FONDEMENTS DU DIALOGUE CULTUREL

TEACHING SANSKRIT TO EUROPEAN RROMANIES (or RROMS)

Marcel COURTHIADE
INALCO Paris
mdro2emc@gmail.com

A living Indo-Aryan language in the very heart of Europe

When hearing the word "Sanskrit", average Europeans picture some taciturn elderly bookworms dealing with huge dusty grammars. Some Europeans are perhaps better informed and the image which comes to their mind is full of wise bearded old men totally immersed in meditation, high above the intellectual pursuits of ordinary mortals. Yet very few are those who are aware of the fact that hundreds of Sanskrit words can be heard in a living language spoken almost everywhere in Europe – and many places of America – and that this language was brought by sons of Bharat who left India, more precisely Kannauj in U.P., one thousand years ago (in 1018). The everyday home use of this language was never discontinued among them over centuries and still now they are the first historical diaspora of India. They call themselves Rroma and their language is Rromani, known in Europe by some 7 to 8 million Rroms and spoken everyday by almost 5 million[1].

The Sanskrit name of the Rromani people

The name "Rrom" is already a Sanskrit word, since it derives from डोम्ब "a percussionist, a musician, a performer, an artist" and the feminine डोम्ब्री is also quite close to Rromni, the feminine of Rrom. Old Indo-Aryan

[1] CIEMEN (Centre Internacional Escarré per a les Minories Ètniques i les Nacions) Barcelone : Alan Viaut report.

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

retroflex consonants developed into Rromani as a specific /r/, which sharply differs in sound from the regular trilled [r], inherited from Old Indo-Aryan [r]: Skr. वर्ष [varṣa] "year" > in Rromani **berś** "id.", द्रव्य [dravya] "object, stuff; medecine" > **drab** "medecine", Skr. क्षुरी [kṣurī] "knife" > **ćhuri** "id.", दूर [dūra] "distant, remote" > **dur** "far" etc… and most varieties of modern Rromani still distinguish sharply between the afore-mentioned [r] originating from India [r] and a different /r/, which can have developed from one of the various Indo-Aryan retroflexes ट [ṭ], ड [ḍ], ड़ [ṛ] (sometimes ण्ड [ṇḍ]). The latter is pronounced as a retroflex [ɽ], a flap [ɾ], a long/strong [r:], a velar [ɣ] or [x], a [ʁ] spoken with a burr etc… and it is written since 1890 with double **rr**[2]. Rromani words like **bar** "hedge" and **barr** "stone", **rani** "lady" and **rrani** "branch" or **ćoripen** "theft" and **ćorripen** "poverty" are distinguished only by the nature of the /r/.

It is therefore of the utmost importance to respect at the same time etymology and pronunciation while writing double **rr** when relevant, including in the name of the **Rromani** people itself. In addition to respect for linguistic truth, the double-**rr** spelling links directly this national name to Sanskrit and India, and it avoids any confusion between Rroms and Romanians, thereby eliminating false etymologies referring to "Rome" (the city in Italy), the Rums (Arabic name for "Christians") or Rama (in Indian culture). Writing with a single **r** is but a consequence of the inability of the Europeans' ear to distinguish between the two sounds, because this contrast doesn't exist in European languages.

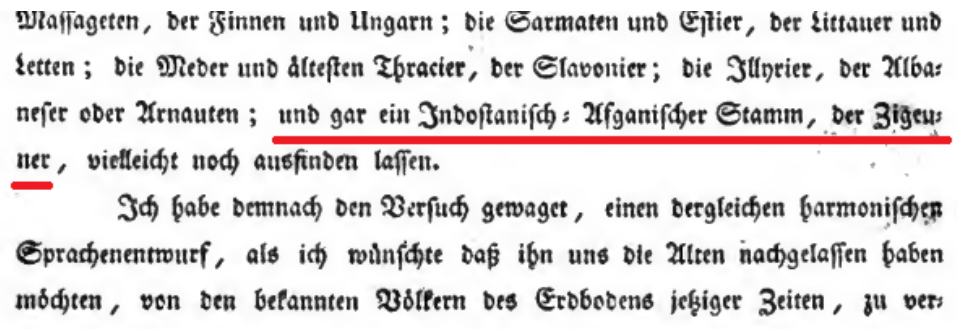
India: remembered or not?

One is led to believe that the Rroms themselves were quite aware of their Indian origin, even after their arrival in Western Europe around 1350, if we take into account that at least 6 documents[3] between 1422 and 1630 mention their Indian origin. Although this concept circulated among the learned, they began to treat it as a serious option only when a few of them began to compare the Rromani language with Indian languages, first a sort of approximative खड़ीबोली [khaṛī-bolī] by Rüdiger around 1780. Yet the

[2] Sztojka Ferencz's Hungarian-Rromani dictionary, Paks (Hungary) 1890.
[3] See *Informaciaqo lil e Rromane Uniaqoro (p. 1) N° 7-9 1992* and more extensively in Courthiade 2012.

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

idea was not entirely new, since in 1771, Christian Wilhelm Büttner had already mentioned in the foreword to a book[4] that in Europe there is even an Indostano-Afghan tribe, the Rroms…



Maſſageten, der Finnen und Ungarn; die Sarmaten und Eſtier, der Littauer und Letten; die Meder und älteſten Thracier, der Slavonier; die Illyrier, der Albaneſer oder Arnauten; und gar ein Indoſtaniſch-Afganiſcher Stamm, der Zigeuner, vielleicht noch ausfinden laſſen.

