Romani as a case study of language codification and revitalisation

Structure of the talk
1 Roma and Romani
2 Sociolinguistic background
3 Self organisation, language standardisation and expansion
4 Examples of language codification projects
4.1 The Austrian Romani project
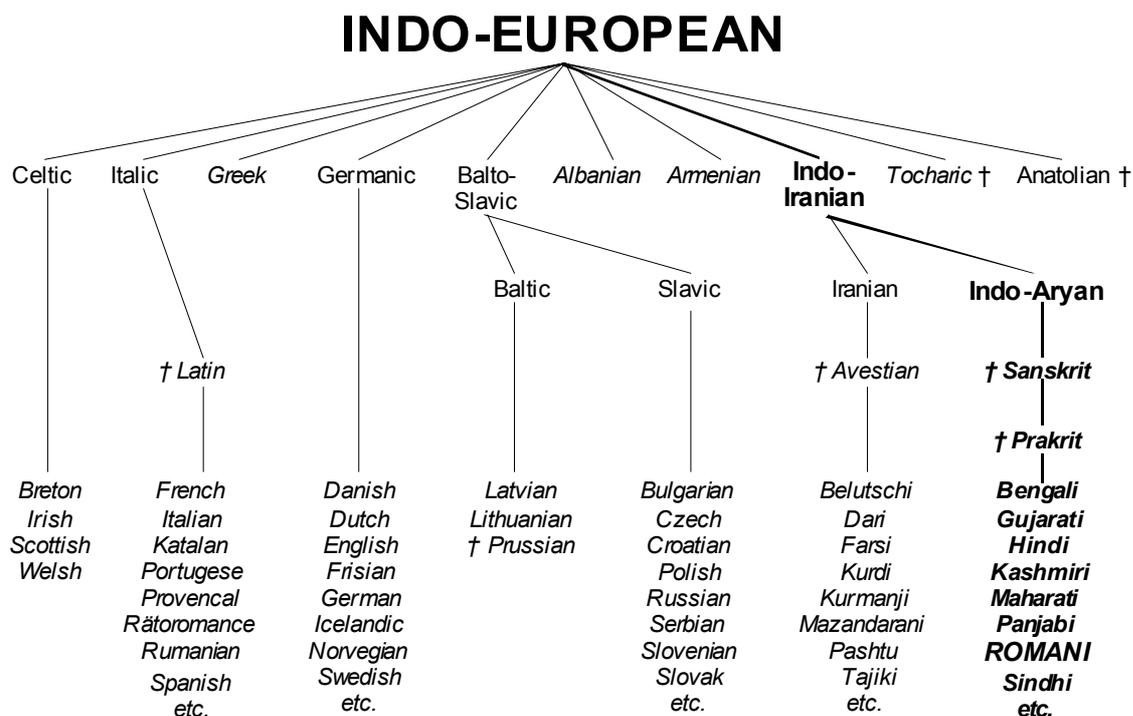4.2 The ROMLEX project

In this talk we aim to illustrate different trends and developments concerning Romani codification, standardisation and revitalisation. The last decade saw an increase in codification efforts and also the promotion of Romani into new functional domains, such as politics, media, school etc. We will discuss this topic on two levels: First we will introduce some general trends in Romani codification and expansion in the last decade. Then we will discuss to projects concerned with Romani codification and revitalisation. The first one, the Austrian Romani project, is an example of a regional, group-specific approach to codification. The second one, the ROMLEX project in our opinion represents a new approach to codification both on a theoretical and practical level.
We will first give a general introduction to Roma and Romani with a special emphasis on the specific sociolinguistic situation of Roma in Europe. Before discussing the two projects we will give a survey of trends in Romani codification of the last decade.

1 Roma and Romani

Romani is the common language of the Roma, the Sinti, the Kale and other European population groups summarised by the pejorative denomination gypsies. It belongs to the Indo-Aryan branch of the Indo-European language family and is the only New-Indo-Aryan language spoken exclusively outside of the Indian subcontinent. As illustrated in table 1 the closest living relatives of Romani are languages like Hindi, Panjabi etc.   table 1

INDO-EUROPEAN

Celtic	Italic	*Greek*	Germanic	Balto-Slavic	*Albanian*	*Armenian*	**Indo-Iranian**	*Tocharic* †	Anatolian †

				Baltic	Slavic		Iranian	**Indo-Aryan**	

† Latin

† Avestian   *† Sanskrit*

† Prakrit

Celtic	Italic	Germanic	Baltic	Slavic	Iranian	Indo-Aryan
Breton	*French*	*Danish*	*Latvian*	*Bulgarian*	*Belutschi*	***Bengali***
Irish	*Italian*	*Dutch*	*Lithuanian*	*Czech*	*Dari*	***Gujarati***
Scottish	*Katalan*	*English*	*† Prussian*	*Croatian*	*Farsi*	***Hindi***
Welsh	*Portugese*	*Frisian*		*Polish*	*Kurdi*	***Kashmiri***
	Provencal	*German*		*Russian*	*Kurmanji*	***Maharati***
	Rätoromance	*Icelandic*		*Serbian*	*Mazandarani*	***Panjabi***
	Rumanian	*Norvegian*		*Slovenian*	*Pashtu*	***ROMANI***
	Spanish	*Swedish*		*Slovak*	*Tajiki*	***Sindhi***
	etc.	*etc.*		*etc.*	*etc.*	*etc.*

From a linguistic point of view Romani can be described as a heterogenous cluster of varieties that share a lexical and grammatical core. There is no homogenising standard. Slide two gives you an impression of the dialectal heterogeneity of Romani. It is a simplified graphical version of the dialect decsription given in Matras (2002).

SLIDE 2 (powerpoint slide in the presentation)

The terminology used for the individual varieties is primarily based on the denominations of the respective groups of speakers, which again are highly heterogeneous: We find geographic definitions (e.g. in the Central varieties and also Greek Vlax) as for instance *Ajia Varvara*, which is a suburb of Athens, and of *Dendropotamos*, which is a suburb of Thessaloniki. Some denominations refer to professions of the respective groups, e.g. *Bugurdži, Čurara, Kalderaš, Lovara, Sepečides*, with the meaning: drill-makers, sieve-makers, tinkers, horse-dealers, basket-weavers. The denominations *Arli* used for Kosovarian and Macedonian Roma as well as *Erli* for a group living in Sofia are indications of the long-lasting settled way of life of these Balkan Roma: the Turkish word *yerli* stands for 'native'. The name *Gurbet* derives from the Arabic word *gharib* 'strange' which has been transmitted via Turkish.

Because of shared conservative features, Northwestern, Northeastern, British and Iberian varieties are sometimes treated as the Northern group of Romani (Bakker 1999). Denominations among these varieties range from geographical definitions to group names and even one language denomination is used: *Rómanes*, derived from an ethnonym adverb, is a widespread language name among the *Sinte*, a group denomination with unclear etymology. *Manuš* 'human being' and *Caló* 'black' are both self-designations among Northern groups. The geographical denominations define the current living space – *Lombard Sinte*, *Finnish Romani*, etc. – as well as the country of origin – *Estrexarja Sinte* (in the case of the *Russian Estrexarja Sinte* the country of origin is the former Austrian Hungarian Empire).

So called Para-Romani varieties are marked by brackets. These are varieties of the respective majority language with Romani lexicon and, if at all, only a few Romani structural features: *Errumantxela* is a variety of Basque, *Caló* is a variety of Spanish, *Angloromani* of English, *Scandoromani* summarises Para-Romani varieties based on various Scandinavian languages.

Estimations of the total number of Roma in Europe range from 6.6 million people to about 12 million people in more generous estimations. The number of Romani speakers in Europe is reported to be about 5 million. As the Roma have always been, and still are, a group which demographically can only be identified with difficulty, all numbers are assumptions and more or less realistic estimations. What is even more problematic is the different basis for such estimations. Many European countries have a Romani population consisting of an autochthonous Romani group, which immigrated centuries ago, members of Vlax groups who came from the late 19th century on and recent migrants from Eastern Europe and the Balkans. In different estimations different parts of these Romani populations are taken into account.

2 Sociolinguistic situation

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 codified standard and, as a consequence, no prescriptive norms. This linguistic situation reflects the socio-political situation of the Roma: politically, economically and culturally marginalized, ethnically stigmatised, discriminated against and persecuted, the Roma could only survive in small groups, that led to the geographical and social heterogeneity that still exists today. Consequently, the people concerned have been in no position to build large political-economic structures or to get their share of political and economic power. Considering that the development of standard varieties generally follows the development of political and economic power structures, it becomes clear why Romani does not have a codified

standard and also that it will not be able to develop a generally accepted standard in the near future. This specific socio-political situation is also the reason why the Roma were always dependent on the socio-economic structures of the respective majority population and – as a consequence – Romani was and is limited to intra-group-communication and thus to to certain domains.

Romani primarily functions as an intimate variety. Nearly all Romani speakers are bi- or multilingual and use the language of the respective majority population(s) for inter-group-communication in public and most often also in informal or partly public domains. As a result, no social stratification can be found within the individual Romani varieties. The dominance in the use of the respective majority language becomes apparent in the abstracted collective repertoire in table 2:

table 2

acrolect	MAJORITY LANGUAGE(S)	Public diatypes that are used in public formal domains when dealing with authorities, at school, in the media etc.
mesolect	MAJORITY LANGUAGE(S) Romani	Diatypes of the social macrocosm that are used in partly public informal domains with acquaintances, at work, etc.
basilect	MAJORITY LANGUAGE(S) ROMANI	Diatypes of the social microcosm that are used in private informal domains in the family and when in contact with friends, etc.

diatype = functionally defined linguistic variety

Tis repertoire displays the full range of functions as, for example, among some Kalderaš groups where Romani dominates the internal communication and is also used when in contact with speakers of other Vlax varieties. More frequently, however, Romani does not function in the social macrocosm but is only used as an intimate variety, as shown in table 3:

table 3

acrolect	MAJORITY LANGUAGE(S)
mesolect	MAJORITY LANGUAGE(S)
basilect	MAJORITY LANGUAGE(S) Romani

These limitations in the functional dimensions together with the lack of a standard and a written language are the major reasons for the fact that Romani has not only very little prestige with the majority population, but also that many Roma consider it inferior as compared to the language of the respective majority population.

The low prestige of the language, reduced domains, multilingualism and the pressure to assimilate on the part of the majority culture make Romani a dominated language whose relationship to the contact languages has always been asymmetric and never bilateral. As a result various phenomena of language contact and language shift occur ranging from lexical borrowings from the majority language to monolingualism in the majority language. In this way some Roma groups have given up Romani without, however, losing their ethnic awareness. Today, Roma living in Romania, Greece Hungary, and Serbia have Romanian as their mother tongue but still feel as Roma. Of course, there are also groups whose ethnic awareness was also lost when language shift occurred.

The status of Romani among Roma has recently experienced some change, however.

3 Codification, Expansion, Functions

The last decade(s) saw the self-organisation of Roma on both a regional and an international level. In the context of self-organisation of individual Roma groups Romani plays a key role: Self-organisation in most cases goes in line with the guidelines of the dominant culture and may be
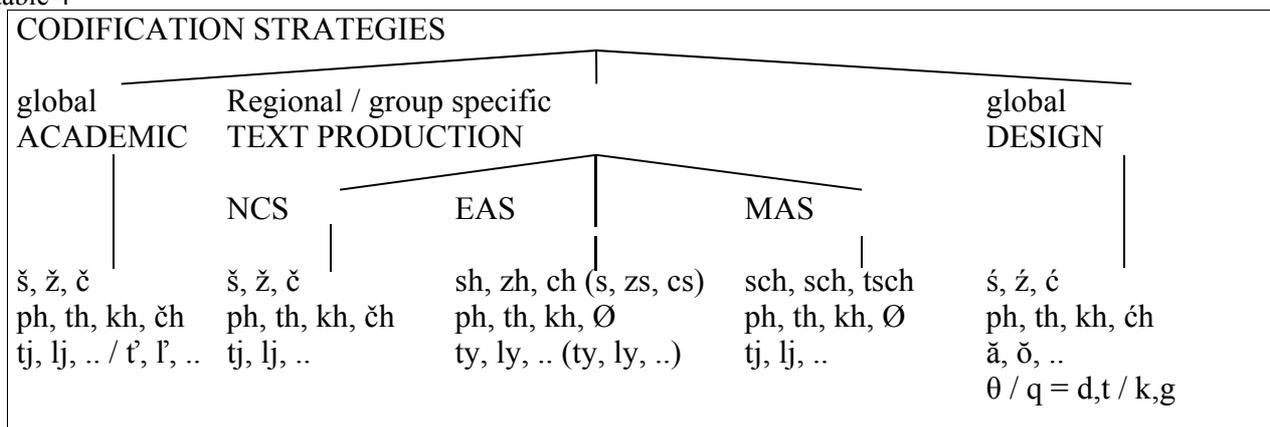
described as emancipation by means of organisational assimilation. It is via this development that the common definition criteria of 'nation', 'ethnic group', and other socio-cultural concepts of the respective majority population are taken over and adopted. In the course of this organisational assimilation the ideology of European national states becomes more and more important and consequently Romani becomes the primary identity factor of the respective Roma group. This process triggers the codification and expansion of Romani and the emblematic and mobilising functions of language use are gaining ground.

3.1 Codification

Codification is to be regarded as the emancipation of Romani from other European national languages. The single attempts of codification lie between the tensions of regional and global approaches resulting from the heterogeneity of Roma and Romani. The regional or group-specific approaches take into account the requirements of the single groups of speakers and therefore are primarily oriented on a communicative basis, whereas the global approaches are more or less politically motivated.

As the missing tradition of a written language is regarded as a major deficit the promotion of a written form is the first step of any codification initiative. According to Matras (1999) in this context there are two types of global codification approaches in contrast to the regional group-specific approaches.

table 4

CODIFICATION STRATEGIES				
global ACADEMIC	Regional / group specific TEXT PRODUCTION			global DESIGN
	NCS	EAS	MAS	
š, ž, č	š, ž, č	sh, zh, ch (s, zs, cs)	sch, sch, tsch	ś, ź, ć
ph, th, kh, čh	ph, th, kh, čh	ph, th, kh, Ø	ph, th, kh, Ø	ph, th, kh, ćh
tj, lj, .. / ť, ľ, ..	tj, lj, ..	ty, ly, .. (ty, ly, ..)	tj, lj, ..	ă, ŏ, ..
				θ / q = d,t / k,g

source: Matras (1999)

The "codification for the purpose of academic documentation of speech (Academic Codification)" and "codification in the service of unification and language engineering projects (Language Design)" are to be considered as global approaches. Whereas, the "codification for the purpose of transposing oral usage into texts directed at audiences (Text Production)" is determined by communicative aspects on a group-specific level.

In view of the academic level, a type of norm has developed in recent years: "Though never conventionalised among Romani linguists, a consensus seems to prevail on the use of wedge-accents, as employed in south-western Slavic alphabets, to indicate palato-alveolars (š, ž, č) and the use of -h to indicate distinctive aspiration on voiceless stops and affricates (ph, th, kh, čh)." The marking of the palatisation with postponed *j* or with accent is one of the few open points of this convention for the written language.

The Language Design Project presented here is classified by Djuric/Cortiade (1990) as polylectal and, as it has been declared to be the standard by the International Romani Union (IRU), is described as the Romani common alphabet. The single common factor with the academic convention is the marking of the aspiration with the postponed –h. For the palato-alveolars the acute accents are used instead of the wedge-accents. The palatisation is marked by the so-called *čiriklo* ('bird') on the vowel. One of the most apparent characteristics of this process to implement a written language is the so-called morphographs which are used to capture the morphophonological alternation of case suffixes in different phonological environments. The locative suffix in *mande* as

well as in *tute* is written by the sign similar to the theta <θ>, for example, the dative suffix in *mange* and in *tuke* is written by the <q>. This codification, which is to be understood as polylectal or as variety embracing, attempts to do justice to Romani in terms of structure but also attempts to emphasise the independent appearance of the written form. Considering that the development of standard varieties generally follows the development of political and economic power structures, it becomes clear why this codification approach has not become a generally accepted standard, although having been declared as such by the IRU. Besides IRU documents the writing system is in use in Romania where approx. 15.000 children are taught Romani in school. The Romanian implementation of this global approach strategy is on the one hand due to a centralised Romani teaching approach by the ministry of education. On the other hand it is made possible by the, compared to other countries, quite homogeneous linguistic situation of Romani in Romania.

The group-specific and communicatively oriented approaches which naturally gain higher, but regionally limited, acceptance lie in a position exposed to tensions between the approach of bringing it in line with global approaches and the taking over of the conventions of the written language of the respective majority language.

Three strategies can be distinguished: in the case of the 'no-compromise' strategy, global strategies for the regionally group-specific text production are adopted. Some publications in Kalderaš-Romani use the academic conventions without any consideration of the conventions of the written language of the respective country of publication (France, Sweden, USA).

The 'elaborate adjustment' strategy adapts, to a certain extent, to the conventions of the respective majority language, on the other hand it stresses also the independence of Romani by adopting global conventions. That is why Hungarian Romani publications are characterised by the adoption of Hungarian conventions as well as by the use of elements of global strategies: in Hungarian the marking of the palatisation is by a postponed *-y* which is global or in this case international as taken over from the convention of written language on the basis of the ASCII-Code; this shows the written form of palato-alveolars. The possible Hungarian way of writing is indicated below in brackets.

In the third strategy, the 'moderate adjustment strategy' graphemes and grapheme combinations that do not appear in the majority language are largely avoided. As an example of the 'moderate adjustment' strategy the written language of Burgenland Romani is mentioned. The marking of palatisation and aspiration correlate the scientific approach because also in German the multigraph and not the diacritica convention is applied. The palato-alveolars are also written according to the German pattern: <sch, tsch>. The voiced-unvoiced dichotomy is neutralized in this variety.

3.2 Expansion

Hand in hand with the process of self-organisation and codification goes the expansion of Romani into new functional domains.

Up to a few decades ago Romani was primarily used as intimate variety. Only with the self-organisation according to the ideal of the majority culture was this limitation perceived as a shortcoming. At the same time the first attempts to use Romani as language in the media were launched, which nurture the necessity of expanding into formal domains. This expansion is most evident on the lexical level as an enormous amount of technolectal designations are missing in Romani. The contrast between global and regional approaches also characterises the lexical expansion of Romani.

The lack of these definitions has not created a communicative problem until the attempt to consciously change the status of Romani. Romani disposes of morphological devices that make it possible to adopt each word from the respective majority language and to integrate it morphologically into Romani. This advantage of Romani is considered as a disadvantage in view of the attempts of emancipation of the language resulting in 'purity demands'.

At the beginning of the codification the representatives of the Burgenland-Roma, for example, requested to substitute all Slavonic, Hungarian and German elements of their Romani variety by

Indian or best by Sanskrit lexemes. Only a laborious demonstration of the etymological heterogeneity of other European national languages could neutralise this demand. In other cases the international value of English is meant to increase the prestige of Romani. Increasing the prestige of the language by the adoption of lexemes from high-prestige languages represents one of the global strategies for the expansion of Romani.

   Another strategy applies derivational and compositional morphology. See two examples from the design Romani by Courthiade:

Instead of using *internacionalo* 'international' the word *maškarthemutno* is used. This lexeme consists of *maškar* 'between' and the adjective *themutno* deriving from *them* 'country, state' and as a consequence it is to be considered a loan translation. In the meantime this positive example for expansion from the substance of the language starts to assert itself on an international level – *maškarthemutne*.

Xurdelin 'kindergarten, nursery school' which similarly has been created from the substance of the language, on the other hand, is an unfortunate neologism. The plural *xurde* used as a noun of the adjective *xurdo* 'tiny' is often used for 'small change, coins', sometimes for 'children'. The derivation suffix *–lin* makes fruit trees from fruits, such as *phabalin* 'apple tree' from *phaba* 'apple'. The question whether 'small change tree' or 'children tree' has asserted itself is easily answered: *Xurdelin* in which meaning it might have, is only used, if at all, by real hardliners of this design-project.

The regional expansion strategies are also based on the integration morphology, but in contrast to the global integration strategy they use the respective primary contact language as the lexical source.

Burgenland-Romani does not use the internationalism *organizacia* 'organisation, association' but *farajn* taken over from German 'Verein'. In contrast to the international use, 'Romani organization' is not named *romani organizacia* but *romano farajn*. Corresponding to the German dialectal use for 'hospital' the lexeme *schpita* taken from 'Spital' is used. Instead of using *schpita* the loan translation *nasvalengero kher* from the inherited substance would also be possible (*nasvale* means 'the sick (pl.)' and *kher* 'house').

Whether the regional strategy dominates the global strategy in lexical expansions depends on a variety of factors. If the initiator of the codification is an international organisation, global strategies will prevail. The same is valid for groups with a strong international spread. Groups which have been settled for a longer period and which therefore are partly assimilated, on the other hand, generally apply regional strategies and mainly borrow from the primary contact language. This distinction does not establish a rule at all, but shows a general tendency since there are enough counter examples.

3.3 Emblematic and Mobilising Functions

In parallel to codification and to lexical expansion Romani has also experienced changes and expansion on the pragmatic level. Due to the new internal status as primary identity marker the emblematic function and the mobilising function, beside the communicative function, come to the foreground for the written use of Romani.

table 5

Pragmatic Functions	
EMMBLEMATIC communicative MOBILIZING	
I Rikeripaske ap u Sinti de Roma, mare Mulenge, gei weian maschke 1933 de 1945 mardo an u Manuschengromarepen.	maśkarthemutne bi-raipné organizàcie ... butĭ vaś-e Manuśikane Hakaja p-o maśkar-themutno nivo ...
In memory of the Sinti and Roma, our dead, who were murdered in the Holocaust between 1933 and 1945.	international non-governmental organizations ... human rights activities at the international level ...

source: Matras (1999)

Matras (1999: 495) defines "emblematic texts which are not intended to enhance the adressee's knowledge in order to action on his part, but rather to elicit emotional identification on the part of the adressee with the aesthetic symbolism of the text in its particular language-external context."

Emblematic function on a micro level is attributed to Romani single words in texts in the majority language. In this context it is worth mentioning the numerous names of newspapers and magazines in Romani, such as the Austrian magazine *Romano Kipo* 'Roma Picture' which is quarterly published and which – apart from the title – almost exclusively is written in German.

A further, longer example for an emblematic text is the sentence mentioned in figure 2 in Rómanes being the Romani variety of the German Sinti. This sentence was shown together with its German translation at the end of a commemorative exposition to the Holocaust. This shows a regional group-specific codification: German conventions for the implementation of a written language are used. It is worth mentioning that the neologism *Manuschengromarepen* was used for 'Holocaust', which literally translated means 'striking dead of people'.

On the macro level of emblematic texts there are Bible translations and among others the translation of the *Ramayana* by Leksa Manuš. Such translations have their effects on the external world and demonstrate to the majority population that Romani is suitable for long complex texts; on the other hand they also have their effects on the internal world, on the Roma, to whom such translations demonstrate the value of Romani. This consequently strengthens the identification with their own language and culture. The primary and from most translators main intention to give the addressees great reading pleasure is mostly prevailed by the emotive-emblematic function.

As mobilising-rallying functions Matras (1999: 496) defines "the shaping of a text in such a way that would demonstrate ideological commitment and political allegiance and identification."

The use of the Design-Romani by Courthiade after its recognition as official Romani standard in the context of the Fourth Romani World Congress in April 1990 in Warsaw has shown its mobilising-rallying function. In the years after, the proponents of this decision have used the conventions defining the criteria for the implementation of a written language set by the Warsaw decision and also the neologisms in internal papers as well as in publicly obtainable publications, such as the *Rromani Uniaquoro Lil* , the newspaper of the Romani union which is published irregularly. The examples of texts shown in figure 2 are taken from the minutes of a meeting of Romani delegates at an international conference in June 1994. By using these conventions, ideological commitment as well as political allegiance and identification with the Romani union, its resolutions and its decision bearers is demonstrated.

4 Examples of Romani language standardisation and codification

4.1 The Austrian Romani project

The Austrian Romani project is neither a purely scientific academic project nor a folkloristic must of language conservation, but it is a scientifically grounded project conducted on behalf of individual Roma groups. The project aims to confront the threatening language death and wants to contribute to the preservation of the culture and the identity of the Roma. In an communtiy approach to language standardisation representatives of Austrian Roma communities co-operated with linguists in developing a written from of their language in expanding the vocabulary of their language to new domains and in developing teaching materials.

At the beginning of the collaboration of Roma and science in 1993, the members of the Austrian Romani project have not at all been aware of the heterogeneity of the Austrian Roma society:

table 6 – demographic parameters of the different groups

	Burgenland-Roma	Sinti	Lovara		Balkan-Vlax-Roma *Kalderaš/Gurbet/...*	Balkan-Muslimic-Roma *Arlije/Bugurdži/...*	various
A	Hungary	S-Germany	Hunga ry	Hunga ry	Serbia / Montenegro /	Macedonia/ Kosovo /	Balkans E-Europe
	15th century	~ 1900		1956	mid 1960s on		late 1980s on
B	E-Austria	cities	larger Vienna area				cities
	rural [urban]	urban					urban
C	roman catholic / protestant / ...				orthodox	muslimic	various
D	autochthonous			allochthonous			illegal

A	country & time of immigration	B	area of living
C	main religious affiliation	D	political status

Without the openness to accept the socio-cultural differences of the single groups and the resulting divergent attitudes, the project could have never been realised in such a way. Both sides had to go through difficult moments of learning and experiencing which then have founded the basis of a successful collaboration.

Apart from the necessary knowledge concerning heterogeneity among the Austrian Roma and the problem arising from the homogeneity requirements, it was absolutely necessary for all scientific collaborators to avoid any kind of »scientific arrogance«. Right from the start it has been important to be open and to approach the task without prejudice and without any pre-established concepts. It was crucial to approach all tasks in an ingenious way and to react to each situation in an intuitive manner. Only by adopting such an approach it is possible to obtain relatively authentic and descriptive results and to avoid the problem of the »self-fulfilling prophecy«. Furthermore, it is necessary to accept the Roma who participate in the project and to consider them experts in relation to their mother-tongue and to take into consideration their ideas regarding the writing, standardisation and realisation in the teaching etc. up to a certain extent. As far as the integration of certain groups is concerned, it is absolutely necessary to allow the participation of as many members of the Roma groups as possible and not just a few representatives of them.

The Romani project has started with the codification of Burgenland Romani threatened by extinction. The starting point of the codification was meant to be the development of a writing system. As today's experience shows it, some incorrect steps were taken in these first phase of the codification process. On the grounds of the idea to homogenise the language and considering it a politically necessary step to create a common writing system, the representatives of the Burgenland-Roma were ready to use the south-Slavic Latin conventions for writing their language. Only when the circle of Roma people participating in the project was extended it became clear that the main part of the Roma preferred a writing system on the basis of the German language. This way of thinking regarding the writing system was supported by a relatively extended study with respect to a spontaneous writing approach. The reason why German was chosen as the basis for writing is the wish to disassociate clearly from other Roma groups and from other minorities of the Burgenland (Croats and Hungarians). The fact that the younger generation of the Burgenland-Roma in particular was taught the German alphabet also supported the idea to base the writing system on the German alphabet. If there had not been an extension of the members of the circle of Roma collaborators – which honestly was not an easy step – the writing of the language would have been realised according to the method that scientifically and maybe also politically was considered the best. Developments in view of the main problems and main desires to preserve the language, however, would not have been achieved at all. This experience has made clear just how important it is to take a large group of Roma collaborators for the acquisition of information in order to obtain results that are practicable. It is true that one-person-studies are of high scientific value. For the realisation of the results in the context of a whole group, however, they might constitute a problem.

The question of writing in view of the collaboration with a Sinti group in a larger town in Austria may be described as follows: in this case too, it was the wish to disassociate from the non-Austrian Slavic alphabet, increased by the fact that the German alphabet was learned, to lead to the creation of writing conventions according to the German model. Additionally, there is the wish to disassociate from the Burgenland-Roma: the Sinti use in their writing system among others the letters <z> and <w> which do not exist within the conventions established for Burgenland-Romani; for expressing the diphthong /a^i/ they write <ei> instead of <aj> as also shown in table 7:

example 7

Bgld-Romani	Rómanes	English	"Romani proper"
cidel	zerdel	'to pull'	cidel
nevo	newo	'new'	nevo
tschaj	tschei	'girl, daughter'	čaj

This way of writing does not correspond to the internationally common conventions, it corresponds, however, to the requirements and ideas of the Sinti concerned. They want to hand down *Rómanes* to their children without causing any sensation. As very often it happens that if only the parents and the children, but not the grand-parents, are living in the same house the handing down of the language from generation to generation is at risk. This is the reason why it would be helpful to have written materials in order to encounter this development. As a lot of Sinti – as mentioned before – regard *Rómanes* as a tabooed in-group-marker it is not in the interest of the proponent to disseminate their materials outside the sub-group. This way of thinking, however, is respected by the scientific members of the project team.

As far as the Lovara are concerned, a writing systen in accordance with international convention has been chosen. There are two reasons for this decision: on the one hand there are no family ties at all, rendering impossible that common decisions are taken and on the other hand the Austrian Lovara are part of an internationally largely extended group and therefore it seems worth producing written materials which are understood not only by the small group of Lovara living in Austria but by all of them.

The collection of texts *Tusa ande akhoren khelos ...* (Cech et al. 2001) is written in a way of writing according to international standards. The tellers of the single stories, who belong to the group of Austrian Lovara immigrated at the end of the 19th century, refuse this kind of writing with the argument that the writing does not correspond to the Austrian Lovara. An Austrian writing system is considered a writing system which corresponds to the German example and which also considers sub-group specific phonetic realisations.

The collaboration with the Lovara is an essential basis for a successful collaboration of Roma and science: if there is no self-organisation supported by the majority of the respective Roma-group, it is nearly impossible to reach results which are acceptable for the majority of the group. A successful self-organisation requires that there are individuals who are generally accepted and that there is an organisational centre, an association or something similar within the group. If these conditions are not fulfilled, as is the case within the Austrian Lovara, no satisfactory results are achieved in relation to the realisation of the results of the codification. On the contrary, the linguistic awareness arisen by the project causes that not only the variety spoken within the group is attributed the level of a »pure Romani« but also the idiolect is considered as such. That does not endanger the mutual understanding of the individual speakers but a group variety which is agreed on, not only from a written but also from a lexical and grammatical point of view is generally impossible.

The ideas regarding the creation of a writing mentioned above is valid for the lexical and the grammatical description, too. Group-specific characteristics are to be considered in the respective codification as also the sub-group-specific characteristics are to be taken into consideration. When taking account of the variety it is guaranteed that a possibly large group of speakers is able to identify with the result of the codification. This implies to walk on a tightrope between group-specific and sub-group-specific characteristics but also between collective and individual or

idiolect-related characteristics. In this process, the Romani as a whole must not be omitted; this aspect within the process of codification requires that a certain control from outside is exerted by an expert of Romani who, however, is not involved in the current works.

Codifying is essential in order to reduce the probability of the language being lost and consequently becoming extinct in the future. Each codification of Romani implies that, in the future, Romani should be taught in school. According to the principle »a language which is only taught in school is a dead language« it has to be said, however, that the main emphasis of counteracting language death or revitalizing language use has to focus on the preservation or the creation of a vivid linguistic surrounding. The most important contribution, the primary socialisation of their children in Romani, is to be fulfilled by the Roma themselves.

Interest in the language and the culture shown from outside and relating activities within the group as a consequence, may positively influence the decision to use Romani as intra-group variety and when talking to the children. This aspect is of increasing importance as the Roma in central and western Europe are subject to high pressure to assimilate on the grounds of economic reasons and the children are brought up monolingual, in the language of the majority population. This monolingual education is reasoned with the argument that the opportunities of the children to participate in the society of wellbeing would increase. If the own culture and the own language do have a positive image the probability for a functioning linguistic surrounding increases which, as a consequence, increases the continuity in the passing on of the language.

The co-operation of Roma and science in the Austrian Romani project surely created a lot of interest in the language and culture from the outside as well as encouraged activities within the speech community. Burgenland Romani for instance, a Romani variety that ten years ago was on the edge of extinction, was revitalised to the extent that today it is not only used in newspapers and in official documents, but is also taught in Austrian schools. Also there is a bilingual publication series in German and different varieties of Romani aimed to promote Roma culture to a wider audience.

Nonetheless, the passing on of the language to the next generation within Romafamilies even today is not guaranteed.

4.2 The ROMLEX project

The ROMLEX project is a international project based at the Universities of Graz (Austria), Manchester (United Kingdom), and Århus (Denmark). ROMLEX is a comprehensive and comparative lexical documentation of the varieties of Romani in electronic form. The database includes Romani entries in various Romani dialects, grammatical and etymological information, examples for usage and translations into English (principal target language) as well as other languages of Europe. The potential of such a dictionary is at least twofold:

1. As an academic dictionary, which will encompass all or most documented dialects of Romani, it can provide a resource that would facilitate and indeed achieve to a considerable degree an historical reconstruction of the language. The outcome – in the domains of semantics and phonology, which are relevant to the lexicon – could be taken to reflect a Common or Early Romani (cf. Matras 2000, Matras 2002). Furthermore, it can facilitate dialect classification and the reconstruction of branching and sub-branching within Romani, by providing a reliable data source for dialect comparison. This applies most obviously to variation in the lexical domain, but also to phonology and partly word-level morphology: lexical composition, plural formation, gender assignment, oblique case and perfective tense formation – all aspects that can be included in the entries of a dictionary.
2. The online database, which is accessible via the web, allows translations between several target languages and several Romani dialects. English translations are available for all the Romani-dialects included in the online database. In order to provide a useful tool, especially for the Romani speech community itself, ROMLEX aims to provide translations of all Romani-dialects

into the majority languages of the countries in which the respective dialects are spoken. In doing so ROMLEX is not intended to be a learning tool for the respective Romani-dialects, but a tool that aids translations, text production and production of teaching materials, intended primarily for Romani-speakers active in these areas.

It also provides a resource for practical use, one that can facilitate language planning by providing an overview of dialectal variation. Moreover, it also documents common neologisms that have recently emerged in institutional contexts of usage (such as *maškarthemutno* 'international', or *manušikane čačipena* 'human rights'), and perhaps also provides some guidelines on spelling. This can be done by adopting a consistent orthography, while at the same time including dialectal variants, thus supporting the trend for pluralism in forms coinciding with the use of compatible orthographies which has been described for recent Romani codification efforts (cf. Kenrick 1996, Hübschmannová & Neustupny 1996, Matras 1999).

ROMLEX as a new approach to language codification

ROMLEX takes a novel approach to theory, methodology, and practice in the field of language standardisation and codification. The dictionary will not attempt to standardise (in the sense of suggesting a High or unified variety), but instead will list dialect-specific entries, linked via a scientifically justified, historical 'base' form. In turn, the web-based resource will not just list individual dialectal variants but will also be able to read various orthographic conventions. In this way the dominant trend in Romani literacy, toward a polycentric approach to codification that accepts variation, can be supported (cf. Hübschmannová & Neustupny 1996, Matras 1999). The particular needs of the Romani speech community and the challenge, which their situation poses, to traditional models of codification can thus be addressed, by re-examining the traditional principles of standardisation and literacy development (selection, elaboration and implementation of a single norm). The result is a new model for the role that descriptive linguistics can play in the implementation of language engineering and language planning efforts.

We would like to end our presentation with showing you ROMLEX live via the internet.

6. References

Bakker, Peter (1999) The Northern branch of Romani, in: D. W. Halwachs and F. Menz (eds.) *Die Sprache der Roma*, Klagenfurt: Drava, 172–209.

Bakker, Peter (2001) Romani in Europe, in: G. Extra and D. Gorter (eds.) *The Other Languages of Europe*, Clevedon: Multilingual Matters, 293–313.

Cech, Petra et al. (eds) 2001 *Tusa ande akhoren khelos. Lovarenge paramiči*, Graz/Wien: Institut für Sprachwissenschaft/Romano Centro.

Djuric, Rajko and Marcel Cortiade (1990) *Konsultacia vaś-i Standardizacia e ćhibaquiri. Materiali. Working paper of the Romani Union*, Warschau, Romani Union.

Halwachs, Dieter W. (1993) Polysystem, Repertoire und Identität, *Grazer Linguistische Studien 39/40*, 71–90.

Manuš, Leksa (1990) *Romani Ramayana vaj Pheniben pal e Ramaste – Romany Ramayana or the tale about Rama*, Chandigarh: Roma Publications.

Matras, Yaron (1999) Writing Romani: The Pragmatics of Codification in a Stateless Language, *Applied Linguistics, 20/4*, 481–502.

Matras, Yaron (2002) *Romani: A Linguistic Introduction*, Cambridge: Cambridge University Press.