

ROMANI | LANGUAGE

Factsheets on Romani Language: General Introduction

0.0

Romani-Project Graz / Dieter W. Halwachs

The *Roma*, *Sinti*, *Calè* and many other European population groups who are collectively referred to by the mostly pejorative term “gypsies” refer to their language as *Romani*, *Romanes* or *romani čhib*. Linguistic-genetically it is a New Indo-Aryan language and as such belongs to the Indo-Iranian branch of the Indo-European languages. As an Indo-Aryan diaspora language which occurs only outside the Indian subcontinent, Romani has been spoken in Europe since the Middle Ages and today forms an integral part of European linguistic diversity.

The first factsheet addresses the genetic and historical aspects of Romani as indicated. Four further factsheets cover the individual linguistic structural levels: lexis, phonology, morphology and syntax. This is followed by a detailed discussion of dialectology and a final presentation of the socio-linguistic situation of Romani.

1_ ROMANI: AN INDO-ARYAN LANGUAGE OF EUROPE

deals with the genetic affiliation and with the history of science and linguistics of Romani and Romani linguistics.

2_ WORDS

discusses the Romani lexicon which is divided into two layers: Recent loanwords from European languages are opposed by the so-called pre-European inherited lexicon. The latter allowed researchers to trace the migration route of the Roma from India to Europe.

3_ SOUNDS

describes the phonology of Romani, which includes a discussion of typical Indo-Aryan sounds and of variety-specific European contact phenomena.

THE OPINIONS EXPRESSED IN THIS WORK ARE THE RESPONSIBILITY OF THE AUTHORS AND DO NOT NECESSARILY REFLECT THE OFFICIAL POLICY OF THE COUNCIL OF EUROPE.

All rights reserved. No part of this publication may be translated, reproduced or transmitted in any form or by any means, electronic (CD-Rom, Internet, etc.) or mechanical, including photocopying, recording or any information storage or retrieval system, without the prior permission in writing from the Publishing Division, Directorate of Communication (F-67075, Strasbourg cedex or publishing@coe.int)

4_ FORMS

is a description of the morphology according to its traditional division into nouns, verbs and particles and again deals with the dichotomy between European and pre-European elements.

5_ SENTENCES

presents the syntax of Romani which is syntactic-typologically classified as an SVO language due to its subject-verb-object sequence in a neutral declarative sentence.

6_ DIALECTOLOGY

deals with the plurality of Romani which manifests itself in a variety of dialects and varieties. The rather static representation of the dialect groups in Factsheet 6.0 is verified by a more dynamic view in Factsheet 6.1.

7_ SOCIOLINGUISTICS

deals with the socio-political and socio-cultural situation of Romani and discusses both its position and its functionality in the collective linguistic repertoire of different groups of speakers.

The open form of factsheets was also selected for the description of Romani to allow extensions and additions both in the thematic areas and other areas, such as to provide detailed descriptions of individual varieties or dialects. As with the other topics covered, this constitutes an initial step and a basic first presentation and is by no means an exhaustive, final presentation of the subject area.

LITERATURE

In addition to Matras, Yaron. 2002. *Romani: A linguistic introduction*. Cambridge: Cambridge University Press, the standard work on Romani and Romani linguistics which is also repeatedly cited and referenced in the individual factsheets, we would like to refer you to the websites of the two Romani projects involved in the preparation of the factsheets:

<http://romani.uni-graz.at/romani/> and <http://romani.humanities.manchester.ac.uk/>

ACKNOWLEDGEMENTS

We are indebted to all the authors who contributed actively to the factsheets and the passive contributors who are cited as sources. We are grateful that we could rely on the expertise and support of Mozes F. Heinschink, who i.a. made the *Romani Project at the University of Graz* possible – and as a consequence also the factsheets – and who has been and still is of great help in all activities related to the topic.

THE LANGUAGE FACTSHEETS' TEAM

editing: *Romani Project Graz* | translation: Ulla & Henry Briscoe | layout and design: Marcus Wiesner | coordination: *Romani Project at the University of Graz* in close cooperation with the *Romani Project at the University of Manchester* and the *Council of Europe project Education of Roma Children in Europe*

ROMANI | LANGUAGE

Factsheets on Romani Language: ROMANI – An Indo-Aryan Language of Europe

1.0

Romani-Project Graz / Dieter W. Halwachs

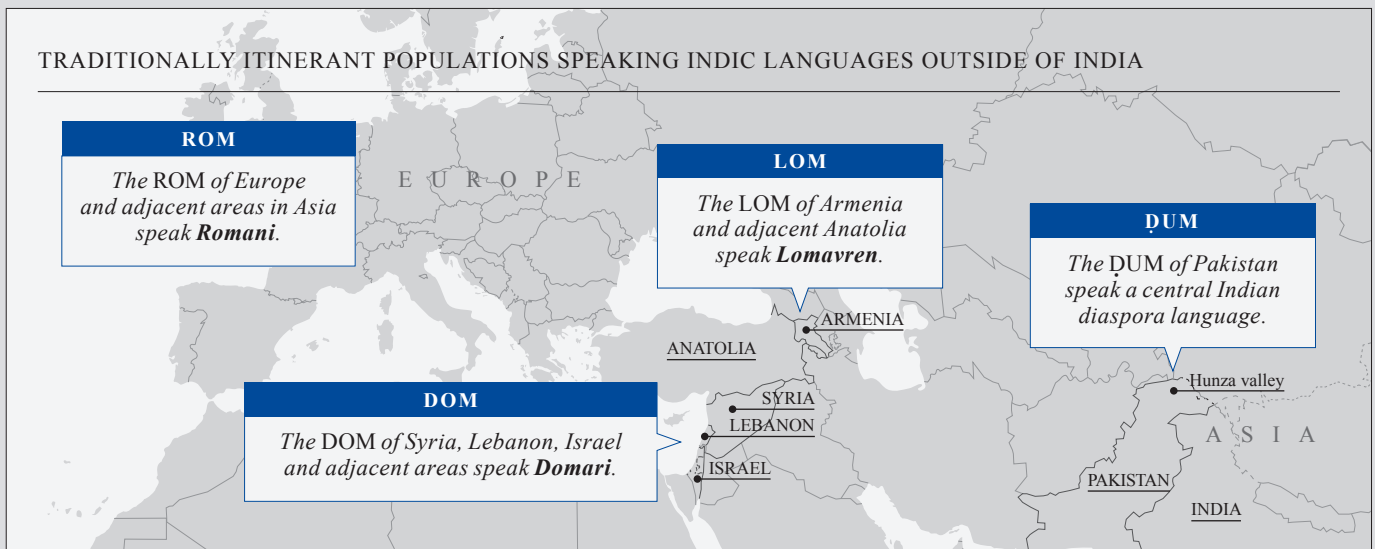
Romani and **Romanes** are the general names for the language of the Roma, Sinti, Kale and all other ethnic groups in Europe who speak or spoke an Indic, or respectively Indo-Aryan language. These population groups are collectively called “Gypsies”, a term mostly used in a derogative sense.

- **Romanes** is derived from an adverb: *Džanes romanes?* ‘Do you know/speak “Romani”?’
Romanes is used almost exclusively in German speaking areas.
- **Romani** is derived from an adjective: *Romani čhib* ‘Romani tongue, Romani language’.
Romani – often spelled Romany in English texts– is used internationally. Moreover, most names in New-Indo-Aryan languages, to whom Romani is genetically affiliated, end in -i: Assami, Bengali, Gujarati, Hindi, Maharathi, Panjabi, etc. The international name thus simultaneously implies its kinship to the language group.
- **Roman** is used by the Roma of Burgenland for their variety, while the Sinti refer to their varieties as **Rómanes** (with a stress on the o) or **Sintitikes**.

In the following we exclusively use the term Romani. Single Romani varieties are labelled as Burgenland Romani, Kalderaš Romani, Lovara Romani, Sepečides Romani, Sinti Romani, Ursari Romani etc.

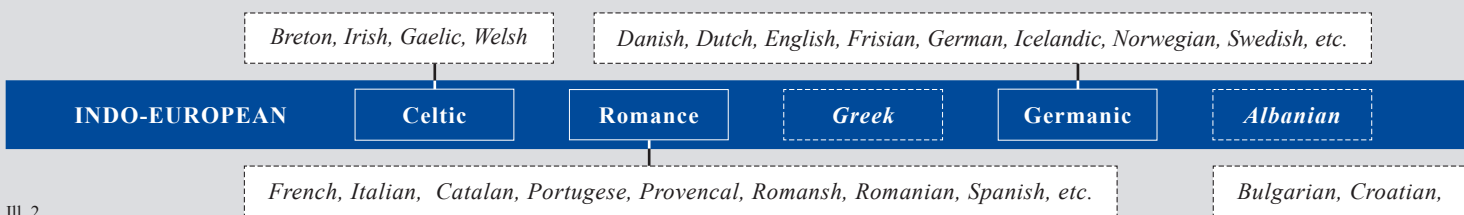
THE OPINIONS EXPRESSED IN THIS WORK ARE THE RESPONSIBILITY OF THE AUTHORS AND DO NOT NECESSARILY REFLECT THE OFFICIAL POLICY OF THE COUNCIL OF EUROPE.

All rights reserved. No part of this publication may be translated, reproduced or transmitted in any form or by any means, electronic (CD-Rom, Internet, etc.) or mechanical, including photocopying, recording or any information storage or retrieval system, without the prior permission in writing from the Publishing Division, Directorate of Communication (F-67075, Strasbourg cedex or publishing@coe.int)



III. 1

ROMANI AS A MEMBER OF THE INDO-EUROPEAN FAMILY OF LANGUAGES



III. 2

ROMA, SINTI, KALE, ETC.¹

The ethnonym *Roma* or *Řoma* – ř = /r/ – is the plural of *Rom* / *Řom* ‘husband, man’. Most groups use *rom* and *romni* as kinship terms to mean ‘husband’ and ‘wife’ respectively, but also as general designations for persons who are members of the group – *rom* ‘(Romani) man’, *romni* ‘(Romani) woman’. As Sinte Romani lacks the latter meanings Sinti do not accept *Roma* as a name for the collective, ‘the Romani people’. To counteract a widespread error it has to be noted that *rom* does not mean ‘human being’. Instead the general use of the noun *manuš* has been observed in Romani.

The designation *Romas* is also frequent. It is based on the wrong assumption of *Roma* as a singular. As *Romas* is also used by competent speakers

of Romani in statements in English, German, etc. the term has come to be generally understood and accepted and is thus to be interpreted as a neologism.

Some groups have adopted other labels as well: *Romaničal*, *Kale*, *Manuš* and *Sinti* are some examples of self-appellations used by Romani-speaking populations. *Sinti* is used by those subgroups who entered the German speaking area at a relatively early point in time. The *Sinti* of France refer to themselves as *Manuš* resp. *Manouche*. *Romaničal* is found primarily among British groups. Some of them also claim *Gypsies* as their self-designation. *Kale* ‘blacks’ is used by the *Calé* who have been living for a long time on the Iberian Peninsular and by the *Kaale* of Scandinavia residing in Finland and Sweden. *Roma* is used as self-designation among all the groups living in central and eastern Europe or,

respectively, by those who emigrated in the 19th and 20th centuries from central and eastern Europe to western Europe and overseas.

Other group-specific names were adopted from other languages. These are often based on traditional occupations, such as in *Kalderaš* ‘tanners’ from Romanian *căldărar*, *Čurari* ‘sieve-makers’ from Romanian *ciurar*, *Ursari* ‘bear-leaders’ from Romanian *ursar*, *Sepeči* ‘basket-weavers’ from Turkish *sepetçi*, *Bugurdži* ‘drill-makers’ from Turkish *bugurcu*, *Arli* or *Erli* ‘settled’ from Turkish *yerli* and *Lovara* ‘horse-dealers’ from Hungarian *ló* ‘horse’.

Non-Roma are usually referred to as *gadže* – *gadžo* ‘non-Romani man’, *gadži* ‘non-Romani woman’. This is an ancient designation for outsiders which is also found among the Middle Eastern *Dom* as *kaddža*, among the Armenian

¹ Parts of this chapter have been taken over from the Homepage of the Manchester Romani Project: Matras, Yaron. History of the Romani Language: Names. 4. Aug. 2008 <http://romani.humanities.manchester.ac.uk/files/11_names.shtml>.

EXAMPLES WHICH ILLUSTRATE TURNER'S ASSUMPTION OF AN EARLY MIGRATION OF THE ROMA WITHIN INDIA

Innovations shared by Romani with central Indic languages such as Hindi: these regular sound changes both suggest a relation to Sanskrit and a longer residence of Romani speakers in the central Indian area:

	'tree'	'blood'
Sanskrit:	ṛ kṣ̣ a	ra kt a
Romani:	ru kh	ra t
Hindi:	rū kh	rā t ā

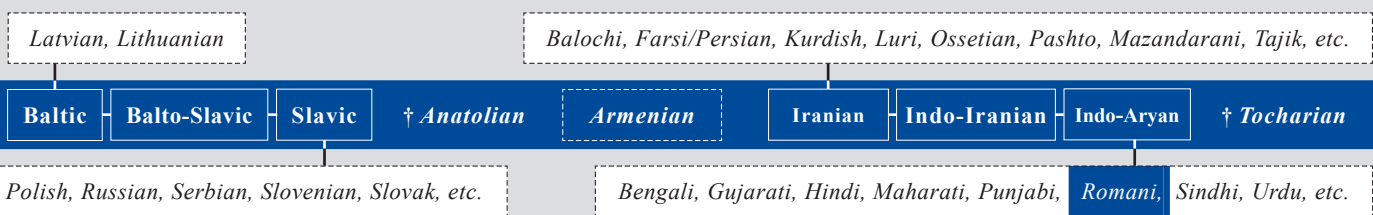
Differences between Romani and central Indic languages: these are conservative features of Romani as opposed to central Indian innovations, which support the assumption of an early emigration from central India:

	'grape'	'good'
Sanskrit:	dr ākṣ̣ā	mi ṣṭ̣ a
Romani:	dr akh	mi ṣt o
Hindi:	d ākh	mī th ā

Parallels to innovations in northwest Indic languages such as Sindhi, which are not found in the languages of central India, suggest that over a longer period of time the speakers of Romani resided in the northwest of the Indian subcontinent:

	'bent'	'tooth'
Sanskrit:	va ṅk uḥ	da nt aḥ
Romani:	ba ng o	da nd
Sindhi:	wi ṅg u	ḍa nd u
Hindi:	bā k a:	dā t

III. 3



Lom as *kača*, and among the different groups of *Ḍom* in India as *kājwā*, *kajjā*, or *kājarō*. In some regions, more specific names are found. For example,

in the Balkans, Muslims (including Turks and Albanians) are referred to by the Roma as *xoraxane* or *koraxane*. Slavs are referred to as *das*, based on

the same original Indic word for 'slave' – a designation inspired by the similarity between Greek *sklavos* 'slave' and *slavos* 'Slav'.

ROM, LOM, DOM²

Romani is the only Indo-Aryan language used exclusively in Europe since the Middle Ages. It is part of the phenomenon of the so called Indic diaspora languages spoken by travelling communities of Indian origin outside of India. The name *Rom* or *Řom* has related cognates in the names employed by other travelling (peripatetic) communities that speak Indic languages or use a special vocabulary derived from Indic: the *Lom* of the Caucasus and Anatolia insert Indic vocabulary into *Lomavren*, their va-

riety of Armenian. The *Dom* of the Near East, originally metalworkers and entertainers, speak *Domari*, one of the most conservative modern Indo-Aryan languages. In the Hunza valley in the north of Pakistan, the population called the *Dum*, who are also metalworkers and musicians, speak a central Indic (i.e. not a local) language. [III. 1]

Based on the systematicity of sound changes attested in these languages, we know with a fair degree of certainty that these names all derive from the Indian term *ḍom*. In various parts of India itself, groups known as *Ḍom* are castes of commercial nomads. References

to the *Ḍom* are made already by a number of medieval Indian writers such as Alberuni (~1020 CE), the grammarian Hemachandra (~1120 CE), and the Brahmin historian of Kashmir, Kalhana (~1150 CE). They all describe the *Ḍom* as a low-status caste whose typical trades included sweeping, making music, juggling, metal work and basket weaving, in some areas also seasonal farm-work. Similar occupations are still reported for the *Ḍom* in modern India. The self-designation *ḍom* > *řom* thus appears to have originally been a caste-designation, used in different regions by different populations with similar types of trades.

2 This chapter has been taken over from the Homepage of the Manchester Romani Project: Matras, Yaron. History of the Romani Language: Origins. 4. Aug. 2008 <http://romani.humanities.manchester.ac.uk/files/11_origins.shtml>.

ROMANI-LINGUISTICS

Up until the end of the 18th century, there was nothing but wildest speculations about the origin and the language of the Roma. In his 1697 paper titled “De civitate Norimbergensis commentario”, Johann Christoph Wagenseil characteristically describes Romani as a mixture of German, Jiddish, Hebrew, and phantasmagorical words, claiming that:

The first Gypsies were German-born Jews.

Even in 1781, the weekly journal “Neueste Mannigfaltigkeiten” published in Temesvár still reads:

Out of the mixing of Ethiopians, Troglodites and Egyptians, there evolved an individual, migrating folk, which has retained something of all three nations and whose descendents can be assumed to be today’s Gypsies.

The development of comparative methods in linguistics helped to clarify the origin of the Roma. Serious discussion of this topic started with Johann Christian Christoph Rüdiger, who in his study of 1782, titled “Von der Sprache und Herkunft der Zigeuner aus Indien”

he proves the relationship between Romani and Indic languages. [Ill. 2]

Criticizing prevailing discriminatory and romanticising prejudices, he calls the miserable living conditions of the Roma

a political inconsistency, which to tolerate our enlightened century should be ashamed of.

In 1783, one year after Rüdiger’s text, Grellmann’s book titled “Die Zigeuner” was published. It became the most widely known and read work of its time and had a significant influence on public opinion. Grellmann continues Rüdiger’s studies on a broad basis; his linguistic explanations are profound. Other than Rüdiger, however, he uncritically adopts the stereotypical and discriminatory prejudices of his time.

Sixty years later, Pott’s study “Die Zigeuner in Europa und Asien” marks another milestone in the linguistic discussion of Romani. Pott specifies its origin and thus the origin of the Roma. Accordingly, Romani is to be attributed to the northern Indic languages and

thus holds a blood relation to the proud Sanskrit.

It is worth mentioning the work of the Slavist Franz Miklosich as another

milestone of Romani linguistics. In two series of articles published between 1872 and 1881, Romani is i.a. classified into different dialects. Miklosich organizes the language into thirteen idioms based on linguistic influences of the various host countries, differentiating between Greek, Romanian, Hungarian and other dialects.

In 1926, Ralph L. Turner publishes an article called “The Position of Romani in Indo-Aryan”, in which he compares Romani, Sanskrit and various New-Indo-Aryan languages, concluding that there must have been an early relationship of Romani to the central group of Indo-Aryan languages. Thus, the ancestors of the Roma must have lived in the central Indian area, from where they migrated to the northwest of India. There, they must have stayed for a longer period of time, as Romani also shares innovations with New-Indo-Aryan languages of the northwest.