Ich habe demnach den Verſuch gewaget, einen dergleichen harmoniſchen Sprachenentwurf, als ich wünſchte daß ihn uns die Alten nachgelaſſen haben möchten, von den bekannten Völkern des Erdbodens jetziger Zeiten, zu ver-

Further research focused mainly upon Sanskrit – which was better known to Europeans since Jones' works than New Indo-Aryan languages – we should remember that this was much before Lallu Lal's times and Indian languages of the North were much more Persianized than today. So Sanskrit again played an important part in restoring historical truth.

Teaching ancient languages in modern Europe

The history of teaching ancient languages has gone through several stages, especially in Europe. One or two centuries ago it was a rule among the elite to master Latin and Ancient Greek, while the learning and knowledge of modern languages was viewed as a dilettante's hobby. After WWII, a new mythology arose in the world, widely dominated by Europe: languages were viewed as useful for business, profit and expansion – first of all English, but in practice, basically only English in many persons' conception. Not only culture and wisdom, including the paramount Indian heritage, but also any glimpse of brotherhood and humanity, were sacrificed to the new gods of the mythology of business. It seems that the most intelligent part of human societies is giving up on this error and linguistic and cultural

[4] *Comparative Tables of the Ways of Writing of Various Peoples of the Past and Present Times des modes d'écritures de divers peuples du passé et du présent* (Götingen & Gotha – 1771)

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

diversity are gaining in popularity again – at least among the intellectual elite and this is quite significant, since other social layers tend to imitate the elite. This new appeal for linguistic and cultural riches encompasses also so-called dead languages: Latin, Ancient Greek but also Ottoman Turkish is now proposed in Turkey. In this respect it seems a seasonable time to promote Sanskrit teaching for Rromani youth. This may sound as a lost cause, taking into account the appalling conditions in which Rroms live in Europe. In fact, this is a biased image of the Rroms, because "only" one quarter to one-third live in bad or very bad conditions – what is anyway disproportionate as compared to other nations like the Poles, the Danes or the Sames, but breaking this image is necessary in order to get out of the so harmful only-social treatment of the Rroms, following centuries of discrimination built on a medieval misunderstanding, which itself was rooted in the denial of the Rroms' Indian heritage and culture, which were in addition underestimated. In this perspective, the rehabilitation of Rromani culture is a major gateway toward social justice regarding this people – and Sanskrit has to play its part in this project of revitalizing global respect toward the first historical diaspora of India.

A specific Rrom-addressed course in Sanskrit

One of the tools in use for this purpose is the on-line course "Restore the European Dimension of Rromani Language and Culture". So far only a comparatively small part is devoted to India (in the History component) but it is our intention to prepare a specific component devoted to Sanskrit. However, one could pose a basic question about the approach to be followed, and indeed, some facts lead us to the conviction that a specific didactic has to be elaborated. True enough, the heavy German or British books of grammar of the past do not fit anymore to our times and the youth's mentality. The publisher of the "Teach yourself" series has circulated alternative, more accessible, books to teach Sanskrit, but in the case of the Rroms, it is not only a matter of pedagogy, but much more of languages. Rromani children, in their mother tongue already have a series of elements which other students of Sanskrit lack totally.

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

The sandhi system in Rromani – as compared with other languages

It is well known that one of the main obstacles for students of Sanskrit is the system of sandhi. More or less all languages have sandhi rules, even when speakers are not aware of it. Let us mention but a few examples from some very different areas of Europe:

1. In modern Greek: an initial stop consonant creates a new one when linked to a final **-n** of a preceding words:

την κοπέλα [tin+kopela] "the girl" > [tiŋgopela]

τον παπά [ton+papa] "the pope" > [tombapa].

2. In French: the mute **s**-ending of the article is again pronounced, as a voiced [z], when the following word begins with a wovel ("*liaison*"):

les parents "the parents" > [le parã]

les enfants "the children" > [lez‿ãfã].

3. In Breton (another Celtic language): the first consonant of a word changes according to specific ruled after another element:

kador "chair" > ***ur gador*** "a chair"

gavr "goat" > ***ar c'havr*** "the goat".

This phenomenon goes even beyond the word in contact and affects the following word:

kador kaer "beautiful chair" > ***ar gador gaer*** "the beautiful chair".

4. In Polish: a voiced final consonant at the end of a word becomes voiceless when not followed by any other element:

wóz "cart" > [wus]

mów "speak!" > [muf].

but in some areas the voiced consonant reappears before a voiced consonant, not in others:

wóz "cart" > [wus]

wóz Ewy "Eva's cart" > [wus evɨ] in Warsaw but [wuz‿evɨ] in Cracow.

As one may observe, these sandhis are very restricted in extension and also in complexity (most of European sandhis are limited to the unvoicing of a final consonant – as in German, Polish, Russian, Bulgarian, etc.).

The situation is quite different in Rromani, where we have:

a) in word formation, there is practically no sandhi and the construction is visible all the time:

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

kin "purchase" > **kindo** "purchased" > **kindikano** "originated from purchasing" > **kindikanipen** "the quality of being a purchasable item, purchasability"[5]

b) in postpositional structures, there is a very complex system of sandhi distributed in three layers: sonority sandhi, visargation sandhi and third sandhi (often of vicedness but not only). Let us see the following examples:

1. first sandhi[6]: after an **n**-ending, all postpositions begin with a voiced consonant:

e manuśenqe [e manuśeŋge] "for the people"

e manuśenθar [e manuśendar] "from the people"

but it is voiceless after any other sound:

e manuśesqe [e manuśeske] "for the man"

e raklăθar [e raklătar] "from the girl"

2. second sandhi: the **-s-** may be dropped as the end of a word or right before a postposition, depending on the concrete vernacular (**k** may also be dropped or geminated):

e manuśesqe [e manuśeske/manuśehke/manuśe^{h}ke/manuśese/ manuśeke/manuśekke] for the man"

3. third sandhi: it can be of various kinds – but in the present case, it is represented by a palatalisation of the dorsal stop before a front vowel, depending on the concrete vernacular. If we take the above example with several degrees of palatalization, we may find the following realisations (written here in IPA):

[5] Hübschmanová rightly points out that "The original 'Indic' words in Rromani, in comparison with borrowed words, have the greatest morphosyntactic potency, which means that it is possible to create a number of other words with the specific suffixes" (Hübschmanová in History and Politics : www.rromani-uni.graz.at)

[6] This sandhi is of special significance, because it is parallel to the same evolution within the root of the words, and it developed under Greek influence in Asia Minor and the Balkan, encompassing also some Arabic, Kurdish, Armenian and Albanian dialects (in his "Traité de phonétique" [Paris, 1933:189], Maurice Grammont describes this evolution of homorganic clusters in Albanian and Syriac). It is in fact a very general phonetic rule, rampant all over the world, and it was Turner's mistake to link it specifically to Dardic languages, while it was widely operating in Asia Minor and the Balkan.

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

manuʃeske	manuʃeskʲe	manuʃestɕe	manuʃestʃe	manuʃestʲe	manuʃeʃtʃe
manuʃehke	manuʃehkʲe	manuʃehtɕe	manuʃehtʃe	manuʃehtʲe	
manuʃeʰke	manuʃeʰkʲe	manuʃeʰtɕe	manuʃeʰtʃe	manuʃeʰtʲe	
manuʃese					
manuʃeke	manuʃekʲe	manuʃetɕe	manuʃetʃe	manuʃetʲe	
manuʃekke					

Note that palatalization is also possible with the voiced equivalent [manuʃeŋge] which develops as below:

manuʃeŋge	manuʃeŋgʲe	manuʃeɳʤe	manuʃeɳʤe	manuʃendʲe

„for the people".

So there is a series of mutual interferences between the various kinds of postpositional sandhi, which has probably no counterpart in any language. From this point of view, the Sanskrit sandhi will probably look as very simple to an Rromani student – provided s/he has been taught about the sandhi in his/her own mother tongue and provided also that the Sanskrit sandhi is presented to him/her as a quite natural phenomenon, not a fortress to conquer…

Romani morphology as compared to some other languages

In-so-far as morphology is concerned, one has to distinguish between the morphology of the verb and the one of the nominal group.

A) morphology of the verb. Let us compare the present tense forms in Sanskrit, English and Rromani of the verb:

		Sanskrit	English	Rromani	Serbo-Croatian	Lithuanian	Greek
sg.	1st	जीवामि	I live	**ʒivav**	živim	gyvenu	βιώνω[7]
	2nd	जीवसि	you live	**ʒives**	živiš	gyveni	βιώνεις
	3rd	जीवति	s/he lives	**ʒivel**	živi	gyvena	βιώνει
pl.	1st	जीवामस	we live	**ʒivas**	živimo	gyvename	βιώνουμε
	2nd	जीवथ	you(guys) live	**ʒiven**	živite	gyvenate	βιώνετε
	3rd	जीवन्ति	they live	**ʒiven**	žive	gyvena	βιώνουν
dl.	123		no dual		no more dual		

[7] Βιώνει means in fact "to experience" rather than "to live" but this is the etymological cognate of जीवति.

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

The similarity between Sanskrit and Rromani is definitely striking, much more than between Sanskrit and Slavic (here Serbian[8]), Greek and even Lithuanian, reputed as especially close to Sanskrit – not to mention English.

Vowel **-a-** in the first person (sg. & pl.) in Rromani corresponds to a long ***-ā-*** in Sanskrit, whereas **-e-** in other persons corresponds to short ***-a-*** in Sanskrit. One may also notice the regular evolution of intervocalic ***-m-*** into **-v-** (1st person sg.[9]). It is worth mentioning that the old Sanskrit **m**-ending of the first person sg. has been retained in the Baltic and North-Russian area, as well as in some Balkan vernaculars – yet only in two verbs: **kamam** (beside **kamav**) "I love, I want" and **tromam** (beside **tromav**) "I dare".

Thus the endings of the Sanskrit present have persisted up to Rromani without almost any other change than normal phonetic changes – to paraphrase Jules Bloch (1914:243) about the Marāṭhī verb:

		Sanskrit	Rromani	Marāṭhī (intransitive, old present)
sg.	1st	जीवामि	**ʒivav**	ʒiṇeṁ
	2nd	जीवसि	**ʒives**	ʒiṇes/ʒiṇas
	3rd	जीवति	**ʒivel**	ʒiṇe
pl.	1st	जीवामस	**ʒivas**	ʒiṇoṁ
	2nd	जीवथ	**ʒiven**	ʒiṇā
	3rd	जीवन्ति	**ʒiven**	ʒiṇat

So in this respect, Rromani is as close to Sanskrit as Marāṭhī is.

Similar comparaisons may be made for other tenses but Middle-Indo-Aryan and New Indo-Aryan (especially Marāṭhī) have to be taken into account. The same cannot be said about medio-passive, since Rromani medio-passive doesn't originate from its Sanskrit cognate, but has been built up in the Anatolia and later in the Balkan under Albano-Greek influence, out of Indian lexical material, especially the one of the *copulæ* (namely **ovel** "to become").

[8] In this case, Serbo-Croatian is the closest to Sanskrit among Slavic languages.
[9] A very old evolution, as mentioned by Pischel 1981:206 (§ 251).

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

B) morphology of the nominal group.

The morphology of the nominal group is quite different, since it has lost both dual and neuter, as well as most Sanskrit cases – a development shared by other Indo-Aryan languages in a way that may be illustrated roughly by the following table (vocative is not taken into consideration):

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)	

One may distinguish between two levels in this table:

a) the melting of Sanskrit Accusative with Nominative, developing into the A-case of Rromani, while Dative melts with Genitive resulting in Rromani B-case – similarly as other New Indo-Aryan languages. In this respect Rromani is closer to other New Indo-Aryan languages than to Sanskrit. The use of postpositions in both Rromani and NIA is also a common point which was revealed as early as 1780 by Johann Christian Rüdiger. Curiously enough[10], this similarity is still widely denied in many Rromani grammars, which follow the Latin-German (or Russian) non-Rromani pattern, as if there were a fear of recognising the Indian identity of the Rromani language and of the 15 millions of people, for whom this language is a crucial and beloved heritage. Nevertheless, if you look at the following table, it seems impossible to reject the Indian postpositional system of Rromani:

[10] In fact, it is not so curious if we observe that this denial arises mainly from persons who refuse to accept the obvious Indian origin of the Rromani people.

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

Direct case		i **bakri** (*singular*)	o/e **bakria** (*plural*)
Indirect case	with no postp.	e **bakria**	e **bakrien**
Indirect case	with postp. **-qe**	e **bakria-qe** [ke] (sing.)	e **bakrien-qe** [ge]
	with postp. **-q**/**o**,**-i**, **-e**	e **bakria-qo** [ko]	e **bakrien-qo** [go]
	with postp. **-θe**	e **bakria-θe** [te]	e **bakrien-θe** [de]
	with postp. **-θar**	e **bakria-θar** [tar]	e **bakrien-θar** [dar]
	with postp. **-ça**	e **bakria-ça** [sa]	e **bakrien-ça** [tsa]
Indirect case	with circump. **bi -qo**	bi **bakria-qo** [ko]	bi **bakrien-qo** [go]

What is the argument against the recognition of postpositions in Rromani? The denial is based on the fact that the second layer of adpositions is postponed in Indian languages (**bakria ke pās** "near the goat/sheep"), while it is anteposed in Rromani (**paś-e bakriaθe** [arch.])….

Let us look at the following table illustrating the Rromani possessive postposition:

Short variant	possessed object singular	possessed object plural
possessed object masc. **kan**	e **bakria-qo** kan	e **bakria-qe** kana
possesses object fem. **jakh**	e **bakria-qi** jakh	e **bakria-qe** jakha

but also (more restricted in dialectal terms in both Rromani and Hindi):

Long variant	possessed object singular	possessed object plural
possessed object masc. **kan**	e **bakria-qoro** kan	e **bakria-qere** kana
possesses object fem. **jakh**	e **bakria-qiri** jakh	e **bakria-qere** jakha

Only liars talking to ignoramuses can maintain that this grammatical system is not Indian. I would also add that this complex system evidences that Rromani is not a pidgin or a lingua franca, which would never have preserved such an elaborated structure, but a genuine – albeit forgotten – Prakrit, to use Pathania's so appropriately coined formula.

b) survival of Sanskrit Locative and Ablative endings of the first declension, namely **-e** (as in देवे) and **-al** (as in देवात with the normal phonetic changes [t] > *[d] > [l]) as in the two last lines of the table above. These remnants of Sanskrit cases are not anymore productive and the substantives, which may take one or both of them are in limited number. Note that ***-at***/**-al** has survived much more in Rromani than in any other New Indo-Aryan language and as a whole Rromani is finally closer to Sanskrit than any other languages of the same family. It seems probably paradoxical that an uncultivated language in diaspora has kept more similarities with the

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

"mother language" than the language ultivated on the Indian soil, but such is the result of History. In morphology, the Rromani comparative in **-eder** is also a Sanskrit vestige.

Phonetic and lexical evolutions or „Is Rromani a daughter of Sanskrit?"

One may read not infrequently "Rromani is a daughter of Sanskrit?" Under this old fashioned expression (languages do not marry and do not have children), one may perceive much more the enthusiastic pride of some Rroms involved in research than a genuine scientific statement. As a matter of fact, Sanskrit was coined by inspired seers (visionaries) to compose the most elevated works ever produced by human mind, taking the dough out of popular old Indo-Aryan basilects in order to express perfectly their teaching. Nevertheless, the basilects didn't disappear but developed further among the people, leaving aside Sanskrit until their speakers lost any ability to understand Sanskrit, fixed once and for all centuries earlier. It was the time when the Śramanic reformers began to teach in Prakrits and the great Aśokə dispatched his *Dhamma* all over the Indo-Aryan area, written on pillars and rocks, regularly read aloud to illiterate people; all these languages arose from local basilects, probably already converging into regional mesolects and anyway enriched by Sanskrit elements. Middle and modern Indo-Aryan languages developed out of this system. One of the Prakrits, namely Śaurasenī, was spoken in the Śūrasena country (शूरसेन – after the name of a ruler of the Yadava dynasty) – which extended probably from Varanasi region to the north of the Vindhya, with Mathurā as its capital[11] (Sircar, 1071:109). Śaurasenī is considered as the nearest to Classical Sanskrit out of all the Prakrits (Woolner 1917:5) but it seems also to be the closest to Rromani, judging by some linguistic features. To mention here but one, the famous Sanskrit verbal ending of 3rd person (present tense, first group of verbs) ***-ati*** develops in MI as ***-aï***, except in Śauraseni, where it preserves a stop consonant was not dropped but voiced (a remnant of an early stage ?) ***-adi***, ***-edi***. Rromani dropped the final ***i*** and

[11] *Studies in the Geography of Ancient and Medieval India* (Sircar, 1071:109) states that Greek writers refer to the Sourasenoi and to their cities Methora (Mathurā) and Kleisopura (possibly Kṛṣṇapura = Gokula).

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

developed this consonant further into the lateral [l]: **-el** (an evolution which would have been impossible out of other Prakrits of ***-aï*** type):

Other examples: ***gata*** > Rr. **gelo** but Śauraseni *gada*, *gata* – while other MI is ***gaa***

gitā, ***gīti*** > Rr. **gili** "song" same Śauraseni gīda, gīdi but other MI gīa "sung"

mr̥ta > **mulo** with MI ***maa***, ***mua*** but Śauraseni ***muda***

ghr̥ta > Māgadhi ***ghaa*** > ***ghī***, but Śaur. ***ghida***, hence Rr. **khil**

śata > Māgadhi ***saa*** > ***sau***, but Śaur. ***sada***, hence Rr. **śel**

marati > Māgadhi ***maraï***, but Śaur. ***maradi***, hence Rr. **marel** "he beats".

It is obvious that ***gata*** may give **gelo**, but ***gaa*** cannot possible restore a consonant between the two **a** (cf. also Marāṭhī ***gela*** "s/he went" – Rr. **gelăs**/**gelo** "id."; cf. also Marāṭhī ***ala*** "s/he came" – Rr. **avilăs**/**avilo**, dial. [Macedonia]: **alo** "id.")

The affiliation to Śauraseni, geographically located around Mathura, corroborates Sir Ralph Turner's conclusions, who put proto-Rromani in the central group of the Indo-Aryan languages, and both elements advocate for the Kannauj thesis, since Turner's area of origin for proto-Rromani is between Awadhi and Braj Bhasa/Bhakha, corresponding more or less to today's state of Uttar Pradesh in Northern India. Unfortunately, Turner took into consideration only stems and not morphological elements, which would have shed more light on this issue. As for the question if Rromani is or not a daughter of Sanskrit and to keep an anthropomorphic alegory, it is better to say that Sanskrit was a kind of grand aunt, who took divine orders and devoted her life, unmarried, to spirituality – and so grew out of her siblings' level, reaching a high cultural life, leaving to their progeny an exceptional immaterial heritage.

Eight hundred Sanskrit roots in Rromani

The high number of Indo-Aryan roots in Rromani language, namely some eight hundred when collected in the core (inherited, disregarding European loanwords) vocabulary of all dialects, is probably the most privileging factor for Rromani students of Sanskrit. It

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

is much more than the 200 Greek, 70 Persian and 35 Armenian roots of the same core vocabulary.

Almost comprehensive lists of this vocabulary have been given by Turner in his etymological Nepali dictionary and his "Position of Rromani in Indo-Aryan" so there is no point to repeat them here. Some are obvious for the speaker, as for example:

jakh "eye"	Skt. *àkṣi*; Hi. *ãkh*
drakh "grape"	Skt. *drākṣā*; Hi. Nep. *dākh*
phak "wing"	Skt. *pakṣàḥ*
rukh "tree"	Skt. *rukṣàḥ*
khino "tired"	Skt. *kṣīṇàḥ*
khil "butter"	Skt. *kṣīràm*
khelel "moves, plays, dances"	Skt. *kṣvelati*
makh(**i**) "fly"	Skt. *màkṣikā*
ćhuri(**k**) "knife"	Skt. *kṣurī*; Hi. *ćhurī*
devel "god, sky, heaven"	Skt. *devàtā* "divinity"

while others can be identified only by researchers:

lolo "red"	Skt. *lòhitam*;
giv "wheat"	Skt. *godhū̀maḥ*; Hi. *gehū̃*
bori "daughter-in-law, bride"	Skt. *vadhūṭī*

and many others.

We meet also cases of "false" Sanskrit etymology: it would seem obvious to link the verb **kamel** "to love, to desire, to want" to the root काम (cf. कामसूत्र) and to a verb *कामति[12] but in fact, due to the afore-mentioned rule (foot-note 8), such a verb would have developed into *kavel. Did this happen? Yes, but very locally indeed, only in some parts of the Carpathian Mountains, where in addition the group -**ave**- develops into -**ā**-[13], giving ***kāl**. This form doesn't exist by itself but only linked to a reflexive pronoun **pe**(s) "itself": **pekāl**, meaning "it is necessary, one must" etc. In-so-far as **kamel** is

[12] The real verb has the causative form कामयति but the meaning of the root.
[13] Other examples: **daravel** "to frighten" > **darāl** "id." or **avel** "to come" > **āl**.

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

concerned, it originated probably from the same stem but the Persian کامستن [kāmistan] induced the conservation of the **m** as in the original word – and therefore the coexistence of **kamel** and **pe-kal**.

The Persian element in Rromani and Hindi

The example above, where ***kām*** is both Persian and Sanskrit, reminds us that there are a number of cases where it is almost impossible to determine which of these languages gave the Rromani word. There are several words of this kind like **kirmo** "worm", **angušt** "finger", **xer** "donkey", **xarr** "pit" etc. One case is of special interest: **kòkalo** "bone" usely ascribed to Greek κόκαλο[14], also a common word in Bulgarian, as кокало (the Slavic word кост is used in Bulgarian only in sacred context). Although it is true that we can find in Sanskrit a similar word: कघ्काल or कंकाल[15], it appears comparatively late (VIIIth cent.), with the meaning "skeleton, structure" (there is no earlier mention, even in Harṣa's play Nagananda, in spite of the story with the revived skeletons). Nevertheless, there are several reasons to ascribe **kòkalo** to Greek and not Sanskrit:

- the frequency of this word is extremely high in Greek (including Medieval Greek of Anatolia) and Bulgarian, while it is almost a hapax in Sanskrit;

- the plural of this word is in **-a** (**kòkala**) as an Greek (same with **pètalo** "horseshoe", pl. **pètala** – an ascertained Greek borrowing), with possible oblique **kokalan-** which would be unexpected for an Indian inherited word;

- and finally Sanskrit कघ्काल left no traces in modern Indo-Aryan languages. It is a matter of etymological method: to link a

[14] According to Andriotis, Modern Greek κόκκαλο continues Ancient Greek κόκκαλος "seeds of conifer cones", from κόκκος "grains, seeds"; cf. also general i.-e. *ko(ŋ/g)k- "grain, bone, shell".

[15] It is practically restricted to Bhavabuthi's two plays: Malatimadhava (5.14) and Uttararāmacharita (3.43) - about Shiva reduced at the state of a skeleton. In addition, Kakkola in Ardha-Magadhi has no link with this word, as it derives from Sanskrit कर्कोट (in both cases name of some snakes and of a people in India).

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

Sanskrit form with an Rromani lexeme, seeming similarity is not enough – one has to take under consideration also other factors[16].

In fact, the position of the Persian vocabulary in Rromani and Hindi is slightly more complicated. On the one hand, Rromani has integrated some 70 Persian roots but this did not happen in today's areas of Farsi, Dari or Tadjik, but in Asia Minor, where the Seldjuk Turks had brought Persian as their language of culture and civic life. Persian remained in Asia Minor the chief public and literary language from 1070 to 1300, while Arabic was used in the Mosche and the Tribunal. Turkish was then only a home language among Turks and Turkmens. This is the period when these 70 Persian words entered Rromani. However, the introduction of Persian words into Indo-Aryan languages occurred in a totally different context among both "Muslims and Hindus, who have been subjected to this influence. Persian words are found everywhere in 'Hindustani', even in rustic dialects" as Jules Bloch highlights (1914:13).

Accordingly, it is not surprising to note that the Persian vocabularies in Rromani and Hindi are not the same. Only in a few cases there is a commonality, but even then the words refer to different realia: in India, the word ***Hodā*** (خدا), means in Urdu as in Persian "Lord, God" whereas its Rromani counterpart **xulaj** means "lord, master". Similarly, ***amrud*** in India is a fruit differing from Rromani **ambrol** – "a pear". As a result, it is very difficult and sensitive to study today's Hindi dialects in so far they do not mirror, by far, the language of early XIth century, when the proto-Rroms were deported to Ghaznī – and I would daresay that Rromani language, in the core of its vocabulary, is closer to Sanskrit than Hindi itself – provided that recent loan-words from Sanskrit into Hindi are not taken into account.

