Romani language

Romani rromani chib	
Spoken in:	The speakers of Romani are scattered throughout the world
Total speakers:	4.8 million

Language family:	Indo-European
	 Indo-Iranian
	  Indo-Aryan
	   Central Zone
	    Romani

Official status	
Official language in:	Shuto Orizari (Republic of Macedonia), officially-recognised minority language of Sweden, officially-recognised minority language of Finland, co-official in 79 rural communes in Romania and one town (Budesti)
Regulated by:	*no official regulation*
Language codes	
ISO 639-1:	*none*
ISO 639-2:	`rom`
ISO 639-3:	variously: `rom` — Romani (generic)

`rmn`	—	[Balkan Romani](#)
`rml`	—	[Baltic Romani](#)
`rmc`	—	[Carpathian Romani](#)
`rmf`	—	Kalo Finnish Romani
`rmo`	—	[Sinte Romani](#)
`rmy`	—	[Vlax Romani](#)
`rmw`	—	[Welsh Romani](#)

This page contains [Indic text](#). Without rendering support you may see irregular vowel positioning and a lack of conjuncts. [More...](#)

Romani or **Romany** (native name: *romani chib*) is the [language](#) of the [Roma](#) and [Sinti](#). The [Indo-Aryan](#) Romani language should not be confused with either [Romanian](#) (spoken by [Romanians](#)), or [Romansh](#) (spoken in parts of southeastern [Switzerland](#)), both of which are [Romance languages](#).

Contents

[[hide](#)]

[[edit](#)] Classification and status

Analysis of the Romani language has shown that it is closely related to those spoken in central and northern India, [Pothwari](#) in particular. This linguistic relationship is believed to indicate the Roma's and Sinti's geographical origin. Loanwords in Romani make it possible to trace the pattern of their migration westwards. They came originally from the [Indian subcontinent](#) or what is now northern [India](#) and parts of [Pakistan](#). The Romani language is usually included in the [Central Indo-Aryan languages](#) (together with Western [Hindi](#), [Bhili](#), [Gujarati](#), Khandeshi, [Rajasthani](#) etc.). It is still debated whether the origin of the name *Sinti* is the same as that of the [toponym](#) for the [Sindh](#) region of southeastern

Pakistan and far western India ([Rajasthan](#) and [Gujarat](#)), around the lower [Indus River](#) or is a European loanword in Romani, recognizable as such in its morphological integration into the language (plural Sinte, feminine singular Sintica). It was primarily through [comparative linguistic](#) studies of the Romani language with various north Indian dialects and languages that the origins of the Roma people were traced back to India.

Romani, [Panjabi](#), and [Pothwari](#) share some words and similar grammatical systems. A 2003 study published in *[Nature](#)* suggests Romani is also related to [Sinhalese](#),[1] presently spoken in [Sri Lanka](#).

In terms of its grammatical structures, Romani is conservative in maintaining almost intact the Middle Indo-Aryan present-tense person concord markers, and in maintaining consonantal endings for nominal case – both features that have been eroded in most other modern languages of Central India. It shares an innovative pattern of past-tense person concord with the languages of the Northwest, such as Kashmiri and Shina. This is believed to be further proof that Romani originated in the Central region, then migrated to the Northwest. Characteristic for Romani is the fusion of postpositions of the second Layer (or case marking clitics) to the nominal stem, and the emergence of external tense morphology that attaches to the person suffix. All of these features are shared between Romani and Domari, which has prompted much discussion about the relationships between these two languages.

The Romani language is sometimes considered a group of dialects or a collection of related languages that comprise all the members of a single genetic subgroup.

While the language is nowhere official, there are attempts currently aimed at the creation of a standard language out of all variants (such as those from [Romania](#), the [USA](#), [Sweden](#)). Also, different variants of the language are now in the process of being codified in those countries with high Roma populations (for example, [Slovakia](#)).

[[edit](#)] History

There are no sure historical documents about the early phases of the Romani language. The language is not directly cited in the epic *[Shahnameh](#)* by the 11th century [Persian](#) poet [Firdausi](#), who wrote about the 10,000 or 12,000 [Desi](#) musicians who were given in the 5th century AD by King [Shankal](#) of [Kanauj](#) (in [Sindh](#)) to [Bahram Gur](#) the [King of Persia](#). Nevertheless, many have suggested that these people are the ancestors of the Roma.

However, research carried out already in the nineteenth century by Pott (1845) and Miklosich (1882-1888) showed this to be unlikely. The Romani language proves to be a New Indo-Aryan language (NIA), not a [Middle Indo-Aryan](#) (MIA), as it would have to be to fit Firdausi's scheme. The principal argument favouring a migration during or after the transition period to NIA is the loss of the old system of nominal case, and its reduction to just a two-way case system, nominative vs. oblique. A secondary argument concerns the system of gender differentiation. Romani has only two [genders](#) (masculine and feminine). Middle Indo-Aryan languages (named MIA) generally had three genders (masculine,

feminine and neuter), and some modern Indo-Aryan languages retain this old system even today. It is argued that loss of the neuter gender did not occur until the transition to NIA. Most of the neuter nouns became masculine while a few feminine, like the neuter   (agni) in the [Prakrit](#) became the feminine ？？ (ag) in [Hindi](#) and *jag* in Romani. The parallels in grammatical gender evolution between Romani and other NIA languages have been cited as evidence that the forerunner of Romani remained in the Indian Subcontinent until a later period, perhaps even as late as the tenth century CE.