During the second half of the 20th century scientific interest in Romani intensifies, manifesting itself in both quantity and quality of the publications. With Yaron Matras’ comprehensive description “Romani: a linguistic introduction”, which was published in 2002, Romani linguistics finally establishes itself as an integral part of modern linguistics.

LITERATURE

Two books available in English which offer an overview of Romani language are proposed for further reading:

Peter Bakker et al. (2000) *What is the Romani Language?* Hatfield: University of Hertfordshire Press, which is an easy to read introduction intended for all those who have an interest in the language;

Yaron Matras (2002) *Romani: A linguistic introduction.* Cambridge: Cambridge University Press, which is a scientific introduction dealing with all details of Romani in the best way possible.

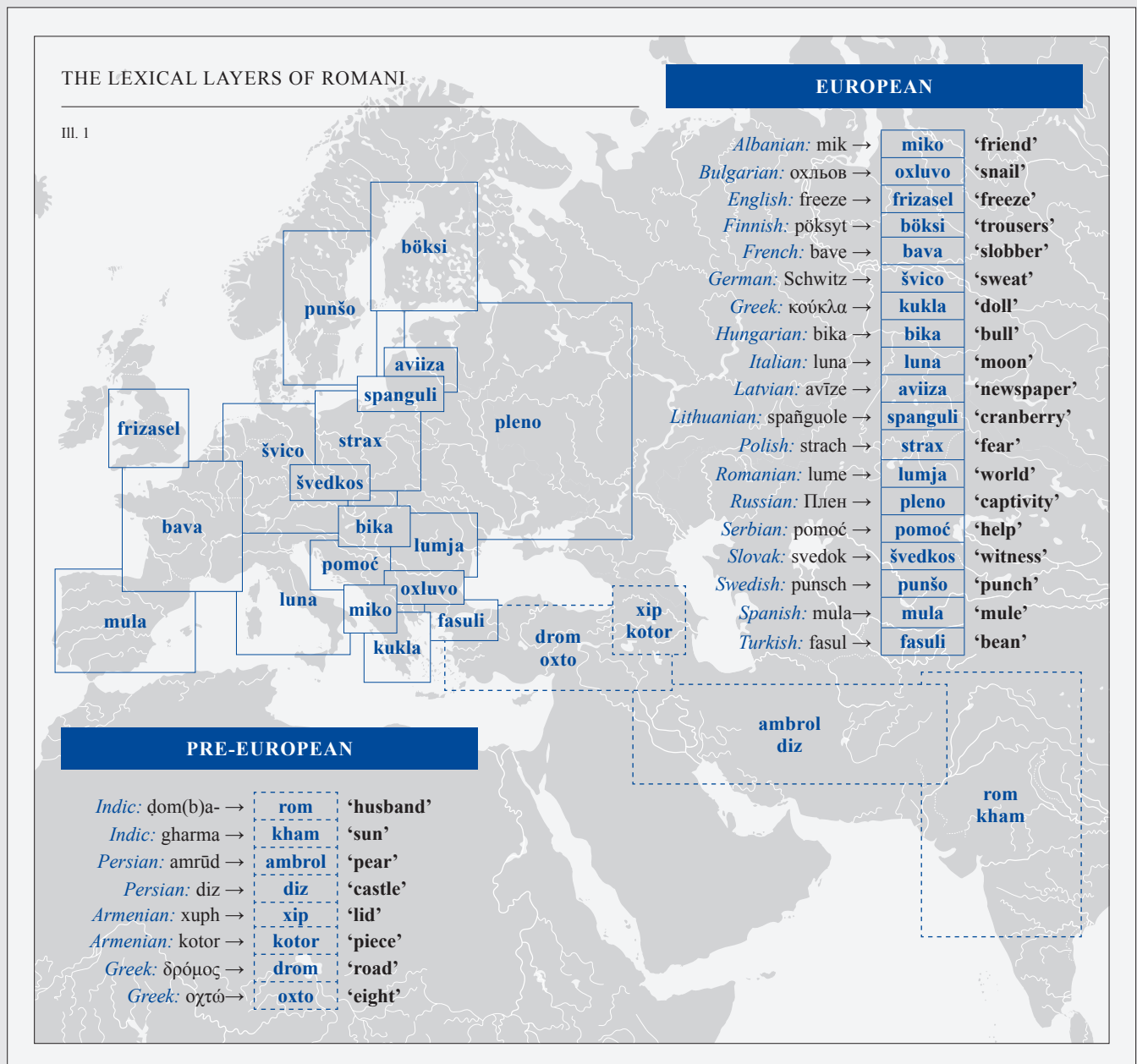
editing: Romani-Projekt at the University of Graz <<http://romani.uni-graz.at/romani>>, coordinated by Dieter W. Halwachs, using materials of the Manchester Romani Project <<http://romani.humanities.manchester.ac.uk>>, coordinated by Yaron Matras. | **translation:** Ines Farcher | **layout and graphics:** Marcus Wiesner

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 <small>Ill. 3</small>

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>
kham	veš	khoni	drom	praxo	lumja	kolopa	berga
‘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

jekh	duj	trin	štar	pandž	šov	efta	oxto	enja	deš	biš	tr(ij)anda	saranda	šel
‘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.¹

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

¹ 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

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

Dialects / Dialectology I.

6.0

Manchester Romani Project / Yaron Matras

There is a tendency in Romani linguistics to identify, tentatively at least, the dialect groups of the Balkan Dialects, the Vlax Dialects, the Central Dialects, the Northwestern Dialects, the North-eastern Dialects, the British Dialects and the Iberian Dialects.

ROMANI DIALECT GROUPS

III.1

BALKAN		VLAX		CENTRAL	
ARLI	“zis-dialects”	Southern	Northern	Southern	Northern
Krim	Bugurdži	Agia Varvara	KALDERAŠ	Roman	Bergitka
Sepečides	Drindari	Čergari	LOVARI	“ROMUNGRO”	Bohemian
URSARI	Kalajdži	GURBET	Mačvaja	Vend	East-Slovak
<i>etc.</i>		<i>etc.</i>	<i>etc.</i>	<i>etc.</i>	<i>etc.</i>

NORTHEAST	NORTHWEST		BRITISH	IBERIAN
Lithuanian	Sinte Manuš	Finnish	Welsh	(Caló)
Lotfika	Estrexarja	(Scandoromani)	(Angloromani)	(Errumantxela)
Polska	Manouche			
Xaladitka	RÓMANES			
<i>etc.</i>	<i>etc.</i>			

CAPITALS = dialect clusters () = Para-Romani varieties¹

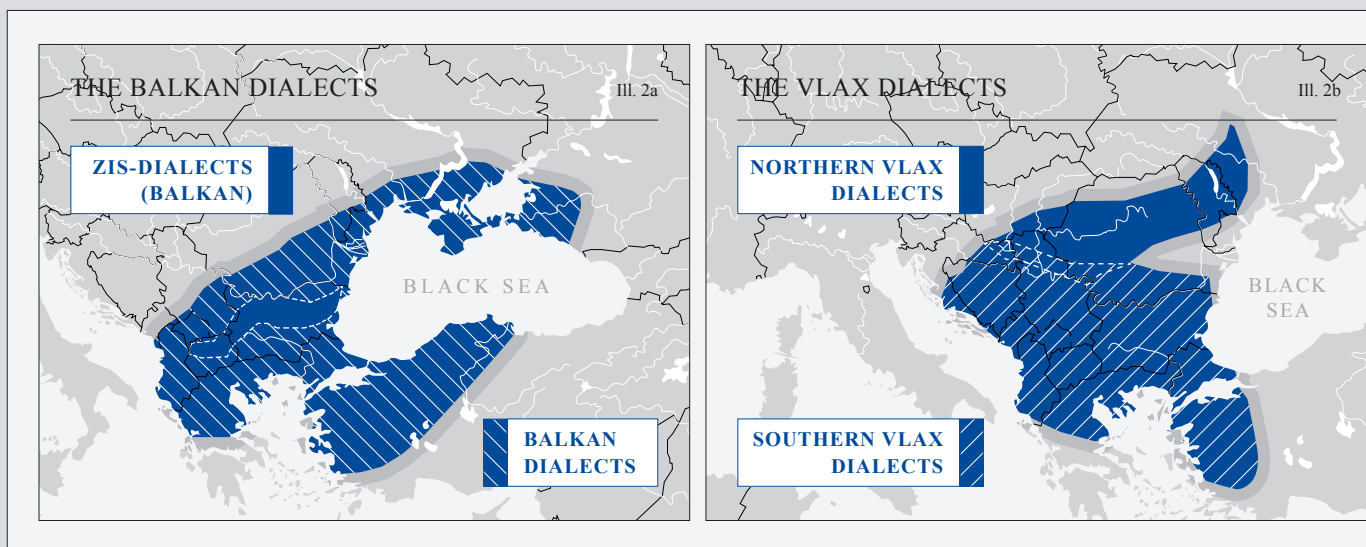
DIALECT CLASSIFICATION

There is no ‘easy’ way to classify dialects. One must first select the criteria on which a classification is to be based. Sometimes dialect classification is based strictly on geography, sometimes it is based strictly on the structural features –

lexicon, phonology, morphology – of the dialects. In the latter case, it is necessary to select those features that are of global relevance and that can be used as a reference grid to compare the different dialects and to determine the relationships among them. Scholars often disagree on which features should be given greater attention as a basis for a classification.