An unexpected transcontinental dimension of the Sanskrit vocabulary

There are two areas in which the student will be amazed to find the words of his own Rromani tongue: European slangs and

[16] This doesn't exclude

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

languages of South-East Asia. Most European languages have borrowed Rromani words, usually through their slang – due to socio-historical contexts, but some have become a real part of the high-register vocabulary in the given tongue. This subject would deserve a special etymological and socio-linguistic study, but we will show here only a few examples (the number of Rromani loan-words in every area is given in the first column):

Language	register	word	Rromani etymon	Sanskrit origin
U.K. *some 10 items*	popular	pal "friend"	**phral** "brother"	bhrata
	Shakespearean	Caliban (a name)	**kaliben** "blackness"	*kalitva (kṛṣṇatva)
	obsolete	lowe "money"	**love** "id."	loha
	childish	lollipop =	**loli phabaj** "red apple"	lohita (ātā)phala
France *some 30 items*	popular	berge "year"	**berś** "id."	varṣa
	slang	chourave, choure "to steal"	**ćorav** "id."	ćorayati (causative)
	obsolete	chourin, surin "knife"	**ćhuri** "id."	kṣurī
	slang	marave "to beat"	**marav** "id."	marayami
	général, depr.	manouche "Rrom"	**manuś** "human being"	mānuśa
	slang	nachave "to go away"	**naśav** "id."	naṣyati
Spain *maybe 100 items*	common	camelar "to want"	**kamel** "id."	kāma
	kaló slang	diquelar "to look at"	**dikhel** "id."	drṣ; Śaur. dekkhadi
	in songs	ducas "sufferings"	**dukha** "id."	duhkha[sya]
	slang	braquí "goat"	**bakri** "sheep"	barkari
	slang	berji "year"	**berś** "id."	varṣa
	kaló slang	cam "sun"	**kham** "id."	gharma
	kaló slang	chachipen "truth"	**ćaćipen** "truth"	satyātva
	general slang	chipe "slang"	**ćhib** "language"	jihva
	slang	chor "thief"	**ćor** "id."	ćora
	kaló slang	debel "god"	**devel** "id."	devata "divinity"
	kaló slang	eray "lord"	**raj** "id."	rājan

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

Hungary *some 15 items*	slang	csóró “poor”	**ćorro** “id.”	kśudra
		dinló “crazy”	**dinilo**, **dilo** “id.”	dīna “depressed, sad”
		dzsuvás “lousy”	**ʒuvalo** “id.”	yukāla
Romania *over 200 items*	slang and everyday register	lovele “money”	**love** “id.”	loha
		mişto “fine, good, well”	**miśto** “id.”	miṣṭa “sweet”
		diliu “crazy”	**dilo** “id.”	dīna “depressed, sad”
		jovaliu “lousy”	**ʒuvalo** “id.”	yukāla
Bosnia *some 50 itens*	šatrovački (market slang)	bakrinka “sheep”	**bakri** “sheep”	barkari
		lovuška “money”	**love** “id.”	loha
		čordisati “to steal”	**ćorel** “id.”	ćorayati (causative)
		džaniška “Mr. Know-it-all”	**ʒanel** “id.”	jāñāti
		šingina “horn”	**śing** “id.”	śṛṇga
Albania *6 items*	urban slang	llovi “money”	**love** “id.”	loha
		shella “hundred”	**śel** “id.”	śata
		nash “get lost!!	**naś** “id.”	naśya

As one may notice, the words *money*, *year*, *crazy* and *to steal*/*a thief* are the most represented in this vocabulary – and some items are perfectly integrated into the mainstream language, like in English *pal* (cf. *penpal* – a synonym of *penfriend* which is used even in Japan as ペンパル), *lollipop* or in French *chourave*, *berge*(*s*), or even cultivated a specific kind of songs, like the *coplas flamencas* in Spain and Catalunya.

Very far from this area, Sanskrit words – whose roots are present in Rromani – are encountered also in the languages of Southeast Asia, as the following handful of examples illustrate (in fact they are hundreds, or even thousands):

Language	Word	Rromani	Origin	
Cambodian	[stha:n] “place”	**than** “id.”	Sanskrit	स्थान “venue”
	[tha:n] “id.”		Pali	ṭhāna “id.”
	[phu:m] “earth; village”	**phuv** “id.”	Sanskrit	भूमि “earth”

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

	[talːu] “palate (anat.)”	**talaj** “id.”	Pali	tālu “id.”
	[nɛəti] “river”	**len** “id.”	Sanskrit	नदी “river”
	[nɛəvoutban] “novelty”	**nevipen** “id.”	Sanskrit	नवत्व “novelty”
	[nitėəkhɛəʔ] “summer”	**nilaj** “id.”	Sanskrit	निदाघ “heat, summer”
	[paɲcaʔ] “5” (bes. [pram]	**panʒ** “id.”	Sanskrit, Pali	पञ्च “five”
Thai	[thev] “god”	**devel** “id.”	Sanskrit Pali	देव “god”
Burmese	[dewatāw] “divinity”	**devel** “god”	Sanskrit Pali	देवता “goddess, deity”
Tagalog	dukhâ “poverty”	**dukha** “pains”	Sanskrit	दुःख “distress, suffering”
	mukhâ “face”	**muj** “face, mouth”	Sanskrit	मूख
Indonesian	bumi “earth”	**phuv** “id.”	Sanskrit	भूमि “earth”
	citra “image”	**ćitrel** “to draw”	Sanskrit	जित्र “picture”
	dana “donation, funds”	**deni(pen**) “gift”	Sanskrit	दान “gift, donation”
	duka “sadness”	**dukha** “pains”	Sanskrit	दुःख “distress, suffering”
	gita “song”	**gili** “id.”	Sanskrit	गीति “singing, balade, song”
	jelara “common people”	**ʒene** “persons”	Sanscrit	जन “person, common people”
	jiwa “life, soul, sanity”	**ʒivi(pen**) “life”	Sanscrit	जी “soul”
	kerja “work”	**kerel** “to do”	Sanscrit	करोति “to do”
	mala “stain, dirt”	**mel** “id.”	Sanscrit	मल dust, impurity”
	muka “face”	**muj** “face, mouth”	Sanskrit	मख “id,”
	nama “name”	(**a**)**nav** “id.”	Sanskrit	नाम “id.”
	pustaka “book”	**pustik**[17] “id.”	Sanskrit	पुस्तक “id.”

