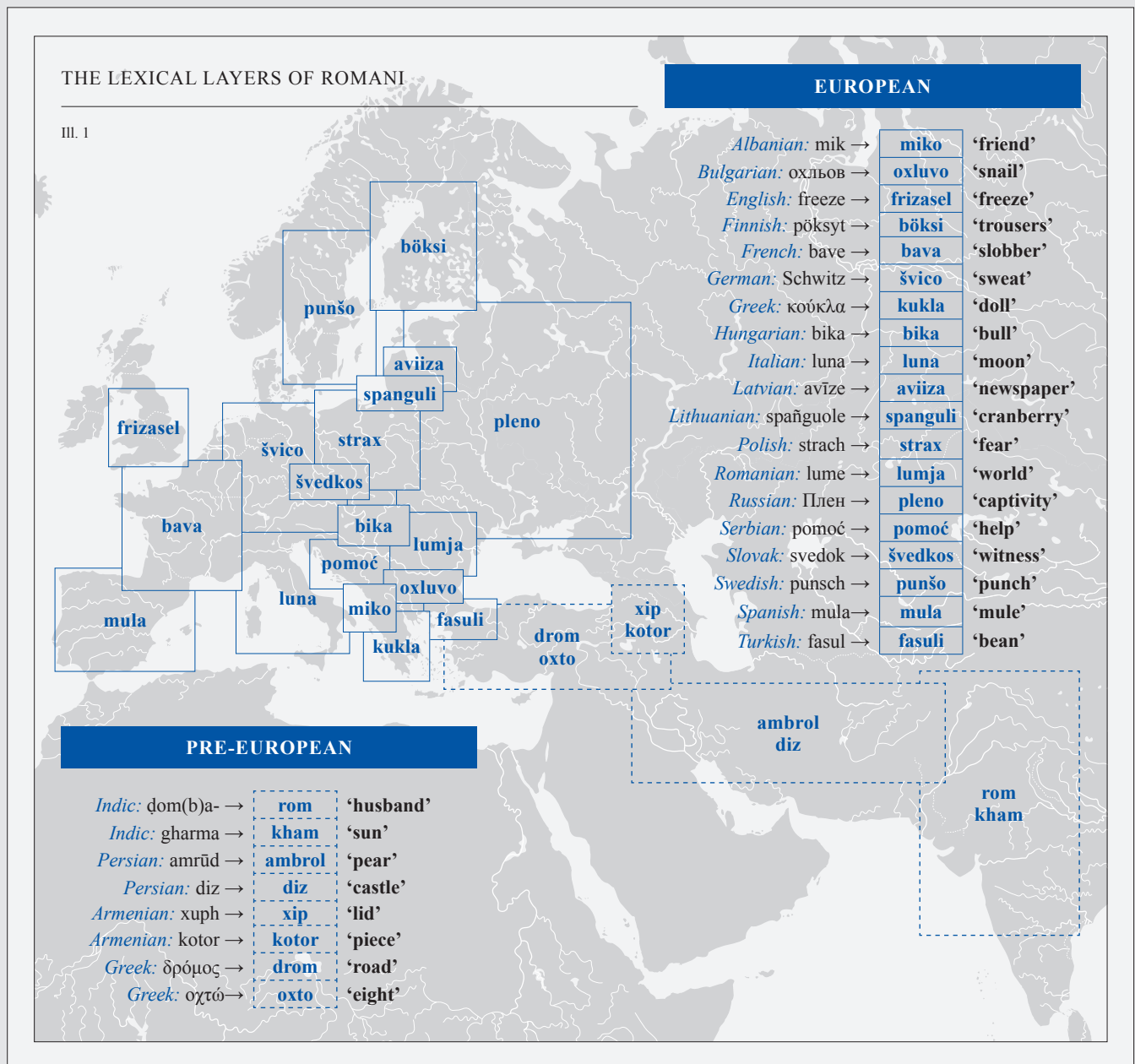


# Words / Lexicon

2.0

Romani-Project Graz / Dieter W. Halwachs

The lexicon of Romani summarises words of Indian origin and of all other languages Romani varieties had contact with. Basically these lexical layers are divided into a pre-European and a European vocabulary.