There are no historical proofs to clarify who were the ancestors of the Roma, and what motivated them to emigrate from the [Indian subcontinent](#), but there are various theories. The influence of the [Greek language](#) (and to a lesser extent of the [Iranian languages](#), like [Persian](#), [Kurdish](#) and of the [Armenian language](#)), points to a prolonged stay in [Anatolia](#) after the departure from South Asia.

The [Mongol invasion of Europe](#) beginning in the first half of the [13th Century](#) triggered another westward migration. The Roma arrived in [Europe](#) and afterwards spread to the other continents. The great distances between the scattered Romani groups led to the development of local community distinctions. The differing local influences have greatly affected the modern language, splitting it into a number of different (originally exclusively regional) dialects.

Today Romani is spoken by small groups in 42 European countries [1]. A project at [Manchester University](#) in [England](#) is transcribing Romani dialects, many of which are on the brink of extinction, for the first time. [2]

[[edit](#)] Modern language

Today's dialects of Romani are differentiated by the vocabulary accumulated since their departure from Anatolia, as well as through [divergent phonemic evolutions](#) and grammatical features. Many Roma no longer speak the language or speak various new [contact languages](#) from the local language with the addition of Romani vocabulary.

A long-standing common categorisation was a division between the *Vlax* (from *[Vlach](#)*) from *non-Vlax* dialects. *Vlax* are those Roma who lived many centuries in the territory of Romania. The main distinction between the two groups is the degree to which their vocabulary is borrowed from [Romanian](#). Vlax-speaking groups include the great number of speakers (between half and two-thirds of all Romani speakers). [Bernard Gilliath-Smith](#) first made this distinction, and coined the term *Vlax* in 1915 in the book *The Report on the Gypsy tribes of North East Bulgaria*. Subsequently, other groups of dialects were recognized, primarily based on geographical and vocabulary criteria, including:

- [Balkan Romani](#): in [Albania](#), [Bulgaria](#), [Greece](#), [Macedonia](#), [Moldova](#), [Montenegro](#), [Serbia](#), [Romania](#), [Turkey](#) and [Ukraine](#)
- [Romani of Wales](#)
- [Romani of Finland](#)

- Sinte: in Austria, Croatia, the Czech Republic, France, Germany, Italy, the Netherlands, Poland, Serbia, Montenegro, Slovenia, and Switzerland
- Carpathian Romani: in the Czech Republic, Poland (particularly in the south), Slovakia, Hungary, Romania, and Ukraine
- Baltic Romani: in Estonia, Latvia, Lithuania, Poland (particularly in the north), Belarus, Ukraine and Russia
- Turkish dialects:
    - Thraki (Thrace) dialect (Thrace, Uskudar, a district on the Anatolian side of the Bosphorus): most loanwords are from Greek
    - Anatolian dialect. Most loanwords are from Turkish, Kurdish and Persian
    - Posha dialect, Armenian Gypsies from eastern Anatolia mostly nomads although some have settled in the region of Van, Turkey. The Kurds call them *Mytryp* (settled ones).

In the past several decades, some scholars have worked out a categorisation of Romani dialects from a linguistic point of view on the basis of historical evolution and isoglosses. Much of this work was carried out by Bochum-based linguist Norbert Boretzky, who pioneered the systematic plotting of structural features of Romani dialects onto geographical maps. This culminated in an Atlas of Romani Dialects, co-authored with Birgit Igla, which appeared in 2005 and plots numerous isoglosses onto maps. At the University of Manchester, similar work has been carried out by linguist and former Romani-rights activist Yaron Matras, and his associates. Together with Viktor Elšík (now of Charles University, Prague), Matras compiled the Romani Morpho-Syntax database, which is the largest compilation of data on the dialects of Romani. Parts of this database can be accessed online via the webpage of the Manchester Romani Project. Matras (2002, 2005) has argued for a theory of geographical classification of Romani dialects, which is based on the diffusion in space of innovations. According to this theory, Early Romani (as spoken in the Byzantine Empire) was brought to western and other parts of Europe through population migrations of Rom in the 14th-15th centuries. These groups settled in the various European regions during the 16th and 17th centuries, acquiring fluency in a variety of contact languages. Changes emerged then, which spread in wave-like patterns, creating the dialect differences attested today. According to Matras, there were two major centres of innovations: some changes emerged in western Europe (Germany and vicinity), spreading eastwards; other emerged in the Wallachian area, spreading to the west and south. In addition, many regional and local isoglosses formed, creating a complex wave of language boundaries. Matras points to the prohtesis of j- in aro > jaro 'egg' and ov > jov 'he' as typical examples of west-to-east diffusion, and of addition of prothetic a- in bijav > abijav as a typical east-to-west spread. His conclusion is that dialect differences formed in situ, and not as a result of differetn waves of migration.

In a series of articles (beginning from 1982), Marcel Courthiade proposed a different kind of classification. He concentrates on the dialectal diversity of Romani in three successive strata of expansion, using the criteria of phonological and grammatical changes. Finding the common linguistic features of the dialects, he presents the historical evolution from the first stratum (the dialects closest to the Anatolian Romani of the 13th century) to the second and third strata. He also names as "pogadialects" (after the *Pogadi* dialect from

[Great Britain](#)) those which have only a Romani vocabulary grafted into a non-Romani language.

A table of some dialectal differences: