevli "I have a basket" (is me one basket)
In the case of possessor expressed by a noun, the latter appears most often in A case (more rarely B case) and it is echoed by a personal pronoun:
o śośoj si les duj bare kana "the rabbit has two big ears" (the rabbit [A] is him etc..)
e śośojes si les duj bare kana "id." (the rabbit [B] is him etc..)
l) Unlike other Indian languages, Rromani has an article which originates from the evolution of an old demonstrative pronouns but was strongly influenced by the Greek article (the only one in the Balkan area to precede the noun).
In addition to these features, one may mention the existence of some 900 lexical elements (stems and affixes).

9 - The relationship between Indic and Persic fund – Persian among Salçuk Turks
The Iranian input into Rromani has been underestimated so far, due to an inappropriate understanding of the Iranian position in the ancient to mediæval world. As a matter of fact, where the Indo-Iranian branch of the Indo-European migrations splitted up after the late Bactro-Margian period (maybe 2,200-,800 BC), Iranian and apparented peoples covered a huge area extended from Serinda (today's Chinese Xīnjiāng) to the Mediterranean sea, including later (1100-1300 CE) the major part of Asia Minor. When in the second half of the first millenium CE various tribes, often known as Turks (litt. "dangerous") or Oğuz (litt. "arrows"), migrated from the banks of the Orkhon river to the Eastern part of the Iranian expansion, they discovered a totally new social system, with strong urban administration and they adopted, along with the new civilisation,

the local form of old Iranian, more adapted to express their new milieu. Therefore the image of Turks living in Asia Minor and borrowing vocabulary from neighbouring Persian must be revised and replaced by the notion of Turks adopting Iranian languages as language of culture as early as the 7th or 8th century CE, while keeping ancient Turkish as they home language, which automatically integrated Iranian vocabulary.

The deportation of the Proto-Rroms from the central valley of the Ganges by the Ghaznavids – a Persian speaking dynasty, to Chorassan, also a Persian speaking area, occured in this context, but the principal background of Persian influence on Rromani was the Salçuk Sultanate of Rūm, in Asia Minor, where Persian remained the official language during more than two centuries, namely since 1071 to 1307 CE.

The most striking Persian element in Rromani is a vocabulary of some 70 to 80 stems, belonging mainly to the nature or domestic economy. The Persian verbal paradigm may account for the identication of the 2nd to the 3rd person pl. of the verb, since a similar evolution occured in Persian:

	literary, archaic	popular
2nd	-id	-en/-in
3rd	-and	-en/-an

Nevertheless, not all evolutions can be explained by Perian influence.

The main input of Persian is the emergence of complex sentences, with subordinated phrases, whereas old and MI languages used widely various impersonal forms, as infinitives, participes, absolutives etc... In this respect, even modern Indic languages adopted some Persian features, like the conjunction of quotation ki, ke "that" (and the use of ʒo for the same purpose). Rromani uses also in this function ka, kaj, so and kë, treated as a Rumanian loan-word, but which can have been influenced by Persian ki, ke or an ealier form (old a developped into i in various Persian words after k). The Rromani conjunction te, after verbs of volition, may also be of Persian origin.

Other Persian structures were introduced in Rromani, like the nominative rection of nouns introduced by words like "full of" "to fill up with" etc... a feature also present in modern Greek and Albanian. However some very pregnant Persian structure, like the expression of possession by means of the ezafe construction, left no footprint in Rromani.

10 - The Armenian element in Rromani

Curiously, the Armenian influence in Rromani is most of the time explained by a round trop of Proto-Rroms to Armenian areas in the Caucase, as if Armenians were not one of the most widespread populations not only in Caucase and Lesser Armenia (Silicia, after the Salçuk victory upon the Armenian Baghratid kingdom) but also all over Asia Minor and beyond. Like the Ossets (an Iranian people descending of the ancient Alans), their lived all over this immense peninsula and were everywhere very active in business, so in contact with all populations, including Rroms. True enough the Armenian influence in Rromani is limited to some 30 loan-words but also probably to some phonological evolutions:

a) the epenthesis of an initial v- in some words beginning with u- or o-: uśt "lip" > vuśt, udar "door" > vudar, ov "he" > vov, vo, oj "she" > voj etc.

b) the sonorization of stop voiceless consonants after nasal in an homorganic cluster of the kind nt > nd (*v. supra*) – a feature widely shared in Asia Minor and the Balkan.

c) possibly the new shape of the consonant system, which changed voiced aspirate stops to their voiceless equivalent, resulting in a system similar to the Armenian one:

	Indic		Rromani		Armenian	
	non asp.	asp.	non asp.	asp.	non asp.	asp.
voiceless	p, t, k	ph, th, kh	p, t, k	ph, th, kh	p, t, k	ph, th, kh
voiced	b, d, g	bh, dh, gh	b, d, g		b, d, g	

As a matter of fact, another language of the area presents a similar system of aspirates, namely Cypriot Greek, where ππ, ττ and κκ are pronounced respectively [p^h], [t^h] and [k^h] (or [c^h/ʧ^h] before front vowels); it is not excluded that at the time of wide Greek expansion in Asia Minor, other Greek dialects of the mainland had this pecularity, but there is no evidence or record of it, only presumptions (yet, Pontic in the North has no aspirate).

Taking on account that the Proto-Rroms' migration from India to Asia Minor didn't last more than two generations, it is quite possible that the loss of the voiced aspirated series and their shift to respective voiceless equivalents occured on the Anatolian soil (note that Western Armenian underwent also a specific restructuration in terms of voicedness in respect to Eastern Armenian). Not that the Armenian system of postposition is to some extend similar to the modern Indian system (expression of possession + postposition, more like Indian compound postpositions

11 - The Greek influence on Rromani

This influence has also been underestimated, due to the fact that the very ancient Greek presence in Asia Minor is surprisingly disregarded by European scholars. As a matter of fact, Greek, alongside with Armenian and Kurd, was the most widespread language all over Asia Minor when the Proto-Rroms arrived in this area and this accounts for the important Greek lexical element (over 200 stems and a few suffixes as -ìtko, -me(n) or maybe -mos) to be found in the core vocabulary of Rromani.

One may distinguish the earliest elements, which didn't take any epenthetic vowel at the end of the loan-words and those which took -o(s) or -i(s), sometimes -e(s). In the first category one may mention skamin (or skamind) "chair" (< Gk σκαμνί "stool", mediæval σκαμνίν) or drom "road" (< Gk δρόμος "id."). It has been attempted, albeit not quite conclusively, to explain these borrowings by the suggestion that the Rroms "discovered" furniture after ages spent on the soil and paved road, sharply differing from the tracks of Asia.

Most other borrowings into Rromani, irrespectively of their origin, exhibit a final -o or -i, sometimes with an -s at the end. The origin has to be found in the Greek mediæval neuter substantives in -ιν [in] (< ancient -ιον), -ης/-ις [is] and -ος [os], which transferred their ending to newly borrowed foreign words (first from Western languages but also soon from Turkish):

early borrowings:

κανονιέρης "gunner"	μπαρμπέρης "barber"
καβαλ(λ)άρης "horseman"	χαλίφης "caliph"
τσοπάνος/τσοπάνης "shepherd"	ντελάλης "crieur public"
βεζίρης "vizier"	ζόρι "force"
σάλι "shawl"	ταλέντο "talent"

more recent borrowings:

πολισμάνος "policeman"	μπολσεβίκος "bolshevik"
κάνναβις/κανναβι "hamp"	μπασκετμπολίστας "basketball player"
καολίνης "kaolin"	μπιφτέκι "beefsteak"
ταλέντο "talent"	ραπόρτο "rapport",

Note that nouns denoting objects may be inflected with or without final -ς, while those denoting male persons always end in -ς.

In the twenties-thirties of the 19th century however, Greek began to borrows foreign words without any epenthetic final added vowel and this practice became a rule in the sixties (with a few exceptions like κήνσορας/κήνσωρ "censor", an early loan without epenthetic ending due to the fact that the stem κήνσος was already present in ancient Greek as a borrowing from Latin and the ending -ωρ also existed).

ρέκορντμαν "champion"	ρεβόλβερ "revolver"
μποξέρ "boxer"	φεστιβαλ "festival"
σεσουάρ "hairdryer"	σέντερ μπακ "back center (sport)"

κοντέρ "storyteller"   μπαργούμαν "barwoman"

The important point for Rromani is that, after a stage of borrowings with no added extra ending, Rromani began to follow the Greek pattern with the addition of -is/-os. The final -s was eventually dropped in most Rromani dialects but remained to date especially in the vernaculars of Bulgaria and Slovakia (postièris "postman") in A case, which is always paroxyton, whereas it is preserved in B case, which is oxyton:
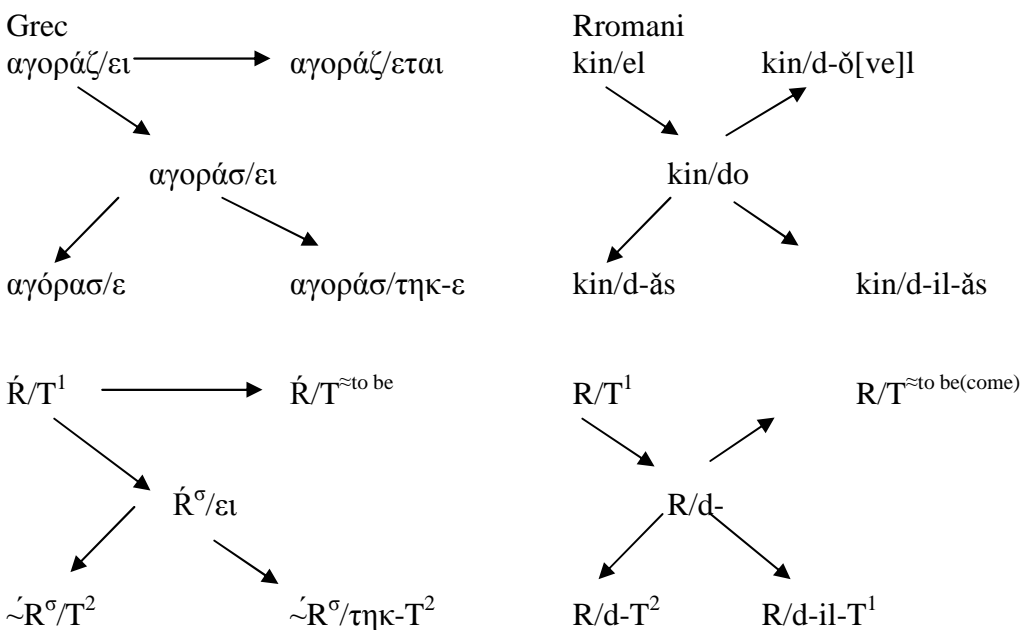
	Bulgarian vernaculars	most dialects
A case	ćobànis, ćobànos	ćobàni, ćobàno
B case	ćoban*is*, ćoban*os*	ćoban*is*, ćoban*os*

The main differences between Rromani and Greek is that Greek didn't loose the final -ς during history on the one hand and that on the other hand Rromani so far still adds an ending practically to all loan-words, even the most recent ones: pùlti "TV remote control".

A final -s is also encountered under Greek influence in the plural of feminine nouns, as attested especially in Bulgaria, but it is difficult to assert if it was first a widely known form of if it emerged locally through analogy and/or Greek influence:

	Rromani			Greek	
	sg.	pl.		sg.	pl.
dir.	**làmpa**	**làmpe**(**s**)	nom.	λάμπα	λάμπες
obl.	**lampa**~	**lampen**~	acc.	λάμπα(ν)	λάμπες

Another major morphological influence is to be found in the verbal paradigm in terms of tense building. Apparently, the MI verbal system was almost all dismantled in Rromani when the speakers arrived and Asia Minor and it was then rebuilt according to a new pattern partly under the influence of Greek which was undergoing a similar restructuration at the same time. Modern Greek conjunction gave up early the dozen of tenses of ancient Greek and only a few of them remained as synthetic forms (other ones being expressed by means of enclitics added to synthetic tenses). There is a broad similarity between Rromani and Greek new verbal systems, as evidenced by the following scheme (we leave aside the imperfect, a synthetic tense in Greek but built by means of the enclitic **sas** "was, were" added to the present tense in Rromani):

Grec   Rromani
αγοράζ/ει ⟶ αγοράζ/εται   kin/el   kin/d-ǒ[ve]l

αγοράσ/ει   kin/do

αγόρασ/ε   αγοράσ/τηκ-ε   kin/d-ǎs   kin/d-il-ǎs

$Ŕ/T^1$ ⟶ $Ŕ/T^{\approx \text{to be}}$ R/T^1 $R/T^{\approx \text{to be(come)}}$

$Ŕ^{\sigma}/ει$ R/d-

$\sim R^{\sigma}/T^2$ $\acute{}R^{\sigma}/τηκ\text{-}T^2$ $R/d\text{-}T^2$ $R/d\text{-}il\text{-}T^1$

Similarities and discrepancies are conspicuous through the following symbols: Ŕ = stem stressed on its last syllable, $\sim R^{\sigma}$ = stem with stress withdrawal on the preceing syllable and sigmatic

feature, T^1 = first system of endings (in Greek: -ω, -εις, -ει, -ουμα, -ετε, -ουν; in Rromani: -av, -es, -el, -as, -en, -en), T^2 = second system of endings (in Greek: -α, -ες, -ε, -αμα, -ατε, -αν; in Rromani: -om/-em, -an, -ǎs, -am, -en, -e), $T^{\approx \text{to be(come)}}$ = ending mirroring the copula (*v. supra*). One should precise that to the South of Danube, the ending of the third person T^2 is not -ǎs in the singular but -o in the masculine and -i in the feminine (basically an adjective ending); in fact we have the same phenomenon in the plural but it is not visible due to the homonymy between the two kinds of endings (genuine verbal in -e and adjectival also in -e).

In addition, one can observe some further similarities in the copula's paradigm, namely:
- a common third person form of the singular and plural of the copula, in both present and past tenses: (**i**)**si**, **hi**, -**j** in the present and sine, hine, sas, has in the past in Rromani; είναι in the present and ήταν in the past, whereas all other Balkan languages distinguish singular and plural;
- the lack of instantaneous future and conjunctive of the copula in Greek and Rromani;
- the concomitant replacement of the copula's future and conjunctive in Rromani by forms of the verb ov- "to become", itself originating from Sanskrit verb bhavāti "to be", through MI hoti (v. supra).
- a sort of paralellism in the formation of the present medio-passive, since the ending of this tense is almost identical in both languages to the present copula in Greek and to the suppletive form ovel in Rromani; as a matter of fact, Rromani integrated the ending of the copula into the medio-passive present, whereas in Greek the ancient copula ειμί [itself corresponding to Sanskrit asmi] had been reshaped already in mediæval times after the ending of the medio-passive verb, which remained unchanged.
- and obviously the same values of medio-passive in Rromani, Greek and Albanian.
The main dissimilarities between the two verbal systems are the following:
- as mentionned above, the imperfect forms are not related between the two languages;
- the center of the system is in Rromani the Past passive Participle which corresponds formally to the Greek participle in -σει but semantically to the Greek participle in -μμενος;
- the future and conjunctive have only one form in Rromani, but two (instantaneous and longlasting) in Greek;
- Finally, it should be pointed out that the Rromani modal opposition between a naked present form on the one hand and the present form + enclitic -a on the other is totally absent in Greek (where the deciduous final α has no grammatical value). In Rromani this contrast expresses somehow different values from one vernacular to the other, including within the same area and between closely related dialects. Since this contrast seemingly doesn't originate from Indian paradigms or from any language in contact, it is likely that it was created inside the Rromani system shortly after the arrival of the speakers in Asia Minor, with eventual semantic specialisation of the forms.

Remaining in the field of morphology, the Rromani article is quite often treated as a Greek borrowing. In fact, its form developed from inherited demonstrative adjectives but their final structure could quite well have been influenced by the Greek article, as suggested by the table below:

Demonstrative pronoun-adjective **akava** "this", **odova** "that" (example from Kelderaś)

	proximal			distal		
	SINGULAR		PLURAL	SINGULAR		PLURAL
	m.	f.	m. & f.	m.	f.	m. & f.
A (dir.)	**kad*o***	**kad*ia***	**kada*le***	**kod*o***	**kod*ia***	**kodo*le***
B (obl.)	**kada*le***	**kada*la***	**kada*le***	**kodo*le***	**kodo*la***	**kodo*le***

The definite article in Rromani and modern Greek

	SINGULAR		PLURAL		SINGULAR			PLURAL		
	m.	f.	m. & f.		m.	f.	n.	m.	f.	n.
A (dir.)	**o**	**i/e**	**o/le**	nom.	**o**	**η**	**το**	**οι**	**οι**	**τα**
B (obl.)	**(l)e**	**(l)e /la**	**(l)e**	acc.	**τον**	**την**	**το**	**τους**	**τις**	**τα**
				gen.	**του**	**της**	**το**	**των**	**των**	**των**

The article itself differs by no means from the inherited demonstrative endings, once the final -a has been deleted from the feminine sg. (note that the masculine sg. ending -**o** developed from the evolution -**ava** > -**o**, a form still in wide use in other dialects; accordingly the latter includes also a final -**a**, while the preceding sound -**v**- [correlated with **u/o**] embodies the masculine feature of the deictic). Note that several forms of the affirmative interjection are also related to these forms : **va**, **ova**, **oja** "yes" (also **e** and local borrowings like **da**, **po**, **tam** [< tamam] etc.).

