

# **ROMANI GRAMMAR**

by Marcel Courthiade

**Volume 1: General Information, Phonology, and Morphology**

**Translated from Albanian by Geoff Husič**

**Originally published in Tirana, Albania, 1989 as  
Gramatika e gjuhës rromë**

## Notes on the digital edition

I originally completed this translation in 1989 and have recently digitized the print edition, the original text file of which is in a now-unconvertable format. Other than the few changes mentioned below, this pdf is identical to the print version. I have changed the original spelling of *Romany* to the more modern spelling *Romani*. Replacing text in a PDF is a fairly clunky procedure so there may be a few resulting format aberrations. It was also not possible to change the spelling in few cases in the images in the text. The name of the author has been changed from *Marcel Cortiade* to his preferred spelling *Marcel Courthiade*.

Geoff Husić  
Slavic & Near East Studies Librarian  
University of Kansas Libraries  
December 4, 2019

## Contents

|   |    |
|---|----|
| Translator's Preface . . . . .  | iv |
| Part I: General Information On Romani. . . . .                                |    |
| The Common Romani Language . . . . .  |    |
| Romany's Relationship To Other Languages . . . . .                            |    |
| The Distribution of Romani and The Dialectal Composition . . . . .            |    |
| Creolized and Atypical Dialects . . . . .                                     |    |
| True Romani Dialects . . . . .  |    |
| .Standardization . . . . .  |    |
| .The Principal of . . . . .   |    |
| <del>Standardization</del> Standard, With Three Levels of Tolerance . . . . . |    |
| The Common Romani Alphabet . . . . .  |    |
| Part 2: Phonology . . . . .   |    |
| Vowels . . . . .  |    |
| Basic Vowels . . . . .  |    |
| Pre-jotizing Vowels . . . . .   |    |
| Vowels Distinctive To Particular Dialects . . . . .                           |    |
| Stress . . . . .  |    |
| Stress in Declined Words . . . . .  |    |
| Stress in Conjugated Words . . . . .  |    |
| Stress in invariable Words . . . . .  |    |
| Conclusion and Practical Application . . . . .                                |    |
| Consonants . . . . .  |    |
| Subsystem of Occlusives . . . . .   |    |
| Subsystem of Laterals . . . . .   |    |
| Subsystem of Sibilants . . . . .  |    |
| The Microsystem of r . . . . .  |    |
| The Microsystem of j . . . . .  |    |
| The Problem of Dividing Lexemes into Parts of Speech . . . . .                |    |
| Part 3: Morphology . . . . .  |    |
| Morphology of Nouns . . . . .   |    |
| The Category of the Noun Class in Romani . . . . .                            |    |
| The Category of Gender . . . . .  |    |
| The Category of Number . . . . .  |    |
| The Category of Case . . . . .  |    |
| The Old Locative . . . . .  |    |
| The Substantive . . . . .   |    |
| [f-sg-a] . . . . .  |    |
| [m-pl-a] . . . . .  |    |
| [f-pl-a] . . . . .  |    |
| [f-sg-b] . . . . .  |    |
| [m-pl-b] . . . . .  |    |
| [f-pl-b] . . . . .  |    |
| Postpositions . . . . .   |    |
| Sandhi in Postpositions . . . . .   |    |
| The Combination of Postpositions . . . . .                                    |    |

|  |     |
|--|-----|
| The Situation in Indic Languages . . . . .   | 65  |
| Together Or Separate . . . . .   | 65  |
| The Pre-postposition <b>bi</b> ...- <b>qo</b> . . . . .  | 66  |
| The Vocative . . . . .   | 68  |
| The Article . . . . .  | 71  |
| Adjectives . . . . .   | 74  |
| The Declension of Adjectives . . . . .   | 75  |
| The Group of "Long" Adjectives: . . . . .  | 75  |
| The Group of "Short" Adjectives: . . . . .   | 77  |
| The Group of Foreign Adjectives . . . . .  | 78  |
| The Comparative Degree of Adjectives . . . . .   | 80  |
| Personal Pronouns . . . . .  | 82  |
| Personal Reflexive Pronouns . . . . .  | 82  |
| Possessive Adjectives . . . . .  | 84  |
| Demonstratives . . . . .   | 86  |
| Demonstrative Forms . . . . .  | 88  |
| Numbers . . . . .  | 91  |
| Other Words With Noun Declensions . . . . .  | 93  |
| The Morphology of Verb Classes . . . . .   | 95  |
| Classification of The Romani Verb . . . . .  | 95  |
| The Citation Form: 3rd Person Singular Present Tense . . . . .   | 96  |
| General information On Verb Conjugations . . . . .   | 97  |
| Personal Endings . . . . .   | 100 |
| Forms of the Participle and the Past Tense . . . . .   | 103 |
| Thematic Verbs . . . . .   | 106 |
| Foreign Words . . . . .  | 107 |
| Verbs With A Stressed Root With Stem Final -d- . . . . .   | 110 |
| Verbs Derived From Adjectives, and Causative Verbs . . . . .   | 112 |
| Imperatives . . . . .  | 115 |
| Formation of Tenses Derived Morphologically . . . . .  | 116 |
| Nominal Lexemes Formed From Verbs . . . . .  | 117 |
| Participles . . . . .  | 117 |
| Verbal Adverbs . . . . .   | 117 |
| Verbal Nouns . . . . .   | 118 |
| The Copula . . . . .   | 119 |
| Invariable Lexemes . . . . .   | 121 |
| Adverbs . . . . .  | 121 |
| Prepositions and Postpositions . . . . .   | 123 |
| Conjunctions . . . . .   | 124 |
| Particles of Negation . . . . .  | 124 |
| The Relationship Between Variants of the Common Language . . . . .   | 126 |
|  |     |
| Diagrams and Tables . . . . .  | 128 |
|  |     |
| Diagram #1: The Position of Romani Dialects in Comparison to<br>Common Romani, and Hindi and Punjabi . . . . . | 128 |
| Diagram #2: The Path of Romani Standardization . . . . .   | 130 |
| Table #1: Relationship Between Nouns and Verbs in Romani . . . . .   | 131 |
| Table #2: Synoptic Overview of Plural Forms of Nouns Ending in<br>A Consonant in the Singular . . . . .        | 132 |
| Table #3: Sandhi System . . . . .  | 133 |
| Table #4: Sandhi System (example of Application) . . . . .   | 134 |



|  |     |
|--|-----|
| Table #5: Comparison of the Postpositional Systems of Hindi<br>and Romani . . . . .            | 135 |
| Table #6a: Declension of Nouns: Types Bakro m. «ram» and<br>Bakri f. «ewe» . . . . .           | 136 |
| Table #6b: Declension of Nouns: Types Rrom m. «rom Man» and<br>Rromni f. «rom Woman» . . . . . | 137 |
| Table #6c: Declension of Nouns: Types Dad m. «father» and<br>*daj, °daj f. «mother» . . . . .  | 138 |
| Table #6d: Declension of Nouns: Types Marro m. «bread» and<br>*pani, °paj m. «water» . . . . . | 139 |
| Table #6e: Declension of Nouns: Types Muj m. «mouth» and<br>Sasuj f. «mother-in-law» . . . . . | 140 |
| Table #7: Correlative words . . . . .  | 141 |

## Translator's Preface

It has been my pleasure to translate this Romani Grammar from Albanian into English. Please observe the following small changes that I have made in the presentation of the grammar.

There are several levels of transliteration used in the Albanian edition of this grammar: 1) Romani in the official orthography, 2) Romani transliterated phonemically using a) Albanian orthography, b) letters from some other languages: Serbo-Croatian, Macedonian, Turkish, Greek, etc. In the Albanian original of this grammar M. Cortiade uses the Albanian orthography in his phonetic and phonemic descriptions.

Since the Albanian values are of little value to the average English reader, in translating this grammar, I decided to substitute the Albanian orthography with the official Romani letters in most cases, since they are just as revealing as the Albanian. In the case of the descriptions of phonetic variants that arise from that palatalization of velars and dentals I have employed a transcription different from Cortiade's. I have done this mainly to avoid using the ambiguous letters *ć*, *ç*, and *q* which serve two different purposes. In phonetic descriptions *ć* represents the sound of the Serbo-Croatian *ć*, whereas the Romani letter *ć* is equivalent to the Serbo-Croatian *č*; *ç* in phonetic descriptions represents *ch* as in *church*, whereas the Romani letter is pronounced either [ts] or [s], depending on environment; and *q* in phonetic descriptions represent the *q* in Albanian (a palatal stop), whereas the Romani letter is pronounced either [k] or [g], depending on environment.

In addition, in phonetic descriptions I have substituted *ĵ* (*j* in *judge*) for the sound designated as [xh] in the original, rather than using the Romani [z], since this letter looks so similar to the IPA (International Phonetic Alphabet) notation for the sound of *z* in *azure*.

I hope these changes help rather than hinder reading this translation.

In the section describing the Romani alphabet I have added the English equivalents (or near equivalents) of these letters and a phonetic description to aid the reader in reproducing an approximate value.

I give below the phonetic/phonemic transliteration I have substituted, with the Romani and Albanian, for those letter which differ:

| Transliteration | Romani letter | Albanian letter |
|-----------------|---------------|-----------------|
| ḱ               | *             | q               |
| ḡ               | *             | gj              |
| ṛ               | *             | ç†              |
| ċ               | ċ             | ç               |
| ḏ               | *             | dž†             |
| dz              | dz            | x               |
| ḵ               | z             | xh              |
| š               | š             | sh              |
| ž               | ž             | zh              |

\*These are phonetic variants only, not phonemes.

†This is not an Albanian letter, but is nevertheless used in the original.

The term "Common Languages" in this work signifies the unified, standard language composed of certain common elements across several dialects which is proposed in this grammar, rather than signifying the "every-day" speech of any one dialect.

There are some problems with the exact definition of what are called "reflexive" verbs in the grammar. These verbs cover several functions, i.e. true reflexive meanings, mutual meanings, as well as passive meanings. Therefore, I have felt more comfortable labeling them "reflexive/passive" to make this issue clear.

I have chosen to transliterate the names of dialects with the Official Orthography, (e.g. Mečkari, etc.), since there really seems to be no consistency in the English renderings of these names.

The following note was tipped into the original Albanian edition. I have included it here for historical interest (See also Translator's preface).

Note: The 4th Romani Congress, which was held in Warsaw, Poland on April 8, 1990, approved the proposal of the Linguistic Committee for an official common Romani alphabet. This alphabet is generally very similar to the alphabet used in this grammar, however there are some differences which are shown below:

- instead of the letter ċ, the letter z is used to represent the Albanian xh. The graphic variant ž is also acceptable.

- the letters ä, ö, and y have been totally abandoned, as has the consonant ɸ (Albanian x).

- instead of 8 for the postpositional grapheme, the letter ̸ is now used (with the value of d after n, and t in every other position).

## PART I: GENERAL INFORMATION ON Romani

### THE COMMON Romani LANGUAGE Romany's Relationship to other Languages

Romani, along with ten languages spoken in India, constitutes a branch of the Indic group of Indo-Iranian languages which, in turn, belong to the large Indo-European family of languages. The most well-known Indic languages are Hindi (and its variant, Urdu), Bengali, spoken in Bengal, Nepali, Singhalese, spoken in Sri Lanka, etc. The close affinity of Romani with these languages was ascertained as early as the 17th century and is supported by the following features:

- 1) Many phonetic correspondences, particularly the presence in Romani of the aspirated consonant phonemes /ph/, /th/, /ch/, and /kh/, which are also characteristic for Indic languages;
- 2) As in the Indic languages, the occurrence of the phonemes /c/, /f/, and /z/ is low in Romani, and are encountered almost exclusively in foreign borrowings. On the other hand, the phoneme /j/ is very frequent in Romani, as it is in the Indic languages;
- 3) The existence of several hundreds of roots common to Romani and the Indic languages, and not only do these roots constitute approximately 80% of the basic Romani lexicon, as reflected by Swadesh's list, but they also exhibit regular etymological correspondences;
- 4) Correspondences between the Romani and Indic nominal systems (i.e. declension based on two forms: A-Form and B-Form) as well as correspondences in a series of nominal and non-nominal endings, e.g. the existence of two classes of adjectives according to case;
- 5) The alternation of pronoun forms according to gender and number;
- 6) The use of postpositions, which are characteristic of Indic languages, alongside with prepositions, which are characteristic of the European languages.
- 7) Certain correspondences in the verbal system.

Of course, during their migration from India to Europe, the Rom have acquired a considerable number of words from languages spoken in areas through which they passed, e.g. from Persian, Turkish, Armenian, Greek, etc., as well as the Slavic languages, and in our century the Rom have acquired many words of an international nature which have now become firmly entrenched in the language.

## THE DISTRIBUTION OF Rōmani AND THE DIALECTAL COMPOSITION

Romani is spoken by Rom dispersed throughout Europe, the Near and Middle East, and North and South America. The overwhelming majority have preserved their mother tongue well, however some clans have lost it, especially in the Iberian Peninsula, as a consequence of the Inquisition which forced them to use the Castilian language, to take Christian names, and adopt the religion of the merciless and "very Catholic" kings. They were punished with death by burning if they showed defiance. It is no wonder that under such conditions the Rom in the Iberian Peninsula have preserved only a few hundred or even tens of words of Indic origin.

On the other hand, the Rom of Central and Eastern Europe (including those who penetrated into Western Europe in the last 100 years) use the same language, which, like all languages, can be divided into different dialects.

Until now there has been no one exact and satisfactory scientific classification for Romani dialects. Nevertheless, all linguists more or less agree that Romani dialects can be divided into the following:

- creolized or jargonized dialects: Ibero-Romani or Kalo (about 100,000 speakers), Anglo-Romani or Pogadi çip (about 30,000 speakers), Armeno-Romani or Bosha (number of speakers unknown), and, according to some authors, Irano-Romani or Zargari and Siro-Romany or Zawari;

- semi-creolized dialects: Manuś and Sinto (several hundred thousand speakers);

- atypical dialects: of Finland, Zakopane in southern Poland, and Wales (several hundred thousand speakers);

- true Romani dialects, which constitute the greatest number of dialects (with several million speakers), which are divided into two branches of dialects:

- Balkan branch, which is characterized by the pronunciation of the phoneme čh as [çh], e.g. čhuri «knife», that is to say, retaining the affricate character. The dialects of this group is found mainly in the Balkan Peninsula, but has also penetrated into Northern Europe to Poland and the Soviet Union, and into North and South America.

- Central European group, which is also call the "Vlah" group, which is characterized by the non-occlusive pronunciation of the phoneme čh as [śj] or [ś]: čhuri [śjuri, śuri].

There has been much discussion about the details of this classification. There are also some very fundamental problems which have yet to be solved. Thus, specialists are not of one mind about the division of dialects, and there is no answer to the question whether we can talk of two dialects (Balkan and Central European) or two branches of dialects (Balkan and Central European) into which particular dialects are divided. The difficulty is a methodological one and arises from the fact that general linguistics has yet to determine what exactly is a "dialect." In the case of Romani, we perceive that we are dealing with two main dialects:

- one which participated in the change of the affricate /<sup>h</sup>ʃ/ and /<sup>h</sup>č/ into [ʒ] and [š]. This is the Vlah or Central European dialect with the subdialects Lovari, Kalderaš, Vižinas, etc.
- the other, which has not participated in this change and retains the original pronunciation as [ʃ] and [č]. This the Balkan dialect with the subdialect Mahazeri, Mečkari, Ursari, Lovari, etc.

#### CREOLIZED AND ATYPICAL DIALECTS

As we have seen above, the Rom of the Iberian Peninsula, who call themselves *hitanos*, now use native words connected by grammatical features from languages of the areas where they live, especially Spanish, forming a kind of jargon, which can not be quite called a language. They use it as a secret language in certain situations. In its social function and its intelligibility it can be compared to the "Dogançe" jargon used by Albanian craftsmen in past times. The Rom of England have also lost their language from the grammatical standpoint. Besides English, they speak a jargon formed of words of Romani origin and English grammar. In contrast, the Rom of Wales have a dialect with many foreign words, but have to a great degree retained Romani grammar. This difference is directly connected with the difference in the position of the English and Welsh in comparison to the Rom, in an English domain.

It is interesting to compare a text in one of the jargons with a translation into both of the true Romani dialects. For this purpose we can cite a popular verse in the Anglo-Romani jargon:

As I was a-jawing to the gav yeck divvus  
I dikk'd on the drom miro Romani chai  
I pootch'd las whether she come sar mande,  
And she penn'd: "you sar wafo rommandis!"

No, sar mande there is keck wafo romady,  
So penn'd I to miro Romani chai  
And I'll ker tute miro tacho romani

If you penn you come sar mande.

As I was walking to the town one day,  
I saw on the road my Gypsy girl.  
I asked here whether she'd come with me,  
and she said, "You walk with other Gypsy women!"

No, I have no other Gypsy woman,  
So said I to my Gypsy girl.  
And I'll make you my true Gypsy girl,  
If you say you'll come with me.

As we can see, this text presents all the characteristics of a creolized dialect:

- a large portion of foreign words, especially grammatical words;
- a fundamental change in the native element as regards pronunciation, (e.g. wafo «another» from aver, vaver), or in meaning (e.g. sar mande «like me» is used with the meaning «with me», probably due to a change in the postposition -çar into a preposition and its confusion with the conjunction sar, similarly the word meripe «death» with the meaning «life» in some dialects;
- a grammatical system with a great number of foreign elements (such as past tense formations with the English -ed: penn'd «said», or the future with the English 'll: I'll ker «I'll do», and the plural with s, e.g. rommadis «women»);
- mistakes, in the foreign elements (e.g. she come), as well as in the native element (miro tacho romni «my true Gypsy girl», where the adjective has the masculine endings even though modifying a feminine noun), the loss of accusative, etc.

#### TRUE Romani DIALECTS

Let us compare now this text with its translation into true Romani, but into the two dialects the farthest apart from each other:

A) in the dialects of the Warsaw Rom (Central European branch):

Kana me zàvas and-o gav jekh dës  
Dikhlëm p-o drom mure romane čhovorra  
Pučhlëm la te avel mança  
Aj voj phendã: "tu zas avere romënça!"

Na, mançar naj nijekh aver romni  
Sa gadã phendëm mure romane čhovorraqe  
Aj kerav tut muri čáci romni  
Te phenësa so aves mança.

B) into the dialect of the Babuxhi of Korçë (Balkan branch):

Kanà me zàvsas an-o gav jekh dives  
Dikhlo m t-o drom me romane çhaja  
Puçhlom la te avel maņçar  
Ta oj phendas: "tu zàsa avere romëņçar!"

Na, maņçar nanaj asjekh aver romni  
Akhhal phendom me romane çhajaqe  
Ka kerav tut mi çáci romni  
Te phenësa so aves maņçar.

It is apparent that the differences between these dialects are minimal, considering the distance between Korçë and Warsaw, while both are distinguishable from Anglo-Romani. In recent years lexico-statistical methods have yielded some valuable data to qualify the "distance" between dialects, not from a geographical point of view, but from a dialectal one. These have verified that Romani dialects, as stated above, constitute one language, however the creolized and atypical dialects do not belong to this language, since they are too different from true Romani to be considered as part of it. Diagram # I represents the mutual positions of the main Romani dialects in light of results of the "measurement" of distance between them (distance 1 is unitary; when the distance between two dialects is greater than one, it is considered that we are dealing with two languages, while when the distance is most small, then they are considered two dialects of the same language).

This is a valuable basis for defining the dialectal range as well as the material base of the common language. We are indisputably dealing with one language, however, like all languages, it manifests dialectal forms. It can likewise be noticed that, except for the creolized and atypical dialects such as Anglo-Romany, Sinto-Manuš, Kalo, etc., the vast majority of dialects constitute one very similar polylectal system which can serve well as the basis for the cultivation of a common Romani language. For a long time at congresses, and various cultural and scientific activities the Rom have expressed their need for a common language which would aid in the further development of a standard Romani language.



## STANDARDIZATION

### The Principal of Standardization

In light of the modern point of view regarding the field of linguistic standardization and while relying on an analysis of the dialectal divergences in Romani it becomes indisputably evident that there are two opposing, but complementary, forces for standardizing the language: codification and normalization. Codification is defined as the accommodation between a language and its graphic code, that is to say, a system of letters, while defining pronunciation for each letter according to its position in the word according to dialect. Codification does not affect the integrity of the spoken language, but only determines the correspondences, not always on a one to one basis, of the totality of phonemes on one hand, and the totality of letters on the other. Normalization is defined as the accommodation of the language between a system of norms which eliminate elements which cause obstacles to mutual understanding between speakers of different dialects. It is accomplished only when complete and easy mutual understanding is not sufficient to be assured by codification.

Romani was first written in the national alphabets of the areas where the Rom resided, so that a word, even though pronounced alike in a different dialect, had a different spelling in a different country. The numbers of spellings increases when the pronunciation of the word varied between dialects. In recent years the so-called "Common Spelling" (maškarutno lekhipe), which allows many different spoken features to be written in one common graphic dressing, has become widespread. This is the only means to rigorously accomplish the resolution of the 1st Romani Congress, held in London in April IV 1971, which recommended: "There is no one dialect that is better than any other, but we are seeking a common normalized language to use at the Congress and to write international papers."

It is precisely this spelling, valid for all Romani dialects, which is used in the manual. The theoretical and historical aspects of the spelling of this graphic inter-system are not of particular interest as regards the goals of this manual, and so it will not be analyzed here. The reader who wishes to delve more deeply into this question can read, among others, the articles published in Koha in Titograd (1986/1 and 1987/2), as well as the work of the congress "Gjuha e kultura e rromëve," held in Sarajevo in June of 1986.

It suffices here to illustrate briefly the principle of comparing a written feature with several pronunciations by means of special rules for each of the dialects, while taking a concrete example:

Example #1: realization of the velar occlusive k

Throughout the history of the language, the Old Romani velar occlusive k has undergone palatalization before front vowels in a series of dialects, so that it is pronounced [k'] in Erli, [k̟] in Mećkari, [tʃ] in Gurbeti, [ć] in Tharegoni, or even [t'] in Ursari of the Soviet Union, while many other dialects have no such change, as for example in Mahazeri.

Example #2: realization of the dental d in past tense endings

The dental occlusive d in Old Romani was palatalized in a different phonetic context, i.e. before pre-jotizing vowels which form the personal forms of the past tense in a some, but not all, of dialects. Thus, the original ending of the third person

singular -dā (probably from -d-ia; ā indicating the diphthong [ja] and is still pronounced today as [dja] in many dialects, while in others it has developed into a form such as [da] in Montenegro (with the loss of pre-jotization), [j̟a] in Mahazeri, [ǰ] in Zambaz, Baruć, and some dialects of Slovakia, [d̟ja] in Erli, and [d̟əa] in many dialects of Romania. In the Gurbeti dialects pre-jotization has completely disappeared to be pronounced [da].

However, if we take as an example the third person of the past tense of the verb «to do», which is written kerdā in the common spelling, then we discover that a large number of realizations of pronunciation is possible, according to the degree of palatalization. These are shown in the table below:

|         |          |         |         |         |          |
|---------|----------|---------|---------|---------|----------|
| kerdja  | k'erdja  | ǰerdja  | terdja  | ćerdja  | t'erdja  |
| kerđa   | k'erđa   | ǰerđa   | terđa   | ćerđa   | t'erđa   |
| kerĵa   | k'erĵa   | ǰerĵa   | terĵa   | ćerĵa   | t'erĵa   |
| kerĝa   | k'erĝa   | ǰerĝa   | terĝa   | ćerĝa   | t'erĝa   |
| kerġa   | k'erġa   | ǰerġa   | terġa   | ćerġa   | t'erġa   |
| kerd̟ea | k'erd̟ea | ǰerd̟ea | terd̟ea | ćerd̟ea | t'erd̟ea |

and with the loss of pre-jotization:

|       |        |       |       |       |        |
|-------|--------|-------|-------|-------|--------|
| kerda | k'erda | ǰerda | terda | ćerda | t'erda |
|-------|--------|-------|-------|-------|--------|

Nearly half of these forms are attested in Romani dialects, but it must be considered that less than 10% of Romani dialects have been rigorously studied, and so the possibility of even a larger number of realizations being attested cannot be precluded.

When applying the common spelling these many forms are identically represented as kerdā in dialects which have preserved

pre-jotization, regardless of the actual pronunciation, whereas in dialects which have loss pre-jotization these are written simply as *kerda*.

Each speaker has at his disposal a separate series of rules for his dialect. Thus, for example, in *Mahazeri*, the speaker will read according to the above rules, specific to his dialect:

- a) *-k-* is always pronounced as [k], regardless of its position;
- b) *-d-* is pronounced as [j] before a pre-jotizing sign.

Consequently he will pronounce: [ker]a].

On the other hand, a reader from *Topansko Pole* (Skopje) will realize [tʷerǵa], since his rules are:

- a) *-k-* is palatalized to [tʷ] before a front vowel;
- b) the group *-d̄* is pronounced as [ǵ] in Macedonian (or *gj* in the Albanian of *Gjirokaster*), and so forth for the other dialects.

It can be clearly seen that this common spelling is simplified in comparison with phonetic alphabet in the narrow sense, which reflect all variants and realizations in the spelling, a topic of interest only to phoneticians, but not for the everyday use of the language. On the other hand, the common spelling presents no particular difficulty to the average reader, since he needs only to assimilate the rules which reflect his own dialect, that is to say, very few rules. He will also read other dialects according to his own rules. Thanks to codification, we easily achieve a unification of the written language, without violating features specific to spoken dialects.

However, there are many cases when the forms are so different that codification is not able to reconcile the differences in spelling. This pertains particularly to lexical differences, as, for example, *nakh* and *ruthuni* «nose», *men* and *kori* «neck», *zuket* and *rikono* «dog», *gilavel* and *bagel* «she, she sings», or *dikarel*, *dićarel*, *dikerel*, *ikerel*, *ićerel*, *ligarel*, *lingarel*, *inkarel*, *lićarel*, *inkerel*, *anćarel*, *rikerel*, *likrerel*, *ankerel*, etc., all meaning «to hold». In this case there is no way to find a codification which will consolidate all these forms, and so the task must rely on normalization, i.e. standardization which will proclaim a "norm," whether or not to regulate the acceptance of the lexemes in the word, and under what conditions. If we proceed from the examples of the above lexemes, "normalization" will propose:

- a) the words *nakh* and *ruthuni* «nose», *men* and *kori* «neck», *zuket* and *rikono* «dog» should be accepted as synonyms in the Common Language, just as Albanian has synonyms with different dialectal origins, e.g. *derr* and *thi* «pig», *gjalpë* and *tlyn* «butter», etc. Eventually a distinction in the sense of these pairs may develop,

as already exists in some dialects, e.g. **nakh** «nose» and **ruthuni** «snout», **men** «neck» and **kori** «nape», **zukul** «dog» and **rikono** «puppy», etc.

b) the situation differs somewhat with the pairs **gilavel** and **bagel**, for the reason that **bagel** has a very limited dialectal distribution and is not intelligible in other dialects, while **gilavel**, being related to **gili** «song», is intelligible even to the Rom that use **bagel**. Thus, we can only propose that the form **bagel** be totally excluded from the Common Language. However, in its defense, there are some pertinent historical factors. The first Rom to compose poetry, **Papuša** (Bronisława Wajs, 1908-1986), frequently used the verb **bagel** in his works. Consequently **bagel** should also be accepted into the Common Language, but as having a poetic nuance, reminiscent of the poetry of **Papuša**. Another task is the selection between the variants **galavel**, **gilävel**, **gilabel**, **gilabal**, etc.

c) in the case of the many forms of the verb «to hold», standardization will select one or two recommended forms for the Common Language, which will consolidate the other superfluous forms, especially since they present no particular interest, nuance, or cultural value.

#### ONE COMMON STANDARD, WITH THREE LEVELS OF TOLERANCE

Codification and normalization play different roles in the four main realms of dialectal divergences: phonetics, morphology, lexicon, and syntax, while examining three levels of tolerance which must be respected in the formation of the Common Language.

These three levels are:

- tolerance in the written language, i.e. more than one referential form is selected as the written standard, as long as it does not violate intelligibility, on the basis of scientific criteria, rather than arbitrarily;
- tolerance in pronunciation, while allowing the reading rules in each dialect to reflect the authentic pronunciation;
- tolerance in speech, which swings between two linguistic levels: a rigid standard on one hand, and the colloquial speech among family on the other, sometimes closer to one, sometimes to the other, according to concrete conditions where the speech occurs. The selection of the level it accomplished by the speaker automatically in an unconscious manner, as in all languages.

A kind of dynamic framework arises here which creates suitable conditions for the language's free development towards a complete consolidation of the dialects, as illustrated in Diagram #2.

It must be emphasized that codification is nothing other than exposure of an internal system of interdialectal relations which had already existed in the language, whereas normalization is a human intervention into linguistic material in order to adjust particular points in cases where codification is incapable of creating a satisfactory compromise. A linguist can improve codification in his own work, for the reason that by means of this process he creates nothing, does not change existing conditions, but only brings out a deep system which forms the common basis of dialects.

In contrast, **normalization** requires an intensive and broad collaboration between linguists from different places and with different specialization (lexicologists, semanticists, etymologists, socio-linguists, scholars of Indic languages, etc), since every difficult topic needs to undergo detailed discussion: elimination of words which violate mutual intelligibility. selection of forms most appropriate for the Common Language, classification of the preferences between them, definition of their semantic and morphological variants, creation of specialized subject dictionaries, etc.

It also becomes clear that, in the process of standardization, codification chronologically precedes normalization and is complemented by it.

## The Common Romani Alphabet

We give below the Romani alphabet along with a transliteration. (See translator's introduction).

|         |                  |   |
|---------|------------------|---|
| a       | [a]              | low central vowel (o in hot)  |
| ä       | [ɨ]              | Middle front vowel. Like ı in Turkish or î in Romanian.   |
| ǎ       | [ja]             | (as well as ǒ for [jo], ũ for [ju])   |
| b       | [b]              | voiced bilabial occlusive (b in bat)  |
| c       | [c] <sup>2</sup> | voiceless dental affricate (ts in bats)   |
| ć       | [ć]              | voiceless palatalo-alveolar affricate (ch in church)  |
| čh      | [čh]             | aspirated palatalo-alveolar fricative (ch in charge but more strongly aspirated)  |
| d       | [d]              | dental occlusive (d in dog)   |
| e       | [e]              | middle front vowel (e in red)   |
| ë<br>or | [ë]<br>[o]       | schwa or middle back vowel (e in the or o in whole)   |
| f       | [f]              | labio-dental fricative (f in fire)  |
| g       | [g]              | voiced velar occlusive (in some dialects palatalized as [ğ], [ǰ] or [j]) before front and pre-jotizing vowels) (g in guard, g in argue, g in age, etc.)   |
| h       | [h]              | glottal fricative (h in hat)  |
| x       | [x]              | velar fricative (ch in Scotts loch)   |
| j       | [j]              | palatal semi-vowel (y in yet)   |
| k       | [k]              | voiceless velar occlusive (in some dialects palatalized as [k'] [k̟] or [ć] before front and pre-jotizing vowels) (k in skip, k in hack, ch in hatch, etc.)   |
| kh      | [kh]             | aspirated voiceless velar occlusive (in some dialects palatalized as [k'h] [tʰ] or [ćh] before front and pre-jotizing vowels) (c in courage, c in care, ch in charity, but more strongly aspirated) |

|    |            |   |
|----|------------|---|
| l  | [l]        | voiced alveolar nonfricative lateral (usually palatalized before front and pre-jotizing vowels) (l in light, but more palatal before a front vowel)                   |
| m  | [m]        | voiced bilabial nasal (m in mom)  |
| n  | [n]        | voiced dental nasal (n in nose)   |
| o  | [o]        | middle back vowel (o in whole)  |
| p  | [p]        | voiceless bilabial occlusive (p in spot)  |
| ph | [ph]       | aspirated voiceless bilabial occlusive (p in pal, but more strongly aspirated)  |
| r  | [r]        | dental flap (r in Spanish pero)   |
| rr | [rr], [r̄] | (nasal), [ɣ], [r̄] (retroflex), [ndr], etc. (many pronunciations. See section on laterals)  |
| s  | [s]        | voiceless dental fricative (s in sail)  |
| ś  | [ś]        | voiceless palato-alveolar fricative (sh in share)   |
| t  | [t]        | voiceless dental occlusive (in some dialects palatalized as [t̄] or [č] pre-jotizing vowels) (t in start, t in British pronunciation of tuesday, etc.)                |
| th | [th]       | aspirated voiceless dental occlusive (in some dialects palatalized as [t̄h] or [čh] pre-jotizing vowels) (t in tour, ch in charge, etc., but more strongly aspirated) |
| u  | [u]        | high back vowel (u in rue, but without w glide)   |
| v  | [v]        | voiced labio-dental fricative (v in van)  |
| z  | [z]        | voiced dental fricative (z in zany)   |
| ž  | [ž]        | voiced palato-alveolar fricative (z in azure)   |
| ž  | [ʃ]        | voiced palato-alveolar affricate (j in judge)   |

Besides these letters, the Common Language also uses three other graphemes, which appear only in postpositions, and which each have two pronunciations. depending on position, i.e. after an n or after another phoneme:

After n

Θ [d]  
ç [c]  
q [g]

After another phoneme

[t]  
[s] [h] [j] or [0] depending on dialect  
[k]

Examples: amenΘe 'by us,' laΘe 'by her,' e rromença 'with the Roms,' me diaça 'with mother,' e çhavenqe 'to the boys,' and tuqe 'to you.'

The letter q undergoes palatalization before front vowels if g and k undergo such a palatalization when pronouncing the word. The group -sq- is realized as [-sk-], [-hk-] or [-k-], as well as with the respective palatalizations. In some dialects it is pronounced as a simple [-s-].

Example: me dadesqe «to my father», me phralesqo kher «my brother's house», lesqoro çhavorro «his son», etc.

The rules which dictate the pronunciation of these signs (graphemes) are different for each dialect and are called sandhi. This word comes from the grammar of the languages of India, which exhibit this phenomenon.



## PART 2: PHONOLOGY

### VOWELS

#### BASIC VOWELS

Romani has five basic vowels which are common to all dialects. They are:

|        | front | central | back |
|--------|-------|---------|------|
| high   | i     | -       | u    |
| middle | e     | -       | o    |
| low    | -     | a       | -    |

The middle open vowels (e and o) represent only one degree of openness. There is no opposition between an open e and o and a closed e and o. Differences that arise are not phonemically relevant.

#### PRE-JOTIZING VOWELS

Besides the above simple vowels we must also mention the pre-jotizing vowels *ǎ*, *ǒ*, *ǔ* (and in a few cases *ě* and *ĩ*) which can be transliterated as [ja], [jo], [ju], etc., in the majority of cases. These are pronounced the same as the simple vowels, however they mark a change in the preceding consonant according to rules which will be laid out below. Only *ǎ* is present in all dialects.

#### VOWELS DISTINCTIVE TO PARTICULAR DIALECTS

Besides the above basic vowels and their pre-jotizing counterparts, many dialects also have the following vowels:

- *ü*, which is pronounced similar to the Albanian *y* or the German *ü* in *über*. This sound occurs only in foreign borrowings from Turkish, Albanian, or Hungarian. In many dialects this sound is replaced with either *i* or *u*. For example the Turkish word *dünya* «world» appears in Romani dialects as either *dînia* or *dûnia*.

Syllabic *r* occurs in a small number of dialects: *patrɿn* «leaf», *brşind* «rain», *vrtinel* «it spins», etc.

- *ä* fulfills three functions:

1) it represents a high, unrounded vowel, like the Turkish *ı* or the Romanian *î*, and occurs, for the most part, in borrowings from Turkish: *adämi* (from Turkish *adım* «pace», or from Romanian:

**gändiv** or **gändisarav** «I think» (from Romanian **mă gândesc** «I think»). This borrowing has of late tended to be replaced by the word **godisarel**. This sound also occurs in some words of uncertain origin: **xänza** «small hole».

2) it also represents the sound which accompanies the syllabic **r**: **patärn**, **bäršind**, and **värtinel**, etc. This sound is found also in the Bulgarian **ъ**: in **вѣртя** «I spin» or in the Romanian **î**: in **vîrtej** «whirlpool».

3) the third function of this grapheme is to represent a much weakened sound, which, in most dialects, does not quite become zero: the words **gäja** «so» (Gurbeti dialect) and **akäja** «this» (in most Balkan dialects) can be transliterated as **[gǰja]** and **[akǰja]**. This sound may also completely disappear in some pronunciations, i.e. **[g-ia]** and **[ak-ia]**, however it is preserved in writing to show that the **j** does not palatized the consonant preceding it.

The letter **ë** does not represent one particular Romani vowel. It is a grapheme which represents several pronunciations depending on the dialect. There are words in which an original **e** becomes labialized into an **o** in some dialects, while in others it might become the central schwa vowel **é**. Consequently, three stages of this development, **e**, **é** and **o** are represented by the one grapheme, **ë**. This development can be seen in the synchrony of some dialects: in the Tharegoni dialect of Prishtina, while the elder generation pronounce **ë** as **e**, the younger people pronounce it as **é** or **o**. However, in Gurbeti and Mahazeri **ë** has stabilized and is pronounced **o**.

This letter grapheme is of particular importance in that it occurs not only in a number of roots, but also in the most frequent verbal endings, especially the third person singular of the present tense, for a large group of verbs: those ending in **-ël** or **-ëla**, which are pronounced as **[-el(a)]**, **[-ël(a)]**, or **[-ol(a)]**.

It should be noted that in these endings, as well as in most other cases, the labialized **l** (like dark **l** English **table**, or **ll** in Albanian **yll**) cannot not follow **e**: in **šël** «hundred», **šëlo** «rope», **devël** «God», **zukul** «dog», etc. Occasionally **ë** is also encountered in another context, as seen in the word **věš** «forest» which is pronounced either **[veš]**, **[věš]**, or **[voš]** according to dialect.

The dots over **ë** can be written in either all dialects, or in just those in which **ë** has become centralized to **[é]** or labialized to **[o]**.

## STRESS

Before we proceed to the chapter on consonants it would be useful to review the system of stress (or accent) in Romani.

Stress in Romani is characterized as being dynamic, and the vowels of unstressed vowels are not reduced, but are pronounced quite clearly, giving Romani a distinctive sound.

The place of stress varies based on grammatical and lexical criteria, that is to say, it is not phonological, and thus Romani is said to have free stress.

In most dialects, as well as in the Common Language, stress can be differentiated in three categories of words, i.e. declined words, conjugated words, and invariable words. This division is understandable since stress in Romani is morphological.

### STRESS IN DECLINED WORDS

This group comprises both native words and foreign borrowings.

a) In words of Indo-Iranian origin the stress falls on the last syllable: **raklò** «non-Romani boy», **manùs** «person», **gurumnì** «cow», **pošomalò** «hairy», **bibaxtalò** «unfortunate», etc.

The word **lindra**, or its variant **lindri** «sleep» is an exception. It is paroxytone. In this chapter we mark stress in the conventional manner, with an grave accent over the vowel).

Like Indo-Iranian words, many borrowings from Greek also are oxytone irrespective of the place of stress in the original Greek: **luludì** «flower», and **stadi(k)** «hat», which are both paroxytone in Greek. Exceptions are **kòkalo** «bone», **pètalo** «horseshoe», and **kakàvi** or **kikàvi** «pot».

There exists also a secondary stress that usually falls on the first syllable of the word (indicated here by the underlining): **pošomalò**, **bibaxtalorrò**, etc. This holds true also for words of two syllables, and in many dialects the stress many fall on either syllable in a number of common words: **marrò** or **màrro** «bread», **panì** or **pàni** «water», **dilò** or **dìlo** «crazy», etc. It is frequently difficult to determine exactly where primary and secondary stress indeed fall.

In words with the suffix **-pe** (and its variants **-pen**, **-pa**, **-ba**, **-be**, **-ben** and **-mos**) the secondary stress falls on the penultimate syllable. In some Balkan dialects (in Mečkari) words with this suffix are proparoxytone: **nasvalipè** or **nasàvlipe** «illness», and **nasvalò** or **nàsavlo**. We are unable to determine from synchrony whether this feature is an old one or whether it is an innovation.

Stress in the plural falls on the same syllable as in the singular: *barlă*, *baxtalè*, and *luludă*, but *lîndre* and *kòkala*, etc.

In the case of words having an added ending in the plural, the stress is thrown onto that ending, and the word remains oxytone: *manuşà*.

Until now we have mentioned words appearing only in the A-Form, that is to say, in the case used mostly after prepositions, as the subject, and, for nouns denoting an inanimate object, for the direct object. However, besides the A-Form, declinable words in Romani also appear in another form, called the B-Form, which is used mainly before postpositions and, for nouns denoting animate beings, for the direct object.

In the B-Form, native Romani words always have the stress on the oxytone: *gurumnă*, *manuşès*, *baxtalè*, *luludă-*, *stadă-*, etc., as well as: *lindrà-*, *kokalès-* (the dash after the words indicated that this is a bound form which cannot exist in isolation and must be bound to a postposition. This is the case in nouns denoting inanimate objects).

b) Foreign words from Turkish or European languages, as a rule, enter Romani with the stress on the word's root: *bòlnica* «hospital», *pòpara* «bread with milk», *sòfra* «driver», *mèsecina* «moon» (the true Romani word is *čhumut*), *krèveti* «bed», etc. If the stress does vary from this rule, then the word becomes paroxytone: *bolnica*, *mesečina*, *krevèti* or *krevàti*, etc.

As we shall see in the section devoted to lexicon, in masculine foreign words ending with a consonant, the vowels *o* or *i* are affixed in Romani. This ending is never stressed: *jezèro* «lake», *pràxo* «dust», *natàri* «key» (a later borrowing which replaced the earlier Greek word *klidi*), *aparàti* «apparatus», etc. The same phenomenon is observed in adjectives: *lùngo* «long», *slàbo* «weak», *zelèno* «green», etc.

Foreign words ending in an unstressed *a* retain this pattern in Romani: *žàpka* «frog», *kumbàra* «godfather», *filizòfia* or *filizofia* «philosophy», etc. In some dialects a number of words have acquired a derivational suffix of unknown origin, i.e. *-va*: *kafàva* «coffee», *malàva* «town quarter», *okàva* «a unit of weight», but this makes no difference in the stress pattern. In the plural the stress in these words always falls on the same syllable as in the singular, even when a syllable is added (e.g. in the case of the plural ending *-ură*: *bòlnice* or *bolnice*, *krevătă* or *krevătură*, *žàpke*, etc.

In the B-Form, as in native words, this stress is always oxytone: *bolnicà-* (sing.), *bolnicèn-* (pl.), *krevatès-* (sing.), *krevatèn-* or *krevatonèn-* (pl.), *žàpka-* *žapkèn-* (pl.), etc.

In the bulletin issued by the 2nd Congress of Romani Studies (Paris, December 1986), Shaip Jusuf gives an interesting interpretation which states that the root of a noun is indeed oxytone, however some endings have an even stronger stress than that of the root. These endings are said to have "strong stress" (in Romani *zoralo stres*) and are distinguished from endings which have a "weak stress" (in Romani *kišlo stres*). It is likewise observed that all endings of the B-Form have strong stress, while the vast majority of endings of the A-Form can be divided into two groups:

- endings with strong stress, which occur in native words;
- endings with weak stress, which occur in foreign words.

Delving deeper into Jusuf's analysis, we can demonstrate this system in the table below:

|             | STRONG STRESS |    |               |       | WEAK STRESS  |  |               |      |
|-------------|---------------|----|---------------|-------|--------------|--|---------------|------|
|             | native words  |    | foreign words |       | native words |  | foreign words |      |
|             | A-Form        |    |               |       |              |  |               |      |
| masc. sing. | -o            |    |               |       |              |  | -o            | -i   |
| masc. pl.   | -e            | -a |               |       |              |  | -ă            | -ură |
| fem. sing.  | -i            |    |               |       | (-a)         |  | -a            | (-i) |
| fem. pl.    | ă             | -a |               |       |              |  | -e            |      |
|             | B-Form        |    |               |       |              |  |               |      |
| masc. sing. | -es           |    | -os           | is    |              |  |               |      |
| masc. pl.   | -en           |    | -ěn           | -onen |              |  |               |      |
| fem. sing.  | -ă            | -a | -a            |       |              |  |               |      |
| fem. pl.    | -ě            |    | -en           |       |              |  |               |      |

Notes:

1) Rare forms are indicated in parentheses, as in the weak ending -a in the native lexicon (this occurs only in the word *lîndra*) as

well as the weak feminine ending -i (this is encountered in very few words such as kakàvi.

2) In most cases the endings -os, -is, and -onen (which correspond to the weak endings -o, -i, and -urã respectively in the A-Form) are used with a postposition. However these are widespread without a postposition especially in the Mahazeri dialect of Prishtina.

It is necessary to emphasize that these postpositions are always enclitic. Since they are connected only to the B-Form, which is oxytone, the syllable which directly preceded the postposition is always stressed.

#### STRESS IN CONJUGATED WORDS

Stress in the verbal system is built on a completely different principle from that of nouns. Nevertheless, while using and expanding Jusuf's classification, we can discern a series of fundamental endings, which we will call "monemes" (the value of these monemes will be discussed in more detail in the section devoted to morphology of the verb classes):

a) present tense pronominal monemes:

a': -av, -es, -el, -as, -en, en  
a": -v, -s, -l, -s, -n, -n

b) past tense pronominal monemes:

In the active voice:

b': -em/-om/-um, -an, -a(s), -am, -en, -e

In the passive voice:

b": -em/-om/-um, -an, -o (fem. -i), -am, -en, -e

c) present passive moneme: -öv-, -üv-, or -iv-

d) past active moneme: -d- or -l-, or rarely -t-

e) past passive monemes:

e': -nd-  
e": -il-

f) derivational moneme: -d- (which apparently derives from the root d- «to give»)

g) foreign active moneme: -in- and isar-

- h) foreign active moneme: **-iz-**
- i) foreign passive moneme: **-isa-**
- j) past tense moneme: **-j-** which appears in a few specific words
- v) **-v-** link

We add to these monemes a series of enclitics: **-a, -sas, -na, -ne, -sine** with their variants, as well as the proclitic **ka-**, and naturally also the verb root itself. From the viewpoint of stress verb roots can be divide into stress-bearing and non-stress-bearing. All stress-bearing roots have a stressed à, belong<sup>2</sup> to group A of verbs, and are naturally verbal (see Morphology of the Verb), while the others represent no particular form and are of various derivations (verbal, nominal, or adjectival). Non-stress-bearing roots are marked as (Rr), while stress-bearing roots are marked are (Rrà).

Stress in the verbal stem can be summarized by the following statements:

According to Jusuf:

- the weak monemes (a"), (c), (d), (g), (i) and (v), and naturally the proclitics and enclitics never take the stress, and never affect where it falls;
- the strong monemes: (a') and (b), as well as the verb root (Rrà) are stressed;
- the monemes which throw the stress back onto the preceding syllable are (e'), (e"), (f), (h) and (j).

When two or three monemes meet in one verbal form, it occasionally occurs that the stresses associated with the monemes compete with each other. We point out these cases below:

1) when two strong monemes meet: This is impossible.

2) when a strong moneme meets a moneme which throws the stress back onto the proceeding syllable: In this case, depending on the dialect, we have two possibilities for the place of stress, i.e. stress falls either immediately before the "throwing" moneme, or on the strong moneme. Some examples are:

- (e')+(b): **mar-d-ilem**
- (h)+(a'): **misl-iz-av**
- (j)+(b): **darà-j-em**
- (f)+(a'): **phùr-d-av**
- (f)+(in-d)+b: **phùr-d-ind-em**

Note: In this form, the first **d** is a derivational moneme, whereas the second **d** is a past tense active moneme. In all these cases,

the first stress is recommended for the Common Language: mårdilem, mìslizav, daràjem, rùjem, phùrdav, kìdinem, etc.

3) when three or four monemes meet: this due to the sum of the stresses. For example, if we have (Rrà)+(e")+(e')+b, three monemes support the stress on the final syllable of the root against only one moneme for the bearer, then the final vowel of the root is stressed, e.g. daràndilom. This is the stress that occurs in almost all dialects and is preserved in the Common Language. In some cases the logical sum of the stresses yields no result, as for example in the structure: (Rr)+(f)+(i)-(e')+(b), where (f) throws the stress onto (Rr), (e') onto (i) and (b) is stressed, e.g. in ki-d-isa-il-om. In this case the stress is accepted by the central element, e.g. kidisàlom or kidisàj lom, or even kidisàlom as a result of reduction.

We have introduced here only the basic, regular system. In various dialects we encounter different patterns. Some interesting forms are found in a group of dialects which include, among others, the Bugurzi of Obilić and the Drindari of Bulgaria. In these dialects the root stress appears to be stronger than that of moneme (a), and as a result causes syncope of the e in the 2nd and 3rd persons singular and plural: phìrsa, phìrla, phìrna for phìresa, phìrela, phìrena «you go (sing.)», «he goes», «you go (pl.)», «they go» (-a is enclitic).

#### STRESS IN INVARIABLE WORDS

This is not the place to analyze the stress system of these words. Suffice it to mention the main elements.

- As stated above, postpositions are always enclitic and the syllable to which it is affixed is always stressed.

- Prepositions containing a vowel, have their own stress. When they are bisyllabic like kotar «from» or andar «about», they are oxytone. An article which follows forms one stress group with it, like an enclitic, which is indicated in writing by means of a dash: o baro foro «the big city», e purane gilā «the ancient songs», but kotar-o baro foro «from the big city», andar-e purane gilā «about the ancient songs».

- Adverbs formed from a noun in an archaic case like khere «at home» (locative case) or kheral «from home» (ablative case), are oxytone.

#### CONCLUSION AND PRACTICAL APPLICATION

We see that in the majority of cases the stress falls on the final syllable, therefore there is no need to indicate it in writing. However, we will offer the following rules below:



- when the stress is oxytone, it is not indicated. It can be assumed that a word without a stress mark is oxytone.

- when the stress does not fall on the final syllable it is indicated with a grave accent above the vowel: à, è, ì, ò, ù.

In addition, the presence of a postpositions indicates that the syllable preceding it is stressed.

Of course the stress systems of contiguous languages have affected, to various degrees, the stress system in Romani. What has been presented above is more or less unaffected by other languages.

In some areas the foreign influence has been so strong that the Romani stress system has almost completely disappeared, e.g. in Poland, where most dialects follow the paroxytone pattern of Polish, or in Czechoslovakia, where the stress is carried on the initial syllable. These stress patterns are not recommended, as they are foreign.

The opposition of long and short vowels, which has appeared in some dialects due to foreign influence (especially Hungarian), is also not recommended.

## CONSONANTS

The Romani consonantal system can be divided, for the sake of ease, into several subsystems, which we will examine each in turn.

- the subsystem of occlusives, which are also of special interest, in the case of the velars (k kh g) belong to this group;
- the subsystem of laterals, which can be considered to be the center of the whole system;
- the subsystem of sibilants, which, along with some other features, is partly the basis for distinguishing between the two large groups of dialects, will be shown in the section on morphology.

Finally, we will examine the microsystem of r (r and rr) and j.

### SUBSYSTEM OF OCCLUSIVES

This subsystem agrees widely with the subsystem of occlusives in Albanian: in addition to the oral and voiceless consonants p t k and the voiced b d g, we must add the nasals m and n. The latter can be realized in several variants, the main of which are [n] and [ɲ]. The realization [ɲ] is a positional variant before k and g and does not constitute a separate phoneme in the Common Language.

A characteristic feature of Romani is the presence of a series of aspirates ph th kh (a fact which demonstrates the relationship with Indic languages), as well as the presence of the aspirated affricate čh, which is pronounced like [čʰ], in the Balkan or non-Vlah group of dialects. In contrast with the Indic languages, Romani has only voiceless aspirates and does not exhibit the voiced series of aspirates (bh dh, etc.).

The treatment of the velar occlusives (k kh g) varies from dialect to dialect. In the Northern or Vlah branch, they are generally not palatalized and are pronounced as [k] [kh] [g] in every position. However in the Balkan group many dialects more or less palatalize velars after front vowels, while other dialects retain a central pronunciation, regardless of the phonetic environment. Consequently we can distinguish "palatalizing" and "non-palatalizing" dialects, as we have seen in the first part of the article (Koha nr 1986/1 : Description of phonemes, p. 87 and the manner of writing, p. 97).

According to this article, non-palatalizing dialects always pronounce these phonemes as [k] [kh] and [g], whereas palatalizing dialects behave similarly to some Romance languages,

e.g. Italian in the words capo, cel, ciglio, corone, gatta, giro, gufo, etc., or Romanian in the words cartea, celul, cine, coroana, galben, gema, gigea, etc.

Although we have given some examples only for velars in initial position, these rules clearly apply in any position. In the middle of a word, we must add also the pre-jotizing vowels which cause palatalization of the preceding consonant with various realizations, depending on the dialect (this phenomenon is explained below):

- in non-palatalizing dialects these are pronounced, as a rule, as a pure vowel with a [j] preceding it: [ja] [jo], etc., without altering the value of the consonant which it follows;

- in palatalizing dialects they are pronounced as a pure vowel, and the pre-jotization produces a form of palatalization of the consonant which it follows, e.g.:

| non-palatalizing dialect |          | palatalizing dialect |
|--------------------------|----------|----------------------|
| šukōl «it dries»         | [šukjol] | [šuk'ol, šu'ol]      |

Occasionally the palatalization also affects dentals, as in the Mahazeri dialect of Prishtina:

|                    |          |                      |
|--------------------|----------|----------------------|
| matōl «gets drunk» | [matjol] | In Mahazeri: [ma'ol] |
| butā «work»        | [butja]  | [bu'a]               |

In Skopje we even notice confusion of palatalized dentals with velars, under the influence of a similar development in Macedonian, i.e. [tj] has yielded [k'], as in [buk'a] for butā.

In this manner, the letters ḳ kḥ g̣ ṭ tḥ ḍ are common to all dialects, however each dialect pronounces them in its own way.

It is necessary to emphasize that the distinction between palatalizing and non-palatalizing dialects is not absolute, neither from the standpoint of the realization of phonemes, nor from the standpoint of their distribution within the word.

From the phonetic standpoint, palatalization of velars can be weak, while producing sounds comparable to those in Southern Tosk in k'itet, g'ethe, or in the Macedonian кибрит, гинее. But it can also be very strong, producing sounds barely distinguishable (Mečkari) or undistinguishable (Tharegoni) from palatal affricates.

In some dialects, in every case where a palatalizing vowel follows a velar within the word, the velar is palatalized (Palatalization type III, fairly widespread).

In other dialects, we must consider also the following question: Does the combination of velar + palatalizing vowel occur within a morpheme or between morphemes? This difference plays an important role for the reason that palatalization appears in some dialects only within a morpheme and not between morphemes (Palatalization type I, a potential type, which has not yet been proven), or even, as in Mahazeri of Prishtina, palatalization appears in intramorphemic combinations in nouns but not in verbs (intermediate type II). In this case we are dealing with a kind of intramorphemic sandhi:

|                           | Non-palatalizing type | Palatalizing type |      |      |
|---------------------------|-----------------------|-------------------|------|------|
| intramorphemic            |                       | I                 | II   | III  |
| <b>kher</b> «house»       | [kh]                  | ([tʰ])            | [tʰ] | [tʰ] |
| intermorphemic            |                       |                   |      |      |
| in nouns:                 |                       |                   |      |      |
| <b>jakh-en-</b> «eyes»    | [kh]                  | ([kh])            | [tʰ] | [tʰ] |
| in verbs:                 |                       |                   |      |      |
| <b>dikh-en</b> «they see» | [kh]                  | ([kh])            | [kh] | [tʰ] |

In other words, we can say that dialects of type II treat combination such as nominal root + plural ending of the B-Form as intramorphemic (**jakh-en-**), whereas the combination verbal root + active ending is treated as intermorphemic (**dikh-en**). Much research into dialectology is necessary in order to gain a clear perspective on this phenomenon.

#### SUBSYSTEM OF LATERALS

One important question in Romani phonology is that of the number and status of laterals. There are two realizations which can be represented by [l] (as Albanian l) and [L] (as in Albanian ll), but this raises the question, whether these are indeed two phonemes or two variants of the same phoneme. This question is quite complicated, but, since it is very important for the language, it merits thorough consideration. We have devoted a thorough study to this question which was published in the journal Lingua posnaniensis, no. 31, in Poland, under the title of "Les latérales en romanès." The question is treated there in detail, and in interdialectal or polydialectal manner, and in connection with neighboring languages. It is sufficient here to present a brief overview, while particularly emphasizing the results of the study.

## I) Definition of the System of Laterals

In order to define the subsystem of laterals, we have elaborated the following:

1) We notice that Romani laterals are pronounced in different ways, not only according to dialect, but also within one dialect. Their pronunciations varies from a hard sound like the Albanian **ll** (even occasionally like Albanian **dh** or Welsh **ll**) to a fairly soft sound like **-l<sup>h</sup>**.

2) In order to identify whether or not these sounds are proper phonemes, we apply the method of minimal pairs, i.e., we look for pairs of words which are identical except for one sound, which is in the first case sound A and in the second case sound B. If the difference in this sound results in a difference in meaning, then this shows that the sounds are distinct phonemes.

For example, in Albanian we notice that **k** and **q** are phonemes, since the pairs **shkuar/shquar**, **ke/qe**, **kap/qap**, **kind/qind**, **shok/shoq**, **bujk/bujq**, **fik/fiq**, **teket/teqet** have different meanings, although there are cases where a difference in the phonetics does not cause a difference in meaning, e.g. **shkelm = shqelm**.

At this stage, we must distinguish two phonemes, which can be designated **/l/** for the soft sound, and **/L/** for the hard sound. These phonemes can be identified by means of the following pairs:

|           |          |           |               |
|-----------|----------|-----------|---------------|
| [liLa]    | «letter» | [lila]    | «he took»     |
| [kaLo]    | «black»  | [kalo]    | «blacken!»    |
| [nasavLo] | «ill»    | [nasavlo] | «become ill!» |
| [LoLo]    | «red»    | [Lolo]    | «become red!» |

3) We discover that this opposition does not occur in all dialects, and except for [liLa/lila], this opposition appears to be quite artificial. They are not very natural, and are not used in every-day speech, unlike the examples given to illustrate the opposition of **k** and **q** in Albanian. This shows that the question needs to be examined more thoroughly, for the reason that a contrast that appears fairly rarely does not have a oppositional function, but rather derives from some other phenomenon.

4) We then look for environments where [l] and [L] appear. Can they appear in one and the same environment regularly or only exceptionally, as in the above examples? If they appear in the same environment regularly, then we can be sure that they are proper phonemes, and that rarity of minimal pairs is purely accidental. On the other hand, i.e., if they are rarely encountered in the same environment, but appear regularly in a specific environment, then we are then obliged to examine this question.

a) In initial position, we have:

| before a front vowel |                | before a non-front vowel |          |
|----------------------|----------------|--------------------------|----------|
| l:                   | [liL] «letter» | [Lon]                    | «salt»   |
|                      | [len] «river»  | [LuLudi]                 | «flower» |

b) In syllable-final position, we have /L/, independently of the vowel which occurs before it, or whether it occurs in absolute final position:

|    |         |           |
|----|---------|-----------|
| L: | [phraL] | «brother» |
|    | [khiL]  | «butter»  |
|    | [deveL] | «God»     |
|    | [moL]   | «wine»    |

non-absolute finals:

|           |            |
|-----------|------------|
| [baLvaL]  | «wind»     |
| [kheLdem] | «I jumped» |

c) inside the word, at the beginning of a syllable, we usually have:

| before front vowels |                       | before a non-front vowel |                        |
|---------------------|-----------------------|--------------------------|------------------------|
| l:                  | [kale] «black» (fem.) | L:                       | [kaLo] «black» (masc.) |
|                     | [Lolipa] «redness»    |                          | [LuLudi] «flower»      |
|                     | [balesa] «with hair»  |                          | [baLa] «hair(s)»       |

But here we also have the cases mentioned above:

| before front vowels |                   | before non-front vowels |                            |
|---------------------|-------------------|-------------------------|----------------------------|
| L:                  | [hamaLi] «porter» | l:                      | [kalarav] «I blacken»      |
|                     |                   |                         | [rakla] «non-Rom girls»    |
|                     |                   |                         | [thuloL] «he gains weight» |

5) We discover that, in general, [l] and [L] occur regularly in different environments ([l] before front vowels, and [L] before non-front vowels), and even appear to be alternates throughout declensions depending on the quality of the vowel of the ending: [kaLo], but [kale], and [baLa], but [balesa].

In contrast, these appear quite rarely in the same context: -[L]- is encountered quite rarely before a front vowel, and then always in a foreign lexeme which has not become assimilated into the language, e.g. [hamaLi]. Therefore it can be stated truthfully that this case does occur at all in this language.

-[l]- is encountered relatively frequently after a non-front vowel, but before a non-front vowel in only a limited number of cases. The syllable-initial position before a non-front vowel, but not at the beginning of a word, is the sole environment where both [L] and [l] can appear, therefore this case must be carefully examined.

6) There is no need to analyze cases where an [L] is followed by a non-front vowel, since this is natural according to the above discussion. On the other hand, we must examine the cases where [l] is followed by a non-front vowel. It can immediately be seen that an [l] which is soft before a, o, or u is encountered only in a few specific, clearly morphological, cases.

I) before the causative suffix -ar- :

from [kaLo] «black» is formed [kalarav] «I blacken»

II) before the passive suffix -ov- :

from [pašLo] «laid» is formed [pašlovav] «I lie down»

III) before an active ending in the past tense. We can give as an example the verb **dikhel** «he sees» in this tense:

|           |               |           |                 |
|-----------|---------------|-----------|-----------------|
| [dikhlum] | «I saw»       | [dikhlam] | «we saw»        |
| [dikhlan] | «you saw»     | [dikhlen] | «you (pl.) saw» |
| [dikhla]  | «he, she saw» | [dikhle]  | «they saw»      |

IV) in some feminine plurals which have a stressed -li in the singular:

from [rakli] the plural is [rakla]

V) in two isolated lexemes: [gilabel] «he sings» and [kholanel] «he angers».

The verbal forms [xulav] «I descend» and [uklav] «I ascend» do not apply here, for they are contrasted with [xulovav] and [uklovav], therefore they belong with the passive forms (class II).

VI) a small number of foreign words, like [filan] «so and so».

7) Our analysis requires us to compare these derivations with parallel derivations, but from a root that does not end in a lateral:

I) [kaLo / kalar-av] [san-o / sanjar-av] «thin» / «I thin»

II) [pašL-o / pašlov-av] [san-o / sanjov-av] «I become thin»

III) [dikhlum] [śundjum, śunġum] «I heard»  
 IV) [rakl-i / rakl-a] [bor-i / bor-ja] «bride» / «brides»

These comparisons cannot be made naturally in classes V and VI, however those that have been made up to now suffice to confirm that we are dealing with a change in the sole lateral L in its soft variant under the influence of a j or, in other words, a pre-jotizing vowel: -ja-, -jo-, or -ju-. This is also confirmed by the plurals of class IV, which in many dialects has in some cases retained a full form of the type: [rakl-ia], either due to an archaism or a gender distinction in the B-Form:

[rakl-ien] «girls» ≠ [rakl-en] «boys», while many Balkan dialects have the common form: [rakl-en].

It is interesting to compare these feminine plurals with the corresponding Indic forms, like, e.g. Sanskrit or in a neo-Indic language (Hindi):

|          | singular                                    | plural   |
|----------|---|--|
|          | -ī  | -īas [-īah]  |
| Sanskrit | devī «divine»<br>nadī «river»               | devī-as<br>nadī-as   |
| Hindi    | -ī  | -iā  |
|          | ĉhurī «knife»<br>bakrī «goat»               | ĉhuriā<br>bakriā   |
| Romani   | -i  | -ia -ja -a   |
|          | ĉhuri «knife»<br>(and ĉurik)<br>bakri «ewe» | ĉhuria, ĉhurā [-rja] or [-jra]<br>(and ĉhurika, ĉhura)<br>bakria, bakrā, bakra |

The change in the consonant before a pre-jotizing vowel is not limited to laterals, but also affects the following:

- occlusive velars (in palatalizing dialects): from [śuk-o] «dry» is formed [śuĉar-av] «I dry» and [śuĉov-av] «I become dried».

- occlusive dentals (in Mahazeri): from [mat-o] «drunk» is formed [maĉar-av] «I make drunk» and [maĉov-av] «I become drunk»; from [but-i] «work» is formed the plural [buĉa] alongside [butja] in other dialects. In the same way [morthi] «skin», and the plural [mortha] alongside [morthja].

The treatment of consonants before pre-jotizing vowels should not surprise us, for we are reminded of the notes by the French linguist Andre Mejes, who wrote before the fifties: "In order to



conform to the rules for consonants j tends to be written into the pronunciation of the consonant which precedes it, so that frequently it loses its own existence while fundamentally changing this consonant." Thus we see that in Romani we only have a special application of a general rule.

## II) Conclusion of the Practical Application

It can be seen here that we are dealing with one lateral phoneme, which is realized as a soft [l] before front and pre-jotizing vowels, and a hard [L] in all other positions.

In general, a special letter (or a diacritical mark used with the lateral) is not needed, for each realization, since the position itself shows how the lateral is pronounced.

- when followed by an i or e, it is pronounced soft;
- when followed by an a, o, u, or a consonant, or when at the end of a word, it is pronounced hard.

There remain cases where the lateral is joined with a pre-jotizing vowel. The writing of j in this case is not a satisfactory solution for the following reasons:

1) The true realization as [j + vowel] is very rare, and in the dialects where it does actually occur it does not occur in every position.

2) Frequently it is realized as [ɛa], especially in Romania, and does not agree with the notation [ja].

3) In Yugoslavia, where the Rom are considerable in number and contribute greatly to emancipation and the affirmation of Rom culture, the reader who has learned to read and write Serbo-Croatian is accustomed to the existence in this language of two phonemes (with a phonological opposition in Serbo-Croatian), which are written as l (or л) and lj (or љ). However it would be illusive to imagine that the reader will be able disassociate himself from the laws he has acquired in relation to the writing of Serbo-Croatian, and surely the combination -lj- in writing would firmly suggest, despite all our assurances to the contrary, that Romani also has such a phoneme. With such a misleading impression, the reader will tend to write this pseudo-phoneme wherever he hears a soft l, which making phonological errors such as: \*ljil for lil «letter», \*ljen for len «luck», etc.

Furthermore, this reader will tend to mark each palatalization he hears with j, creating such absurd forms as \*gjilji for gili, or khjeljel, kjheljel, or kjhjeljel for khelel in the presence of palatalization before i and e, as we have see above. Not only are these forms appearing in writing bad, they are also ugly and inaccurate, for they introduce into the writing irrelevant types

of pronunciations, and, while emphasizing non-semantic dialectal differences, they consequently work against the unity of the written language, something which could cause great harm to its development.

Thus, only the pre-jotizing vowel needs to be supplied with the sign of pre-jotization, which is quite logical. A mark, called a *ćiriklo* «bird» in Romani, is used for this purpose: *ă*, *ǫ*, and *ũ* and applying it to the cases discussed above, we have:

I) *kalărav* (and *thulărav* «I fatten», *paślărav* «I spread», etc);  
II) *paślǫvav* (or *suslǫvav* «I get wet», *śutalǫvav* «I become sour», etc.);

III) *dikhlŭm*, *dikhlăn*, etc., in all dialects that have *-l-* as a moneme in the simple past tense;

IV) *raklă* (and *berlă* «wasps», *rovlă* «sticks», *belă* «rods», etc.

V) *gilăbav* and *xolănel*.

As far as foreign words are concerned, these are being replaced with Romani words, e.g. *filan* with *đumano*, *kotaruno*, or *ilaç* «medicine» with *đrab*, etc.

In dialects which do not exhibit this phenomenon, the pre-jotizing sign is not used: *kalarav*, *paşlovav*, *dikhlom*, *dikhlan*, *gilabav*, etc. In the majority of forms such as *raklă*, all the dialects we have discussed so far exhibit pre-jotization.

The use of this sign naturally extends beyond the cases of contact with a lateral, and is always written in the following morphemes: in the suffixes *-ăr-* and *-ǫv-*, in the endings for the active past tense, and in the plural *-ă*, etc. In this way the question of the change in some occlusives, a phenomenon we have mentioned above, is also solved. These are thus written: *śukărav*, *śukǫvav*, *matărav*, *matǫvav*, *bută*, *morthă*, etc.

Consequently we will use the sign *˘* also above *i* and *e*, usually not in connection with laterals, but rather with the dental occlusives *t*, *th*, and *đ*, in cases where they are palatalized:

- either before some special occurrences of *i-*, e.g. [*đive*] «day» in Gurbeti, in contrast to [*dikhav*] «I see» in the same dialect. We thus write: *đive*;

- or before some endings with *-e* in the *-a* paradigm, e.g.

[*phirđe*] «they walked» alongside [*phirđa*] «he walked» - then we write *phirđe* to contrast *phirđă*.

[buʎen-] «work» (B-Form) alongside [buʎa] «work» (A-Form) - then we write *butěn*] in contrast to *bută*.

We have analyzed fairly superficially the question of Romăni laterals, for the reason that it constitutes a central question in this language and, at the same time, pervades the entire phonological system.

The question of laterals is also interesting from the dichronic point of view. We have given an dichronic overview of the development of l in the article "Gjuha rrome, fonetika, fonologjia, shkrimi," published in the journal *Përparimi* 1984/4. This shows that we are dealing with a variational distinction which probably arose in Turkey and developed further towards an opposition of phoneme/phoneme under the influence of rephonologization undergone by several Balkan dialects at the same time. Thus the interest in the question transcends the domain of Romăni studies, and the comparative analysis of the systems of laterals in the dialects of the Balkan languages (not of the standard forms) can contribute many useful elements to Balkan studies. Furthermore, this phenomenon also has special importance for phonological theory, for this shows that the distinction between oppositional relations (phoneme/phoneme - a semantic contrast) and variant relations ((variant/variant - a non-semantic variant) in many cases is not clear-cut. It cannot be established directly except for in a few minimal pairs, especially when we take into consideration the language in its totality and in its social function, that is to say, in its natural situation. We are dealing here, for example, with the appearance of a opposition which has the further tendency to develop or to disappear, depending on the influence of the phonological systems of the neighboring languages. The written form with the haček is supported by the deep structure of the Romăni language, and thus does not only express its essence, but is also a situation which unites the various realizations into one common written form without distorting the special pronunciations of the dialects.

## SUBSYSTEM OF SIBILANTS

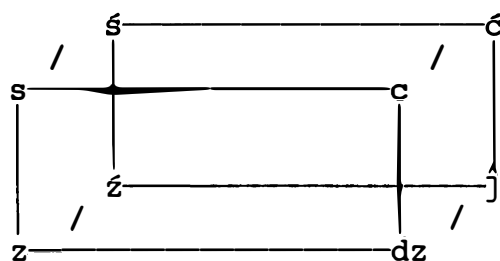
As we have seen above, sibilants constitute their own system. This system differs fundamentally between the two groups of dialects, and is in fact an important feature for classifying the dialects.

1) The system of sibilants in the Balkan group of Romani dialects:

This system agrees almost completely with the system in Albanian, so a description here is superfluous. It suffices for us to show the correspondences graphically:

| Balkan Romani            | Albanian  | Translator's transcription <sup>1</sup> |
|--------------------------|-----------|---|
| f      v                 | f      v  | f      v                                |
| s      z                 | s      z  | s      z                                |
| c      (dz) <sup>2</sup> | c      x  | c      dz                               |
| š      ž                 | sh    zh  | š      ž                                |
| č      ž                 | ç      xh | č      j                                |

We must add the phoneme *h*, which is pronounced as the standard Albanian *h*, and the letter *x*, which is pronounced like the *ch* in German *rauchen*, or the *ch* in Scottish English *loch* or as the Albanian *th* in the speech of the Tirana's villages. The group of phonemes which correspond to the letters *s z c x sh zh ç* and *zh* in Albanian are especially interesting. They form a very coherent system which can be represented according to their three features: compactness, voicing, and affricateness) in the following chart:



The top surface represents the feature [voicing-], while the lower surface represents [voicing+].

---

<sup>1</sup>Please see Translator's introduction.

<sup>2</sup>The letter  $\varpi$  originally used for this sound has been abandoned in the current orthography. [Translator]

The front face represents the feature [compactness+], while the back face represents [compactness-].

The right-hand side represents the feature [affricateness+], while the left-hand side represents [affricateness-].

The phonological features are reflected in a consistent manner by means of graphic features: [voicing-] is marked in writing in a circular form (s ś c ć), while [voicing+] is marked in writing by means of a broken line (z ź dz ź). The feature [compactness+] is marked by the absence of an accent mark above (s z c dz), while [compactness-] is marked with an accent above (ś ź ć ź). [Affricateness+] is indicated by means of letters which include only open space, while [affricateness-] is marked by means of letters with two open spaces within.

In fact, the sound [dz], which is designated by means of x in Albanian, is quite rare in Romani. Nevertheless, not only does it satisfy the symmetry of the system, but is also encountered in a few Romani words (beside foreign words), e.g. handza «a hole in bread or in a wall», or dzindzōla «stretches out».

2) The system of sibilants in the northern branch of Romani dialects.

While the Balkan system can be defined as a 2x2x2 system which can be visualized in three dimensions, the northern system presents only two dimensions, each having three grades:

- instead of two positions of pronunciation, we have three positions which can be represented by means of the symbols: s z c, s' z' c', ś ź ć.

The series of voiced affricates is completely lost (except for in a few rare cases which have been caused by some paradoxical foreign words), and the phonemes which remain for each position of pronunciation can be arranged as follows:

| row I | row II | row III |
|-------|--------|---------|
| s     | s'     | ś       |
| z     | z'     | ź       |
| c     | c'     | ć       |

However, this fundamental difference creates a great obstacle to the unification of Romani writing. Thus we must further analyze the question of divergence among the Balkan and northern dialects.

We notice that some phonemes agree completely among the dialect groups, such as, e.g. s (sivel «he sews» in the two groups) c (cidel, cadel, cadel) «he throws» according to dialect, but

always with c), **ś** (**śaj**) «to be able» in the two groups), and many other cases.

Only two pairs of phonemes differ systematically. The phonemes **čh** and **z** of the Balkan group always correspond to **s'** and **z'** in the Northern branch as result of the mutation of affricates which the northern branch of dialects has undergone.

| Balkan group | Northern group |
|--------------|----------------|
| <b>čhavo</b> | <b>s'avo</b>   |
| <b>mačho</b> | <b>mas'o</b>   |
| <b>zanel</b> | <b>z'anel</b>  |
| <b>gazo</b>  | <b>gaz'o</b>   |

This rule has no exceptions. When comparing the two systems side by side one particular element nevertheless remains: **c'**, but this element immediately enters the system when we see that it is a northern realization of **t** (and sometimes of **k**) followed by a pre-jotizing vowel:

| Balkan group   | Northern group  |
|----------------|-----------------|
| <b>matărel</b> | <b>mac'arel</b> |
| <b>tõ</b>      | <b>c'o</b>      |

The sound **dz** does not appear in the Northern group except for in a few words in the dialect of the Ukraine, where it is only a variant of the phoneme **z**: **dzen** «saddle», **dzet** or **zet** «oil».

There is thus no obstacle for the realization 2x2x2 or 3x3 to be written in a common manner. The basis is the Balkan group which is more archaic and displays a deep level of the language. We apply two systems of rules of pronunciation to this structure, one system for the Balkan group (the graphemes **čh** and **z** are pronounced like [**çh** (i.e. **čh**)] and [**xh** (i.e. **ĵ**)] in Albanian), and another system for the northern group (the graphemes **čh** and **z** are pronounced like [**sh** (i.e. **ś**)] and [**zh** (i.e. **ĵ**)] in Albanian). The rules for the pronunciation of **t** followed by a pre-jotizing vowel will be given in the respective chapter.

#### THE MICROSYSTEM OF **r**

A number of dialects have only one kind of **r**, which is pronounced like the **r** in Albanian, while from the **r** (which appears in all positions of the word, e.g. **rovel** «he cries», **phirel** «he walks», **dur** «far») some dialects distinguish another **r**, which is realized in a different manner, depending on dialect: a strong [**ṛ**], a velar [**R**] like in modern German or French, a retroflex [**ɻ**], as in Indian languages, [**x**], [**ɣ**], nasal [**ɹ̃**] or a pure [**r**] with nasalization of the preceding vowel. In addition it can be pronounced [**nr**], [**rn**], or [**ngr**].

It is not so much the manner of pronunciation which is important but rather the mere fact that two r's contrast one another. The pure [r] is written r, whereas to mark the second we use the digraph rr. Some authors have proposed the use of the diacritic ř (Kenrick) or r (Barthémély).

The lexemes which represent such an r in many dialect are realized in speech as a voiced and nasalized dental occlusive instead of an r. In such a case either rr or the phonetic form nd can be written.

The correspondence between r, rr, and nd is not yet completely clear for all dialects. Let us take, for example, the realization of the word marro «bread»:

[maño], [maRo], [maño], [maɣo], [maxo], [maño], [mâro] which are all written marro;

[maro], which is written maro;

[manro], [marno], [mando], which are written marro or phonetically as manro, marno, or mando (the rules are not yet fixed).

The opposition r ≠ rr is relevant in many dialects, as the following examples illustrate: **ćoripe** «theft» versus **ćorripe** «poverty», **bar** «stone» versus **barr** «fence», etc. Dialects which have lost this opposition preserve the lexical distinction by other means: **ćoripe** «theft» versus **ćororipe** or **ćorolipe** «poverty», **bar** «stone» versus **bur** «fence».

In the Common language the use of the greatest possible distinction between the two r's is advised, regardless of the concrete pronunciation of rr.

#### THE MICROSYSTEM OF j

We must discern the features of j in different positions:

- in word-initial position (=jV-) appears before the vowels a (jag «fire», jakh «eye», javer «other», or more frequently aver), e (jekh «one», more rarely ekh), o (jov «he», or vov and more frequently ov), and i (jiv «snow», or more frequently iv, jilo, or ilo or more frequently vilo). In words in which it is stable in all or almost all dialects it is advisable to write j: jag, jakha, jekh. When it does not appear (this occurs in many dialects: ov, iv and aver are more prevalent than jov, jiv, and javer), then it is advisable to not write it. It does not occur before u.

- word-final position (-Vj=): j does not present a problem and is written in all cases: *daj* «mother», *šej* «thing», *šošoj* «hare», *muj* «mouth». It does not occur before *i*.

- in word-internal position between vowels (-VjV-): j is written except when one of the vowels is *i*, for the reason that in these cases *j* has no phonological value: *muja* «mouths», *roja* «spoons», *akaja* «this», etc., but plain «mountain», *komsia* «neighbor», *odia* «that».

- after a vowel and before a consonant (-VjC-): this must be examined to verify that this does not arise from the metathesis of a jot from a pre-jotizing vowel in the following syllable, e.g. [*bojra*] for [*borã*] (see above). In this case it is written into the respective syllable by means of a haček. Otherwise, it is written *j*, but such cases are rare. This occurs for the most part in foreign words: *hãjde* «Let's go!» or as a variant of *i*: *nasvãjlo* or *nasvãilo* «ill», *thoimasqe* or *thojmasqe* «to wash», etc. These can then be written either *j* or *i*.

- after a consonant and before a vowel (-CjV-): in most cases we are dealing with a pre-jotizing vowel, and the sound *j* is then written with a haček, as we have seen above. In other cases, especially in foreign words, the use of the grapheme *i* to represent [j] is recommended for the following reasons:

- in most cases there is no relevant opposition between [i] and [j] in this position since [i] and [j] are free variants in foreign words as well as in native words (*obièkta* or *objèkta*, *romnia* or *romnã*, *deriav*, *derjav*, *derãv* or *derav* «sea» a very old foreign word from the Persian *dariã*).

- Likewise the contact of the grapheme *j* after a consonant grapheme is also avoided, because in many non-Rom traditions this contact constitutes a digraph with a different value, and this could confuse the Rom in reading and writing. See above for the undesirable consequences of the use of the Serbo-Croatian digraph - *lj*- in Romani. Just as undesirable is the appearance of the combination -*gj*-, which is a digraph in Albanian but not in Romani, as are also the combinations: -*dj*-, -*nj*-, etc.



## THE PROBLEM OF DIVIDING LEXEMES INTO PARTS OF SPEECH

The description of morphology implies the division of linguistic material into parts of speech. The concept of the definition "word" and "parts of speech" has not yet found a satisfactory correspondence for all languages, and therefore this will be studied here only from the viewpoint of Romani. At first, conditions seem to simplify the task, since Romani is an Indo-European language, and the traditional division into parts of speech, which has been created on the basis of the study of Indo-European languages, will be suitable for Romani. Unfortunately, when we proceed from the division of nouns and verbs, we encounter some undeniable difficulties.

This does not mean to imply that the distinction between nouns and verbs does not exist in Romani. "If there is ever an opposition which is never violated, undoubtedly it is the distinction between verbs and nouns," so wrote Zhak Fëjè, but the dividing line between these two classes of words is not completely clear cut. The existence of true verbs, on the one hand, and true nouns, on the other, cannot be refuted, neither in Romani nor in the other Indo-European languages. Thus, for example **phirel** «he walks» is undoubtedly a verb, whereas **bakri** «ewe» is a noun. This is attested not only by the semantic criterion that **phirel** expresses an action, whereas **bakri** denote a thing, but also by the inflections of these two words, which differ completely from each other, not only in form, but also especially in their inflectional category. This criterion, being morphological, naturally has more force than the semantic criterion.

However, we must emphasize the difficulty of defining these two classes while examining some loopholes, as follows:

Let us conjugate, for example **phirel** «he walks» in the past tense (singular):

|        |                  |                                |                  |
|--------|------------------|--------------------------------|------------------|
| 1st p. | <b>phirdŭm</b>   | (or <b>phirdem, phirdum</b> )  | «I walked»       |
| 2nd p. | <b>phirdăn</b>   | (or <b>phirdan</b> )           | «You walked»     |
| 3rd p. | <b>phirdă(s)</b> | (or <b>phirda</b> , see above) | «He, she walked» |

These are true verbal forms which are encountered in many verbs (we are not considering the variants **-em** and **-om** in the 1st p.)

**kerel** «he does»

**dikhel** «he sees» conjugated:

|        |                 |               |                  |               |
|--------|-----------------|---------------|------------------|---------------|
| 1st p. | <b>kerdŭm</b>   | «I did»       | <b>dikhlŭm</b>   | «I saw»       |
| 2nd p. | <b>kerdăn</b>   | «You did»     | <b>dikhlăn</b>   | «You saw»     |
| 3rd p. | <b>kerdă(s)</b> | «He, she did» | <b>dikhlă(s)</b> | «He, she saw» |

On the other hand, verbs such as **bešel** «he sits» and **nakhel** «he passes» display the following past tense forms:

|        |               |           |                |              |
|--------|---------------|-----------|----------------|--------------|
| 1st p. | <b>bešlūm</b> | «I sat»   | <b>nakhlūm</b> | «I passed»   |
| 2nd p. | <b>bešlān</b> | «You sat» | <b>nakhlān</b> | «You passed» |
| 3rd p. | <b>bešlo</b>  | «He sat»  | <b>nakhlo</b>  | «He passed»  |
|        | <b>bešli</b>  | «She sat» | <b>nakhli</b>  | «She passed» |

We notice here not only a special ending, but also coincidence according to gender, exactly as in the case of the nouns: **bakro** «ram» and **bakri** «ewe». Adjectives also demonstrate this coincidence: **kalo** «black (masc.)» and **kali** (fem.). In any case, from a syntactical point of view we have a parallelism between the use of adjectives and forms such as **nakhlo**, **bešlo**, etc., since in many dialects there is no copula to join the predicative with the subject, e.g. **o čhavo tikno** «The child (is) small» and **o čhavo nakhlo** «The child passed by» (the question of the structure of types such as **o čhavo tikno tano** and **i čhavo nakhlo tano** will be discussed in the section devoted to syntax).

Of course, for all the formal similarity between the past tense with **-o/-i** and adjectives such as **kalo** and nouns such as **bakro**, the verb nevertheless can be distinguished for some of the following reasons: on the one hand, **bakro** and **kalo** are frequently encountered in the B-Form with or without a postposition (e.g. **bakres** «ram (direct object)», **kalen-çar** «with the black ..»). For an explanation of the terms A-Form/B-Form and postposition see below). On the other hand, the third person of the past tense of verbs of the type **nakhel** or **bešel** are found in the A-Form. This opposition seems satisfactory, but if we take a different lexeme, such as, e.g. **merel** «he dies», then we see that it occurs frequently in the B-Form with or without a postposition:

**merel** «he dies»

|        |              |            |
|--------|--------------|------------|
| 1st p. | <b>mulūm</b> | «I died»   |
| 2nd p. | <b>mulān</b> | «you died» |
| 3rd p. | <b>mulo</b>  | «he died»  |

A-Form **o mulo** «the dead man», B-Form **e mules** «the dead man (direct object)», **e mules-qe** «to the dead man», **e mulen-çar** «with the dead men», etc.

**muli** she died

A-Form **i muli** «the dead woman», B-Form **mulā** «the dead woman (direct object)», **e mulā** «by the dead woman», etc.

We also encounter, although rarely, but in a true grammatical fashion, the B-Form of lexemes of the type **našlo**, **pašlo**, etc.: **čordile e šeja e našlenqere** «They stole the things of the

fugitives», del dùm a pašlençar «He speaks with those who had laid down», etc. The rarity of these forms depends more on semantic or dialectal factors than grammatical ones. Along the path of these forms we can gradually move from the class of nouns towards verbs. There is a similar path, as, for example, in words formed by means of the ending -pe (or its variants -ba, -pa, etc.). These words retain many characteristics both of a noun and a verb.

We can represent the forms which we have analyzed above in a table, while arranging them in columns according to the traditional classification.

### PART 3: MORPHOLOGY

From the standpoint of combinatory morphology, it seems that we have two classes: that of conjugated words (verbs) and that of declinable words (nouns), which partially overlap and do not agree completely with the traditional classes. The question is complicated by the fact that some verbs agree with the two forms in the 3rd person past tense, e.g. *našel* «he flees» or *bašel* «it sounds» with their variants *našlã* beside *našl/o*, *-i* and *bašlã* beside *bašl/o*, *-i*, frequently in the same dialect. In fact, a number of central European dialects have lost the variant with *-o*, *-i* and form the 3rd person analogically with the ending *-ã* or *-a*: *mulã* «he, she dies», *bešlã* «he, she sat», etc.

In regards to the alternation of forms with *-o*, *-i* on one hand, and *-ã* on the other, Romani grammars state that some transitive verbs (and in some cases transitive verbs) are lacking the 3rd person past tense forms which are replaced with a participle. This interpretation is purely a practical formula, but naturally it need not be accepted as a linguistic explanation.

We have only raised here the question of the status of forms of the type *nakhl/o*, *-i*, *mul/o*, *-i*, etc., in order to emphasize the difficulty of categorically dividing words into parts of speech and the weakness of the traditional classification. Nevertheless, this does prove to be of practical use and will be applied here while keeping in mind the reservations we have expressed. "Words" will be placed together into three groups according to their morphology:

- declinable words (nouns in the broad sense of the word);
- conjugated words (verbs and copulae);
- invariable words.

Within this grouping we have the traditional classification.

MORPHOLOGY OF NOUNS  
The Category of the Noun Class in Romani

The most obvious type of word which is declined is the substantive itself, which in Romani has two genders (masculine and feminine), two numbers (singular and plural), and two cases (the direct and the oblique, which are called in Romani **A-čham** and **B-čham**, literally face A and face B. We will use the terms A-Form and B-Form.

Some authors speak incorrectly about the category of definiteness in Romani. However, from a practical stance it is better to say that the article, which is prepositive in Romani (as in Greek and English), should not be considered as a part of the noun, but as a distinct word, which will be discussed later.

As an example we can take the paradigm of the noun **bakro** «ram» and **balo** «boar» (variants are given in parentheses):

| BAKRO «ram», BAKRI «ewe» |           |        |                   |                                  |
|--------------------------|-----------|--------|-------------------|----------------------------------|
|                          | masculine |        | feminine          |                                  |
|                          | sing.     | pl.    | sing.             | pl.                              |
| A-Form                   | bakro     | bakre  | bakri             | bakria<br>(bakră)                |
| B-Form                   | bakre(s)  | bakren | bakria<br>(bakră) | bakrien<br>(bakriën)<br>(bakriă) |
| BALO «boar», BALI «sow»  |           |        |                   |                                  |
| A-Form                   | balo      | bale   | bali              | balia<br>(bală)                  |
| B-Form                   | bale(s)   | balen  | balia<br>(bală)   | balien<br>(balen)<br>(balăn)     |

1) The Category of Gender

Genuine nouns for which gender is a purely morphological category are rare, that is to say, a category for which a semantic opposition masculine/feminine is represented by means of an alternation in the ending. Gender is more frequently a lexical category, i.e. it is usually fixed as an inseparable attribute of the lexeme in the word. The number of lexemes with gender as a morphological category is relatively greater than in many other languages. Besides **bakro** and **balo**, which we have declined above, we can also mention **rikono** «male dog», **buzno** «billy goat», **zeno** «male person», **raklo** «non-Rom boy», **gazo** «non-Rom man», **khuro**

«colt», **ćiriklo** «male bird», **balićho** «male piglet», **kirvo** «godfather», **xelado** «male soldier», **phivlo** «widower», **rićhino** «male bear», **ćoxano** «male ghost», **gilamno** «male singer», etc. Such words have a feminine form which are formed regularly from the masculine form: **rikoni** «bitch», **buzni** «nanny goat», etc. The number of these groups grows if we examine neologisms with the meaning of "agent of an action" formed with the suffix **-no**, **-ni** according to the model **gilamno** «male singer», **gilamni** «female singer», such as **sikamno** «male teacher», **sikamni** «female teacher», **arakhno** «male guard», etc. These are considerably numerous if we add adjectives with a nominalized meaning: **dilo** (or **dilino**) «crazy man», **geralo** «bald man», **ćhorvalo** (or **ćhoralo**) «bearded man» and many more. Composites of the type **barebalenqo** «long-haired» or **kalejakhengo** «dark-eyed» form a rich class of adjectives with this feature.

There also exists a series of less regular changes which nouns undergo to reflect the concept of gender, by means of other suffixes, e.g. masc. **grast** «stallion», fem. **grasni** «mare», masc. **rrom** «Rom man», fem. **rromni** «Rom woman», masc. **ćhavo** «non-Rom boy», fem. **ćhaj** (variant **ćhej**) «non-Rom girl», masc. **dad** «father», fem. **daj** (variant **dej**) «mother». In many cases there is no formal connection between the two genders, e.g. masc. **bašno** «rooster», fem. **kaxni** (variant **khani**) «hen», masc. **zamutro** «groom», fem. **bori** «bride», etc. We find this at the boundry between morphology, word formation, and lexicology.

Besides these cases, Romani, like every language, has a large number of nouns for which the category of morphological gender has no semantic value, but is imposed without a synchronic motive, as, for example in **kham** «sun» (masc.), **diz** «city» (fem.), **pani** «water» (variant **paj**) «water» (masc.), **phuv** «land» (fem.), etc. These words are used only in one gender. There also exist a limited number of words, the gender of which varies from one dialect to another, as, for example **bokh** (variant **bukh**) «hunger» or **ćang** «leg», which are usually feminine, but which appear as masculine in some dialects in the Soviet Union. Rarely there are nouns, the gender of which varies within one dialect, e.g. **ambrol** «pear» in Mećkari.

## 2) The Category of Number

This category has a more morphological and less lexicological nature than that of gender. Almost all nouns have both a singular and plural form. The limitation in a noun to a specific number is usually for more or less semantic reasons. The words **kham** «sun» or **lindra** «sleep» are rarely used in the plural, while **lima** «snot» or **pixa** «rheum» are rarely used in the singular. In fact, many words which represent waste fluids have a quantitative meaning and express a suggestion of contempt, for example **ćik** «mud», pl. **ćika**, **muter** «urine», pl. **mutera**, **ćhungar** «spit», pl. **ćhungara**, etc. In some cases the difference between the singular

and plural is lexicalized: **thuv** «smoke», pl. **thuva** (variant **thuvã**) «dust» in Mečkari of Myzeqe, or **devel** «heaven, God», pl. **devla** «cloud» in Drindari of Bulgaria. We can mention also **sastrn** «iron», pl. **sastrna** «hardware» in many dialects. This phenomenon is a lexical characteristic.

Words with the suffix **-pe** (with its variants **-pa**, **-pen**, **-be**, **-ba**, **-ben**, **-mos**) which express an action, form the plural with **-mata** in many dialects, but not in all. This plural, which is lacking especially in the Balkan group of dialects, is also used in the Common Language: **mudaripe** «murder», pl. **mudarimata**.

### 3) The Category of Case

The category of case is of the most morphological character, not only because it encompasses all nouns without semantic or lexical distinction, but also because it is conditioned only by the structure of sentence, that is to say, by syntax, in an automatic fashion, without examining any particular semantic element. Every noun in modern Romani has two forms: the simple form or **A-Form** and the oblique form or **B-Form**. The role of the latter is described in the section on syntax, and so here we need only give some general information concerning this question:

The A-Form is the form in which the subject of the sentence and words appearing after a preposition are found.

The B-Form is the form used before a postposition.

The direct object is expressed in the A-Form when it denotes an inanimate object, and in the B-Form when it denotes a living thing. This is the rule in the Common Language, but there are many exceptions in dialects.

In many descriptions of Romani dialects we find noun declensions with seven or eight morphological cases. As we have seen above, Romani possesses only two case forms: the A-Form and the B-Form. Whence then does this large number of cases come?

It must be stated at the beginning that, since we are studying morphology, the term "case" is used here with a purely morphological meaning. It would be a mistake to confuse morphological cases, which concerns the study of forms, with semantic cases, which concerns a different plane. From the semantic standpoint all languages have cases in the broader sense of this word, since all languages must express, for example, the concepts of dative, ablative, instrumental, etc. In the morphological domain we are interested not in this universal characteristic of languages, but rather only in the manner of expressing these concepts by means of morphological cases.

In the Indo-European languages the most prevalent systems for expressing semantic cases are the following: case system (morphological cases, like Sanskrit), prepositional system (like English), declensional/prepositional system (like the Slavic languages and Latin), and postpositional system (like Hindi). If we compare these with Romani, at first glance it appears that Romany's system most resembles the case system:

|       | Sanskrit | English    | Serbian    | Hindi    | Romani   |
|-------|----------|------------|------------|----------|----------|
| nom.  | bhus     | earth      | zemlja     | bhumi    | phuv     |
| acc.  | bhum     | earth      | zemlju     | bhumi    | phuv     |
| inst. | bhuva    | with earth | sa zemljom | bhumi se | phuvaça  |
| dat.  | bhuve    | to earth   | zemlji     | bhumi ko | phuvage  |
| abl.  | bhuvas   | from earth | od zemlje  | bhumi se | phuvaθar |
| gen.  | bhuvas   | of earth   | zemlje     | bhumi ka | phuvaqo  |
|       |          |            |            | ki       | -qi      |
|       |          |            |            | ke       | -qe      |
| loc.  | bhuvi    | in earth   | u zemlji   | bhumi me | phuvaθe  |

However, the existence of a series of prepositions indicates that we cannot speak of a declensional/prepositional system. Thus we have, besides the forms given above, the possibility to link these words with a preposition:

**kotar-i phuv** «from the land» = **e phuvāθar**

**and-i phuv** «in the land»

**upr-i phuv** «on the land» = **e phuvāθe**

As we can see, the forms with the endings -θar and -θe can be replaced with a prepositional construction, while **phuvāça** «with land» and **phuvāqe** «to the land» cannot be replaced at all with such constructions.

In order to determine if we are truly dealing with a declensional/prepositional system it is not sufficient to compare the inflection of a noun. It is essential that we compare, side-by-side, two groups of nouns in Romani with a language of the true declensional-prepositional type, as for example Serbo-Croatian.



| Serbo-Croatian   | Romani   |
|--|--|
| moja mlada snaha<br>moju mladu snahu<br>moje mlade snahe<br>mojoj mladoj snasi<br>s mojom mladom snahom<br>o mojoj mojoj snasi | mi terni bori<br>me terne boră<br>me terne boră- <i>qo</i> , - <i>qi</i> , - <i>qe</i><br>me terne boră- <i>qe</i><br>me terne boră- <i>ça</i><br>me terne boră- <i>qe</i> |
| and further:   |  |
| kod moje mlade snahe<br>od moje mlade snahe  | me terne boră- <i>Ōe</i><br>me terne boră- <i>Ōar</i>  |

One can immediately see that in Serbian-Croatian each element of the noun group is declined individually in most cases, while in Romani we have one two forms: *mi terni bori* and *me terne boră*, which are respectively, the A-Form and the B-Form. We then add the particle *-qo*, *-qe*, *-ça*, *-Ōe*, etc., which are postpositions, as in Hindi. The difference is based on the fact that *bori* represents an oblique form after a preposition, whereas *bhumi* in the above table does not change. In fact, Hindi also exhibits the same alternation between a simple form and an oblique form in many words, e.g. *laRka* «boy», oblique *laRke*, while in feminine nouns like *bhumi*, these forms are homonyms.

If declensions exist as in Serbo-Croatian we would have forms such as *\*măqe ternăqe borăqe* «to my young bride», *\*măça ternăça borăça* «with my young bride», etc., which are utterly impossible in Romani.

In addition, if we compare the singulars and plurals of the forms of particles, which are frequently presented as "case endings," we see that they are very similar for each "case": [-*te*] in the locative singular opposite [-*de*] in the locative plural, [-*tar*] in the "ablative" singular opposite [-*dar*] in the ablative plural, [-*sa*] in the instrumental singular opposite [-*ca*] in the instrumental plural, etc. However, in languages with true declensions we do not have such a similarity as the comparison of endings for singular and plural in different cases in Serbo-Croatian illustrates: locative *snasi/snahama*, instrumental *snahom/snahama*, genitive *snahe/snaha*, etc. It seems clear that the only difference between the singular and plural of particles which characterize "cases" in Romani arises from the contact of their first phoneme which becomes voiced after the nasal *-n*, while remaining unvoiced after a vowel or an *-s*. The change of the phonological type of a phoneme when in contact with another is termed *sandhi*. *Sandhi* of voicing characterizes the contact of a Romani postposition with the stem, except in the case of the postposition *-ça*, which does not take on a voiced character in

contact with -n, but rather becomes semi-occlusive: [c], which also happens in many other languages in the environment n + s. Besides sandhi of voicing there is also general sandhi, which explains the often confusing situation which is encountered at the boundary between stem and postposition.

It is interesting to note that some Rom, when writing their language, intuitively treat the postposition as a distinct word and write it separately, as is done in Urdu. For example on cassettes by Esma Redžepova, we find postpositions which denote gender detached from the stem: Esma kiri gili «Esma's song», etc.

A number of researchers of Romani have interpreted this structure in a quite unique fashion very early on, as, for example, did Sampson and Sowa. In 1926, John Sampson wrote in his sizable work The Dialect of the Gypsies of Wales that the cases are formed

"from the oblique form of the singular and plural, by affixing a postposition; the oblique form is identical with the accusative form." He is of the opinion that the contemporary case system in Romani arose before the 10th century of our era in Persia, for the reason that at that time the Rom of Europe split off from the Rom of Asia, and this system is found in both of those groups. When examining declension in Romani, he repeats for each "case" the explanation by means of postpositions. He writes, for example, about the dative: "The dative is formed from the oblique stem of the singular and plural of both genders by means of the ~~postposition -ke~~, (i) ~~en~~ masc In-~~the~~ plur ~~en~~. ~~the~~ ~~en~~ ~~ke~~ as written as -en in the oblique case -en-ke, but is pronounced -en-ge".

However, this explanation holds true for other "cases" as well. Several authors distinguish a "primary case", which is the direct form, and a "secondary case", which are formed on the basis of the oblique case by means of postpositions, which some call "case suffixes." Nevertheless, a number of other researchers do not recognize this characteristic, because they have not examined the declension of a group of nouns, but rather only one isolated noun, and have usually decline a masculine noun like rrom or manuš. These words present a rather complex sandhi in a number of dialects, which frequently masks the fact that the postposition is affixed to the noun, and thus it appears that the endings are declined. In addition, the milieu of Latin, Slavic languages, or Sanskrit has obscured the fact that Romani possesses a quite unique system.

This explanation leads to the conclusion that Romani can be defined as a language of the "declensional/postpositional/prepositional type". In fact, this leads to yet another construction, which can be termed pre-postpositional. This concerns the manner of expressing the "lack of something" (Albanian uses the prefix pa), as, e.g. bi kheres-go «homeless», or bi me terne borā-go «without my young bride». The particle bi...-go can be termed a pre-postposition (circumposition). Until

now all of the Romani constructions have been found to express "semantic cases", but we can say that this type is mixed. From the purely morphological point of view, there are only two cases: the A-Form and the B-Form, which is characteristic of neo-Indic languages.

In ancient times Romani had a true declension with several cases, among them the locative and ablative. These two cases are now limited to a small number of vestigial expressions with a strictly adverbial meaning.

a) The Old Locative

has been retained in the words:

|                   |                             |       |                                 |
|-------------------|-----------------------------|-------|---------------------------------|
|                   |                             | from: |                                 |
| <b>khere</b>      | «at home»                   |       | <b>kher</b> «house»             |
| <b>jak-thane</b>  | «together»                  |       | <b>than</b> «place»             |
| <b>aver-thane</b> | «elsewhere»                 |       | " "                             |
| <b>belvele</b>    | «in the evening» (Kalderaś) |       | <b>belvel</b> «evening»         |
| <b>gave</b>       | «in the country»            |       | <b>gav</b> «village»            |
| <b>kurke</b>      | «Sunday» (Kalderaś)         |       | <b>kurko</b> «week»             |
| <b>nilaje</b>     | «in summer»                 |       | <b>nilaj, milaj</b><br>«summer» |
| <b>jevende</b>    | «in winter»                 |       | <b>jevend</b> «winter»          |
| <b>jek-čhane</b>  | «likewise»                  |       | <b>čhan</b> «manner»            |
| <b>averčhane</b>  | «otherwise»                 |       | " "                             |
| <b>divese</b>     | «during the day»            |       | <b>dives</b> «day»              |
| <b>ziberše</b>    | «next year»                 |       | <b>berš</b> «year»              |

The forms **diveça** or **disara** «during the day» are more common. In addition the locative appears also in two sexual expressions **del bule** and **del minze**. It should be mentioned that the Indic languages which have retained this case, like Punjabi, also from it with the ending **-e: käre** «at home». The use of the ablative in Punjabi is also limited.

b) The Old Ablative has been retained in the words:

|                  |                          |       |                     |
|------------------|--------------------------|-------|---------------------|
|                  |                          | from: |                     |
| <b>kheral</b>    | «from home»              |       | <b>kher</b> «house» |
| <b>dromal</b>    | «on the way»             |       | <b>drom</b> «road»  |
| <b>trujal</b>    | «all around»             |       | <b>truj</b> «wheel» |
| <b>dondal</b>    | «all around» (In Turkey) |       | ?                   |
| <b>(po)ćorāl</b> | «secretly»               |       | <b>ćor</b> «thief»  |
| <b>mujal</b>     | «upside down»            |       | <b>muj</b> «mouth»  |

(in the expression **del mujal** «to overturn»)

We must add a series of prepositions which become adverbs when the locative ending is added:

| PREFIX     |                | ADVERB      |              |
|------------|----------------|-------------|--------------|
|            |                | Locative -e | Ablative -al |
| angl-      | before         | angle       | anglal       |
| pal-       | after          | pale        | palal        |
| tel-       | under          | tele        | telal        |
| opr-, upr- | above, upon    | opre, upre  | opral, upral |
| andr-      | inside         | andre       | andral       |
| avri       | outside        | avri        | avrial       |
| paš-       | near, close to | paše        | pašal        |
| maškar-    | between        | maškare     | maškaral     |
| prd-       | beyond         | prde-       | prdal        |

Notes: The word **maškar** also exists as a noun with the meaning «middle» and «backbone», e.g. in the expression noted by Mann: **dav maškare** «I interfere».

The adverb **dural** «from afar» is formed from the adjective/adverb **dur** «far».

Besides having an ablative meaning, words that end in **-al** can have an indefinite locative meaning. We can add to these the diminutive suffixes: **-ör-**, **-örr-**, e.g. **palöral**, **palörral** «a little behind», etc.

Adverbs are also formed from demonstratives: **ajal**, **akhal**, **akhajal**, which do not have an ablative meaning, but mean «thus».

The morphological cases, the locative and the ablative, are dead categories in the Romani language, and their meanings are now expressed by means of the postpositions **-Ōe** and **-Ōar**, or the prepositions **k-**, **an-**, **kotar-**, etc. If we take the word **kher** «house», the locative and ablative meanings can be expressed in three different ways:

|          | with archaic cases | with prepositions | with postpositions  |             |
|----------|--------------------|-------------------|---------------------|-------------|
| locative | <b>khere</b>       | <b>kheres-Ōe</b>  | <b>an-o kher</b>    | «at home»   |
| ablative | <b>kheral</b>      | <b>kheres-Ōar</b> | <b>kotar-o kher</b> | «from home» |

If these forms with postpositions are interpreted as "morphological cases", then it is difficult to determine the

nature of true cases which appear in the words **khere** and **kheral**. These endings arose etymologically from the ancient Indo-European cases, while postpositions have no connection with them at all.

The suffix **-al**, which is now accepted as having an adverbial value, can be revived in order to interpret international adverbs of the type **normal**, **natural**, **tradicional**, **general**, **original**, etc. This class of adverbs can be enriched with forms such as **automatikal** «automatic» or **ačaral** «as a rule» (from **ačar** «law»).

It can be noted here that the endings **-e** and **-al** are even encountered in the ancient Indic epic the Rig-veda, as e.g. the loc. **devé** and the abl. **devat** from the word **devas** [-ah] «god», or the loc. **prijē** and abl. **prijat** from **prijas** «likeable» (this root is related to the Slavic **prija-telj** «friend»). The development of the Indic **t** (-at) into **l** in Romani (-al) is the rule: Sanskrit **gīti**, **gītá**, Hindi **gitī**, Romani **gili** «song».

A vestige of the ending **-at** is also found in Old Persian (Avestan): **durat** «from afar», and, according to McDonnal, also in the most ancient Latin inscriptions in the form **-uod**: **Gnaiuod** (later **Cnaeo** «from Knaeus»), and in the adverb **τῶδ(ε)** «then» in the old dialect of Crete (not to be confused with the Dorian form **τῶδε** for **τοῦδε**, the genitive of **ὄδε**).

The categories of gender, number, and case encompass all elements of noun classes, while the vestige of cases are limited to the examples which have provided above. We will study in turn the various morphological elements of noun classes, while proceeding from the most characteristic one, the substantive.

THE SUBSTANTIVE  
(along with postpositions)

Here we will present the manner of forming a feminine noun from a masculine one (when possible), the manner of forming the plural for both genders, and the formation of the B-Form for both numbers and both genders. Throughout this study we will use abbreviations in brackets to mark gender, number, and the case of the word, e.g. [f-sg-B] means: "feminine, singular in the B-Form".

The Formation of the Feminine from a Masculine Noun [f-sg-A]

1) By morphological means:

a) in some cases formed with the endings: -o into -i:

|                |               |                |                  |
|----------------|---------------|----------------|------------------|
| <b>raklo</b>   | «non-Rom boy» | <b>rakli</b>   | «Rom girl»       |
| <b>ćiriklo</b> | «bird»        | <b>ćirikli</b> | «female bird»    |
| <b>sikamno</b> | «teacher»     | <b>sikamni</b> | «female teacher» |

(the same for **bašalno** «musician», **gilamno** «singer», etc. A list of these words is given above).

b) by means of the ending -ni:

|             |           |               |                 |
|-------------|-----------|---------------|-----------------|
| <b>rrom</b> | «Rom man» | <b>rromni</b> | «Rom woman»     |
| <b>ruv</b>  | «wolf»    | <b>ruvni</b>  | «she-wolf»      |
| <b>xer</b>  | «donkey»  | <b>xerni</b>  | «female donkey» |

The word **manuš** has the feminine form **manušni** in some dialects, as well as in the Common language, while **muršni**, from **murš** «male», is used in the meaning «courageous»: **muršni sar i Ganimèta** «as courageous as Ganimede» (Tërbeshi).

c) sometimes additions of -ni causes some other changes:

the loss of final -j after a vowel:

|              |          |               |             |
|--------------|----------|---------------|-------------|
| <b>raj</b>   | «Sir»    | <b>rani</b>   | «Madam»     |
| <b>rašaj</b> | «priest» | <b>rašani</b> | «priestess» |

**xoraraj**, **khoraj** «Moslem» (in dialects also «Turk») has the feminine forms **xoraxani**, **khorani**.

|              |                       |               |                     |
|--------------|-----------------------|---------------|---------------------|
| <b>xulaj</b> | «master of the house» | <b>xulani</b> | «lady of the house» |
|--------------|-----------------------|---------------|---------------------|

the simplification of a consonant cluster:

|              |            |                                 |        |
|--------------|------------|---------------------------------|--------|
| <b>grast</b> | «stallion» | <b>grasni</b> or <b>grastni</b> | «mare» |
|--------------|------------|---------------------------------|--------|

the reemergence of a old form:

**guruv** (from **gurum**) «bull»                      **gurumni** «cow»

d) in some words the singular direct form is the same in both genders:

**xenamik** «in-law»                                      **xenamik**

e) the addition of an **-i** is quite rare:

**zukul** «dog»    **zukli** «bitch»  
**devel** «god»    **devli** «goddess»

2) By lexical means:

We will give here only a few examples where the feminine noun is not formed from the corresponding masculine noun in the expected fashion:

|                |              |              |                   |
|----------------|--------------|--------------|-------------------|
| <b>dad</b>     | «father»     | <b>daj</b>   | «mother»          |
| <b>phral</b>   | «brother»    | <b>phen</b>  | «sister»          |
| <b>zamutro</b> | «son-in-law» | <b>bori</b>  | «daughter-in-law» |
| <b>bašno</b>   | «rooster»    | <b>kaxni</b> | «hen»             |

3) In some cases of foreign nouns the feminine are homophones with the masculine:

**kòmšia** «neighbor»                                      **kòmšia** «female neighbor»

In such as case the use of the differentiated form with the Slavic prefix **-ka**, which has developed in many dialects, is recommended: **komšika**.

#### The Formation of the Plural of Masculine Nouns [m-pl-A]

We can distinguish nine groups of masculine nouns in Romani based on their endings and the formation of the their plural:

1) All nouns ending in stressed **-o** (which shows that these are native words) change this ending into **-e** to form the plural:

|              |         |              |
|--------------|---------|--------------|
| <b>ilo</b>   | «heart» | <b>ile</b>   |
| <b>ćaro</b>  | «dish»  | <b>ćare</b>  |
| <b>bakro</b> | «ram»   | <b>bakre</b> |

while including here the neologisms with the suffix **-tro** or **-itro**, which indicate an agent of an action:

**šunitro** «listener»                      **šunitre** (from **šunel** «to listen»)

2) Only one masculine noun ends in stressed **-i**:

|             |         |             |
|-------------|---------|-------------|
| <b>pani</b> | «water» | <b>pană</b> |
|-------------|---------|-------------|

3) Masculine nouns ending in unstressed **-o** or **-i** (which shows they are foreign words), generally form plurals with **-ă** or **-ură**:

|                  |            |                       |
|------------------|------------|-----------------------|
| <b>lăfi</b>      | «word»     | <b>lăfă, lafură</b>   |
| <b>partizăni</b> | «partisan» | <b>partizănă</b>      |
| <b>kômbo</b>     | «knot»     | <b>kômbă, kombură</b> |
| <b>kongrêso</b>  | «congress» | <b>kongrêsă</b>       |

The word **kôkalo** «bone» and **pêtalo** «horseshoe» are exceptions, having the plurals **kôkala** and **pêtala**.

Some very old foreign words, alongside forms with a final vowel (e.g. **drôm** «road» and **văkto** «time») have a preferred form which have become assimilated into the language: **drom**, **vaxt**, etc.

4) Foreign nouns ending in a velar or a dental, to which an epenthetical **-o** or **-i** is added, form the plural with the suffix **-ură** (with the less wide spread variant **-ora**):

|               |           |                  |
|---------------|-----------|------------------|
| <b>brêgo</b>  | «hill»    | <b>brêgură</b>   |
| <b>miting</b> | «meeting» | <b>mitingură</b> |

Internationalisms of the type **dialèkti**, **elemènti**, have also entered this class, but vary in their final vowel (unstressed **-o** or **-i**) according to dialect (**dialèkto dialèkti**). Recently the tendency to enter the Common Language in the feminine gender with an unstressed **-a** has been noticed: **i dialèkta**. In this manner the variation in form in both the singular and plural is avoided. The plural is then **dialèkte** (following the type **tràsta**, see below). In fact, the word **dialèkta** is feminine in Greek, the language from which it was borrowed. Internationalisms vary from masculine to feminine, as shown in the table below, in which we have chosen just a few languages:



|          | Alb. | Rus. | Pol. | Rom.          | Serbo-C | Maced. |
|----------|------|------|------|---------------|---------|--------|
| poem     | f.   | f.   | m.   | a.*           | f.      | f.     |
| problem  | m.   | f.   | m.   | f.            | m.      | m.     |
| telegram | m.   | f.   | m.   | f.<br>(or a.) | m.      | f.     |
| group    | m.   | f.   | f.   | a.            | f.      | f.     |
| episode  | m.   | m.   | m.   | a.            | f.      | f.     |
| control  | m.   | m.   | f.   | a.            | f.      | f.     |
| program  | m.   | f.   | m.   | a.            | m.      | f.     |
| montage  | m.   | m.   | m.   | a.            | f.      | f.     |
| variant  | m.   | m.   | m.   | f.            | f.      | f.     |

\* Note: the neuter gender in Romanian has a masculine form in the singular and a feminine form in the plural.

The Rom, depending on where they reside, have borrowed these words with the gender of the intermediary language. As Lumia Osmani has shown, there is not particular reason why they should be masculine in one language and a different gender in another language, which creates difficulties for Rom who have learned a word with a particular gender in their dialect, which may not agree with other dialects. For this reason, all these words are unified in the feminine gender, which is easier to transmit. Word classes of this type are continually enriched by new lexemes: **elemènta, konsonànta, konstànta** and many other international words.

5) Some nouns which denote a person end in **-o** (stressed or unstressed, depending on dialect) and form the plural with the suffix **-vã**:

|                   |                  |                       |
|-------------------|------------------|-----------------------|
| <b>pàpo, pàpu</b> | «grandfather»    | <b>pàpovã, pàpuvã</b> |
| <b>kàko, kako</b> | «maternal uncle» | <b>kàkovã</b>         |

This is the rule for internationalisms with velar + unstressed o:

|                 |               |                   |
|-----------------|---------------|-------------------|
| <b>etnològo</b> | «ethnologist» | <b>etnològovã</b> |
|-----------------|---------------|-------------------|

In some dialects of the Gurbeti group, the ending **-urã** is encountered also in nouns of this category, e.g. **etnològurã**.

6) Many foreign words which end in unstressed **-a** change this into **-e**, always unstressed in the plural:

|                |            |                |
|----------------|------------|----------------|
| <b>kòmšia</b>  | «neighbor» | <b>kòmšie</b>  |
| <b>kumbàra</b> | «sand»     | <b>kumbàre</b> |

This category includes many names for members of various clans:

|                 |   |                 |
|-----------------|---|-----------------|
| <b>Arlìa</b>    | a clan in Prishtina, Skopje, etc.                 | <b>Arlìe</b>    |
| <b>Aškalià</b>  | Rom who don't speak Romani (especially in Kosova) | <b>Aškaliè</b>  |
| <b>Topanlià</b> | residents of Tophan quarter in Skopje             | <b>Topanliè</b> |

This category includes names of members of professions ending in -ista:

|                    |              |                    |
|--------------------|--------------|--------------------|
| <b>specialista</b> | «specialist» | <b>specialiste</b> |
|--------------------|--------------|--------------------|

It must be noted that in all examples until now the stress remains on the same syllable in the singular and the plural.

7) All other native nouns form the plural by adding a stressed -a to the stem, i.e. the stress continues to be final, while passing from the final syllable of the stem to plural ending:

|                     |         |                       |
|---------------------|---------|-----------------------|
| <b>sap</b>          | «snake» | <b>sapa</b>           |
| <b>kan</b>          | «ear»   | <b>kana</b>           |
| <b>manuś</b>        | «man»   | <b>manuśa</b>         |
| <b>grast (graj)</b> | «horse» | <b>grasta (graja)</b> |

A limited number of very old foreign words follow this pattern:

|             |        |              |
|-------------|--------|--------------|
| <b>drom</b> | «road» | <b>droma</b> |
|-------------|--------|--------------|

Besides the plural in -a, the nouns **dive(s)** «day», **masek** or **čhon** «moon», **berś** «year», have a plural identical with the singular when used when the noun is accompanied by a number: **duj dive** «two days», **šov masek** «six months», etc. However, the form with -a is not wrong.

In some dialects, especially the northern dialects, the plural in -a also appears in the nouns **rrom** «Rom», **phral** «brother», **manuś** «man», **bal** «hair», etc. In the Common Language the full form is preferred: **rroma**, **phrala**, etc.

Two special cases must be added here:

8) Nouns with the abstract suffix -pe (with the dialectal variants -pa, -be, -ben, -ba, -mo, -mos) are rarely used in the plural by virtue of their meaning. Nevertheless, the plural does exist in many dialects; it is either identical to the singular, or is formed with -mata:

|                |            |                             |
|----------------|------------|-----------------------------|
| <b>phućipe</b> | «question» | <b>phućipe or phućimata</b> |
| <b>naśimos</b> | «stroll»   | <b>naśimata</b>             |

The place of stress varies depending to dialect.

9) In the Southern Balkan dialects, some nouns of a Turkish origin with a meaning alluding to Ottoman times, have preserved the Turkish plural suffix -lâr, to which the Romani -ă is added.

|      |                 |          |
|------|-----------------|----------|
| paša | «pasha»         | pašalàră |
| hoža | «Moslem priest» | hožalàră |

#### The Formation of the Plural of Feminine Nouns [f-pl-A]

Three main groups of feminine nouns are distinguished, according to their endings and the plural form, which depends on the singular form:

1) Nouns ending in a stressed -i (which shows that it is native), form the plural with -ia or ă:

|                |                 |                 |
|----------------|-----------------|-----------------|
| rromni         | «Rom woman»     | rromnia, rromnă |
| bori           | «bride»         | boră            |
| rakli          | «non-Rom girl»  | raklia, raklă   |
| gazi           | «non-Rom woman» | gazia           |
| ćuci           | «breast»        | ćucia, ćucă     |
| buti, buti     | «work»          | bută            |
| morthi, morthi | «skin»          | morthă          |
| luludi, luludi | «flower»        | luludă          |

In a series of dialects, especially among the Rom who live or have lived in Romania, the ending -ă is pronounced [ɛa].

Palatalization of dentals before -ă occurs in those dialects where this is the rule:

|        |        |                |
|--------|--------|----------------|
| buti   | «work» | bută [-t̪a]    |
| morthi | «skin» | morthă [-t̪ha] |

The collision of -ă with the affricates ć and ȝ frequently softens the pronunciation of these consonants: [ćuc'a], [gaz'a].

The archaic bisyllabic plural -ia appears in many dialects, especially after the consonants -r, -ć, and -ȝ, and are used extensively in poetry for rhythmic reasons (see above).

Some dialects of the Gurbeti type form the noun plural ending in -i in the singular by changing this to -a without prejotization:

|        |             |        |
|--------|-------------|--------|
| rromni | «Rom woman» | rromna |
| bori   | «bride»     | bora   |

The spread of this plural is irregular from dialect to dialect, and from word to word. Therefore the plural form -ă or -ia is preferred in the Common Language.

2) Feminine nouns ending in a consonant in the singular can be divided into two subgroups:

a) those which form the plural with **-a** in all dialects:

|              |          |                     |
|--------------|----------|---------------------|
| <b>jag</b>   | «fire»   | <b>jaga</b>         |
| <b>phakh</b> | «wing»   | <b>phakha</b>       |
| <b>ćik</b>   | «mud»    | <b>ćika</b>         |
| <b>baj</b>   | «sleeve» | <b>baja</b>         |
| <b>pošom</b> | «wool»   | <b>pošoma, etc.</b> |

b) those which form the plural with **-ă** or **-a**, depending on dialect:

|             |            |                         |
|-------------|------------|-------------------------|
| <b>jakh</b> | «eye»      | <b>jakhă, jakha</b>     |
| <b>thuv</b> | «smoke»    | <b>thuvă, thuva</b>     |
| <b>ćhib</b> | «language» | <b>ćhibă, ćhiba</b>     |
| <b>diz</b>  | «city»     | <b>diză, diza</b>       |
| <b>por</b>  | «feather»  | <b>poră, pora, etc.</b> |

In the Common Language the systematic use of **-ă** is recommended, whenever possible.

In a series of dialects some nouns denoting inanimate objects and which end in **-i** have developed into **-ik**. Each form has its own the plural and are used in the Common Language:

|                        |         |  |
|------------------------|---------|--|
| <b>ćhuri, ćhurik</b>   | «knife» | <b>ćhuria, ćhură, ćhura, ćhurika</b>     |
| <b>kangli, kanglik</b> | «comb»  | <b>kanglia, kanglă, kangla, kanglika</b> |

3) Feminine nouns ending in unstressed **-a** in the singular. These are foreign words, and have **-e** in the plural, with no change in stress:

|                  |              |                        |
|------------------|--------------|------------------------|
| <b>tràsta</b>    | «bag»        | <b>tràste</b>          |
| <b>temèla</b>    | «basis»      | <b>temèle</b>          |
| <b>lahùta</b>    | «violin»     | <b>lahùte</b>          |
| <b>kemàna</b>    | «violin»     | <b>kemàne</b>          |
| <b>filozòfia</b> | «philosophy» | <b>filozòfie, etc.</b> |
| or               |              |                        |
| <b>filozofìa</b> |              | <b>filozofie</b>       |

This penetration of foreign words of the type **dialèkta**, which has entered Romani as a feminine noun, while being masculine in Albanian:

|                   |             |   |
|-------------------|-------------|---|
| <b>dialèkta</b>   | «dialect»   | <b>dialèkte</b> (also feminine in Greek)          |
| <b>elemènta</b>   | «element»   | <b>elemènte</b>                                   |
| <b>konsonànta</b> | «consonant» | <b>konsonànte</b>                                 |
| <b>variànta</b>   | «variant»   | <b>variànte</b>                                   |
| <b>telegràma</b>  | «telegram»  | <b>telegràme</b><br>(also feminine in Macedonian) |

4) Four nouns has special plurals:

|       |                 |                      |
|-------|-----------------|----------------------|
| daj   | «mother»        | dia (or daja, dà)    |
| čhaj  | «non-Rom girl»  | čhia (or čhaja, čhà) |
| sasuj | «mother-in-law» | sasa                 |
| asvin | «tear»          | asva or asia         |

## The Formation of the B-Form

### The Formation of the B-Form of Singular Masculine Nouns

1) Masculine nouns with the singular the A-Form ending in stressed -o form the B-Form with -es:

|       |               | B-Form |
|-------|---------------|--------|
| čhavo | «Rom boy»     | čhaves |
| raklo | «non-Rom boy» | rakles |
| bakro | «ram»         | bakres |
| ilo   | «heart»       | iles-  |
| šelo  | «rope»        | šeles- |

The B-Form of nouns denoting inanimate objects is used only when linked to a postposition and is never found alone. In grammars and dictionaries this particularity is indicated by means of a dash: *iles-*, *šales-*, etc., which shows that they occur only with postpositions, as, for example, *iles@ar*, *šalesqo*, etc., while words such as *čhaves*, *rakles*, *bakres*, which indicate living creatures, are encountered both with and without postpositions. These are given without a dash.

In the majority of the Balkan dialects final -s is not pronounced: [čhave], [rakle], etc. When linked to a postposition -s is not final and therefore is usually pronounced in the usual fashion, although in a series of dialects it changes to [h]: [raklehke] for *raklesqo* (also [raklehʔe] [raklehke], etc. according to the palatalization grades). This is one of the sandhi rules which are distinctive for each dialect. The spelling is the same, but because of different rules for reading, each Rom will use a pronunciation natural for his dialect.

In a series of dialects (e.g. Mahažeri), masculine nouns ending in -o and which denote inanimate objects have -os in the B-Form:

|                 |         |        |
|-----------------|---------|--------|
| ilo (also vilo) | «heart» | ilos-  |
| šelo            | «rope»  | šelos- |

e.g. *zoralevilosqo(ro)* «brave»

2) Nouns ending in a consonant form the B-Form with the ending -es:

|        |                    |                                    |
|--------|--------------------|------------------------------------|
| ruv    | «wolf»             | ruves                              |
| grast  | «horse»            | grastes                            |
| (graj) | «horse»            | grajes (in some northern dialects) |
| xulaj  | «man of the house» | xulajes                            |

In the form of singulars ending in -el, the -e before l disappears throughout all forms of the B-Form:

|       |       |        |
|-------|-------|--------|
| zukul | «dog» | zukles |
| devel | «god» | devels |

3) Foreign nouns ending in unstressed -i or -o usually change this to -es in the B-Form:

|                 |                  |                   |   |
|-----------------|------------------|-------------------|---|
| <b>làfi</b>     | «ghost»          | <b>lafes-</b>     | (there is the tendency to replace the Turkism làfi with the Rom word lav) |
| <b>vampìri</b>  | «vampire»        | <b>vampires</b>   | (the genuine Romàni word is <b>čovano</b> )                               |
| <b>dèvero</b>   | «brother-in-law» | <b>deveres</b>    | (replaced by the Romàni word <b>gešt</b> , variant <b>zešt</b> )          |
| <b>kongrèso</b> | «congress»       | <b>kongreses-</b> |   |

In some dialects, as for example in Mahažeri, the final vowel is retained, and a -s is added:

|               |           |                      |
|---------------|-----------|----------------------|
| <b>làfi</b>   | «word»    | <b>lafis-</b>        |
| <b>papùri</b> | «train»   | <b>papuris-</b>      |
| <b>kòso</b>   | «yoghurt» | <b>kosos-</b> , etc. |

The word **papu** or **papo** usually form the B-Form as above: **papo**, **pàpo**, **papu**, **pàpu** «grandfather», B **papos**, **papus** (both variants are accepted in the Common Language), whereas the word **kak** (variants **kàko**, **kako**) «maternal uncle» is treated, depending on dialect, as a Romàni word or a foreign word, i.e. with the B-Form as **kakes** or **kakos**.

Three words have a special the B-Form:

|             |          |                               |  |
|-------------|----------|-------------------------------|--|
| <b>muj</b>  | «mouth»  | <b>mos-</b>                   | (an older form, tends to be replaced by the analogous form <b>mujes-</b> ) |
| <b>pani</b> | «water»  | <b>paněs-</b>                 | (or <b>bark</b> in some dialects)  |
| <b>vogi</b> | «spirit» | <b>voges</b> or <b>vogies</b> |  |

Besides the form **vogi**, the variants **voži** with the B Case **vozes** are also encountered.

#### The Formation of the B-Form of Singular Feminine Nouns [f-sg-B]

Feminine nouns can be divided into two groups:

1) The group of foreign nouns ending in -a, which form the B-Form, while throwing the stress from the root to the ending -a:

|                 |           |                         |
|-----------------|-----------|-------------------------|
| <b>tràsta</b>   | «bag»     | <b>trasta-</b>          |
| <b>màčka</b>    | «cat»     | <b>mačka</b>            |
| <b>žàpka</b>    | «frog»    | <b>žapka</b>            |
| <b>dialèkta</b> | «dialect» | <b>dialekta-</b>        |
| <b>variànta</b> | «variant» | <b>varianta-</b>        |
| <b>elemènta</b> | «element» | <b>elementa-</b> , etc. |

2) All other nouns have the same form, both in the plural and singular of the B-Form:

|            |                 |                  |
|------------|-----------------|------------------|
| bori       | «bride»         | boria, boră      |
| buti, buti | «work»          | bută-            |
| rakli      | «non-Rom girl»  | raklia, raklă    |
| jag        | «fire»          | jaga-            |
| phuv       | «land»          | phuvă-, phuva-   |
| čhib       | «tongue»        | čhibă-, čhiba-   |
| daj        | «mother»        | dia, daja, dà    |
| čhaj       | «Rom woman»     | čhia, čhaja, čha |
| sasuj      | «mother-in-law» | sasa             |

As in the plural, the recommended forms for the B-Form of the words daj and čhaj are dia and čhia.

#### The Formation of the Plural B-Form of Masculine Nouns [m-pl-B]

The number of plural forms is much less than that of singular forms, both for masculine and feminine nouns:

1) Nouns ending in -a or -e in the A-Form have -en in the plural, including those few nouns which have a Ø ending in the plural of the A-Form:

|        |                |         |
|--------|----------------|---------|
| čhave  | «Rom boys»     | čhaven  |
| rakle  | «non-Rom boys» | raklen  |
| ile    | «hearts»       | ilen-   |
| phrala | «brothers»     | phralen |
| grasta | «horses»       | grasten |
| berś   | «years»        | berśen- |

2) However, nouns ending in -ă change this to -ěn:

|        |          |           |
|--------|----------|-----------|
| papùră | «trains» | pampurěn- |
| lăfă   | «words»  | lafěn-    |

It must be mentioned that the endings -ură in the plural of the A-Form usually corresponds to -onon in the B-Form:

|        |         |                                |
|--------|---------|--------------------------------|
| lăfură | «words» | lafonen- (more rarely lafurěn) |
|--------|---------|--------------------------------|

It is recommended that the foreign word lăfi be replaced in the Common Language with the Romani word lav.

#### The Formation of the Plural B-Form of Feminine Nouns [f-pl-B]

1) Nouns ending in -a or -e in the plural A-Form have the ending -en, as in the case of masculine nouns:

|      |           |        |
|------|-----------|--------|
| jaga | «fires»   | jagen- |
| dia  | «mothers» | dien   |



|          |                  |                 |
|----------|------------------|-----------------|
| sasa     | «mothers-in-law» | sasen           |
| tràsta   | «bags»           | trasten-        |
| dialèkte | «dialects»       | dialekten-      |
| variànte | «variants»       | varianten-      |
| elemènta | «elements»       | elemeten-, etc. |

2) In Romani of the Balkan branch nouns ending in -ă or -ia take the endings -ě̃n or -ien, while in the Central European branch the ending -ě̃n is used (-ian):

|               |                |   |
|---------------|----------------|---|
| raklă, raklia | «non-Rom girl» | Balkan variant raklě̃n, raklien<br>C. Eur. variant raklăn |
| boră, boria   | «bride»        | Balkan variant borě̃n, borien<br>C. Eur. variant borăn    |

Words such as čhuri, čhurik «knife» with the plural čhură, čhurika «knives» have a number of variants for the B-Form:

čhurě̃n-, čhurien-, čhuriken-, čhurăn-, čhurian-, čhurikan- ...  
pustě̃n-, pustien-, pustiken-, «books», etc.

3) The plural of the B-Form for abstract nouns with the suffix -pe (and its variants -pa, -ba, -be, -pen, -ben, -mos) is formed by with the ending -mata for the A-Form (the form with the ending -pa, which occurs in the Balkans, is also acceptable):

|         |        |   |
|---------|--------|---|
| xoxaipe | «liar» | pl. A xoxaimata<br>pl. B xoxaimaten- (or xoxaipan-) |
|---------|--------|---|

In a series of dialects, especially among the Rom who live or have lived in Romania, the ending -ă is pronounced [ɛ̃a].

#### POSTPOSITIONS

At the beginning of this chapter we mentioned the concept of the postposition, while comparing the nominal system of Romani with that of a language of India, namely, Hindi. Although postpositions are invariable morphemes, except for their possessive meaning, they are dealt with here because they are intimately connected with the nouns.

While Hindi has no prepositions, Romani does, and in some cases they are synonyms with postpositions. In most dialects and the in the Common Language a semantic distinction has been created between postpositions and prepositions, and between these and the vestigial case forms. This distinction can be illustrated with the word kher, one of the rarities which retain case endings:

|                |                       |                       |
|----------------|-----------------------|-----------------------|
| case endings:  | khere                 | «at home», «homeward» |
| postpositions: | kheres <sup>o</sup> e | «in the home»         |
| prepositions:  | and-o kher            | «inside the home»     |

In comparison, a case can also be formed with -al opposite the postposition -@ar and the preposition kotar-.

As we have mentioned above, the B-Form of all nouns is used in conjunction with postpositions. In addition, the B-Form of nouns which denote living things are also used alone.

In Romani, there are five postpositions:

|        | meaning              | pronunciation as per sandhi I |                    |
|--------|----------------------|-------------------------------|--------------------|
|        |                      | after -n-                     | in other positions |
| @e     | in, on, at           | [de]                          | [te]               |
| @ar    | from                 | [dar]                         | [tar]              |
| ça(r)  | with                 | [ca(r)]                       | [sa]               |
| qe     | to                   | [ge]                          | [ke]               |
| qW(rW) | (possessive meaning) | [gWrW]                        | [kWrW]             |

We use here the symbol W to indicate the vowels o, i, e, which depend on the gender, number, and case of the noun modified. Some examples are given below:

| possessor | example                        |                    |
|-----------|--------------------------------|--------------------|
| [m-sg-A]  | W=o e rromesqoro <u>kher</u>   | the Rom's house    |
| [f-sg-A]  | W=i e rromesqiri <u>kemàna</u> | the Rom's violin   |
| others    | W=e e rromesqere <u>çhave</u>  | the Rom's children |

This postposition apparently "declines" like an adjective. Compare:

|          |               |                  |
|----------|---------------|------------------|
| [m-sg-A] | o baro kher   | the big house    |
| [f-sg-A] | i bari kemàna | the big violin   |
| others   | e bare çhave  | the big children |

Naturally, the symbol W is not a Romani letter. It will be used only in grammatical tables to bring together the forms -qoro, -qiri, and -qere, in the manner of an abbreviation.

In many dialects, this postposition has a different form with -a, for f-sing-B:

|          |                               |                       |
|----------|-------------------------------|-----------------------|
| [f-sg-B] | W-a e rromesqara <u>diaça</u> | with the Rom's mother |
|----------|-------------------------------|-----------------------|

There is still no consensus on the place of these forms in the Common language. However they have more informational value than the form -qere (common to m-sing-B, m-pl-A, m-pl-B, f-pl-A, and f-pl-B), which is quite rare among dialects and overburdens the sentence.

Besides the full forms of the postposition -qWrW, there also exist three short forms:

-qerW very limited, in Slovakia and Hungary: -qero, -qeri, -qere  
-qrW especially in Central European dialects: -qro, -qri, -qre  
-qW in all dialects: -qo, -qi, -qe

Thus one says e rromesqoro kher, e rromesqero kher, e rromesqro kher and e rromesqo kher, always realizing q as [k], which does not occur after -n-. Many dialects use two forms: -qW and qrW or -qW and -qWrW, while giving the second form (-qrW or -qWrW) a emphatic or nominalized function, whereas the first form has a purely possessive meaning. In the grammar tables the common symbol -qW(rW) encompasses all variants.

### Sandhi in Postpositions

Each postposition demonstrates two levels of sandhi:

- a sandhi common to all dialects (Sandhi I)
- special sandhi, depending on dialect (Sandhi II and III)

Sandhi I is frequently called voicing sandhi. It appears in the postpositions -Θe, -Θar, -qe, and -qWrW. In accordance with this type of sandhi the phonemes Θ and q are realized as voiced after an -n, and unvoiced in all other positions (see above). Nevertheless, in the case of the postposition -ç, there is no voicing, since Sandhi I regulates the pronunciation of -ç as the affricate /c/ after n, but as /s/ after any other phoneme. Sandhi I is common to all dialects and is obligatory in the Common Language.

Sandhi II concern the treatment of final -s in the B-Form of the masculine singular. Depending on dialect, it is pronounced as [s], changes to [h], or is not pronounced at all. In many dialects (e.g. Tharegoni), Sandhi II, on the contrary, makes mute the phoneme [k], and the cluster -sk simplifies into [-s-]. In the case of the postposition -ça, Sandhi II regulates the ultimate pronunciation of /s/ into [j] or [h] or makes it mute.

In Table #3 we have provided a - where there is no Sandhi II, and a + where it operates:

Sandhi III can have two natures:

- a simple palatalization, which operates in those dialects which are characterized by the palatalization of velars before front vowels, such as e and i. Naturally, Sandhi III operates only before front vowels.
- the assimilation of the clusters nasal + occlusive into a nasal (e.g. the realization of the cluster /-nd/ as /-n-/).

This system of Sandhi is reflected in Table #3.

The forms in brackets are all of those found among the dialects. It must be added that in the case of the muting of ç between two

identical vowels, in some dialects these vowels each are long, whereas in other dialects they unite into one long vowel.

In order to illustrate this type of sandhi we can give the forms of the 3rd person personal pronouns on «they», oj «she», and ov «he» in the B-Form, and, respectively, len (example with -n), la and les (examples without -n, as illustrated in Table #4).

#### The Combination of Postpositions

One peculiarity of Romani is its ability to combine two postpositions, and, rarely, even three:

**e thanesqerençar** «with the local people» which can be analyzed as: **e than/es-qer/ençar**.

Shaip Jusuf has cited a word in the Džambaz dialect of Skopje which has three postpositions: **me čhavenqerenqerestar** «from the one that is of those of my children». Of course, such a form is not encountered frequently, and can occur only in a very specific context where it is known who is «he» and who are «they». Forms of this type are heard especially in metalinguistic discussions among Rom, however they are theoretically possible.

In some dialects, e.g. Mečkari, the possessive forms always appear in conjunction with postpositions when they come after a word with a postposition: in contrast to the form **me amalesqe phenāça** «with my friend's mother» (only one postposition with the possessive come first), we have the form **e phenāça me amalesqerāça**, with the possessive element at the end.

Of course, this difference can only be explained by considering that Romani does not have eight cases, but rather a system of postpositions. Case endings cannot be added to each other, whereas postpositions can be combined quite well, as can prepositions in other languages (e.g. in Albanian **për në Prizren** «(heading) for Prizren».

#### The Situation in Indic Languages

It would be interesting at this point to make a comparison between the forms of the possessive postpositions in Romani and their equivalents in the languages of India. We give the forms in classical written Hindi, then comes a transcription in Romani letters (capital R represents the retroflex r, the dash next to the vowel indicates length, and the dot next to the vowel indicates nasality), the parallel translation is in Romani and English:

#### Together or Separate

In the European tradition prepositions are written separate from the noun, in Romani as well as in Albanian: **kotar-o gav** «from the village», **and-o veś** «into the forest», etc. When we consider that

postpositions play the same role as prepositions, they should be written separately: e gaves ̸ar, e veś ̸e, etc., as is done in Urdu. However, there are four reasons for writing them together with the noun:

1) nothing may come between a postposition and a noun, whereas there are many cases where something intervenes between a preposition and a noun, as we can see below:

|                                      |   |                                      |
|--------------------------------------|---|--------------------------------------|
| kotar-o gav                          | = | e gaves ̸ar                          |
| kotar odova gav                      | = | odole gaves ̸ar                      |
| kotar odova baro gav                 | = | odole bare gaves ̸ar                 |
| kotar odova baro thaj<br>barvalo gav | = | odole bare thaj baxtale<br>gaves ̸ar |

(«from the village, from that village, from that big village, from that big and rich village»)

2) Nouns and postpositions constitute one stress unit, and the Rom speaker perceives it as such. In the Common Language the original Romani stress is in effect, i.e. the stress always falls on the final syllable of the B-Form, and consequently on the syllable before the postposition. This is one of the functions of the postposition; it indicates the place of stress.

3) The presence of the graphemes ̸, q, and ç, which are characteristic of postpositions, i.e. of the place where postpositional sandhi operates, as explained above.

4) Most Indic languages write the postposition together with the noun.

#### The Pre-postposition bi ...-qo

As we have mentioned above, Romani has a pre-postposition with the concept «without». The prepositional part is written separately, since another element can intervene between it and the noun, while the postpositional part is written together with the noun:

|                         |                             |
|-------------------------|-----------------------------|
| bi kheresqo             | «homeless»                  |
| bi limoresqo            | «without a grave»           |
| bi me dadesqo           | «without my father»         |
| bi akale purane gadenqo | «without those old clothes» |

In conjunction with a personal pronoun, the respective form of the personal pronoun is usually used. Compare:

me «I», mança «with me», miro «my, mine», bi miro «without me»  
tu «you», tuça «with you», tiro «your, yours», bi tiro «without you»,  
etc.

The form with **bi ...-qo**, which is encountered in very few dialects (**bi manqo**, **bi tuqo**), is not recommended in the Common Language.

Some dialects have replaced the pre-postposition **bi ...-qo** with the preposition **biz**, which probably originates from the Indo-Iranian **bi/be**, but with contamination from the Slavic preposition **bez** (**biz-o kher** «homeless», etc.) The proper Romani form **bi ...-qo** is preferred in the Common Language.

In Hindi the concept «without» can also be expressed otherwise by means of a pre-postposition: **binā ... kā**. Compare:

|         |                     |                 |
|---------|---------------------|-----------------|
| Hindi:  | <b>binā pani kā</b> | «without water» |
| Romani: | <b>bi paněsqo</b>   | «without water» |

|         |                     |                |
|---------|---------------------|----------------|
| Hindi:  | <b>binā paer kā</b> | «without legs» |
| Romani: | <b>bi piresqo</b>   | «without legs» |

We find this same structure in other languages of India, such as Punjabi: **bina ... te/de**. In the Indic languages the element **binā** can be used after the substantive (in Nepali, for example, it is used only after the substantive), whereas in Romani **bi** can not be used as a postposition.

## THE VOCATIVE

Romani is characterized by a considerably developed system of vocative forms. In fact, like the imperative in the verbal system, the vocative in the nominal system forms a kind of marginal subsystem, and many linguists do not accept its status as a true case. It is a form which is used outside of sentence structure, for the purpose of calling a person or an animal, or also a thing in a poetic or allegorical context.

I) In most native words the vocative is expressed by means of a particle which is added to the A-Form or the B-Form of the noun:

a) for singular masculine nouns the particle is **-a** or **-ana** and is usually added to the B-Form without the final **-s**:

|                                    |        |
|------------------------------------|--------|
| <b>rrome!a</b> or <b>rrome!ana</b> | «Rom!» |
| <b>čhave!a</b> or <b>čhave!ana</b> | «Boy!» |

In some nouns ending in a consonant which denote a person, the particle can also be added to the B-Form, but then only the variants **-a** or **-e** are used:

|                     |           |
|---------------------|-----------|
| <b>rrom!a</b>       | «Rom!»    |
| <b>manuś!a</b>      | «Man!»    |
| <b>dad!e, dad!a</b> | «Father!» |

This form is quite rare for nouns of the type **čhavo**, but can be found, especially in the dialects of the Soviet Union. In this case the final **-o** is lost in the A-Form: **čhav!a**.

The noun **dad** «father» has a considerable number of vocative forms. Besides those we have mentioned above, there exist also the forms **dad!o** and **dad!ole**. The forms **babi**, **babo**, **tato**, etc., are also used in this meaning.

As can be noticed the exclamation point [!] is inserted in writing to connect the vocative particle with the main word, while representing the emphatic intonation which accompanies the vocative in speech. The exclamation point serves to distinguish the vocative particle from the homophonous plural ending of the postposition **-ça** in those dialects where it is pronounced [-a], e.g.:

|                |                |  |
|----------------|----------------|--|
| <b>manuśa</b>  | «men»          | [manuśa]                               |
| <b>manuś!a</b> | «Man!»         | [manuśa]                               |
| <b>čhaveça</b> | «with the boy» | [čhavea] or [čhaveja] in many dialects |
| <b>čhave!a</b> | «Boy!»         | [čhavea] or [čhaveja]                  |

This also indicates the vocative in words that have no special form and do not take the vocative particle (see below).

We can say that, analogous with the postpositional system, the exclamation point has two functions:

- a morphological function, which is also semantic here, in that it indicates that the word is the vocative forms;
- a phonological function, since when it occurs between vowels it is pronounced [j] in many dialects: **phralea** [phraleja], etc.

Unlike in postpositions the vocative does not indicate a place of stress, for the reason that it enjoys much freedom in connection with intonation and the emotional change which these forms characterize.

b) for singular feminine nouns, the particle is **-e** or **-ene** which is added to the A-Form:

|                                       |           |
|---------------------------------------|-----------|
| <b>rromni!le</b> or <b>rromni!ene</b> | «Woman!»  |
| <b>bori!e</b> or <b>bori!ene</b>      | «Bride!»  |
| <b>daj!e</b> or <b>daj!ene</b>        | «Mother!» |

An **i** can also be interposed:

|                                   |                                       |
|-----------------------------------|---------------------------------------|
| <b>phen!e</b> or <b>pheni!ene</b> | «Sister!» (from <b>phen</b> «sister») |
|-----------------------------------|---------------------------------------|

c) for plural masculine and feminine nouns, the particle is **-len** or **-le**, and is added to the A-Form:

|                     |              |
|---------------------|--------------|
| <b>rroma!le(n)</b>  | «Rom! (pl.)» |
| <b>phrala!le(n)</b> | «Brothers!»  |

Under the influence of analogy, nouns ending in **-o** in the singular usually have the ending **-a** before the vocative particle:

|                    |         |
|--------------------|---------|
| <b>čhava!le(n)</b> | «Boys!» |
|--------------------|---------|

In dialects of the Soviet Union the particle **-le(n)** is also sometimes encountered in the singular, especially after **-o**: **dad!ole** «Father!», **bib!ole** «Aunt!»

II) Foreign masculine nouns usually do not change their form for the vocative, however they may drop the final epenthetic vowel, especially when this is **-i**:

|                                   |           |
|-----------------------------------|-----------|
| <b>Ismet!</b>                     | «Ismet!»  |
| <b>Nikol!</b>                     | «Nikola!» |
| <b>doktori!</b> or <b>doktor!</b> | «Doctor!» |
| <b>trajo!</b>                     | «Life!»   |

Foreign feminine nouns ending in unstressed **-a** in the singular, usually take the ending **-o** in the vocative (which is characteristic of the vocative in Slavic languages), to which **-ne** may also be added:



|             |                           |
|-------------|---------------------------|
| Dudo!       | «Duda!»                   |
| Irko!       | «Irka!»                   |
| maćko!ne    | «Kitty!»                  |
| kalujero!ne | «Sister! (i.e. to a nun)» |

This ending sometimes appears in words of Romani origin:

|                     |                                       |
|---------------------|---------------------------------------|
| bib!o or bibi!o(ne) | «Aunt!» (from <b>bibi</b> «aunt»)     |
| pheni!o             | «Sister!» (from <b>phen</b> «sister») |

Feminine proper nouns frequently lose the final -a or replace it with -e:

|         |           |
|---------|-----------|
| Zejnep! | «Zeynep!» |
| Lumnie! | «Lumnie!» |

The -a may also be retained:

|        |          |
|--------|----------|
| Arifa! | «Arifa!» |
|--------|----------|

The diminutive suffix is also used with a vocative value, so that some nouns can display a series of different forms. The vocative of the name **Demiràna** can have the following vocative forms: **Demirano!** - **Demirani!le** - **Demirane!** - **Demiran!** - **Demiran!ko** - **Demiran!ke**, etc., as well as **Demirana!** Nevertheless, some forms do not occur, e.g. **\*Demirano!ne** or **\*Dud** (from **Dùda**). It can be stated that, for practical purposes, there are no clear-cut rules for the formation of a system so vital in the language and dependent on the emotion of the moment as is the vocative. Besides this diversity, in some places the vocative of the language from which the proper noun is derived is often used, as, for example, in some masculine and feminine names: **Dragane!** **Bobane!** in Yugoslavia, or **Trajane!** in Romania.

### III) The Vocative in Phrases

In the colloquial style, phrases consisting of two or three words do not use the vocative forms at all, especially after pronouns:

|                            |                                       |
|----------------------------|---------------------------------------|
| hàjde mi phen!             | «Come my sister!»                     |
| mo laćho amal!             | «My good friend!»                     |
| sov mi cinoni, sov mi ćhaj | «Sleep my little one, sleep my girl!» |

The vocative is widely used in poetry, and sometimes also after pronouns:

|                          |                    |
|--------------------------|--------------------|
| kher!a mirre!a           | «My house!»        |
| berše!a bibaxtale!a      | «Unlucky year!»    |
| daj!e amari!e            | «My mother!»       |
| phuv!ene amari!e         | «Our land!»        |
| o borile, lele tu šuži!e | «Beautiful bride!» |

## THE ARTICLE

The Romani article comes before the noun. Several studies have proposed from its position and form that it is a foreign borrowing for a Greek dialect. However, without discounting the importance of Greek in the development of the article, we cannot exclude the possibility that it was formed on the basis of Romani linguistic material, more concretely, from a short form of the demonstrative pronoun. A comparison of demonstrative forms and forms of the article reveal traits similar to the demonstratives.

Like the noun, the Romani article shares certain forms with the noun it accompanies. Its forms differ greatly from dialect to dialect, but we can discern four main types, which are really just variants of each other. Firstly, the dialects of Central Europe differ from those of the Balkans in that the B-Form of the article is characterized by the presence of the l- (le, la, ol), while in the Balkans the only form in the B-Form is e. Secondly, in both groups the plural of the A-Form differentiates two subgroups on the basis of their comparison with the A-Form (group I) or with the B-Form (group II) of the masculine singular:

|                  |        | Group I |     |      |     | Group II |     |      |     |
|------------------|--------|---------|-----|------|-----|----------|-----|------|-----|
|                  |        | masc.   |     | fem. |     | masc.    |     | fem. |     |
|                  |        | sg.     | pl. | sg.  | pl. | sg.      | pl. | sg.  | pl. |
| Balkan           | A-Form | o       | o   | i    | o   | o        | e   | i    | o   |
|                  | B-Form | e       | e   | e    | e   | e        | e   | e    | e   |
| Central European | A-Form | o       | o   | e    | o   | o        | le  | e    | le  |
|                  | B-Form | le      | le  | la   | le  | le       | le  | la   | la  |

### Notes on variants:

- 1) In Group I, two variants are encountered: i and e for [f-sg-A]. In the Common Language the variant with i is recommended, because it is not a homonym with the plural and thus is more informative than the variant e;
- 2) In Hungary some dialects prefix a stressed o- to the B-Form of the article: òle, òle, òla, òle, which giving it a kind of emphasis.
- 3) In some dialects of the Ukraine, the B-Form of the feminine singular, which is la in the Central European dialect, loses its special distinction and is identified with other B-Forms, which are le. This simplification is not accepted in the Common Language;

4) In Poland some dialects exhibit a form intermediate between the two main dialects: the B-Form of the feminine singular is distinct from other forms, however it does not have the initial -l, so that the forms are: e, e, a, e;

5) The Drindari dialect of central Bulgaria (Sliven) has undergone a change in most words whereby o and e have become u and i. This phenomenon also affects the article: **Ti dikhil i däj u šeroru, i Aliesqu šeroru upr-u dromoru** «Then the mother saw the head, Ali's head, on the road».

6) Some dialects of the Central European group elide the article's vowel when it occurs after another vowel: **sa l rrom(a)** «all the Rom», **angla l deš (càsură)** «before 10 o'clock», etc. There are cases where the article is pronounced [äl] or [ol] before a vowel: **ol averenča** «with the others». These variants are purely vocal and need not be expressed in writing, which always uses **le** (except in poetry, where it shows elision): **sa le rrom, le averençar**.

The recommended forms for the article are:

|                          |        | masculine |     | feminine |     |
|--------------------------|--------|-----------|-----|----------|-----|
|                          |        | sing.     | pl. | sing.    | pl. |
| Balkan variant           | A-Form | o         | o   | i        | o   |
|                          | B-Form | e         | e   | e        | e   |
| Central European variant | A-Form | o         | le  | e/i      | le  |
|                          | B-Form | le        | le  | la       | le  |

Note: some studies express the opinion that the final unstressed -o, which characterize foreign masculine nouns in many dialects, is a vestige of the Macedonian article -o(t): **brègo** «hill», **trùpo** «body», **kongrèso** «congress», etc. This hypothesis is interesting, and is under discussion.

In its favor we can mention the fact that some dialects have foreign words with a final -va: **kafàva** «coffee», **basràva** «fairy-tale», **knigàva** «book» (in the Common Language **pust/i, -ă, fem.**), **lulàva** «pipe», etc., while keeping in mind that Macedonian has the three feminine articles: -ta, -va, and -na. However, the etymology based on the Romanian article -ua, which occurs in oxytone feminine nouns, is more persuasive, e.g. **basma** «kerchief», **basmaua** «the kerchief».

There are other factors against this hypothesis: the ending -o is encountered throughout Europe, even among the Rom who have never resided in Macedonia. Not only is -o found, but -os, as well, in dialects which have not lost final -s: **trùpos** «body», **pràxos** «ash, dust». We find this ending not only in the Balkans, but also further

beyond, as far as Poland and Sweden, where, for example, the Hungarian word *àgy* «bed» is borrowed as *jàzos*. On the other hand, many genuine Macedonian words have not been borrowed with the -o ending, but rather with an unstressed final -i. In fact, some accept both variants: *trùpo* or *trùpi* «body», *bròdo* or *bròdi* «ford», *zakòno* or *zakòni* «law», etc. We could also attempt to explain this phenomenon as Albanian influence, by comparing the Albanian article forms: -i or -u. However, this explanation also does not hold, for the reason that the Albanian and Romani endings do not usually correspond, as, for example, in *kòso* «yoghurt» (Albanian has the article -i: *kosi* «the yoghurt») or *bubrègo* «kidney» (in Albanian has -u: *bubregu*), etc.

In any case, from a functional standpoint, these purely epenthetic endings have no synchronic relation to articles and can in no way be considered articles.

## ADJECTIVES

Romani adjectives can be divided into four groups:

I) A group of adjectives with endings which look like those of nouns: **bakro** «ram». This group, which comprises the majority of adjectives, is called "long," e.g. **buxlo** «wide», **lačo** «good»;

II) A group, of limited number, formed from adjectives which end in a consonant in the singular of the A-Form, e.g. **tang** «narrow», **nasul** «wicked». This group is called «short»;

III) A group of foreign words, especially of Slavic origin, with initial stress: **zèleno** «green», **slàbo** «weak»;

IV) A very small group of oxytone foreign words, belonging to the international lexicon, primarily of French origin: **sote** «sauteed», **dore** «gilt», **bordo** «bordeaux colored», and **gri** «grey». As in Albanian, these adjectives are invariable, where as the others decline.

As a rule, when accompanying a noun, adjectives follow a pure declension, with a common B-Form for both genders and numbers, whereas when they are nominalized, they follow a noun declension. In particular, the distinction between the genders of the plural of B-Form consists of the opposition of **ě** (masc.) / **ã** (fem.) in the Central European dialect branch, whereas in the Balkan branch this distinction is expressed by means of the opposition **e** (masc.) / **ie** (fem.). The feminine **i** can be changed into a pre-jotizing vowel, whereby we then have the opposition: **e** (masc.) / **ě** (fem.), or the opposition may disappear completely by neutralizing the gender expression. However, it is recommended that distinction between genders be preserved in the Common Language.

## THE DECLENSION OF ADJECTIVES

I) The group of "long" adjectives:

**BARO** «big», «important»

Balkan dialect branch: Dependent declension (i.e. in which the adjective modifies a noun):

|      | masculine                    |               | feminine                          |                |
|------|------------------------------|---------------|-----------------------------------|----------------|
| Form | o baro rom «the big Rom man» |               | i bari rromni «the big Rom woman» |                |
|      | singular                     | plural        | singular                          | plural         |
| A    | o baro rom                   | e bare rroma  | i bari rromni                     | e bare rromnă  |
| B    | e bare rromes                | e bare rromen | e bare rromnă<br>e bară rromnă    | e bare rromnën |

(It is given that the variants -nia and -nă coexist, as do -nien and -nën, although we show only the short forms in the table).

- independent declension (i.e. used when the adjective is nominalized):

|      | masculine |         | feminine |         |
|------|-----------|---------|----------|---------|
| Form | singular  | plural  | singular | plural  |
| A    | o baro    | e bare  | i bari   | e bară  |
| B    | e bare(s) | e baren | e bară   | e barën |

Central European dialect branch: Dependent forms:

|      | masculine                     |               | feminine                          |                 |
|------|-------------------------------|---------------|-----------------------------------|-----------------|
| Form | o baro rrom «the big Rom man» |               | i bari rromni «the big Rom woman» |                 |
|      | singular                      | plural        | singular                          | plural          |
| A    | o baro rom                    | le bare roma  | i bari rromni                     | le bare rromnă  |
| B    | le bare romes                 | le bare romen | la bară rromnă                    | le bare rromnăn |

- Independent forms:

|      | masculine  |          | feminine |          |
|------|------------|----------|----------|----------|
| Form | singular   | plural   | singular | plural   |
| A    | o baro     | le bare  | i bari   | le bară  |
| B    | le bare(s) | le baren | la bară  | le barăn |

PURANO «old»

Balkan dialect branch:

In the Balkan dialects, types such as purano do not differ from the type baro.

In the Central European dialect branch, the dependent forms are:

|      | masculine                     |                   | feminine                    |                  |
|------|-------------------------------|-------------------|-----------------------------|------------------|
| Form | o purano kher «the old house» |                   | i puraji piri «the old pan» |                  |
|      | singular                      | plural            | singular                    | plural           |
| A    | o purano kher                 | le purane khera   | i puraji piri               | le purane piră   |
| B    | le purane kheres-             | le purane kheren- | la puraja piră-             | le purane pirăn- |

The independent forms are:

|      | masculine    |            | feminine  |            |
|------|--------------|------------|-----------|------------|
| Form | singular     | plural     | singular  | plural     |
| A    | o purano     | le purane  | i puraji  | le puraja  |
| B    | le purane(s) | le puranen | la puraja | le purajan |

The correspondence of -ji or -j in the northern dialects, as opposed to -ni in the other dialects, is rather widespread, not only in adjectives as shown here, but also in nouns: paj versus pani «water», kuj versus kuni «elbow», zej versus zeni «female», etc.

In dialects with palatalization of the velars k, kh, and g to [k̟], [kh̟], and [g̟] or [č], [čh], and [j] before front and pre-jotizing vowels, the endings -i, -e, and -ă cause this palatalization in words such as šuko «dry», lokho «light», and

bango «crooked».

|          |                    |                           |                    |
|----------|--------------------|---------------------------|--------------------|
| [m-sg-A] | ending in -i       | ending in -e              | ending in -ă*      |
|          | [f-sg-A]           | [mf-pl-A]<br>[mf-sg/pl-B] | [f-sg-B]           |
| śuko     | śuki [-ki, ci]     | śuke [-ke, -ce]           | śukă [-ka, -ca]    |
| lokho    | lokhi [-khi, -chi] | lokhe [-khe, -che]        | lokhă [-kha, -cha] |
| bango    | bangi [-gi, -ji]   | bange [-ge, -je]          | bangă [-ga, -ja]   |

\*) The palatalization of the stem root consonant by the ending -ă of the B-Form of the feminine singular [f-sg-B] appears in several dialects (e.g. Mahažeri) also in adjectives, the root of which contain a dental t or d: mată, the oblique form of mati «drunk», pronounced in these dialects as [maća] or [maka]. The latter pronunciation is most widespread in Macedonia, but is also found in Bulgaria.

II) The Group of "Short" Adjectives:

**GODĂVER «clever»**

The Balkan branch: Dependent forms:

|      | masculine                            |                        | feminine                             |                       |
|------|--------------------------------------|------------------------|--------------------------------------|-----------------------|
| Form | o godăver čhavo «the clever Rom boy» |                        | i godăver čhaj «the clever Rom girl» |                       |
|      | singular                             | plural                 | singular                             | plural                |
| A    | o godăver<br>čhavo                   | e godăver(e)<br>čhave  | i godăver<br>čhaj                    | e godăver(a)<br>čhia  |
| B    | e godăver(e)<br>čhave(s)             | e godăver(a)<br>čhaven | e godăver(e)<br>or godăveră<br>čhia  | e godăver(a)<br>čhien |



- Independent forms:

|      | masculine     |              | feminine  |              |
|------|---------------|--------------|-----------|--------------|
| Form | singular      | plural       | singular  | plural       |
| A    | o godäver     | e godäver(a) | i godäver | e godäver(a) |
| B    | e godävere(s) | e godäveren  | e godäver | e godäverën  |

Central European dialect branch: Dependent forms:

|      | masculine                            |                      | feminine                             |                     |
|------|--------------------------------------|----------------------|--------------------------------------|---------------------|
| Form | o godäver čhavo «the clever Rom boy» |                      | i godäver čhej «the clever Rom girl» |                     |
|      | singular                             | plural               | singular                             | plural              |
| A    | o godäver čhavo                      | le godäver čhave     | i godäver čhej                       | e godäver(a) čha    |
| B    | le godäver(e) čhave(s)               | le godäver(e) čhaven | la godäver(a) čhia                   | le godäver(än) čhan |

A very limited number of adjectives follow this type of declension. The most common are: godäver «clever», aver «other», šukar «good», nasul «wicked», kuć «costly», nakuć (or bikuć) «gratis» (in the Soviet Union), đur «far», but «numerous», tang «narrow», xor «deep», as well as adjectives with the comparative ending -er.

III) The Group of Foreign Adjectives

SLÀBO «WEAK»

Balkan branch: Dependent forms:

|      | masculine                        |                                      | feminine                           |                                    |
|------|----------------------------------|--------------------------------------|------------------------------------|------------------------------------|
| Form | o slàbo čhavo «the weak Rom boy» |                                      | i slàbo čhaj «the weak Rom girl»   |                                    |
|      | singular                         | plural                               | singular                           | plural                             |
| A    | o slàbo čhavo                    | e slàbo čhave<br>or<br>e slàbā čhave | i slàbo čhaj<br>or<br>i slàbi čhaj | e slàbo čhia<br>or<br>e slàbā čhia |
| B    | e slàbo(ne) čhave(s)             | e slàbo(ne) čhaven                   | e slàbo(ne) čhia                   | e slàbo(ne) čhien                  |

The independent forms are identical with the long dependent forms.

Central European branch: Dependent forms:

LÛNGO «LONG»

|      | masculine                    |                                      | feminine                      |                                |
|------|------------------------------|--------------------------------------|-------------------------------|--------------------------------|
| Form | o lùngo drom «the long road» |                                      | i lùngo baj «the long sleeve» |                                |
|      | singular                     | plural                               | singular                      | plural                         |
| A    | o lùngo drom                 | le lùngo drom(a) or le lùnzi drom(a) | i lùngo baj                   | le lùngo baja or le lùnzi baja |
| B    | le lùngo(ne) dromes-         | le lùngo(ne) dromen-                 | la lùngo(ne) or lùngoja baja- | le lùngo(na) bajen-            |

The ending -i of the plural A-Form also often causes a change in the final consonant of the adjective's stem, especially in the Central European dialects. This phenomenon is also observed in nouns in these dialects, e.g. amperàto «king», pl. amperàci (this word is now replaced by the Romani word thagar, with the regular plural thagara).

The main changes are as follows:

| singular | plural |
|----------|--------|
| -to -ko  | -ci    |
| -do -go  | -zi    |

as seen in the above example, the singular lùngo has the corresponding plural lùnzi.

It is worth noting that in this case the foreign word lùngo (from Romanian) can be very well replaced with the Romani word uéo «high, long» or dur «far, long», as these expressions demonstrate: uće baja «long sleeves», isi jekh drom but dur «there is a long road», while the word of Slavic origin slàbo can be replaced by the Romani words: kiślo, śuko, bazoralo, etc.

Participle with the endings -mo or -me (which derive from foreign verbs ending in -in- personal ending) appear with invariable endings in some dialects of both groups. In the southern group, e.g. in Mahažeri, Tharegoni, etc., these adjectives are also often treated like "long" adjectives with the declensions o

[m-sg-A], i [f-sg-A], and e for all forms, while in the northern group there exist, besides the invariable form *-me*, a form *-mene*, which appears in the B-Form of both genders and numbers, as seen below:

|                         | invariable variant | variable variant                                    |
|-------------------------|--------------------|---|
| Balkan branch           | želime «married»   | [m-sg-A] želimo<br>[f-sg-A] želimi<br>others želime |
| Central European branch | ansurime «married» | A-Form ansurime<br>B-Form ansurimene                |

#### THE COMPARATIVE DEGREE OF ADJECTIVES

The comparative degree is formed with the native suffix *-der* which is encountered in Romani dialects from the Balkans to Scandinavia, however, it is limited to a small number of adjectives, of which the most frequent are: *bareder* «bigger», *tikneder* «smaller», *terneder* «younger», *phureder* «older», etc. The suffix *-der* is also used with adverbs: *angleder* «more ahead», *paleder* «more behind, afterwards». The adjective *lačho* «good» has a special comparative *fededer* or *feder* «better» in Central Europe, while in Skopje Shaip Jusuf notes the regular form *lačheder*.

The comparative degree is also formed analytically with a foreign particle, as, e.g. *maj* «more» (in Central Europe, from Romanian), or *po* (in the Balkans, from Bulgarian), which precedes the adjective: *maj baro* or *po baro* «bigger». This particle is variously written with the adjective or separate, but it is recommended to write the particle and adjective together, since no other element may intervene between them, and they form one stress unit. In some dialects, e.g. Mećkari, there also exists a comparative which is expressed only by an object and without a morphological affix: *ov isi barvalo man* «he is richer than I», *to rrom si terno mires* «your man is younger than mine». This archaism is not recommended in the Common Language.

The superlative degree is formed in most cases by placing an article before the comparative form, however some dialects have the special particle *naj*, also of Slavic origin, while others add the Turkish particle *em-* to the positive or comparative form of the adjective. The number *jekh* «one» is also sometimes used in this function, e.g. in some dialects of Czechoslovakia: *jekhbuder* or *jekhbutar* «more numerous».

|      | Comparative                                  | Superlative   |
|------|--|---|
| baro | <b>bareder</b>                               | o bareder<br>o embaro<br>o embareder<br>o jekhbareder |
|      | pobareder<br><b>majbaro</b><br><b>pobaro</b> | o majbaro<br>o najbaro<br>o pobaro                    |

(The preferred forms in the Common Language are those in bold).