This “designed” sentence of Lovara Romani demonstrates the Asiatic part of the migration as well as the following way on the Balkan peninsula into Walachian bondage: Ill. 2 source: Heinschink (1999: 178)



‘man’	[+ROMANI]	rom	NEUTRAL	manuš	muř	[-ROMANI]	gadžo
‘woman’		romni		manušni	džuvli		gadži
‘boy’		čhavo		[manušoro]	[muřšoro]		raklo
‘girl’		čhaj		[manušori]	[džuvlori]		rakli <span style="float: right;">Ill. 3</span>

## INTRODUCTION

The lexicon of Romani consists of several layers that can be sub-grouped into a pre-European and a European part. The so-called Indian “words of origin” and the either “earlier” or “later” loans from Persian, Armenian, and Byzantine Greek make up the pre-European lexicon. These “inherited words” (Boretzky 1992) compri-

se about 700 roots from Indian, likely no more than 100 roots from Persian and other Iranian languages, at least 20 from Armenian and up to 250 from Greek. This total of more than 1000 lexemes, however, does not exist in any variety in its full amount. “Recent” loans adopted at a later point in time stem from a range of different European languages of contact (Ill. 2). Among these, loans from Southern Slavic form

the last general layer of the Romani varieties spoken in Europe today. Therefore, the notion of a common lexicon in Illustration 3 is valid up to *praxo*, with all further lexemes being variety-specific: *lumja*, a loan from Romanian, pertains to the lexical inventory of Kalderaš Romani; *kolopa*, stemming from Hungarian, is used in Lovara Romani, and *berga*, which was adopted from German, is used in Sinti Romani.

## THE LEXICON AS ROAD MAP

The pre-European loan strata of Romani have made possible a reconstruction of the migratory route followed by Romani speakers. After their emigration from the northwest of the Indian subcontinent, the first sustainable language contact took place in what was at the time Sassanide Persia. As a consequence, there are elements in the Romani lexicon that can be traced back to middle Persian Pahlevi. It is impossible to define the duration of this contact. In fact, it is unclear whether Romani speakers actually dwelled in the region for a longer period

of time or whether they were engaged in a slow process of transition. Since Romani does not dispose of any Arabic loans at all, it can be assumed that Romani speakers must have left the Persian region before its Arabisation, that is, before the hybridisation of the Iranian and Arabic cultures. Most likely, they moved on via Armenia into the Byzantine Empire, where they stayed for a longer period of time. This assumption is supported by loans from Armenian on the one hand, and a strong influence of Byzantine Greek that by far exceeds mere lexical loans, on the other. This heavy influence on Romani is also reflected in the cardinal numbers listed below, which, along

with numerals of Indian origin, only comprise Greek loans (Ill.5).

The fact that there are no Turkish loans found among the Romani varieties of speakers who immigrated into Europe via the Balkans leads to the assumption that their emigration from Asia Minor must have taken place before the region was Turkicised, that is, before the hybridisation of the Arabic-Iranian-Islamic and Byzantine-Greek cultures under Ottoman political dominance. The Roma living in Europe today did not take part in this process. The varieties spoken by the Roma that remained to dwell in the Balkans and who were later influenced either directly or indirectly by Osmanic-

THE LEXICAL STRATA OF ROMANI

III.4

PRE-EUROPEAN				EUROPEAN			
<i>Indic</i>	<i>Persian</i>	<i>Armenian</i>	<i>Greek</i>	<i>Slavic</i>	<i>Romanian</i>	<i>Hungarian</i>	<i>German</i>
<b>kham</b>	<b>veš</b>	<b>khoni</b>	<b>drom</b>	<b>praxo</b>	<b>lumja</b>	<b>kolopa</b>	<b>berga</b>
‘sun’ ↑ <i>inc.</i> gharma	‘wood’ ↑ <i>per.</i> veša	‘suet’ ↑ <i>arm.</i> khoni	‘road’ ↑ <i>grc.</i> drómos	‘dust’ ↑ <i>sla.</i> prax	‘world’ ↑ <i>ron.</i> lume	‘hat’ ↑ <i>hun.</i> kalap	‘mountain’ ↑ <i>ger.</i> Berg

THE CARDINAL NUMBERS OF ROMANI

III.5

<b>jekh</b>	<b>duj</b>	<b>trin</b>	<b>štar</b>	<b>pandž</b>	<b>šov</b>	<b>efta</b>	<b>oxto</b>	<b>enja</b>	<b>deš</b>	<b>biš</b>	<b>tr(ij)anda</b>	<b>saranda</b>	<b>šel</b>
‘one’ ↑ <i>inc.</i> ekka-	‘two’ ↑ <i>inc.</i> d(u)vā	‘three’ ↑ <i>inc.</i> trīni	‘four’ ↑ <i>inc.</i> catvāra	‘five’ ↑ <i>inc.</i> pañca-	‘six’ ↑ <i>inc.</i> šaś/šat	‘seven’ ↑ <i>grc.</i> éftá	‘eight’ ↑ <i>grc.</i> óxtó	‘nine’ ↑ <i>grc.</i> énnjá	‘ten’ ↑ <i>inc.</i> daša-	‘twenty’ ↑ <i>inc.</i> vimšati	‘thirty’ ↑ <i>grc.</i> triánta	‘fourty’ ↑ <i>grc.</i> saránta	‘hundred’ ↑ <i>inc.</i> šata

Islamic culture, of course also dispose of Osmanic-Turkish loans. However, these loans are to be defined as pertaining to the European part of the lexicon along with all other loans from Slavic onwards, which in numbers dominate in all Romani varieties.

The European loan strata of Romani varieties provide clues as to the further migratory route taken by individual Romani groups in Europe. In the

case of the Romani variety spoken by the Finnish Kaale, for instance, lexemes from German suggest an early contact with the German language and most likely also point to a period of time spent in the German speaking area. Romanian elements in the lexicon of many present Romani varieties worldwide, such as Kalderaš, Gurbet and Čurara Romani, call to mind a common past characterized by bondage and slavery in

Walachia and Moldavia shared by these groups. Consequently, the latter are summarized as Vlax-Roma; the Romani varieties spoken by them are accordingly labelled Vlax Romani. As to the Kalderaš, elements from Russian that are found in present Swedish, French, North and South American Romani varieties, point to the fact that the group must have crossed Russia in the course of their migration.

**BASIC VOCABULARY**

The lexicon of individual Romani varieties by a great majority consists of European loans. Moreover, each word of the respective contact language is also a potential Romani lexeme that can be integrated if necessary. However, as the so-called basic vocabulary – that is designations expressing existentially important entities, states, and processes – of each individual Romani variety by a gre-

at majority consists of words of Indian origin, Romani can be characterised as a New Indo-Aryan language from a lexical perspective as well.<sup>1</sup>

The basic vocabulary covers existentially important basic domains; these are areas close to the human being concerning life and the environment. According Romani terms such as personal designations, for instance, can be traced back to Indian (III 3).

What is striking in this respect, is the differentiation according

to ethnic criteria, marked in the table by the feature of [± romani]. For ethnically neutral terms, the pair of *murš* / *džuvli* focuses gender difference, while *manuš* / *manušni* emphasize the “human” aspect. The neutral terms for ‘boy’ / ‘girl’ given in parenthesis are more or less common diminutive forms of the corresponding terms for ‘man’ / ‘woman’.

Designations for human beings, essentially of Indo-Aryan origin, also function as kinship terms. Accordingly, terms

<sup>1</sup> A similar situation holds true for English. Even though only about one third of the English vocabulary is “Western Germanic” by origin, English is classified as a Western Germanic language, due to the fact that its basic vocabulary largely pertains to this third.