Among the various other common points which may be pointed out between the two languages, the following ones are of special significance:
a) the use of the adverb of location kaj "where" with the value of relatice pronoun, like Greek που;
b) the use of double accusative, actually not restricted to Greek, but quite widespread in this language and to some extent in Rromani (ex.: dikhlŏm tut suno / σε είδα όνειρα "I dreamt of you");
c) the use of nominative as a complement to words meaning "full" (*v. supra* in Persian).
d) the frequency of nominal sentences;
e) suffixes like -isar- (to coin in Rromani verbs on the basis of loan-words), from two Greek suffixes -ιζ+αρ-;
f) the structure used to express the age: numeral in B case + berś "year" + possessive postposition, corresponding to genitive plural in Greek.

As a matter of fact, the most significant Greek influence is probably to be found in the phonologic area, since the sonorisation of soundless stop consonants after nasal in Greek, as well as in other languages of Asia Minor, probably stimulated the same phenomenon in Rromani (*v. supra*). The apheresis of initial unstressed vowels possibly also was generated partly by a Greek influence, although this evelution is currently attested mainly in the Northern branch of Rromani (vernaculars of Poland, Russia and Germany – Sinto) than in the Balkan themselves (except Albanian Mećkar, which is in its turn influenced by popular Albanian).

12 - European influence in Rromani
Greek has been treated here as a language of the Byzantine and later Ottoman empires since it actually belong to this area during most of the time of contact with Rromani and its influence affected all Rromani varieties, as spoken all over the world. There is no point here to treat one by one the influences of European languages, since all had partial impacts upon Rromani, most of the time very dissimilar between them and requiring very specific and detailed explanation of every single language involved.

13 - The formation of the main Rromani dialects
A dialect is defined in linguistics as a group of vernaculars which share a series of similar linguistic – phonemical, phonological, morphological, lexical or other, features and differ in this from other groups of vernaculars. The imaginary line separating groups of vernaculars (dialects) is called an isogloss. In order to define dialects, one has first to identify the possible features which can be relevant for this classification and after that make a hierarchy among them in terms

of relevance: it is clear that an evolution that can occur any time in any language (like palatalisation of consonants before front vowels or reduction of vowel aperture), is much less relevant for dialectal distinction than a very specific evolution, not likely to appear spontaneously in whatever language.

A) The dialectal analysis of Rromani has lead to the conclusion that it arrived as a unique language to Asia Minor and that the first division (isogloss) occured probably on its soil (although an earlier date is not excluded), being the contrast between o and e in the first person of the copula (*v. supra* Σ-**om** ≠ Σ-**em**) and of the ending of the past of the verbs (**phirdom** ≠ **phirdom**, **phirdŏm**, **phirdŭm** "I walked"). This contrast (materialised by an isogloss) is indeed relevant due to two reasons:

- it is a very rare contrast among all the languages of the word and therefore it is not likely to appear at any time;

- it is accompanied by a series of some 15 other contrasts which make up a body of isoglosses separing sharply dialects, called in Rromani respectively O-superdialect and E-superdialect.

From a strictly dialectological perspective, the **o**/**e** isogloss between the two superdialects of Rromani corresponds to the following contrasts:

a) the **o**/**e** contrast occurs also in the stem of the verb "to leave" **muk**[h]**el**/**mek**[h]**el** (*mokhel is not attested) and in the plural of the article in A case: **o**/[l]**e** ; it should not be excluded that the contrast **tordŏl**/**terdŏl** "to stand" might be ascribed to the same phenomenon, albeit the scarcity of these forms in present day Rromani doesn't allow to identify definitively their dialectal belonging;

b) the forms **daj** "mother" and **ćhaj** "daughter" of O-superdialect correspond to **dej** and **ćhej** in E-superdialect;

c) the ending **-ni** (or **-i**) of O-superdialect corresponds in most cases to substantives (and sometimes adjectives) ending in **-j** in E-superdialect: **pani** "water" – **paj**, **khoni** "grease" – **khoj**, **endani** "tribe, clan" – **endaj**, **balani** "sink, handbasin" – **balaj**, **kuni** "elbow" – **kuj**, **ʒeni** "female" – [ʒuvli]-**ʒej** , **ćeni** "earring" – **ćej** etc...

d) clusters **dl** and **tl** characterise O-superdialect, whereas their velar equivalent **gl** and **kl** are typical for E-superdialect: **gudlo** "sweet" – **guglo**, **katli** "spindle" – **kakli** etc...

e) forms **tikno** and **cikno** "small" belong respectively to O- and E-superdialects;

f) E-superdialect add quite often a prosthetic **a-** to words like **aśunel** "to hear, to listen to", **abăv**, **abav** "wedding" vis-à-vis **śunel**, **biav** etc...

g) O-superdialect has preserved a comparatively wide use of the old Indic comparative ending in -**eder**, mainly for frequent adjectives, whereas it is almost entirely replaced by an analytic form in E-superdialect: **terneder** "younger" – **maj terno**, **bareder** "bigger" – **maj baro** etc...

h) the third person of both sg. and pl. of the copula in the past is in O-superdialect **sine**, with negative form **na sine** (**na hine**), whereas the form is **sas** (var. **has**), with negative form **nas** in E-superdialect;

j) the negative adverbs **na** on the one hand and **ni/i** on the other belong respectively to O- and E-superdialects (**na daral** – **ni daral**, **ći daral** "s/he is not afraid");

k) the prohibitive adverbs **ma** on the one hand and **na** on the other belong respectively to O- and E-superdialects (**ma dara** [locally **na dara**] – **na dara** "don't be afraid";

l) the past of verbs in -**a**- stem ending is built with -**andil**- on the one hand and in -**aj**- respectively to O- and E-superdialects;

m) the integrating morphemes of verbal loans sound -**in**-/-**on**- on the one hand and -**isar**-/-**osar**- on the other (with local variants -**iz**-, -**oz**- etc.) respectively in O- and E-superdialects.

n) the postposition of possession has in many vernaculars of O-superdialect both a short form -**qo**, -**qi**, -**qe** and a long form -**qoro**, -**qiri**, -**qere** (with variants -**qero**, -**qeri**, in Bulgaria [where it is often the only form], Macedonia, Kosovia, Northern Hungary and Slovakia and

-**qro**, -**qri**, -**qre** in Baltic countries, Russia and Germany – Sinto) but E-superdialect has only the short form (about the long form of this postposition in the Indian vernaculars of the Kannauj area, *v. supra*). Some Rromani dialects of Greece and Albania have also an alternative long form, used in a very specific context, when itself inflected with a secondary postposition after another noun to which it is postponed for the sake of emphasis: **e ćhavençar e Dudaqerençar** "with Duda's children" (litt. "with the children, those of Duda"), as emphatic alternative to **e Dudaqe ćhavençar**.

o) beside the causative morpheme -**av**- (as in **daravel** "to frighten" < **daral** "to be afraid", *v. supra*), Rromani has another causative appearing, according to the dialect, in both long and short form, namely -**akăr**-/-**aker**- and -**ăr**- in the O-superdialect but only in the short form in E-superdialect. As stated above, the long form is characteristic of mainly in the Balkan and in Russia and Baltic countries, including Poland.

p) the formation of numerals above twenty (with -**u**- as a connector [of Iranian origin] as in the 11-19 numerals/with **thaj** as a connector).

q) the distribution of the apical trill [r] and the various realisations originating from the Indian retroflex (v. supra), as well as the pronounciation of the various forms originating from the Indian group -ṇḍ- differ between O- and E- superdialects.

r) the lexical difference between the Rromani superdialects has often been overestimated by researchers, but actually it concerns a limited number of lexemes, like for example the following (in O-superdialect): **ćulal** "to drip", **puzgal** "to slip", **xalŏl** "to understand", **parrunel** "to bury", **muśtek** "content of the palm (of the hand)", **tamlipen** "darkness", **tasia**(**ra**) "tomorrow", **dumo** "back (of the body)" etc… corresponding respectively to (in E-superdialect): **pićal**, **istral**, **hakărel**, **praxosarel**, **pàlma**, **tuniàriko**, **tehara**, **zeja** (**dumo** exists also in E-superdialect but means "shoulder") etc…

In geographical terms, the O-superdialect is widespread all over Europe according to comparatively clear migrations: Rroms first arrived from Asia Minor in the beginning of the 14th century to the Balkan (we disregard here some individuals or families who could have arrived earlier in Venice but melted into the local population). A good part of them settled in the Balkan but others went Northward to the Carpatic basin (some of them from Bulgaria remaining as slaves in the two Danubian principalties of Muntenia and Moldova as early as the middle of the 14th century) and then a significant group turned to German-speaking countries in which the Sinto subgroup lost contact with other, migrating more and more to the West, while others moved to Poland, the Baltic countries and Russia. Smaller groups migrated also from the Balkan to the South of the Russo-Ukrainian area. These groups are to date all speakers of the O-superdialect. In the same time, some groups moved to Spain, some by sea from Greece, some on dry land and others simply crossed the Adriatic sea to Southern Italy, when it was a possession of the Aragón crown. As stated above, it is not clear where and when the split between the two superdialects occured, but one can observe that the speakers of the most archaic forms of this superdialect are mainly concentrated in the central Balkan.

B) The second significant split among Rromani vernacular is called as "mutation of alveolar fricatives", a term meaning simple the evolution of the fricatives consonants ħ and ʒ into a soft sound, without plosive element (as rendered in English and International Phonetic Alphabet):

basic sound (phoneme)	Indian letter	pronounciation before mutation		pronounciation after mutation	
		English	IPA	English	IPA
ćh	छ	ch'h	[tʃh]	very soft sh	[ɕ]
ʒ	ज	dj	[dʒ]	very soft zh	[ʑ]

This new evolution created an isogloss within the O-superdialect, separating the main trunk, unaffected, from a small group of mutational vernaculars spoken mainly in Eastern Hungary, Northern Rumania and adjacent areas of Ukraine, but within the E-superdialect as well, dividing

it in two numerically almost equal parts. The newly created structure may be represented as follows:
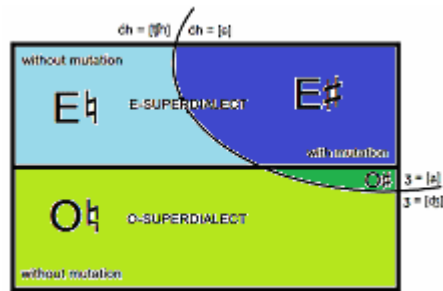
The origin of this mutation is not clearly identified, yet a comparable, albeit not entire similar, mutation can be observed in some peripheral dialects of Rumanian, especially Banatian but also Moesian, spoken to the South of the Danube in Serbian territory. The mutation overlaps with some other features, mainly lexical, but the body of isoglosses ssociated with it is much lesser than the ones associates with the **o**/e isogloss.

These four dialects are usually refered to as O-bi (O-Superdialect without mutation), O-mu (O-Superdialect with mutation), E-bi (E-Superdialect without mutation), E-mu (E-Superdialect with mutation).

As a matter of fact, if the distinction between O-bi and both forms of E-Superdialect is quite sharp, the position of O-mu is much more unclear. If we compare the features mentionned above as exhibited in these various dialects, we can set up the following table with typical O features on a yellow background, typical E features on a purple one and specific evolutions on a green background. True enough, there are less O features than E features in O-mu, but the former are more archaic while the latter correspond often to the loss of a linguistic element (synthetic comparative, long form of possessive postposition and causative etc...):

O-bi	O-mu	E-bi	E-mu
phirdŏm, phirdom	**phirdŏm**	**phirdem**	**phirdem**
o Rroma	**o Rroma**	**e Rroma**	**(l)e Rrom(a)**
mukhel	**mukhel**	**mekel**	**mekel**
tordŏl	**tordŏl**	non extant	**terdŏl**
daj, **ćhaj**	**dej**, **ćhej**	**dej**, **ćhej**	**dej**, **ćhej**
pani	pai	**paj**	**paj**
-dl-	-ll-, -l-	-gl-	-gl-
tikno	**cikno**	**ćikno**	**cïkno**
biav	**biav**	**abav**	**abǎv**
śunel	**śunel**	**aśunel**	**aśunel**
terneder – po terno bareder – po baro etc..	**maj**	**maj**	**maj (-eder)**
sine, hine	**has**	**sàsa, sèsa**	**sas**
na sine (**na hine**)	**na has**	**navlòsa, navlìsa, navlèsa**	**nas**
na daral	**na daral**	**ni daral**	**ći daral**
ma dara!	**na dara!**	**na dara!**	**na dara!**
-andil-	**-aj-**	**-aj-**	**-aj-**
-in-/-on- (locally -**iz**-, -**oz**- etc.)	**-in-/-on-** & -**isar**-/-**osar**-	**-isar**-/-**osar**-	**-isar**-/-**osar**-
-qo, -qi, -qe & -**qoro, -qiri, -qere** (locally -**qero, -qëro**)	**-qo, -qi, -qe**	**-qo, -qi, -qe**	**-qo, -qi, -qe**
e ćhavençar e *N***esqerençar** emphatic, besides **e** *N***esqe ćhavençar**.	**e** *N***esqe ćhavença**	**e** *N***esqe ćhavrrença**	**le** *N***esqe ćhavorrença**
caus. morph. -**av**- & -**akǎr**-/-**aker**- & -**ǎr**-	caus. morph. -**av**-	caus. morph. -**av**-	caus. morph. -**av**-
yx [y > 2] = y-**u**-x	yx [y > 2] = y-**taj**-x	yx [y > 2] = y-**haj**-x	yx [y > 2] = y-**thaj**-x

[ɽ], [ɾ], [r]	[rr]		[rr], [ʀ], [ʁ], [ɣ], [x]
[ndr], [ŋgr], [rr], [r]	[nrr]	[nr], [rn]	[˜r], [nrr], [˜γ], [˜x]
ćulal	pikăl	pićal	pićal
puzgal	istral	istral	istral
xalŏl	hakărel	hakărel	hakărel
parrunel	praxosarel	praxosarel	praxosarel
muśtek	pàlma	pàlma	pàlma
tamlipen	tuniàriko	tuniàriko	tuniàriko
tasia(ra)	tehe	tehàra	tehàra
dumo	dumo	dumo	dumo

The evolution seems to have followed the scenario suggested below:
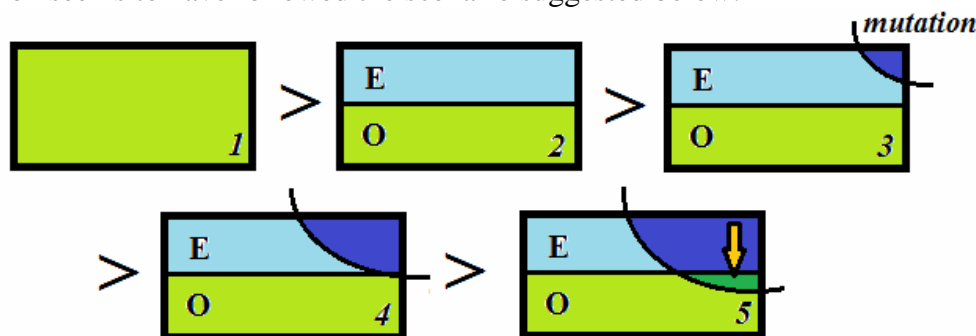
1 > 2 : split into two superdialects
2 > 3 : beginning of the mutation process in the E superdialect
3 > 4 : extension of the mutation to a high number of E vernaculars and some O vernaculars
4 > 5 : strong influence of the E forms onto the mutational O dialect (vertical arrow).
This last influence is operating further in Hungary due a higher prestige of the E-dialect forms, perceived as urban (Budapest, Pécs etc...) contrasting with the O-dialect forms perceived as rural. Some isoglosses (**phirdŏm/phirdem**, **o/(l)e** plural article, verbal formant -**in**-/-**isar**-, **ʒuvel** [for **ʒivel**]/**trail** etc.) confirm the inclusion of O-mu in the O superdialect, whereas others (mainly lexical, therefore comparatively superficial) link it to the E group.

C) Beside these evolutions, which can be described as classical in terms of dialectology, some vernaculars underwent very specific developments, leading to the emergence of so-called Para-Rromani varieties. These are of two kinds:
a) peripheral dialects, which followed specific evolutions which, albeit often drastic, were always progressive and continuous. This is the case of Northern Sinto, spoken mainly in countries historically of Germanic languages (Austria, Germany, Alsace in France) and of Southern Sinto, spoken in the North-East of Italy. Despite of their common name "Sinto", there is no evidence of specific links between the two dialectal groups. Northern Sinte also live in central France, but most of them are deculturated to French – whereas those of Alsace preserved quite well their mother tongue.
Other significant peripheral dialects are Abruzzese (Italy), Dolenjski (Slovenia), Welsh Kalo (Wales, now extinct), Kaalenqi ćhimb of Finland (also extinct since WW2 and Erromano of Argentina. The total of speakers of all peripheral vernaculars is probably less than 100.000 persons, out of 15 millions Rroms at the world level.
b) so-called paggerdilects, which emerged in very specific social conditions: due to merciless persecutions (especially in Spain but also in England and Austro-Hungary), some groups switched to the mainstream language as home language and the children grew up speaking only a domestic form of respectively Spanish, English or Hungarian. However, adults were use to speak some Rromani between them, or at least to introduced some vocabulary for various purposes into the local language. As a result, adolescents could "capture" some words of the ancestral tongue, which they introduced on purpose into their practice of the local language, formerly for reasons

of secrecy, but later more and for as an identity marker. The following examples can give an idea of this kind of speech:

Rromani	Mangav/kamav te ʒav and-o gav te xav xarri manro.
Spanish Kalo	Camelo chalar al gao pa jallar gulu tato/maro.
	[ka'melo ʧa'lar al 'gao pa xa'jar 'gulu 'tato/'maro]
English Paggerdi	I kamel to jall to the gav and haw some marrow.
English	I want to go to the village and eat some bread.

The example is indeed amusing but it doesn't mirror the real use of these paggerdilects in true life. In fact, genuine speakers don't use more than one or two words scattered in a few minutes of speech, just for the sake of collusion, fun or ethnic pride and all attempts of translating parts of the Bible into paggerdilects are exercices in futility. No wonder that none of them was carried out by Rroms.

Speakers of Kalo (there are three main varieties of Kalo: Spanish, Catalan, Euskara – now extinct, with local varieties) are around one million in the Iberic peninsula and Latin America, and those of Anglo-Rromani possibly fifty thousands. Be it as it may, all Rromani peripheral vernaculars and paggerdilects developped from the O-superdialect.

C) The third way of evolution is related to the recent massive oblivion of Rromani vocabulary, especially among some younsters, due to new social contexts, with the loss of traditional objects (tools, food, products etc.), notions and values, concomitant with the intrusion of new technological and administrative objects. In some cases, even if the object still exists, it is so rarely mentioned in everyday life that its name is supplanted by the mainstream language equivalent. However, one should not overesteem these factors, because the European trend to use specialised terminologies in daily speech has not penetrated most of Rromani families and one can hear all over Europe millions of Rroms speaking quite well their particular vernacular. In addition, the Rumanian government has promoted lessons of, and in, Rromani which beneficiate to more than 30,000 pupils and students per year in this country.

To sum up, one should distinguish three types of differentiation among the Rromani and para-Rromani idioms:

a) the strict dialectological division, with two crucial isoglosses, namely the O/E contrast (accompanied by a lexical differentiation of a few dozens items) and the mutational contrast. These contrast are not damaging for the unity of Rromani, because O/E concerns but a reduced segment of the language (one verbal ending, the plural of the article and some associated but pubctual features), whereas the mutation is not always perceived by the ear; in addition both are quite systematic and rigorous. One should add some punctual lexical discrepancies, not related with the aforementionned isoglosses and involving a very low number of lexemes: korr/men "neck", gilabel/bagal "he sings" etc…

b) the socio-linguistic level, with two major types of scenarios for the formation of peripheral idioms and paggerdilects. Their users are not very numerous (some 10% of the total number of Rroms) and as a result the unity of Rromani is not very much affected.

c) the level of local or regional oblivion of lexical items (including lack of develop-ment due to life conditions: rural surrounding language poorer than Rromani, marginalization etc… This does not concern the language itself, just the way it is used in certain areas, and therefore – if an efficient didactic effort is developed in a context of language valorization – oblivion could be compensated by lexical reacquisition and the problem could be solved.

14 - The so-called "dialectal issue" and the use of Rromani mother tongue among Rroms
The so-called "dialectal" disparity of Rromani should be renamed "oblivional" disparity because two Rroms of different dialectal backgrounds understand each other, while each speaking his

Rromani dialect, far better than two Rroms of the same native dialect, who have not properly acquired their mother-tongue. This is related to the fact that the properly Rromani (Asiatic) element in Rromani is amazingly uniform in all dialects, and this fact points at the uniqueness of language of the Rroms' Indian ancestors.

The following comparison has been used to express this:

- the core of the Rromani language is basically the same for all dialects, as the human body is basically the same for everybody (as a result, the terms of anatomy are shared more or less by all dialects, since they refer to common natural concepts);

- the European borrowing differ among the Rroms, just like garments differ among countries (as a result, the terms referring to non-Rromani life (garments, administration, food etc.) differ among Rroms, since they refer to artificial concepts);

- when a Rromani word has been lost, it is replaced by a non-Rromani one, just like an organ/limb missing is replaced by an artificial one but this is by no means a model of life;

- when other dialects can supply a word missing, this solution is preferable, just like transplant is preferable to artificial limbs – but it needs more sophisticated skills.

The first four questions have been answered and we can conclude that, if the common Rromani vocabulary, gathered all over Europe and sorted according to the phonological rules of the various dialects, is circulated again (used in public life and taught to persons who have forgotten it), there no reason to claim that Rromani differs from other European languages in terms of dialectal splitting up.

The second main issue is related to the actual use of Rromani among Rroms. Before discussing this point, one should recall that most immigrant languages are totally lost within four generations. Quite often one can observe young Albanians, born in Albania or Cossovia, speaking French rather than Albanian among themselves. On the other hand the vigor of Rromani, after almost one thousand years of migration, arouses the unanimous admiration of all observers: "Every visit in a Rromani family shows that the children learn first Rromani, their mother tongue, and only then the language of the host country" (Reinhard 1976:III).

15 - How can we give an objective assessment of the situation?

Although it is conspicuous that at the European level Rromani is far more vivid in everyday life than stated in many so-called "objective descriptions", signs of decline have indeed been growing alarmingly in the last decades. One should accordingly explore the reasons why Rromani is in decay, probably sharing the fate of most minority languages in wide urban settlements. Sociolinguists have pointed out that the greater the degree to which an exiled population consists of mixed social backgrounds, the stronger and the longer it will carry on transmitting its original language. The manifold social structure of the Rroms' ancestors when they left India can account for the phenomenal survival of Rromani – as opposed to the situation of most other migrants' languages. One should emphasize that Rromani successfully overcame the drastic changes of cultural context when the Rroms were deported from northern India to Afghanistan and Persia, and later moved to Asia Minor and various European countries, where each time they faced totally unknown civilizations. As emphasized above, the mixed character of this population was probably a factor of preservation. The fact that most Rromani communities in Europe are now reduced to homogenous poor groups quite similar to other migrants' communities at the time of their arrival put Rroms in a similar risk of linguistic acculturation. However the will of forwarding Rromani to upcoming generations is widely expressed in all declarations.

Beyond the gradual weakening of the Rromani language presence and the Rroms' declared commitment to preserve it, it is essential to emphasize that language survival is far less a matter of declaration than of motivation. Since language, as a social phenomenon, has two mains faces: communication and identity, the motivation to keep it alive may be twofold. As an expression of identity, it is supported by everybody aware of this social function but as a mean of

communication, some Rromani speakers find it indeed inappropriate to convey modern messages – an opinion originating from several misunderstandings.

16 - The real social functions of a language and family response

The first task of linguistic practice is not to convey highly sophisticated information but to create a friendly and warm space of *divàno* between people who want to express their feelings to each other, but also all kinds of common-place utterances, worth nothing in terms of information but highly significant for the community's psychological comfort. One should not forget the fact that children learn their mother tongue and culture not from sophisticated scholarly utterances but from family chit-chat. All dialects of any language are suitable for this purpose. The problem arises from the fact that, under the influence of schools and media, the mainstream languages have recently developed a kind of pseudo-intellectual slang, even in the sphere of everyday life. In addition school and media circulate the image that language is a matter of terminology – and learning a matter of school ! Minorities tend to imitate the style of their surrounding majority language, but they do not succeed because their mother tongue has not enjoyed the same special care, which has developed a so-called "high" style in official languages. This leads minorities to under-esteem their mother-tongue and to switch more and more to the majority language. This occurs because they have ceased to think in their mother-tongue and it is far easier to express the majority way of thinking in that language than in their mother-tongue, anyway totally ignored, if not despised, by the majority society spaces of activity: media, school, movies, public places, shops, sporting, games, etc... which leads them to think in the language of all these activities. The family circle constitutes a kind of private sanctuary hosting the last stages of use of a linguistic relic.

In this respect, one can understand that many parents do not perceive the values of Rromani (even if they claim they want it to be transmitted to their children – by others):

● Rromani needs (and has the right) to be formally valorized in public life and accessible at any time of the day: media, school, games, sporting, etc. on an equal footing with the main language(s), which brings also revalorization of the Rroms themselves. Mainstream societies, but also Rroms themselves, have a duty toward the truth to publicly restore respect not only for Rromani but also for all elements of the Rromani heritage, which have to be treated as belonging to a non-territorial nation, not to a formless amalgamation of socially marginalized groups.

● Awareness-raising campaigns should be conducted in school and the media about the importance of all mother-tongues, among others for human feelings of internal solidarity; the idea that language is not only an instrument of communication but also of identity and intellectual development has to be taught to everybody. In this respect, the importance of *gnossodiversity* (diversity of cosmovisions) beyond *glottodiversity* (linguistic diversity) should be pointed out, as well as the role of language in expressing non-material heritage.

● Education in Rromani should be provided to teach how to express modern messages in a more accurate way in Rromani (to raise up from "the analysis were bad" to more accurate "his blood sugar level is so much %" – true enough, this involves also minimal education in physiology, but also in administration, law, politics etc. This would be true empowerment).

● At the same time, modern terminology should be presented as a device of secondary order, as compared with the genuine Rromani expressivity in terms of images, typical lexical resources, proverbs and similar spiritual wealth. The dogma that the modern urban technological and administrative way of expression is a universal ideal should be combatted and the worth of human-to-human communication promoted instead.

● The lack of formal education in the mother-tongue leads to *diglossy*, which means that the mother-tongue is viewed as an instrument devoted to express less and less acurately a sinking world, while the host language conveys all the positive values of modernity, social integration and success. This split leads to the death of the minority language, even if it can go through a

stage of artificial respiration thanks to school classes addressing children who have already lost any native proficiency in their former mother-tongue.

One should at this points think about why language communication is so effective: we may use one word of a few phonemes and understand immediately its meaning: "dog", "house", "son", "father" etc… just because such a group of phonemes has been associated through education to the object concerned. In the cases of these words, the meaning is simple and immediately accessible but for more sophisticated concepts, each culture first creates the image of the concept itself before expressing it through a set of phonemes according to pretty strict rules of derivation, analogy, borrowing etc. This explains the efficiency of language communication and why at the same time language is the inner mirror of our society and cultural references. In this respect, the affirmation of all mother tongues, and for us especially of Rromani, is of the upmost significance. Accordingly, giving up one's language is a response of naive people who are aware only of its informative function (and of its alleged incapacity to fulfill it – and indeed, as they conclude, if it is totally inappropriate, why should they transmit it to their children?). It does not take into account its power of mirroring an entire universe; this sacrifice misleads them to a foreign world they need years, maybe generations, to be integrated in and at the same time, they lower the level of sophistication of the newly adopted language – as it has been evidenced for foreigners' English, which developped into the non-language called Globish. The strategy set forth above, including education in how to perceive language(s), provides a powerful motivation for an effective use of mother-tongue in all circumstances of life. Classes of linguistic recovering have only a symbolic function and can by no means preserve a language alive if the other prerequisites are not met. It is meaningful to observe that a lot of money is devoted to such futureless classes, whereas nothing is done to keep healthy and develop Rromani where it is in regular use as a home language – or to say it other terms, such classes are useful only as a collateral measure and if the native speaking population constitutes a solid reference contingent enjoying the four basic measures developed above. This is a reply to two further questions.

17 - The standardization/modernization issue

Here again we are facing quite confused concepts about the idea of standardization.

Some people still stick to the romantic conception that action upon languages is impossible. Modern linguistics has evidenced that "there exist no 'natural languages', free of any regulation or of any normative process aiming at meeting somehow the needs of their linguistic community […]. As a matter of fact, either at the micro~ or at the macro-linguistic level, language building inevitable and all degrees are possible". Therefore the idea of "improving", "engineering", "standardizing" or "modernizing" Rromani should not be rejected *a priori* as many observers do. They insist on keeping Rromani outside any evolution (except lexical impoverishment, which they admit as a fatality) but they would never accept this for their own everyday language – hiding their discriminatory approach under the colors of respect. Yet it is now clear that if you do not enlarge the expression abilities of the language, you are condemned to use it only for trivial purposes and you create *diglossy* leading to language total extinction or symbolic fossilization (with possible mascotization), which postpones the final outcome but does not change it. The real problem is not "if" but "how" it is possible to have an effect on a language in order to help it optimize its social roles of communication and identity.

Many people mix up standard language and written language. These are two different concepts and we will deal further with *graphization* (*means of writing, созданые алфавита*) of Rromani. When thinking of standardization, they imagine a unique model, as in most "established" languages. In Rromani on the contrary, the traditional feeling of mutual respect among various *endaja* have lead to the very democratic London decision that "no dialect is better than anyone else but we need an international form of language allowing us to understand each other in international conferences and literature" (First Rromani Congress – London, 8 April 1971). So

apart some bystanders who claim that there is no need, no profit or no possibility of standardization in Rromani (Rroms are extremely rare, if any, in this camp), other people are divided between those who want a unique model imposed for all Rroms of their country, as in majority languages (these usually lack any European perception of the Rromani nation) and those who yearn for a flexible European Rromani language, respectful of dialectal cultural riches but easy to use at a wider level of communication.

Some people still believe that Rromani dialects are so dissimilar that a common language is a dream. As a matter of fact, a European Rromani language already exists in the mouth of people having a good command of their native dialect. In the light of the rectified definition of dialects, as given above, of a systematic review of the Rromani *endajolects* and of a clear concept of their mutual relationships, one can conclude that almost all of them are suitable as parts of the basis for common Rromani – provided that one uses their non-forgotten variety. Only peripheral groups (like speakers of very atypical south Italian Rromani, Finnish *Kaalenqi ćhimb* or Welsh Rromani – both now extinct, *vide supra*) show out-of-the-way features but they represent hardly 2-3 % of all speakers. The method of linguistic elaboration consists in:

● collecting as much as possible of all the genuine vocabulary and forms of Rromani all over Europe, including local items (except provincial loan-words, which break mutual intelligibility without bringing a cultural benefit); this task has been carried out by now probably for up to more than 99% of vocabulary from interviews and publications – sometimes very old ones; For example the word **berno** (masc. noun) "circle" has been found in a 16th century Latin text; other example: the word **trom** (fem. noun) "boldness, courage", now forgotten, has been evidenced in a letter written by Radics Lajos from Miskolc to Archduke Joseph von Habsburg in 1888 (while the verb **tromal** "he dares" is widely known).

● sorting this material after the various dialects and looking for equivalents in others;

● considering inter-dialectal borrowings if possible, but only in cases of lexical gaps, since dialectal consistency is encouraged;

● considering resources like derivation, reutilization of obsolete words or semantic extension to widen the language's abilities of expression, as needed by actual present day communication but avoiding to follow, every time it is possible, the strict pattern of foreign models; this method is advisable when all European languages have different words for a specific modern object and it is impossible to chose a pan-European cover-term.

● considering the benefit of borrowing foreign words, mainly for notions related to technical spheres with no emotional dimension; as far as these spheres are concerned, it is common sense to produce common neologisms for all the Rromani varieties: if all Europe says planèta for "planet", there is no point to say bojgòvo in Hungary, just because Hungarian for "planet" is *bolygó* [bojgo]. Be it as it may, borrowing is a natural necessary phenomenon in language evolution allowing getting free of the conceptual ghetto of the past.

● avoiding ambiguous borrowings, especially when they create problems in communication: in some dialects glàso means "glass" (< Germ. *Glass* "id.") and in others "voice" (south Slavic *glas*); it is rather unproductive to use glàso instead of Rromani taxtaj "glass" and krlo "voice". The same may be said for nìpo "people" (< Hung. *nép*) and "grand-son" (< Alb. *nip*). Interesting enough, the over-whelming majority of such ambiguousness is due to loan-words, not to Rromani inherited items.

● checking the given neologisms have an appropriate morphological pattern;

● proposing the concerned forms to wide circulation, with explanation if the context is not sufficient to make the meaning clear, keeping in mind that only practice can confirm the use if specific expressions. This is a major aspect of language affirmation, since it is of no benefit to propagate words if the notions standing behind them are not defined: not only "new" (or "modern") concepts of law, medicine or journalism but also traditional Rromani cultural concepts, which more and more often have lost their natural way of transmission.

The IRU Commission for language and linguistic rights has been active in this field for more than 20 years, through cooperation of dozens of members. However the results of this collective pan-European work is under-esteemed due to the lack of financial resources to make them know and to the obstructive attitude of some "friends of the Rroms" who consider that Rromani will loose its "Gypsy" identity if its European dimension is restored and promoted in addition to local varieties. Keeping Rromani locked in it "genuine" former rural form is an incitation to use it only for songs and folklore and to speak majority languages about important subjects of life.