As a result, is not unusual to find different classification models. There is also an objective difficulty: Some dialects may share ‘typical’ features with two distinct dialect branches. Such transitional dialects are part of any linguistic landscape. It is therefore almost impossible to postulate clear-cut divisions between dialect groups or ‘branches’.

¹ Para-Romani varieties are ethnolects of the respective majority language with (mostly lexical) elements from Romani; e.g. Angloromani is a variety of English with Romani elements. Speakers of such varieties often label their ethnolects “their Romani language”.



Several factors are responsible for dialect differentiation in Romani:

- The migration of Romani-speaking populations throughout Europe, in different periods
- The geographical spread of structural changes, creating ‘isoglosses’
- The influence of contact languages
- Specific changes that are limited to the structure of individual dialects

THE BALKAN DIALECTS

Balkan dialects of Romani (also called: ‘Southern Balkan’, ‘Southern Balkan I’) are spoken in Turkey, Greece, Bulgaria, Macedonia, Albania, Serbia (Kosovo), Romania, Ukraine, and Iran. Dialects belonging to this group include Arli (Macedonia, Kosovo, Greece), Erli (Bulgaria), Mečkar (Albania), Sepeči (Greece, Turkey), Ursari (Romania), Crimean Romani (Ukraine), Zargari (Iran), and others. [III. 2a]

The diagnostic of characteristic features of this group of dialects include:

- *sine* for ‘he/she was’
- Third person pronouns *ov, oj, on*
- Shortened possessive pronouns *mo* ‘my’, *to* ‘your’
- Demonstratives *akava* and *adava*, sometimes *akavka*
- Loan verbs are usually adapted with *-in-* (but *-iz-* appears in the Black Sea region)
- 2nd pl. past tense *-en* in *tume kerden* ‘you.PL did’
- Future tense in *ka*, sometimes *ma*

THE BALKAN ZIS-DIALECTS

These dialects (also called: ‘Drindari-Bugurdži-Kalajdži group’, ‘Southern Balkan II’) are spoken in northern and central Bulgaria and Macedonia. They include the dialects of the Drindari/

Drindari, Kovački, Kalajdži and Bugurdži (but note that these are occupational designations, and are sometimes also found among groups speaking other types of dialects). [III. 2a]

The diagnostic of characteristic features of this group of dialects include:

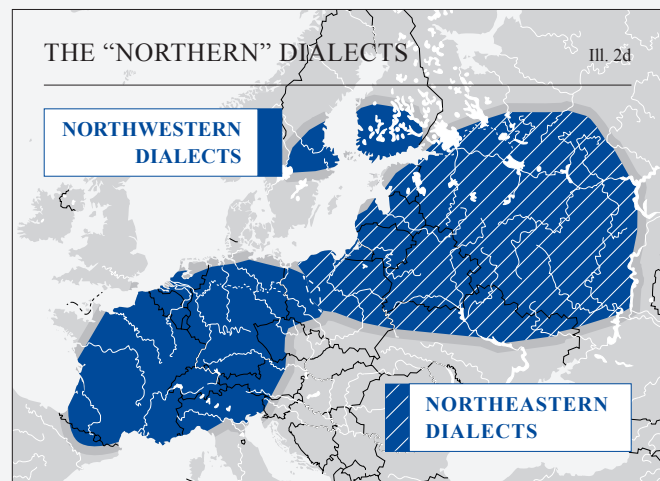
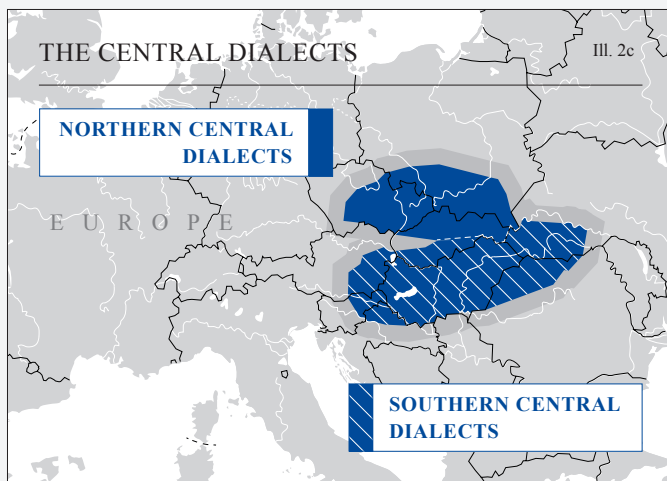
- Demonstratives *kaka* or *kəka*
- Loan verbs are adapted with *-iz-*
- *z-* in *zis* ‘day’ and *zi* ‘soul’
- *c-* in *buci* ‘work’ and *cin-* ‘to buy’
- reduction of *-e-kerela* > *kerla* ‘s/he does’

THE SOUTHERN VLAX DIALECTS

These dialects are spoken in Serbia and Montenegro, Croatia, Bosnia-Herzegovina, Macedonia, southern Romania, Bulgaria, Greece, Albania, and Turkey. They include the dialects of the Gurbet or Džambazi and groups known by other names such as Kalburdžu and Čergar. [III. 2b]

Their diagnostic features include:

- *seha* or *sesa* for ‘he/she was’
- Third person pronouns *vov, voj, von*
- Possessive *mənro* ‘my’, *čo* ‘your’
- Demonstrative *gava*
- Nominal plural endings in *-uri, -ura*
- Loan verbs are adapted with *-isar-*



- Past tense and copula 1.SG in *-em* as in *sem* 'I am'
- 2nd pl. past in *-en* in *tume kerden* 'you.PL did'
- Future tense in *ka*
- Negation in *in* or *ni*
- *dž-* in *dživeh* 'day'
- *-č-* in *buči* 'work'
- *-rn-* in *marno* 'bread'
- *-ej* in *dej* 'mother'
- *a-* in *ašunav* 'I hear'

NORTHERN VLAX DIALECTS

These dialects are spoken in Romania, Moldova, Hungary, Serbia, as well as in migrant communities worldwide. The most widespread and well-known Northern Vlax dialects include Kelderaš (Kalderaš), Lovari, Čurari and the dialect of the Mačvaja. [III. 2b]

Their diagnostic features include:

- *sas* for 'he/she was'
- Third person pronouns *vov*, *voj*, *von*
- Possessives *muro* 'my', *čiro* 'your'
- Nominal plural endings in *-uri*, *-ura*
- Loan verbs are adapted with *-isar-*, *-osar-* or shortened forms *-i-*, *-o-*.
- 1st person past tense in *-em*
- *džes* for 'day'
- *-č-* or *-kj-* in *buči*, *bukji* 'work'
- Negation in *či*
- *ž-* in *žanav* 'I know'
- *š-* in *šavo* 'child'
- *-nr-* or *-nř-* in *manro*, *manřo* 'bread'
- *-ej* in *dej* 'mother',
- *khanči* for 'nothing'

SOUTHERN CENTRAL DIALECTS

These dialects are spoken in Hungary, Slovakia, northern Slovenia, eastern Austria, Ukraine, Romania. They include the dialects of the Romungri, Vend and Burgenland Roman. [III. 2c]

Their diagnostic features include:

- *sina* for 'he/she was'
- Third person pronouns *ov*, *oj*, *on*
- Demonstratives in *ada*
- Loan verbs are adapted with *-in-*
- 2nd sg. past and copula *-al* as in *sal* 'you are'
- *kereha* for 'you shall do'
- Imperfect in *-ahi*
- *-h-* in *leha* 'with him'
- Loss of final *-s* as in *dive*, *di* 'day', *va* 'hand'
- *-r-* in *maro* 'bread'

NORTHERN CENTRAL DIALECTS

These are spoken in Slovakia, Czech Republic, Poland, Ukraine. They include East Slovak Romani and the dialect of the Bergitka Roma of Poland. [III. 2c]

Their diagnostic features include:

- *ehas* or *has* for 'he/she was', *hin* for 'he/she is'
- Third person pronouns *joy*, *joj*, *jon*
- Demonstratives in *kada*
- Loan verbs are adapted with *-in-*
- 2nd sg. past and copula *-al* as in *sal* 'you are'
- *kereha* for 'you shall do'
- Imperfect in *-as*
- *-h-* in *leha* 'with him?'; *-r-* in *maro* 'bread'

NORTHWESTERN DIALECTS

Spoken in Germany, Austria, France, Italy, Netherlands, Belgium, Finland. They include the Sinti-Manuš dialects of Germany, France and surrounding regions, as well as Finnish Romani or Kaale dialect. [Ill. 2d]

Their diagnostic features include:

- *his* for ‘he/she was’ (Sinti only)
- Third person pronouns *jov/job, joj, jon*
- Demonstratives in *kava*
- Loan verbs are adapted with *-av-* or *-ar-*
- 2nd sg. past and copula *-al* as in *sal/hal* ‘you are’
- *-h-* in *leha* ‘with him’
- *h-* in *hom* ‘I am’ and *ho* ‘what’ (Sinti only)
- *-r-* in *maro* ‘bread’; long vowel in *diives* ‘day’

NORTHEASTERN DIALECTS

Spoken in Poland, Lithuania, Latvia, Estonia, Russia, Belarus, Ukraine. They include the Polska Romani dialect, the North Russian or Xaladitka Romani dialect, the dialects of Latvia (Lotfitka), Lithuania, etc. [Ill. 2d]

Their diagnostic features include:

- *isys* for ‘he/she was’
- Third person pronouns *jov, joj/jej, jone*

- Demonstratives in *adava, dava*
- 2nd pl. past in *-e* as in *tume kerde* ‘you.PL did’
- 3rd pl. past in *-ine* as in *jone kerdine* ‘they did’
- 3rd sg. past in *-a* as in *jov kerdja* ‘he did’
- Loan verbs are adapted with *-in-*
- *pšal* for ‘brother’,
- *-r-* in *maro* ‘bread’