[17] **Pustik** is a neologism, borrowed from India in the ’70 by Czechoslovak Rroms and anew independently by Yugoslav Rroms, some 10 years later.

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

Here again, some concepts are represented in priority, but they differ from the other list above: *face*, *pain*, *song*, *place*/*earth*, *deity*. The point here, in terms of pedagogy, is not a list of possibly useful words, to be used on the spot, but the vision of a common linguistic and cultural heritage, which underpins human groups as different and remote as those speaking urban slang in Europe and common people – *jelara*, in South-Eastern Asia – 18.000 km from one another, all this present also in Rromani language.

Conclusion

To close, I would like only to emphasize that teaching Sanskrit to young Rroms would fulfill a need and a wish at the same time – among so many young Rroms who are in search of a way to substantiate their Indian origin. It would also make no sense to extend to Rromani students the teaching method of Sanskrit, as used with mainstream students, due to the Indian baggage Rroms have already integrated together with their mother tongue – and due to the fact that Rromani is probably the living language, which is currently closest to Sanskrit. Pupils just have to become aware of this innate baggage – and some games on line and/or some poems incorporated into the learning process would be of great profit for this purpose. Teaching Sanskrit to young Rroms means teaching it to people who are eager for cultural links with the "Baro Than[18]" and it will make vivid this language within Europe, a significant step forward in bringing closer European civilisation to the inexhaustible source of the wisdom of Indian culture, more than ever needed in our times. This goal could be achieved in completing the university on-line course of Rromani language and culture "Restore the European Dimension of Rromani language and culture" with a specific component "Sanskrit language and culture" (www.red-rrom.com – password: r3drr0m).

[18] "Baro Than" is a phenomenon of naïve paronymic attraction and reinterpretation of one of the names of India (Bharat, unknown among Rroms) according to Rromani words and phonotactism: Bharat > Baro Than "great place".

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)

Bibliography

Ανδριώτης, Ν. Π., 1971, *Ετυμολογικό λεξικό της κοινής νεοελληνικής*. Thessaloniki.

APTE, Vamanrao S., 1989, *The Practical Sanskrit English Dictionary (revised and enlarged edition)*, Delhi: Motilal Banarsidass Publ.

BLOCH, Jules, 2010, *The Formation of the Marāṭhī Language)*, Delhi: Motilal Banarsidass Publ.

COULSON, Michael, 1976, *Teach yourself Sanskrit*, Oxford: Hodder & Stoughton.

COURTHIADE, Marcel, 1991, „Hauptarten der morphologischen Anpanssung der romani Lexemen in der serbokroatischen Gaunersprache "šatrovački" von Bosnien und Herzegowina„ in: *In the margin of Rromani: Languages in Contact*, ed. by Bakker, Peter & Marcel Courthiade. Publikaties van het Instituut voor Algemene Taalwetenschap, Amsterdam: Universiteit van Amsterdam.

COURTHIADE, Marcel, 2012, „Próba bilansu dotychczasowych hipotez o pochodzeniu narodu rromskiego", In: *Studia romologica Nr 5*, Tarnów.

***, 2001, *Diccionario de la Real Academia Española* (*DRAE*), 22nd edition, Madrid.

ГОРГОНИЕВ, Юрий А., 1975, *Кхмерско-русский словарь*. Moscow.

ITU, Mircea, 2001, *Manual de limba sanscrită. Editura Orientul latin*, Braşov.

КАУШИК, Рама Кришна, 2014, *Санскритско-български речник*. Sofia.

КОЧЕРГИНА Вера А., 1978, *Санскритско-русский словарь*. Moscow.

MASICA, Colin P., 1993, *The Indo-Aryan Languages*, Cambridge: Cambridge Language Surveys.

MAX, Frédéric, 1991, „Apports tsiganes dans l'argot français moderne", in: *In the margin of Rromani: Languages in Contact*, ed. by Bakker, Peter & Marcel Courthiade. Publikaties van het Instituut voor Algemene Taalwetenschap, Amsterdam: Universiteit van Amsterdam.

MONIER-WILLIAMS Monier, 1976, *A Sanskrit-English Dictionary*, New Delhi: Sharada Publ.

МИНИНА, Галина Ф., 1976, *Бирманско-русский словарь*. Moscow.

PISCHEL, Richard, 1981, *Grammar of the Prakrit Language*. Delhi: Banarsidass Publ.

BDD-A26081 © 2017 Editura Muzeul Literaturii Române
Provided by Diacronia.ro for IP 216.73.217.25 (2026-09-15 08:21:04 UTC)