GENERATION → GRADE ↓	+2	+1	Ø	-1	-2
DIRECT	‘grandson’	čhavo	phral	dad	papu(s)
	‘granddaughter’	čhaj	phen	daj	mami
INDIRECT		‘nephew’	‘male cousin’	kak(o)	
		‘niece’	‘female cousin’	bibi	

III. 6

<b>ivend</b>	← <i>inc.</i> hemanta	‘winter’	<b>rat</b>	← <i>inc.</i> rātrī	‘night’	<b>kurko</b>	← <i>grc.</i> kuriaké	‘week, sunday’
<b>nilaj</b>	← <i>inc.</i> nidāgha	‘summer’	<b>berš</b>	← <i>inc.</i> varša	‘year’	<b>ciros</b>	← <i>grc.</i> kairós	‘time’
<b>dives</b>	← <i>inc.</i> divasa	‘day’	<b>masek</b>	← <i>inc.</i> māsa	‘month’			

III. 7

describing direct relatives of the same generation – *rom / romni* ‘husband’ / ‘wife’, and *phral / phen*, ‘brother’ / ‘sister’ – as well as those designating relatives of the following and directly previous generations – *čhavo / čhaj* ‘son’ / ‘daughter’ and *dad / daj* ‘father’ / ‘mother’ – have Indian origins.

In contrast, terms for the grandparent generation are loans from Greek – *papus / mami* ‘grandfather’ / ‘grandmother’. Terms designating indirect relatives of the first previous generation, that is the generation of parent siblings, also belong to the early loans and most likely stem from Persian – *kak / bibi* ‘uncle’ / ‘aunt’. All of the other relational

terms are either variety-specific loans from European contact languages or paraphrases. Illustration 6 summarizes the kinship system and the lexical layers of the according terms from an individual’s point of view (III.6).

The human body is another basic domain with a great majority of terms with Indic origin: these terms comprise the body parts, functions, movements, physical and mental states, etc. Numerals (III.5), as well as terms describing nature – landscape, weather, plants, animals, etc. – terms for shelter, tools and basic foods along with terms describing professions and social functions, also belong to the basic vocabulary. As demonstrated by the

examples from the domain of time, a great majority of according items is composed by lexemes of Indo-Aryan origin (III.7).

Similarly to almost all other basic areas, this domain also contains some pre-European loans along with terms of Indic origin. As these loans frequently originate from Byzantine Greek, Romani speakers can be assumed to have stayed in Asia Minor for a longer period of time characterized by intense language contact. The resulting influence of Greek on Romani goes far beyond the lexical domain and is thus further subject matter in the presentation of Romani morphology and syntax.

**BIBLIOGRAPHY:** The most extensive lexical resource of Romani is ROMLEX: <http://romani.uni-graz.at/romlex/>. Valuable information is i. a. provided in: **Boretzky, Norbert (1992)** *Zum Erbwortschatz des Romani*, Zeitschrift für Phonetik Sprachwissenschaft und Kommunikationsforschung 45, p. 227–251. | **Boretzky, Norbert / Igla, Birgit (1994)** *Wörterbuch Romani-Deutsch-Englisch für den südosteuropäischen Raum: Mit einer Grammatik der Dialektvarianten*. Wiesbaden: Harrassowitz. | **Heinschink, Mozes (1999)** *Sprachen der Sinti und Roma*, In: Ohnheiser, Ingeborg / Kienpointner, Manfred / Kalb, Helmut (eds.) *Sprachen in Europa*. Innsbruck: p. 177–190. | **Matras, Yaron (2002)** *Romani: A linguistic introduction*. Cambridge: Cambridge University Press. | **Sampson, John (1926)** *The Dialect of the Gypsies of Wales. Being the Older Form of British Romani Preserved in the Speech of the Clan of Abram Wood*. Oxford: Clarendon Press. [reprint 1968]