OTHER DIALECT GROUPS

Some additional dialects show their own distinct features. This is due either to a period of isolation from other dialects, or to the development of features shared with several different dialect branches. As separate groups we can define the following:

- British Romani, including English Romani and Welsh Romani (now extinct, and surviving as a vocabulary only, known as ‘Angloromani’)
- Iberian Romani, including Spanish Romani, Catalanian Romani, and Errumantxela (Basque Romani), (all extinct, and surviving as a vocabulary only, known as ‘Caló’)
- The Romani dialects of southern Italy, including Abruzzian and Calabrian Romani
- The Slovene Romani dialect (also known as Istriani, Hrvati or Dolenjski)
- The Romani dialects of Iranian Azerbaijan, Zargari and Romano (although these have very close connections to the Balkan dialects, see above)

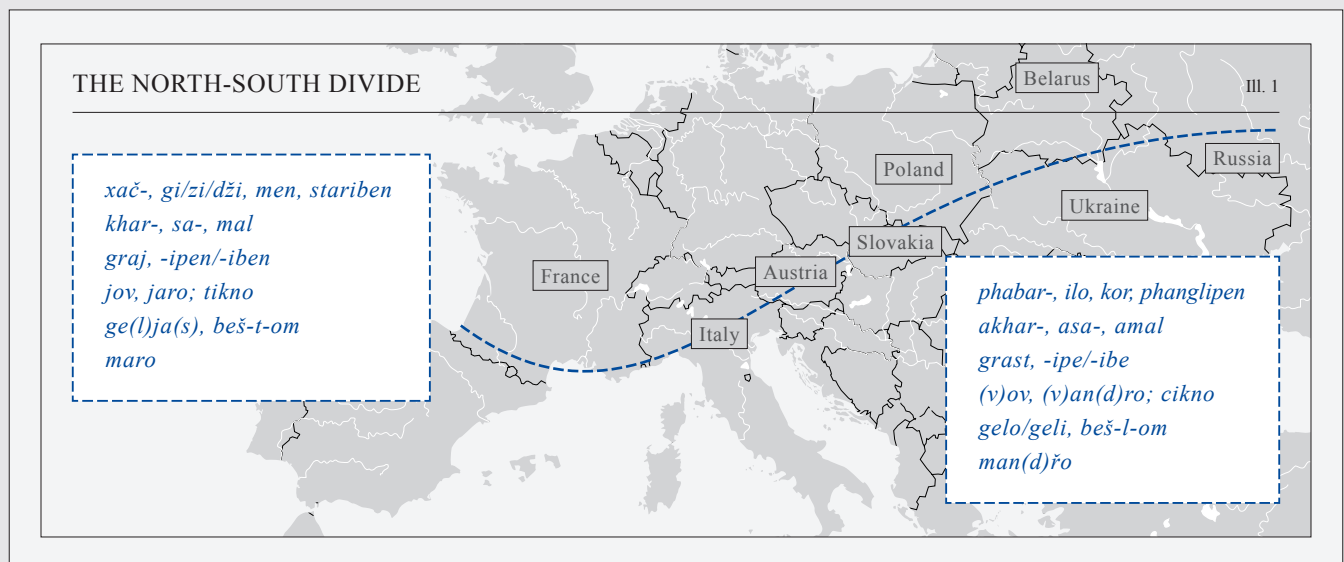
BIBLIOGRAPHY: Bakker, Peter, and Matras, Yaron. 1997. Introduction. In: Matras, Yaron, Bakker, Peter, and Kyuchukov, Hristo. eds. *The typology and dialectology of Romani*. Amsterdam: John Benjamins. vii-xxx. | Boretzky, Norbert. 1999a. Die Gliederung der Zentralen Dialekte und die Beziehungen zwischen Südlichen Zentralen Dialekten (Romungro) und Südbalkanischen Romani-Dialekten. In: Halwachs, Dieter W., and Menz, Florian. Eds. *Die Sprache der Roma: Perspektiven der Romani-Forschung in Österreich im interdisziplinären und internationalen Kontext*. Klagenfurt: Drava. 210-276. | Boretzky, Norbert. 1999b. *Die Verwandtschaftsbeziehungen zwischen den Südbalkanischen Romani-Dialekten. Mit einem Kartenanhang*. Frankfurt am Main: Peter Lang. | Boretzky, Norbert. 2000b. South Balkan II as a Romani dialect branch: Bugurdži, Drindari, and Kalajdzi. *Romani Studies*, fifth series, 10: 105-183. | Boretzky, Norbert. 2003. *Die Vlach-Dialekte des Romani. Strukturen – Sprachgeschichte – Verwandtschaftsverhältnisse – Dialektkarten*. Wiesbaden: Harrassowitz Verlag. | Kaufman, Terrence. 1979. Review of Weer Rajendra Rishi, Multilingual Romani Dictionary. *International Journal of the Sociology of Language* 19: 131-144. | Matras, Yaron. 2002. *Romani: A linguistic introduction*. Cambridge: CUP. (chapters 2 and 9). | Matras, Yaron. 2005. The classification of Romani dialects: A geographical-historical perspective. In: Schrammel, Barbara, Halwachs, Dieter W. & Ambrosch, Gerd, eds.: 7-26. | Schrammel, Barbara, Halwachs, Dieter W. & Ambrosch, Gerd, eds. 2005. *General and Applied Romani Linguistics*. Munich: Lincom.

Dialects / Dialectology II.

6.1

Manchester Romani Project / Yaron Matras

The divisions between the dialects can be plotted in the form of lines on the map, each line or 'isogloss' representing a difference in the realisation of a particular structural feature.



INTRODUCTION

Romani dialects have been grouped mainly on the basis of their geographical location: The conventional classificatory grid recognises a Northwestern, Northeastern, Central, Vlax (centred around Romania and neighbouring regions) and Balkan group, of which the latter three are each further sub-divided into a northern and a southern sub-group. This suggests that the divisions between the dialects can be plotted in the form of lines on the map, each line or 'isogloss' representing a difference in the realisation of a particular structural feature. How do such geographical divisions in the form of isoglosses come about?

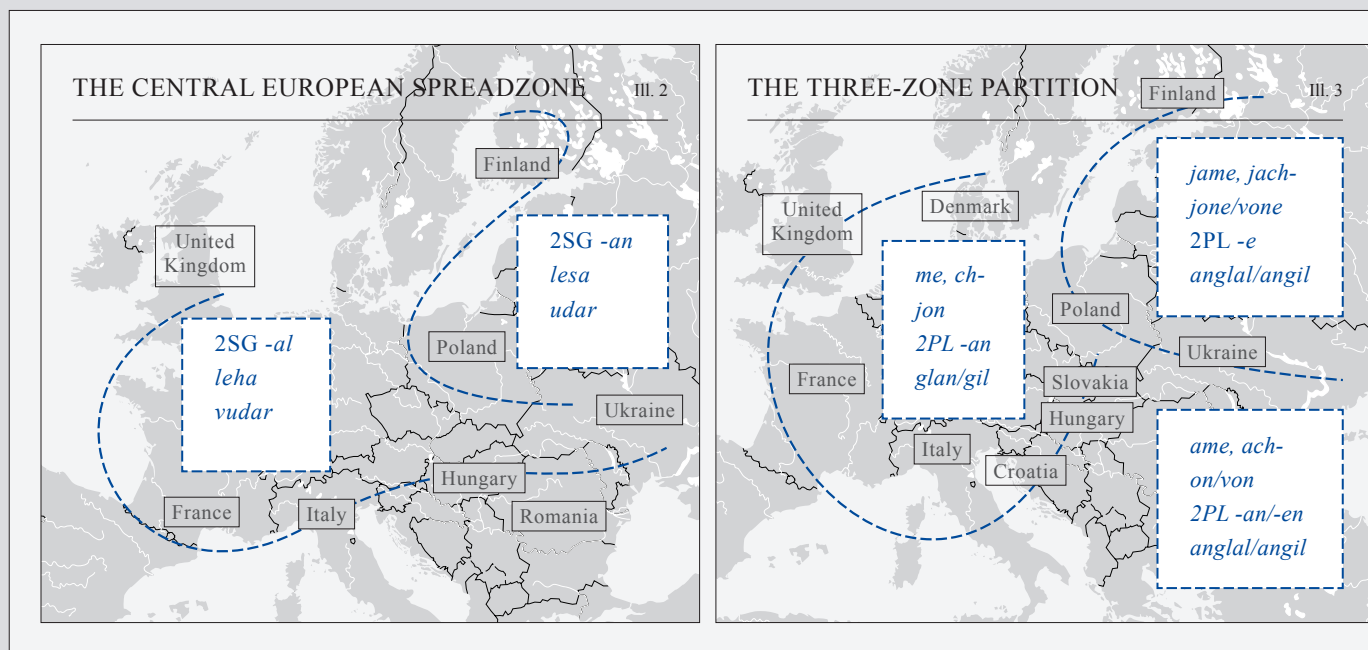
It is likely that the speech forms of different Romani families

and clans differed only slightly before they migrated into Europe between the late fourteenth and fifteenth century. Although they were often known as Travellers due to their specialisation in itinerant trades, most Roma did not habitually travel long distances but remained in familiar regions, interacting with a familiar population of settled clients. They acquired local languages, took on local religions, and adopted a role in local economies. Hybrid identities developed as each Romani population accommodated to its new environment while maintaining its own language, beliefs and customs.

The period that followed settlement in the individual regions in the sixteenth and seventeenth centuries was a period of rapid change du-

ring which distinct regional Romani identities emerged. This period left its mark on the speech forms of Roma in various locations. Each community developed its own structural preferences and adopted influences from the new contact languages. Documentation of Romani proliferated in the early eighteenth century, with scholars taking a keen interest in the language. By this time, Romani dialects were already as diverse as we know them today.

The divisions between the dialects are largely the result of changes that accumulated since the dispersion of Romani populations throughout the European continent. Some of the changes were local, limited to the speech form of several households or a group of closely related clans in a



small region. Roma continued to maintain contact networks with other Roma after settlement, and many changes were passed on to other communi-

ties. The passing of structural innovations from one community to another is known as ‘diffusion’. When plotting the spread of structural features on the

map we are therefore reconstructing the path of their historical diffusion among population groups and so across geographical space.

THE NORTH-SOUTH SPLIT

The differentiating features that capture our attention and are most relevant to a general classification of dialects are those that separate the entire Romani-speaking landscape into identifiable zones. In relation to several prominent features in phonology, morphology and lexicon, there is a tendency toward a north-south split, with innovations occurring on both sides of the divide. This division line tracks the older (sixteenth-seventeenth century) frontier zone between the Habsburg Monarchy and the Ottoman Empire. The political boundary prevented contacts between Roma on either side and blocked the diffusion of innovations, creating a dense and conspicuous cluster of isoglosses [III. 1].

In the north, syllable truncation is triggered in all likelihood by a shift to word-initial stress as a result of Romani-German bilingualism. We find *mal* ‘friend’ for *amal*, *khar-* ‘to call’ for *ak-*

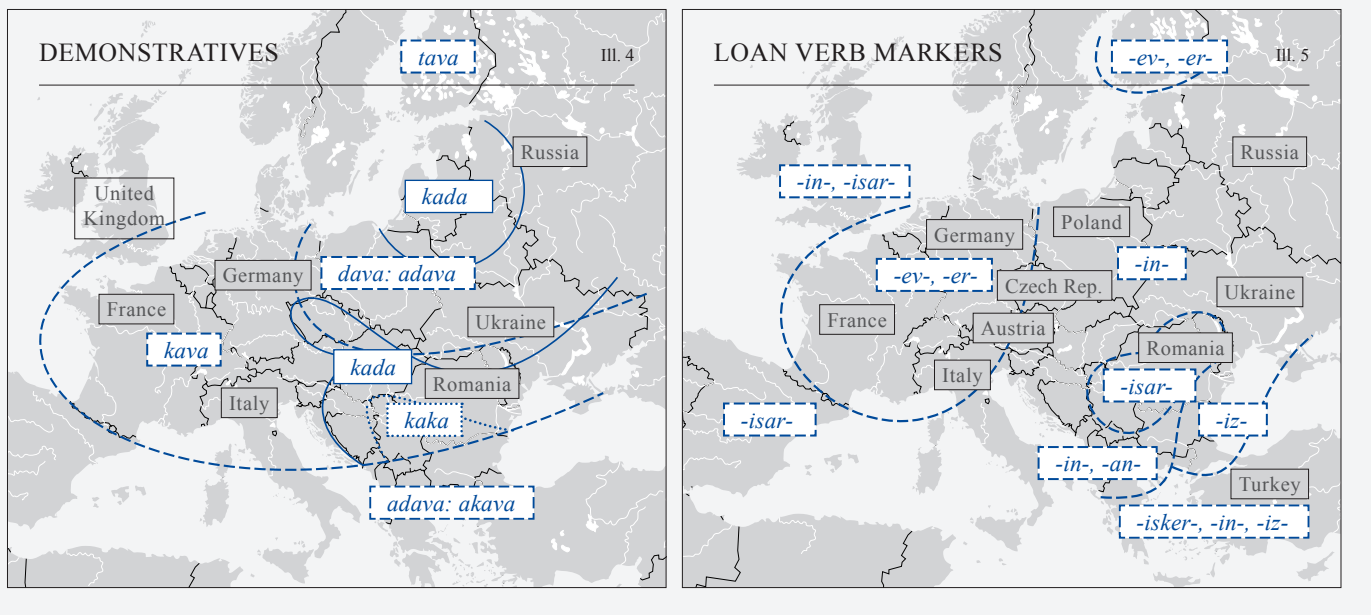
har-, *sa-* ‘to laugh’ for *asa-*, and more. There is also a preference for initial jotation in selected words, among them *jaro* ‘egg’ and the 3rd person pronouns *jov* ‘he’ etc., and the simplification of the historical cluster *nd* to *r* in words like *jaro* ‘egg’ and *maro* ‘bread’. The south, by contrast, maintains non-jotated forms and consonant clusters, as in *(v)ov* ‘he’, *an(d)ro* ‘egg’, *man(d)ro* ‘bread’.

The remarkable coherence of the northern area, from Britain to Finland, the Baltics and northern Russia, might lead us to believe that the individual dialects split away from an earlier group that had settled around the German-Polish contact area. Note that the Romani dialects of the Iberian peninsula tend to remain conservative with respect to these features, indicating that they were not part of the network of contacts that enabled their diffusion in the north. A number of developments fail to reach Finland and appear to have been adopted after the breakaway of the Scandinavian sub-group. They include the loss of the

preposition *katar* ‘from’, which is retained in Finnish Romani, and the assimilation of verbs of motion and change of state into the dominant verb inflection and disappearance of gender-inflected past-tense forms of the type *gelo* ‘he went’, *geli* ‘she went’ (equally retained in Finnish Romani).

A series of lexical preferences spread throughout the north, while inherited variation often continues in the south. The north has *xač-* ‘to burn’ (in the south *phabar-*) and *stariben* ‘prison’ (*phanglipe* in the south, but also in Finnish Romani), as well as *angušt* ‘finger’ (*naj* in the south), derivations of *gi* for ‘heart’ (*ilo* in the south), and *men* ‘neck’ (*kor* in the south).

In the south, the epicentre of innovation appears to be Romania and adjoining regions. Prominent southern innovations include the loss of the nasal segment at the end of the nominalising suffix *-iben/-ipen*, and affrication in *tikno* ‘small’ > *cikno* predominates in the south, though the southern Balkans



show a mixed region. Verbs belonging to the perfective inflection classes that had retained a perfective augment *-t-* are re-assigned to the class of verbs with an augment *-l-* (originally representing verb roots ending in vowels): *beš-t-jom* ‘I sat’ > *beš-l-jom*. Conservative forms occur occasionally in isolation in the south, especially along the Black Sea coast. The north-south divide is complemented by a further divide between a (north) western zone with its centre in Germany and northeastern zone comprising the Baltics and North Russia [III. 2, 3]. The 2SG past-tense and present copula conjugation marker *-al* was probably the older historical form (going back to the 2SG oblique enclitic pronoun **te*). In Early Romani it appears to have competed with *-an*, an analogy to the 2PL marker. The form in *-al* is generalised in the western innovation zone in Germany and spreads eastwards into central Europe to include the Romani dialects of

historical Habsburg Monarchy and on to some of the dialects of Trans-Carpathian Ukraine, but leaves out the entire western periphery (Britain and Spain) as well as northern Poland and the Baltic areas. A very similar diffusion pattern is found for the predominance of *-h-* over *-s-* in grammatical paradigms and in particular in intervocalic position such the singular instrumental/sociative case endings (*leha* ‘with him’ vs. *lesa*). Here too, the variation appears to go back to Early Romani. Note that *s/h* alternation is found in a wide transition zone encompassing the continental side of the Adriatic and stretching all the way to Transylvania. Finnish Romani matches this western-central diffusion zone for both items, indicating that the development preceded the separation from the continental dialects.

The shortening of *anglal/angil* ‘in front’ to *glan/gil*, of *ame* ‘we’ to *me*, and of the verbs *ačh-* ‘to stay’ and

av- ‘to come’ to *čh-* and *v-* (as examples for numerous other items affected by the process) remain limited to Romani varieties spoken within the German-speaking area and neighbouring regions. The areas south of the Great Divide remain unaffected by these developments, while in the northeastern zone jotation appears consistently so that *ame* ‘we’ becomes *jame*, and the verbs *ačh-* ‘to stay’ and *av-* ‘to come’ become *jačh-* and *jav-*. A partition similar in shape emerges around analogies in the past-tense marker of the 2PL. The original *-an* prevails in the northwest as well as in a central belt connecting Germany all the way with the Romanian Black Sea Coast. The innovation centres are once again the northeastern zone, where the predominant form is *-e* (by analogy to the 3PL), and the southern periphery, from southern Romania through to the Mediterranean coast of France, where a partial analogy renders the form *-en*.

CORE VS. PERIPHERY

Many developments spread following a pattern of core vs. periphery. In the case of the word for ‘flour’, Early Ro-

mani appears to have had at least the two variants, with an without initial *v-* (*ařo* and *vařo*). In the northern regions the pressure toward initial jotation affected the word, which became *jařo*. The general absence of initial

segments in the south shifted the balance in favour of a generalisation of the more conservative form *ařo*. But in the geographical periphery, in the absence of pressure in any particular direction, the more innovative of the

two Early Romani variants *vařo* was selected.

Often the periphery remains conservative. The original Early Romani demonstrative opposition set in *adava* : *akava* (with corresponding forms in *-o-*) is retained in the geographical periphery comprising Britain, Spain, Italy, and the southern Balkans [Ill. 4]. The core, by contrast, shows various innovation zones where the original forms are simplified or reinforced to create opposition pairs such as *adava* : *dava*, *kada* : *kaka*, *kava* : *kavka* etc. Though zones partly overlap due to the many forms that can become part of the paradigm, a rough geographical split can be identified between a zone in northern Bulgaria and Romania (*kaka*), a central zone around Hungary and Slovakia (*kada*), a northeastern zone comprising Poland and Russia (*dava* : *adava*) with a unique retention sub-zone in the Baltics (*kada*), a major zone stretching from the Black Sea coast to the North Sea (*kava*), and a Finnish zone (*tava*).

A conservative periphery is also encountered in the retention of Greek-derived verb inflection markers, used in Romani as a means of adapting loan verbs from Greek and subsequent contact languages [Ill. 5]. Romani dialects of present-day Greece show a proliferation of forms. Several forms are retained in Welsh Romani too. Crimean and Zargari Romani keep *-isker-* and *-isar-* appears in Romania-Moldavia and in Spain. The distribution of other forms shows a German-Scandinavian dialect group with *-er-/ev-*, a Black Sea coast group with *-iz-*, and a central-eastern zone from the Baltics and all the way down to western Bulgaria and southern Italy, with *-in-* (primarily, with additional vocalic variation in the Balkans).

Note that each isogloss has its own unique pattern of diffusion. The fact that we are able to review a set of numerable such patterns mirrors the historical fact that networks of social contacts between Romani communities remained stable for considerable

periods of time, allowing the diffusion of several distinct innovations to follow similar pathways, while divisions between groups – through political borders, migrations, or simply through a collapse of social contacts – set demarcations boundaries that contained the diffusion. The result is an accumulation of a complex matrix of different diffusion patterns, yet not without overlap of a number of prominent isoglosses.

When consideration is given to the various bundles of isoglosses representing prominent structural features – such as essential vocabulary items, salient lexico-phonological developments, and especially the organisation of recurrent morphological paradigms – then we obtain a picture that is quite similar to the prevailing reference grid of dialect classification. The classification is thus inspired by the reality of clusters of isoglosses, which in turn are the accumulated result of the diffusion of structural innovation among populations and across geographical space.

For more information on dialect differentiation in Romani consult the online Romani Morpho-Syntax Database and its interactive map-generating function:

<http://romani.humanities.manchester.ac.uk/rms>

BIBLIOGRAPHY

Boretzky, N. 2002. *Die Vlach-Dialekte des Romani: Strukturen - Sprachgeschichte - Verwandtschaftsverhältnisse - Dialektkarten*. Wiesbaden: Harrassowitz. | **Boretzky, N. & Iglá, B. 2004.** *Kommentierter Dialektatlas des Romani. Vol. 1: Vergleich der Dialekte. Vol. 2: Dialektkarten mit einer CD Rom*. Wiesbaden: Harrassowitz. | **Boretzky, N. & Cech, P. & Iglá, B. 2008.** *Die Südbalkanischen Dialekte (SB I) des Romani und ihre innere Gliederung*. Graz: GLM. (= Grazer Linguistische Monographien 26) | **Elsik, V & Matras, Y. 2006.** *Markedness and language change. The Romani sample*. Berlin: Mouton de Gruyter. | **Gilliat-Smith, B.J. 1915.** *A report on the Gypsy tribes of North East Bulgaria*. JGLS, new series, 9: 1-54, 65-109. | **Matras, Yaron. 2002.** *Romani: A linguistic introduction*. Cambridge: CU P. (chapters 2 and 9). | **Matras, Yaron. 2005.** *The classification of Romani dialects: A geographical-historical perspective*. In: Schrammel, Barbara, Halwachs, Dieter W. & Ambrosch, Gerd, eds.: 7-26. | **Schrammel, Barbara, Halwachs, Dieter W. & Ambrosch, Gerd, eds. 2005.** *General and Applied Romani Linguistics*. Munich: Lincom.



© Council of Europe. All rights reserved. No part of this publication may be translated, reproduced or transmitted in any form or by any means, electronic (CD-Rom, Internet, etc.) or mechanical, including photocopying, recording or any information storage or retrieval system, without the prior permission in writing from the Publishing Division, Directorate of Communication (F-67075, Strasbourg cedex or publishing@coe.int) <http://www.coe.int>



PROJECT EDUCATION OF
ROMA CHILDREN IN EUROPE
<http://www.coe.int/education/roma>



PROJECT

<http://romani.uni-graz.at/romani>
<http://romani.humanities.manchester.ac.uk>

Sociolinguistics

7.0

Romani-Project Graz / Dieter W. Halwachs

The sociolinguistic situation of Romani reflects the sociopolitical and the sociocultural status of its speakers. Consequently Romani has to be described as a primarily oral, functionally restricted, dominated, stateless diaspora language with no monolingual speakers.

REPERTOIRES OF ROMANI SPEAKERS	PRIVATE varieties in the social microcosm that are used in informal private domains: with the partner, in the family, when in contact with close friends, etc.	EVERYDAY LIFE varieties in the social macrocosm that are used in domains of everyday life: with acquaintances (at work, at school, etc.), with strangers, when shopping, etc.	PUBLIC varieties used in formal public domains: in the media, in religious contexts, in higher education, when dealing with authorities, etc.
REPERTOIRE 1: language use of most Romani speakers	MAJORITY LANGUAGE(S) Romani	MAJORITY LANGUAGE(S)	MAJORITY LANGUAGE(S)
REPERTOIRE 2: language use of speakers of vital (Vlax) varieties	majority language(s) ROMANI	MAJORITY LANGUAGE(S) Romani	MAJORITY LANGUAGE(S)
REPERTOIRE 3: language use of international Romani activists	MAJORITY LANGUAGE(S) Romani	MAJORITY LANGUAGE(S) Romani	MAJORITY LANGUAGE(S) Romani

A linguistic repertoire is the set of distinguishable linguistic varieties used in different social contexts by a particular speech community. III. 1

Romani is a language that until recently has not existed in a written form and has exclusively been passed on orally. It has not developed a generally accepted written standard and, as a consequence, has no prescriptive norms. This linguistic situation reflects the sociopolitical situation of the Roma: Politically, economically and culturally marginalised,

ethnically stigmatised, discriminated against and persecuted up to genocide, the Roma could only survive in small groups which led to the geographical and social heterogeneity that still exists today. Consequently, Roma have never been in a position to build large political-economic structures or to get their share of political and economic

power. Considering the fact that the development of standard varieties generally follows the development of political and economic power structures, it becomes clear why Romani has not developed such a variety. Furthermore, the outlined sociohistorical situation explains why Romani is labelled a stateless diaspora language.

LANGUAGE USE

For most Roma their respective Romani variety is reduced to intra-group communication, and thus limited to the private domains. Romani primarily functions as an

intimate variety in the social microcosm. Adult Romani speakers are always bi- or multilingual and use the language(s) of the respective majority population(s) for inter-group communication in the public domains and more often in the domains of everyday life. Consequently, Romani spea-

kers may be described as non-monolingual. Compared to the use of Romani, the dominance in the use of the respective majority language(s) is shown in the abstracted linguistic repertoire of Romani speakers.

Repertoire 2 above displays the full range of functions as, for example,

in *Kalderaš* Romani which dominates the internal communication and is also used when in contact with speakers of other Vlach varieties. More frequently, however, the functions of Romani as the inter-group variety in everyday life are extremely limited. In many cases it only functions as an intimate variety in the social microcosm, and even in these domains majority languages are often more frequently used than Romani (see repertoire 1 in Ill. 1). This dominance of majority languages in the linguistic repertoires of Romani speech communities – not only in the public sphere but also in everyday life and in private contexts – characterises Romani as a dominated language.

This asymmetrical relationship between Romani and the majority languages results not only in the functional restrictions mentioned, but also in strong

influence or rather pressure of majority languages on Romani. This influence results in lexical loans as well as in the replication of patterns of the majority languages, which account for the differences between individual Romani varieties. These lexical and structural differences are often perceived as obstacles to inter-group communication and sometimes even lead speakers of a particular Romani variety to value other varieties as different languages. Essentially, problems in inter-group communication are caused by the functional restriction of Romani to intra-group communication in the private sphere. If there is a constant need to use Romani in inter-group communication and in public life, Romani will adapt to these new situations. The ability to meet the communicative needs of its speakers is inherent to any language of the world. Although Ro-

mani has been written for some decades now, the constant communicative need to functionally expand it into all domains of everyday and public life has not generally been given until now. In inter-group communication Romani is primarily used by activists who are able and willing to adapt to each other linguistically. The same applies to the oral use of Romani in formal domains (see repertoire 3 in Ill. 1).

Formal written Romani, above all, has symbolic functions with only marginal communicative ones. The overwhelming majority of texts are translations from majority languages into Romani. Their main purposes are to highlight the ability of Romani to function in these contexts, to support the struggle for sociocultural equality of the Roma, to symbolise the will, need or demand for the sociopolitical integration of the Roma, etc.

SOCIOPOLITICAL STATUS

The vast majority of Romani speakers still use Romani exclusively for intra-group communication, and majority languages in all other domains. As outlined in the last paragraph, this has nothing to do with the linguistic insufficiencies of Romani, but is sociolinguistically rooted. The functional restrictions of Romani are reflected by its sociopolitical status. Romani is marginalised in the media, marginalised in education, irrelevant in public life, and neglected in administration.

Romani is present in almost all types of media. Apart from daily and weekly newspapers, Romani is used in journals, brochures and books. There are radio and television broadcasts on public and private channels and even a few private radio and television stations broadcasting exclusively in Romani. Radio and television are also present on the internet, as are websites, mailing lists and chat rooms. Especially print publications, but also radio and television broadcasts, are often bilingual, thus reflecting both the linguistic repertoires of Romani speakers and the sociolinguistic situation of Romani as a dominated

language. But despite its presence in the media, compared to dominant languages, the impact of Romani media on Romani speakers is insignificant. Romani speakers are above all exposed to the mainstream media of dominant languages and Romani media products and broadcasts are in most cases symbolic, as is written Romani.

In most cases the demand to use Romani in education is part of the political agenda of the struggle of Roma for equal rights and equal opportunities. Resulting activities range from grassroots level actions via national and regional measures to European-wide initiatives. The latter mostly are international treaties or recommendations of supranational organisations which in a few cases are accompanied by concrete measures. One of these rare cases is the Council of Europe's Curriculum Framework for Romani and its corresponding Language Portfolios. These tools were developed according to the standards of the Common European Framework of Reference for Languages (CEFR) which was recommended by the European Union as an instrument in setting up systems of validation of language competences. The most prominent example of an international treaty with an impact on Romani teaching is

the European Charter for Regional or Minority Languages which will be dealt with in the following chapter.

Both recommendations and treaties are top-down instruments that are often accompanied by national or regional measures which – at least legally – make Romani teaching possible. In most cases such measures are embedded into the legal framework for the protection of (national) minorities of a particular country or region and are formulated in the accompanying regulations for minority language teaching. In the case of Romani these top-down measures are, almost as a rule, not actively implemented by the authorities. They just provide the possibility for Romani teaching but leave the implementation to NGOs. This has to be seen in connection with the plurality of Romani, educational authorities are used to dealing with homogeneous languages with a standard that serves as the norm in teaching. Furthermore, most top-down measures are triggered by bottom-up demands.

Without grassroots initiatives, most of the top-down measures would not be brought into force or would remain ineffective declarations of good will. Only a productive cooperation between NGOs and authorities offers the possibility that

Romani teaching becomes part of the education system. But being part of the system does not automatically mean that Romani is integrated into the regular curriculum. On the contrary, for the most part Romani is taught in extracurricular lessons, often only in the framework of lessons on Romani history and culture. Romani as a language of instruction is even more marginalised than Romani as a subject. If a teacher is competent in Romani – which is quite exceptional – it might be used with children whose mother tongue is Romani and who have a low competence in the majority language. In such exceptional cases Romani functions as an auxiliary language for the purpose of acquiring the majority language.

The outlined situation is most probably related to the fact that Romani teaching, and up to a certain point, minority language teaching in Europe in general is less a pedagogical than a political matter. Romani NGOs see Romani teaching as part of the political struggle for emancipation from the majority population and their dominant culture and language. Representatives and authorities of the majority try to value the language and the culture of the Romani minority by declarative acts which grant Romani a marginal role in mainstream education. Depending on the prevailing conditions, extracurricular Romani lessons which also discuss culture and history contribute to the empowerment of Romani children and counteract, at least to some extent, the pressure to linguistically and culturally assimilate into the majority population. Yet despite these side effects, the main function of Romani teaching remains, once again with the background of primarily political motivation, on a symbolic level. There are no known cases where Romani

is used systematically to teach literacy to children who have acquired it as their mother tongue during their primary socialisation. Such a systematic approach to Romani teaching would be the most suitable reason to include it into mainstream education. This would be in line with a recommendation of UNESCO (1953) that the best way to teach literacy is to use the mother tongue of the learner. The marginalisation of Romani in education, again, is a direct result of the sociolinguistic situation of a dominated diaspora language with almost no tradition in literacy. As public life as well as administration are always connected with dominant languages with a profound tradition in literacy, it becomes obvious why Romani is irrelevant in public life and neglected in administration.

The sociolinguistic status of Romani outlined so far explains that changes in the situation of Romani will only result in improvements in the status of its speakers, the Roma. Although it is obvious that Romani will most probably never reach a balanced relationship with the dominant national languages of Europe, the ongoing emancipation process is already effecting changes in its status. Romani is perceived by the majority population as a primary cultural identity factor, public opinion more often attributes it the status of a language. The previously dominant opinion that regarded it as gibberish, as the jargon of fringe groups and as the idiom of crooks is slowly losing strength. This change in opinion results, inter alia, in moderate official attention attributed to Romani as a European minority language. Furthermore, the use of Romani among activists on the international level has an effect on both its functions and structures. Although limi-

ted to a small subgroup of speakers, as a means of communication of the political movement, Romani is functionally expanding into formal domains. This expansion results in lexical enrichment as well as in structural changes. Romani is developing the vocabulary needed to discuss legal, administrative, scientific, etc. topics, as well as structures that enable its users to reflect, write and publicly talk about any relevant topic. Due to its communicative use in formal domains among Romani representatives, Romani has entered the stage of development from a vernacular to a standard language. This development in no case follows the traditional standardisation pattern – i.e. the imposition of a codified variety by law through education – but has to be described as harmonisation, by trial and error in actual communication processes using all linguistic resources at hand. Therefore, translations and standardisation products with primarily symbolic functions as well as communicative experience, repertoire resources, etc. of all speakers involved contribute to this harmonisation process, which is slowly resulting in something like an international Romani. The higher the numbers of Romani speakers participating in this process, the more this international variety will spread and contribute to overcome the communication obstacles between speakers of different Romani varieties. Preconditions for the further development of this international variety are further improvements of the sociopolitical situation of the Roma. Current conditions not only limit the development of Romani, but above all hamper the integration of Roma as equal citizens of their native countries and, consequently, also as European citizens.

ROMANI AND THE CHARTER FOR REGIONAL OR MINORITY LANGUAGES

Despite the ongoing socioeconomic marginalisation and sociopolitical stigmatisation of the Roma, the status of Romani has – as indicated above – improved over

the last decades. This is above all due to the ongoing emancipation process which would not have been possible without changes in the general approach towards minority languages at an overall European level. These changes were initiated by the representatives of traditional linguistic minorities of western Europe – Frisians, Irish, Welsh, etc. – during the second half

of the 20th century. The most important instrument created in this context is the European Charter for Regional or Minority Languages. The Charter was adopted as a convention by the Committee of Ministers of the Council of Europe in 1992, and entered into force on 1 March 1998. On the web page of the Charter its purpose is described as follows:

The Charter for Regional or Minority Languages is a convention designed on the one hand to protect and promote regional and minority languages as a threatened aspect of Europe's cultural heritage and on the other hand to enable speakers of a regional or minority language to use it in private and public life.

III. 2

The Charter defines regional or minority languages as *different from the official language(s) of a state and not a dialect of an official language*, as *traditionally used by nationals within a given territory* (= territorial languages) or by *nationals within the territory of that state* (= non-territorial languages), and as *not a language of (recent) migrants*.

This basic blue print for the definition of a European minority language fully complies with Romani. On the background of the criteria listed, Romani has to be described as different from all official languages of Europe and as used all over Europe since the Middle Ages by nationals of all states of Europe.

Consequently, Romani should be protected as a minority language by the Charter in the territories of all countries which have ratified it up to 2011. But not all countries have recognised Romani and the majority of these opted for the minimum protection as a non-territorial language.

Among the states that excluded Romani are small countries like Liech-

tenstein, which declared that there are no regional or minority languages in the sense of the Charter spoken on its territory, and also countries with quite a number of Roma like Croatia. In this case non-territorial languages are excluded from the ratification in general, thus avoiding the official recognition of Romani. This does not mean that Croatian authorities neglect Romani. There is support for Romani speaking communities both in education and in the media. But the Croatian non-ratification of the Charter for Romani is a symptom of the low sociopolitical status of the Roma in general.

To declare the whole Romani population of a particular country as recent migrants is another way to neglect the necessity to recognise Romani as an official minority language. To counter-argue such an assertion and to prove that Romani is spoken traditionally in a specific European country is sometimes almost impossible: Because of social exclusion Roma have been prevented from owning land and property. Furthermore, quite often their settlements have not been registered properly.

Another consequence of marginalisation and discrimination is ongoing migration which, again, is used to make a case against the autochthonous status of Roma. The differentiation between autochthonous, or indigenous Roma, and allochthonous, migrant Roma, is another symptom of the low support for Romani as a minority language. For instance, Austria declared Romani in its ratification of the Charter as a non-territorial language on the territory of Burgenland, which is the easternmost province bordering Hungary. Therefore, the recognition of Romani is legally limited to the variety of a minority among the Austrian Romani population. The Romani varieties of recent migrants are not only theoretically excluded from protection by the Charter, but also of other autochthonous groups. Practically, the other speakers are supported by the authorities as well. Nevertheless, the fact that this distinction is made in a legally binding convention not only demonstrates the low sociopolitical status of Roma and, consequently, Romani, but also the reservations about the status of Roma as one linguistic minority.

Literature

Friedman, Victor A. (2005.) The Romani Language in Macedonia in the third millenium: Progress and Problems; in: Schrammel, Barbara / Halwachs, Dieter W. / Ambrosch, Gerd (eds.) *General and Applied Romani Linguistics. Proceedings from the 6th International Conference on Romani Linguistics*, Munich: Lincom Europa: 163-173. | **Halwachs, Dieter W. (2003)** The Changing Status of Romani in Europe, in: Hogan-Brun, Gabrielle / Wolff, Stephen (eds.) *Minority Languages in Europe. Frameworks, Status, Prospects*, Houndmills/New York: Palgrave: 192-207. | **Hancock, Ian (ed.) (1979)** *Romani Sociolinguistics*, The Hague: Mouton. (= *International Journal of the Sociology of Language* 19). | **Matras, Yaron (2005)** The future of Romani: Toward a policy of linguistic pluralism, *Roma Rights* 1/2005: 31-44. | **Matras, Yaron (1999)** Writing Romani: The pragmatics of codification in a stateless language, *Applied Linguistics* 20/4: 481-502. | **UNESCO (1953)** *The Use of Vernacular Languages in Education* <<http://unesdoc.unesco.org/images/0000/000028/002897EB.pdf>>