In addition, one can see quite odd projects of "standard Rromani", improvised here and there but unsuitable to actual use, mainly for the following reasons:

● The promoters of such projects do not have the patience, will and competence to check the real resources all over Europe for as many dialects as possible. Instead of elaborating the Rromani language on the basis of proficient speakers, they take as initial corpus the knowledge of a random speaker (often a local self-proclaimed leader who, due to his biography and low needs of communication, has a poor command of Rromani) and try to reconstruct the whole language out of his scarce remainder of Rromani, while borrowing massively from neighboring languages and sticking closely to mainstream language and style of thinking.

● If such resources (vocabulary, expressions etc…) are supplied to them, their refuse to take them into consideration under the pretext of dialectal chauvinism but in fact chiefly out of laziness. They view the European dimension of Rromani as irrelevant or just believe they can force their construction on millions of Rroms. By doing this each of them contributes in splitting an existing language into feeble individual projects of idioms remaining to be constructed on uncertain basis.

● Crude grammatical mistakes are even quite common in their speech: kodo buti "this work" (kodo is masc. but buti is fem.), na śaj "he cannot" (correct form: naśti), na si "it is not" (correct forms: naj, nane, nanaj, more seldom naj si). When commented on, these neo-speakers just pretend it is their dialect, which is but another manipulation of the word "dialect", increasing the erroneous impression of a dialectal split up of Rromani. In this case, one should rather speak of "fantasiolects".

● In many cases they do not need great accuracy in their discourse. They just copy empty main language declarations, as one may observe in numerous associative meetings, e.g.: Anda kodo kritićno kontèksto, amaro sociàlno projèkto śaj popravil i ekonomìćna situàcia e Rromenqi thaj lenqe problème (anda, kodo, amaro, śaj, thaj, lenqe "in", "this", "our", "may", "and", "their" are Rromani & popravil "improve" is Slavic – no need of translation for the rest of the sentence). It is even impossible to retranslate such sentences into regular Rromani, due to the vagueness of the content; all interpreters know how it is difficult to translate if the original text is too vague, except if the target language has developed a similar vague phraseology as it is the case among most "modern" languages. Rromani has not fallen into this kind of political cant and it is also a matter of culture; it is maybe a paradox but it is true to say that this gap is an asset, because it compels to a more concrete analysis of quite important problems.

As a matter of fact, Rromani is able to express far more than many people could expect, even in its current stage, provided that the analysis of the whole meaning is done through a Rromani cultural sieve instead of trying to stick Rromani words to a foreign conceptual pattern. This is the reason why, when speaking among Rroms in Rromani (at a kris for example), one can solve many problems far better than when speaking a foreign language or shadow-Rromani. This shows the close connection between language and culture. When giving up all-European genuine Rromani and its approach of reality, be it out of ignorance, unawareness, chauvinism or laziness, we are ruining a treasure far more valuable than the language itself – and without which the language is just a lexicon: the Rromani cosmovision.

It is a pity to hear in many meetings how Rromani activists say, after greetings in Rromani: "Well I do not have the words in Rromani, I will continue in gaӡikanes" – although you can chat during

hours in Rromani with them. This demonstrates a deep misunderstanding about the notions of language and culture.

Similar mistakes often occur when undertaking the translation of a non-Rromani word list into Rromani, with obvious good intentions but also with a serious risk for the language if the task is not carried out carefully. There are cases in all countries with the vogue of children's picture-books (*slikovnice*).

As a rule these cute books for children present exclusively the Western modern urban rich way of life with standard houses, emblematic objects (various pieces of clothing, furniture and accessories of all kinds, meals etc.), typical activities (sporting, games, gardening, fitness, entertainment, employment, feasts etc.) and the concerns linked with them. This is quite legitimate from the majority standpoint but should not be turned into an absolute model of life — as also pointed out by teachers in post-colonial countries. Other ways of life are equally entitled to visibility but *de facto* they do not enjoy it. As a result, when translating these picture books into Rromani, many words seem to be missing but in reality they do not exist just because they are not needed out of this very specific society, where their presence is dictated mainly by market rules. Translating into Rromani such books, issued at the same time in dozens of urban languages, has a threefold outcome:

● on the one hand, it is a good opportunity to fix many words of great usefulness in everyday life and to develop new vocabulary for real social needs (like school supplies, health care, administration etc.),

● but at the same time you face a number of notions basically useless in a non-mainstream society for the following reasons:

    - because they do not exist in your sphere of practice (and you can do without them very well);

    - because if they exist, they are not so crucial as to need a specific word for them (you may use a phrase);

    - because it is often more efficient, mainly with items deprived of any cultural value, to borrow the corresponding word from English (but there remains the question of grammatical adaptation). Unfortunately, due to economical factors, only translations of such books into Rromani have been affordable, and no genuine creations in the language itself.

● be it as it may, all the specificity of Rromani vocabulary, as a mirror to Rromani cultural, social and spiritual values, is dropped as unknown by all gaʒikane readers or albums.

Accordingly such publications reinforce the erroneous image of deficiency attached to Rromani (or the impression of artificiality when the translator imagines all kinds of solutions), while concealing the genuine conceptual wealth of the language. The use of host language to Rromani dictionaries can be only a part of the linguistic strategy and the affirmation of Rromani should rely mainly on genuine texts (including Rromani to host language dictionaries), produced directly in Rromani from a Rromani perspective because they contain many words and expressions with non counterpart in host languages, together with the feelings, allusions and connotations these words and expressions convey. The foreword of a recent dictionary (2004) mentioned as examples: **manralo** "covered with remainders of fresh bread dough", **baśakărèla** "to provoke a sound", **muzgonèla** "to coat with a kind of adobe", **lokoćinèla** "to prepare mud", **dìpi** "center (bottom) of the basket [weaver's term]", **xonòta** "particular smell of the earth after the rain", **paparinŏla** "to lose one's qualities while soaking in water", **phućivèla** "to lay [eggs] without shell", **źambàla** "kind of ritual collective game during Herdelèzi feast (6 May)", **uźdàga** "specific stick of the Rlìa tribe" – to quote but a few; hundreds idiomatic expressions should be added to this list. Promoting this heritage is far more significant than promoting Rromani lists of words elaborated on a foreign pattern.

To conclude with, Rromani development can be achieved only through additive capitalization. In order to reduce the lexical distance between the Rromani varieties, one has to make widely known the existing vocabulary and to produce common neologisms for new concepts, if and only if needed. This may be called "additive capitalization" and it has been the main device of modernization in all languages. On the contrary, the subtractive approach of eliminating all the vocabulary which is not immediately understandable by individual activists (the "lowest denominator method") leads to the loss of 90% or more of the lexical funds. The resulting impoverishment challenges speakers to replace the lost wealth by artificial constructions and loan-words creating a Rromani shadow-language, deprived of any cultural density. Special efforts have to be done to reactive all the means of expression of the Rromani conceptual heritage.

18 - The spelling issue

When considering spelling strategy, one has to bear in mind the following crucial differences between the mechanisms of oral and written understanding:

a) When enlarging their use through the emergence of a written form, all languages face a series of new exigencies: they lose significant extralinguistic elements, such as intonation, gesture and the presence in the visual field of objects referred to, but at the same time they need to express some more complex, more precise and also often more abstract ideas; they also lose the opportunity the receiver has in oral exchanges to ask if he does not understand; from a prolix and redundant style limited to a small amount of familiar topics, they shift to a dense and economic expression which treats the most varied matters; they have to construct longer sentences with more rigorous articulations; they are supposed finally to confront the trial of time, for as one knows *verba volant, scripta manent.* As a result they have not only to compensate for the lost means of expression by new ones but also to elaborate extra devices of meaning consolidation. A written language is never a transcribed oral language. A transcribed oral text is readable only if the content is extremely simple (this is why demagogic texts, mainly insults, are understandable in written even if just transcribed from oral utterances – in such cases, oral and written registers overlap to a great extend) but the more a text is meaningful, the widest the gap between oral and written registers is.

b) On the other hand, the mental system of understanding is quite different in oral communication, which is natural and relies on innate abilities, and reading, which is artificial and relies on acquired skills. Our mental system is able to compensate widely dialectal discrepancies when hearing a speaker of a different background, through automatic familiarization to his/her dialectal structure but such a "decoder" does not exist in reading. One has to make up for it through specially elaborated writing strategies. A major difference between oral and written codes is that all phonetic realizations which may occur in a given place of the chain ("sounds": [a], [e], [i], [m], [b] etc…) constitute a continuum, with no sharp distinctions between them (the language habits create the distinction in order to identify phonemes which "make sense"), whereas the distinction of their equivalents in writing (the letters) is very sharp, especially in print. In addition, the phonetic features of these "sounds" intersect partly and this gives further flexibility to oral communication, a quality lacking in print. When you hear an indistinct sound (or seemingly indistinct for your dialectal structure, while it can be quite clear for the speaker's dialectal structure), your brain will give this sound a specific value, according to the context and prior experiences. Yet when one writes down the equivalent, one has to choose between clearly distinctive letters and put on the paper some differences which maybe are not relevant in the original dialect, while overlooking other features, essential in the original dialect, but unknown to the reader's variety.

This is the case when everyone is writing in the majority language spelling of his country. The first situation may be exemplified by the mutation mentioned above: in oral communication dialects with or without mutation are mutually quite intelligible and mutation looks rather like an

accent, with no incidence on meaning: [ʧhavo] / [ɕavo] "boy", [ʧhib] / [ɕib] "tongue" etc.). However, writing the two kinds of pronunciation according to non-Rromani spellings (not to mention that it is impossible to render properly the sounds [ʧh] and [ɕ] in almost all European languages) creates a huge difference between them and the reader has to think over the word and its context in order to understand it properly, if spelled according to foreign pronunciations. Reading becomes a puzzle.

The second situation may be exemplified by the two kinds of r-sounds: [ʧoripen] "theft" / [ʧorripen] "poverty", which often are not distinguished by non-Rromani ears and therefore written the same way (some scholars have even drawn moral conclusions of what seemed to them a total homonymy). Note that the second r-sound appears at the beginning of the words Rrom itself. Even young Rroms who learn Rromani from books (or the internet) believe "theft" and "poverty" are homonymous in Rromani.

This system, called "diasystematic", has been established by the 4th Rromani Congress in Warsaw in 1990 (after years of consultations) and it is the most efficient so far proposed, bearing in mind that no spelling is absolutely perfect and that choosing a spelling means often to choose between various disadvantages. The principle is that everybody has to make a little effort to stick to a common spelling in order to save great efforts to all other users who want to read. People write more or less the same way and everybody read the way he/she has learnt from his/her family. The entire system may seem complicated and indeed it is to some extend but this is only the linguist's concern, since every user has to know his/her own dialectal rules of spelling and reading, which are not more complicated than in Italian or Spanish.

Some politicians and observers have appealed for a simplification of this spelling. What does this mean? This means they want a spelling which they can use immediately, without the hour or two of learning, which is necessary for a correct command of the European spelling. This means that the spelling they suggest has to be the one they have personally learnt in their respective schools in majority languages. For Bulgarians, this means to bulgarize the spelling, for Hungarians to hungarize it, for Poles to polonize it[1] etc… every time breaking down the unity of Rromani for the sake of conformity to local languages. The argument is generally that Rromani children are not able to learn a specific spelling of their own (when other minorities' children are able). The moderate wing suggests the Croatian alphabet for everybody, but they do not realize that such an alphabet seals and perpetuates superficial differences of pronunciation, which do not impede oral communication but, once written, make reading very opaque. Yet the radical wing sticks to local alphabets and even promotes several alphabets for one country like for example in Austria where the same sentence "the woman said she knows the truth" is spelled:

i dschuvli pentscha so dschanel o tschatschipe (in Fenĕtiko or Vend dialect [O□F], German spelling)

and

e žuvli phenda so žanel o čačimos (in Lovari dialect [E#L], Croatian script). One should notice that this last sentence is spelled, in Lovari also but beyond the Hungarian border, the following way:

é zsúlyi phéndá szó zsánél o csácsímó (same dialect but Hungarian spelling), while both sentences look in common spelling as below:

i ʒuvli phendă so ʒanel o ćaćipe (Fenĕtiko dialect)

and

e ʒuvli phenda so ʒanel o ćaćimos (Lovari). This demonstrates that what could be a simplification at the regional level is much of a puzzle even within one country, and all the more at the

---

[1] *Zis iz az if aj uer rajting ingliš zis uej* in Jugoslavija, *youzing razeur zisse euzeure oueille* in France and *stil anăzăr uan, laic zis for instăns* in Romania, *нот ту меншън зъ уан ай ўуд чуз фор България*… instead of regular English spelling. One can understand more or less any short sentence of a given language (here English) written in any spelling, not a real text intended for meaningful communication.

European level. Such a treatment would involve texts circulating only in national areas. As a matter of fact, one or even two hours of training is just nothing as compared with the advantages of maintaining pan-European a language of continental size and with the dimensions of the heritage made accessible this way to millions of Rroms. The fact that over 32,000 Rromani pupils attend Rromani classes with this script in Romania every year demonstrates that this problem is forged.

Some users claim that the letters are not available on their key-board. This is not true because several fonts have all the Rromani letters (the most widespread font are Arial Unicode and Times New Roman and there is an executive drivre, called EuroUniv which allows to write Rromani – as well as all other Latin based European languages using a regular British keyboard; in lack of such driver, it is very easy to ascribe a shortcut to the various special letters on the keyboard). In addition, when the Same (or Lapp) language, spoken by less than 40.000 persons, enjoys 9 keyboards of its own on any recent Microsoft set, isn't it a striking discrimination that Rromani with millions of speakers is not even taken on account? This is a clear example of underestimation, this means discrimination, of our language and therefore people.

Thinking globally but acting locally is also true for Rromani. When a village teacher says "Why should I write the European way for my pupils in my remote mountains?" – this does not mean that Rromani spelling is difficult, just that she has not understood the European dimension of the Rromani language, culture and nation and how much the pupils lose while sticking in written to the local pronunciation, perceived through the local non-Rromani spelling system.

18 - Current problems and needs

Actually the main problems are the following:

1. Lack of commitment and money to publish and circulate as much material as possible in a common graphic cloak and in the original dialectal variety (edited, as in all other languages, in order to avoid troubles in understanding) – but also on other supports, like films, tapes, electronic games etc…

2. Lack of motivation, awareness and sometimes industriousness of some Rroms who are reluctant to spend one or two hours in training the common spelling and further leisure to acquire genuine Rromani words forgotten in their community but alive elsewhere in Europe.

3. Lack of consciousness of some Rromani translators who content themselves with most incoherent translations, just to meet an obligation and be paid for it (this is the case with many political documents of the Council of Europe or even literary books, like the first Rromani translation of "The Little Prince"). As I was commenting this to one of them, he answered with a cynical smile: "Anyway Rroms do not read and Gaʒes do not understand". Such publications demoralize the potential readers, misuse scarce funds and torpedo the healthy affirmation of Rromani as a modern European language.

3. Lack of education and motivation by the surrounding world, which still ignores Rromani as it ignores the Rromani people itself, the Rromani genocide, the Rromani contribution to universal civilization, the Rromani part in history etc…

4. Endemic despise of the intellectual abilities of the Rroms (see above).

5. Last but not least: one can observe a clear obstructive attitude of some non-Rroms, a kind of fear to view a so far despised people, numerous and living within Europe, other than as "Rromani communities" (formerly "tribes") but instead as "one Rromani nation" with a great diversity of visages and Rromani not as "clusters of dialects" but instead as "one Rromani language" with also great riches of cultural means of expression. It looks as if they fear to have their own national identity, based on a territorial state, weakened if they recognize the unity of language and identity to a non-territorial huge people. In other words they go out of their depth when national unity is recognized, beyond political borders and cultural diversity, to a people basing his specificity upon a traditional feeling of otherness, a common language (whether it is actually practiced or just

remembered as a symbol of the past) and a common historical fate – all this without any compact territory. This reminds one of the great Sanscritist Jules Bloch's statement: "the Rroms view themselves as a unique people, in spite of the dispersion of their groups and their lack of uniformity. This shared feeling of community allows considering them as one nation, although they lack precisely what has become for us the symbol of a nation, namely unified institutions and a defined territory. Lot of people remain Rroms, even […] persons who have lost the use of the hereditary language".

Illiterate and marginalized Rroms have kept so far the Rromani language alive but they are more and more acculturated to majority languages, except for the heirs of a robust oral culture, who still maintain their love and pride for their mother-tongue. In contrast with the common belief, it is not at all natural to cultivate one's ancestral language: a strong awareness and motivation are required to fight against inertia leading to acculturation. In present time Europe, minority languages have a chance to survive only thanks to volunteerism in the elite, provided that this elite remains in brotherly contact and fruitful exchanges with the rest of the Rromani population. If Rroms benefit from correct aids for studies and can develop their European elite, this elite will hopefully act as a model (a kind of national middle class) for other Rroms and incite them in middle term future to reactivate the language some are currently neglecting – on their own initiative or under somebody else's influence.