The object of the comparative is usually marked with the postposition -@ar: **bareder man@ar** «larger than I», but when contrasting two nouns which bear this postposition the conjunctions **sarobor** or **akatar** are used: **ov thulilo maj but e soimas@ar sarobor e xamas@ar** «he recovered more from the standing water than from the food». These conjunctions can also be used to avoid ambiguity or to emphasize the comparison: **sarobor man@e, nane manus baxtaleder** "there is no man happier than I», **o ś vakārel pes cārra popalal akatar-o than e s-go** «ś is pronounced further back than s», **śaj phenes e vokāle pozorale krleça akatar-e konsonānte** «we can pronounce vowels louder than we can consonants».

## PERSONAL PRONOUNS

Personal pronouns follow the same declension as do nouns, however the 1st and 2nd persons in both the singular and plural do not take the possessive postposition (-qWrW). This form is substituted by possessive adjectives.

As far as the pre-postposition **bi ... -go** is concerned, it can be used with the personal pronouns, e.g. **bi mango** «without me», **bi amengo** «without us», however the use of the particle **bi** with the possessive pronoun is usually preferred: **bimiro** «without me», **biamaro** «without us». In the 2nds person singular only the possessive pronoun in the long form is possible: **bitiro** «without you». The postpositional structure **biz-o** or **biz-u** are also encountered.

Forms of the personal pronouns:

|                 | Form                                  | singular                  | plural                              |
|-----------------|---------------------------------------|---------------------------|-------------------------------------|
| 1st p.          | A-Form<br>B-Form<br>with postposition | me<br>man, ma<br>man-     | amen, ame<br>amen, ame<br>amen-     |
| 2nd p.          | A-Form<br>B-Form<br>with postposition | tu<br>tut, tu<br>tu-      | tumen, tume<br>tumen, tume<br>tume- |
| 3rd p.<br>masc. | A-Form<br>B-Form<br>with postposition | ov<br>(o)le(s)<br>(o)les- | on<br>(o)len<br>(o)len-             |
| 3rd p.<br>fem.  | A-Form<br>B-Form<br>with postposition | oj<br>(o)la<br>(o)la-     | on<br>(o)len<br>(o)len-             |

The presence of a **v-** or **j-** before the A-Form of the 3rd person personal pronouns is very widespread: **vo(v)**, **vo(j)**, **von**, **jov**, **joj**, **jon**. The B-Form with an initial **o-** in the 3rd person usually conveys emphasis.

## PERSONAL REFLEXIVE PRONOUNS

Romani can express the reflexive in several different ways:

a) by the conjugation of a true reflexive verb: **mudardövvav** «I am being killed» (see Verb Morphology below);

b) by the use of an accusative reflexive pronoun as a direct object after the verb: **mudarav ma(n)** or **mudara'ma(n)** «I kill myself»;

c) by the use of the reflexive pronoun **pes** for all persons (this pronoun lacks the A-Form): **mudarav pe(s)**. This structure is quite widespread in several dialects, and seems to follow the Slavic model;

d) some dialects can combine types a and b to emphasize the indefiniteness of the concept (only in the 3rd person singular), e.g. **na žandöl pe(s)** «it is not known».

|        | active         | reflexive         |                     |                    |
|--------|----------------|-------------------|---------------------|--------------------|
|        |                | type a            | type b              | type c             |
| 1st p. | <b>mudarav</b> | <b>mudardövav</b> | <b>mudara'ma(n)</b> | <b>mudarav pes</b> |
| 2nd p. | <b>mudares</b> | <b>mudardös</b>   | <b>mudares tut</b>  | <b>mudares pes</b> |
| 3rd p. | <b>mudarel</b> | <b>mudardöl</b>   | <b>mudarel pes</b>  | <b>mudarel pes</b> |

The use of these different types is regulated by custom, however, in general, the reflexive verb form has a passive meaning, whereas the forms with the pronouns have a reflexive meaning: **mudardövav** «I am killed» as opposed to **mudara'man** or **mudardövav** «I am killing». This conceptual distinction needs to be clarified in the Common Language.

In the case of the 3rd person plural, besides the form **pes**, there is an equivalent plural form **pumen**, e.g. **mudaren pumen** «they kill themselves», the use of which is usually advised. In Central Europe in place of **pumen** the short variant **pes** is used.

## POSSESSIVE ADJECTIVES

The variety of forms for possessive adjectives is quite larger throughout the Romani dialects. In most dialects we have two series of forms for the 1st and 2nd persons singular, which are:

a) the long form, which, like the long form of the postposition **-qoro**, is used in either a invariable manner (denominative), or to emphasize possession, as apposition to the noun, e.g.:

- the invariable usage: **sakon rodel plo** «he, who is seeking what is his own»

- emphasizing possession: **akala čavorre mirre** «my (own) children» (the usual form is **me čavorre**).

The short form varies little among the dialects. The main difference is the presence or absence of palatalization of **t** in the 2nd person.

In the plural there is only one form for each person.

| Form  | singular                              |                                  | plural           |                   |
|-------|---------------------------------------|----------------------------------|------------------|-------------------|
|       | 1st p.<br>(me)                        | 2nd p.<br>(tu)                   | 1st p.<br>(amen) | 2nd p.<br>(tumen) |
| short | mo                                    | to, tō                           | amaro            | tumaro            |
| long  | miro<br>mārro<br>mrro<br>mlo<br>murro | tiro, tī<br><br><br>tlo<br>torro |                  |                   |

In some Balkan dialects there exists also a form **ko** (short) and **klo** (long) for the 2nd person, but it is not recommended in the Common Language.

The role of the 3rd person possessive adjectives is played by the association of the personal pronoun in the B-Form with the possessive postposition **-qoro** (we give here examples in the singular A-Form, i.e. with the ending **-o**):

masc.: (o)lesqo(r)o / olesqo, lesqo, lesqoro, olesqiri, lesqiri, lesqoro, olesqi, lesqre, etc.

fem: (o)laqo(ro) / olaqo, laqo, laqoro, olaqiri, laqero, olaqe, olaqi, laqre, etc.

The reflexive possessive adjectives are formed in the same manner, on the basis of the reflexive pronoun **pes**, i.e. **pesqo(ro)**

/ **pesqo, pesqoro, pesqiri, pesqi**, etc., alongside another analogous form the 1st and 2nd person singular possessive adjectives, which is **po** in the short form and **plo, piro** in the long form. Besides these reflexive forms which can be used in the singular as well as the plural, there are also solely plural forms:

- on the basis of the reflexive possessive adjective: **penqo(ro) / penqo, penqoro, penqiri, penqero, penqi, penqre**, etc.;

- by analogy with the 1st and 2nd persons plural: **pumaro**.

The 3rd person forms of the possessive adjectives, when the thing possessed is in the masculine singular, i.e. [m-sg-A] with the ending **-o**, are as follows (we give here the long forms):

|    | singular      |           | sg. or pl.        | plural  |        |               |
|----|---------------|-----------|-------------------|---------|--------|---------------|
| m. | lesqoro       | pesqoro   | po<br>plo<br>piro | penqoro | pumaro | lenqoro       |
| f. | laqoro        |           |                   |         |        |               |
|    | non-reflexive | reflexive |                   |         |        | non-reflexive |

There also exist forms of the non-reflexive possessives with initial **o-**: **olesqoro, olaqoro, olenqoro**, etc.



## DEMONSTRATIVES

Romani is very rich in demonstrative adjectives and pronouns. The forms vary, but always are connected with basic types which are found in all dialects.

The demonstrative pronoun is a simple denominative of the demonstrative adjective. This denomination is accomplished as in other adjectives (we will give examples later). Thus we can speak generally about "demonstratives" rather than about demonstrative pronouns and demonstrative adjectives.

Each demonstrative is formed from a stem and an ending. Occasionally the stem and ending are confused so that the boundary between them is not clear, however in most cases the constituent elements are distinguishable.

The most frequent stems are: *ada-*, *odo-*, *aka-*, *oko-*, *kada-*, *kodo-*, *kado-*, *ka-*, *ko-*, *ga-*, *go-*, *kako-*, and *kuko-*. The stems *dou-* and *kou-* are also encountered. These contain certain characteristic elements which are in opposition and serve to show proximity or distance of the object modified:

**a/o:** in two series of demonstratives, one with the characteristic vowel *o* or *u*, and the other with the characteristic vowel *a*, the one with *o* indicates a distant object, whereas the one with *a* indicates a near object.

**k/d or k/g:** in two series of demonstratives, one with the characteristic consonant *d* or *g*, and the other with the characteristic consonant *k*, the one with *d* or *g* indicates a distant object, whereas the one with *k* indicates a near object.

The demonstrative agrees in gender, number, and case with the noun it modified. This agreement is indicated by the endings, which are numerous for all demonstratives, as seen below:

|        | masculine |        | feminine |        |
|--------|-----------|--------|----------|--------|
|        | singular  | plural | singular | plural |
| A-Form | -va       | -la    | -ja      | -la    |
| B-Form | -le       | -le    | -la      | -le    |

It can be seen that the plural endings are the same for both genders: *-la* for the A-Form and *-le* for the B-Form. In fact, some dialects, e.g. *Çurar* of Romania, have only one form (*-le*) for both plural forms, and greatly resemble the inflection of the article (keeping in mind the resemblance between *o* and *-va* [probably from *u+a* or *o+a*] on one hand, and between *i* and *-ja* [probably from *i+a*] on the other). We give the forms of the

article in brackets:

|        | masculine |          | feminine |          |
|--------|-----------|----------|----------|----------|
|        | singular  | plural   | singular | plural   |
| A-Form | -va [o]   | -le [le] | -ja [i]  | -le [le] |
| B-Form | -le [le]  | -le [le] | -la [la] | -le [le] |

Some dialects have lost the final vowel of the stem in the demonstrative before the feminine singular Form-A endings [f-sg-A], which is then pronounced [-ia]. This endings does not palatalize the occlusive which precedes it, which is indicated in writing by two dots above the vowel which is not pronounced. This letter is also adopted in the Gurbeti dialect, where a does not disappear, but has a back pronunciation, like the Romanian î or some cases of ë in Albanian: akäja vs. akaja, gäja vs. gaja, etc., whereas in dialects where the ending -ia changes to -ä, this palatalizes the consonant which precedes it (naturally only in dialects that have the palatalization rule): e.g. kodä - pronounced [kodja] or [koja] according to dialect. We also have kadä [kadja] or [kaja], kakä [kakä] or [kaka, kaća], kukä [kuća, kuća], etc.

In the Gurbeti and Kalderaš dialects the demonstratives, which are shorter than in other dialects, may loose the final vowel (a or o), yielding: kav vs. kava, gol vs. gola, etc. In many cases, especially in the northern group, the ending -va is lost in the masculine singular A-Form [m-sg-A]: ada vs. adava, kodo vs. kodova, etc.

A number of dialects of Romania, Bulgaria, and Greece present, besides the usual series of demonstratives, another series with k which intervenes before the final vowel ending (a or e): -vka, -lka, -lke and -ika. The form with the intervening -k- is, as a matter of fact, the only forms in Erli of Sofia. One should note that this dialect, as a rule, losses final stem vowels before the ending -ia and its variant -ika. This can be indicted by two dots above the letter which is not pronounced, as we have said above, or is not indicated at all, for the reason that in this dialects there is no palatalization of velar occlusives by -i, and the presence of i after a velar cannot be read in a mistaken manner.

## Demonstrative Forms

| 1. In the Balkan Branch |                    |                   |                |                   |
|-------------------------|--------------------|-------------------|----------------|-------------------|
|                         | masculine          |                   | feminine       |                   |
|                         | singular           | plural            | singular       | plural            |
| A-Form<br>B-Form        | akava<br>akale(s)  | akala<br>akale(n) | akaja<br>akala | akala<br>akale(n) |
| A-Form<br>B-Form        | okova<br>okole(s)  | okola<br>okole(n) | okoja<br>okola | okola<br>okole(n) |
| A-Form<br>B-Form        | adava†<br>adale(s) | adala<br>adale(n) | adaja<br>adala | adala<br>adale(n) |
| A-Form<br>B-Form        | odova<br>odole(s)  | odola<br>odole(n) | odoja<br>odola | odola<br>odole(n) |

†The form *adava* is relatively rare.

Note: the forms in the parentheses are the full forms used when the demonstrative functions as a pronoun (see below). In this case the form [f-pl-B] can also have an *i* in the ending: *alakien*, *okolien*, etc.

Example with an intervening *-k-*:

|                  | masculine           |                     | feminine        |                     |
|------------------|---------------------|---------------------|-----------------|---------------------|
|                  | singular            | plural              | singular        | plural              |
| A-Form<br>B-Form | akavka<br>akalke(s) | akalka<br>akalke(n) | akaja<br>akalka | akalka<br>akalke(n) |

also *okovka*, *odovka*, etc.

| 1. In the Central European Branch |                                    |  |                              |  |
|-----------------------------------|------------------------------------|--|------------------------------|--|
|                                   | masculine                          |  | feminine                     |  |
|                                   | singular                           | plural   | singular                     | plural   |
| A-Form<br>B-Form                  | <b>kad(av)a</b><br><b>kadle(s)</b> | <b>kadle</b><br><b>kadle(n)</b>                    | <b>kadia</b><br><b>kadla</b> | <b>kadle</b><br><b>(kadlan)</b>                    |
| A-Form<br>B-Form                  | <b>kado</b><br><b>kadales</b>      | <b>kadale</b><br><b>kadala</b><br><b>kadale(n)</b> | <b>kadž</b><br><b>kadala</b> | <b>kadale</b><br><b>kadala</b><br><b>(kadalan)</b> |
| A-Form<br>B-Form                  | <b>kako</b><br><b>kakales</b>      | <b>kakale</b><br><b>kakala</b><br><b>kakale(n)</b> | <b>kakā</b><br><b>kakala</b> | <b>kakale</b><br><b>kakala</b><br><b>(kakalan)</b> |
| A-Form<br>B-Form                  | <b>kuko</b><br><b>kukales</b>      | <b>kukale</b><br><b>kukala</b><br><b>kukale(n)</b> | <b>kukā</b><br><b>kukala</b> | <b>kukale</b><br><b>kukala</b><br><b>(kukalan)</b> |

The common short form (characteristic of Gurbeti) is:

|                  | masculine                                   |   | feminine                                 |   |
|------------------|---|---|--|---|
|                  | singular                                    | plural                                      | singular                                 | plural                                      |
| A-Form<br>B-Form | <b>kava</b><br><b>kav</b><br><b>kale(s)</b> | <b>kala</b><br><b>kal</b><br><b>kale(n)</b> | <b>kaja</b><br><b>kaj</b><br><b>kala</b> | <b>kala</b><br><b>kal</b><br><b>(kalan)</b> |

also **kova** or **kov**, **gava** or **gav**, and **gova** or **gov**.

Particularities of the demonstrative when used independently:

When used independently the demonstrative functions as a demonstrative pronoun. It is formed by adding an **-s** and an **-n** to the masculine B-Form in the singular and plural respectively (which is indicated by parentheses in the above table), while replacing the ending **-le** of the feminine plural B-Form [f-pl-B] with **-lan** or **-lān** in the Central European branch, and with **-lien** in the Balkan branch. The endings are as follows:

|        | masculine |        | feminine |                       |
|--------|-----------|--------|----------|-----------------------|
|        | singular  | plural | singular | plural                |
| A-Form | -va       | -la    | -ja      | -la                   |
| B-Form | -les      | -len   | -la      | -lien<br>-lan<br>-lăn |

The extension of the ending **-le** into **-lien**, when a pronoun is formed from an adjective, is not attested in popular speech, but is proposed here analogously with the extension of the ending **-en** into **-ien** in adjectives when they are denominative, as well as for the informational value of the forms.

Adverbs are also formed on the basis of **aka-** : **akajal** (with its short variant **ajal** «thus» and **akana** «now»).

## NUMBERS

Romani, like other languages, has two series of numbers, cardinal and ordinal, which are given in the table below:

|             | cardinal |                                  | ordinal    |                 |
|-------------|----------|----------------------------------|------------|-----------------|
|             | A-Form   | B-Form                           | A-Form     | B-Form          |
| 1           | jekh     | m. jekhe(s)<br>f. jekha<br>jekhe | jèkhto†    | jekhtone(s)     |
| 2           | duj      | dujen<br>don                     | dùjto      | dujtone(s)      |
| 3           | trin     | trinen                           | trito      | tritone(s)      |
| 4           | star     | staren                           | stàrto     | startone(s)     |
| 5           | panz     | panzen                           | pànzto     | panzone(s)      |
| 6           | sov      | soven                            | sovto      | sovtone(s)      |
| 7           | efta     | eftan                            | eftàto     | eftatone(s)     |
| 8           | oxto     | oxton                            | oxtòto     | oxtotone(s)     |
| 9           | enia     | enian                            | eniàto     | eniatone(s)     |
| 10          | deś      | dešen                            | dèšto      | deštone(s)      |
| 11          | deśujekh | deśujekhen                       | deśujèkhto | deśujekhtone(s) |
| 12          | deśuduj  | deśudujen<br>deśudon             | deśudùjto  | deśudujtone(s)  |
| 13,<br>etc. | deśutrin | deśutrinen                       | deśutrìto  | deśutritone(s)  |
| 20,<br>etc. | biś      | bišen                            | bìšto      | bištone(s)      |
| 100         | šel      | šelen                            | šèlto      | šeltone(s)      |

†also angluno and avguno are used

The cardinal number **jekh** «one» is declined for gender, number, and case as a "long" adjective (ending in a vowel). The ordinal form has disappeared with time is replaced with an analogous creation, **jèkhto**, with the foreign borrowings **prvo**, **ipàri**, etc., or the derivation **angluno** «fore», and **e sevtuno** (from the Turkish **sevte** «forward»).

Besides the regular B-Form **duje(n)**, the number **duj** «two» has also preserved the rare archaic B-Form **do(n)** or **done(n)**: **duje šingençar** or **do(ne) šingençar** «with two horns».

The basic numbers in the B-Form are frequently used with the prepostposition **bi ... -go** «without» to indicate the number of minutes before the hour: **deš bi panženqo** «ten minutes to five», etc.

Compound numbers are formed according to the pattern: **X šel Y-vardeš u/ta Z**, where **X** represents the number of hundreds, **Y** the tens, and **Z** the single digits, e.g. **štar šel eftavardeš u trin** «four hundred seventy three».

In many dialects some Greek borrowings have displaced the designation for tens of the type **B-vardeš**. In all dialects the forms are **triànda** «thirty», **sarànda** «forty», but only certain dialects have the form **peinda** or **pinda** «fifty», and **eksinda** «sixty». The numbers **triànda** and **sarànda** are ingrained in the language, and it would be artificial to recreate the lost forms **\*trinvardeš** and **\*štarvardeš**. On the other hand, in the case of the numbers «fifty» and «sixty», the form with **-vardeš** is recommended: **panzverdeš** and **šovardeš**, which are the most widespread in genuine Romani.

The word for «thousand» is expressed only by foreign borrowings: **mia** or **xiliada** «thousand», **duj mie** or **duj xiliade** «two thousand», etc. The Indic root which would agree with the Sanskrit form **sahasra-** or **sahassa** in Pali is not found in contemporary Romani, nor does **\*dešešel**, the form which would correspond to the Sanskrit number **deśaśata-**.

The word **var** «time», which is used to form the tens, can also appear separately: **kobor var(a) phendem tuqe? Šel var(a)** «How many times have I told you? A hundred times!» It is also used to form some other words: **dujvar** «twice» (usually simplified to **duvar**), **trinvar** or **trivar** «thrice», **butvar** «often», **nijekvar** or **niekvar** «not even once», and **nivar** «never». The word **butvar** «often» is differentiated from **but vara** «many times».

## OTHER WORDS WITH NOUN DECLENSIONS

The interrogative pronouns **ko(n)** «who» and **so** «what» are declined according to the noun pattern, acquiring the following respective B-Forms: **kas** or **kos**, and **sos-**.

**Sosqe** and **sos@ar** (sometimes also **sos@e**) are widely used with the meaning «why». It is worth making a distinction between a nuance of result for **sos@ar** and a nuance of intention for **sosqe**, as in the German, as in the difference between the German **warum** and **wozu**.

**Kasqo(ro)** has the function of the Albanian **i kujt** «whose»: **kasqi(ri) @haj san tu?** «Whose daughter are you?»

As a negative answer to the question **kon** or **so** Romani has two words **khonik** (or **konik**) «no one» and **khan@** «nothing». **Konik** is declined either globally: the B-Form **konikes**, or in two parts **kani-kas**. The B-Form of the word **khan@** is completely regular: **khan@es-**. **Konik** and **khan@** in many cases are replaced with the phrases **niekh manu@** and **niekh @humuni** respectively.

In general, the word **manu@** <person> can function as the indefinite pronoun «someone» in affirmative statements and questions, e.g. **ma khuv andre, se isi manu@** «Do not enter, there is someone inside», **avilo manu@?** «Has anyone arrived?», and when accompanied with a particle negative **na** such as **na, ni, @i, ma** it expresses the opposite meaning: **na dikhlem manu@es** «I didn't see anyone», **ma te ku@el les manu@!** «Don't let anyone insult him!»

The word **khan@** can also express the concept «something» in affirmative statements or questions, whereas in statements containing a negative particle it has the opposite meaning: **r@дав khan@** «I am looking for something», **xal@n khan@?** «Have you eating anything?», **n arakhlem khan@** «I haven't found anything». The noun **@humuni** (masc. or fem.) can also assume this function. It should be noted that in affirmative statements it is frequently accompanied by the number **jekh** «one», especially when referring to a countable object: **r@дав jekh @humuni** or **xal@n @humuni?** (food is not a countable object), **na arakhlem @humuni**, etc. However, the lexeme has been lost in many Romani dialects, but its ancient origin is attested by its present in groups of dialects that have been separated from each other for centuries, e.g. **Me@kari** of **Muzeqe**, in the Sinti of the Italian Piemonte, and in the Romani of Wales.

Various words are used in question for qualification In Albanian **@far@** «which», but the most widespread and preferred is **savo**, which is declined according to the "long" type, and frequently has the meaning «which»: **savo vordon isi les?** «Which wagon does he have?»



Savo, so and ka(j) are widely used as conjunctions, e.g. o thud so/ka pilān «the milk that you drank», naj sosqe te xal, manuś ka/so/savo na kerel buti «a man who does not work has no need to eat», etc. Admirative words, especially so are also used: so i lačo akava cikno! «How nice this little one is!». The negative of the word savo is nisavo: nisavi buti na žanel te kerel «He doesn't know how to do anything», etc. It often has the meaning «vain, useless»: nisavo phralipe «vain brotherhood». From savo are also formed sarsavo or savone meaning «an kind (of)». In sarsavo the last part in declined: sarsavi, sarsave, which in savona the first part is declined: savina, savena.

The qualitative demonstrative is formed by adding the prefix a- (or ga- in the Gurbeti dialect) to the adjective savo: asavo, gasavo «such».

The concept «the whole» in Roman i is expressed by using with the word sare or savorre, which is declined according to the long adjective type (usually in the plural), or in its short, invariable form sa, e.g. sa(vor)re manuśa meren «all people die». The form sa is used before an article: sa e manuśa phrala «all men (are) brothers».

The root sar (from Old Indic sārva) combines with the interrogative pronoun kon «who» in the form sarkon «each»: sarkon pi buti dikhel «each sees his own work». However, in many dialects the similarity with the equivalent Slavic word (Serbo-Croatian svako, Bulgarian всекој; and the Macedonian секој, has exercised an attraction and has caused the change of the form sa(r)kon into sàko, which is declined like a foreign adjective (B-Form sarkones). There is also hybrid forms such as erkon or sekon (the elements er- and se- derive from the Turkish her and Macedonian respectively). The form sàko can be retained in the Common Language in an adjectival function in order to distinguish it from the pronoun sa(r)kon, e.g. sa(r)kon ròdel plo «everyone seeks what is his own» vs. sàko dialèkta si lači «every dialect is good».

## THE MORPHOLOGY OF VERB CLASSES

The Romani verb is defined and distinguished from other parts of speech by the fact that it is conjugated for person for tense. For the latter we distinguish two basic tenses (the present and the past), which are expressed morphologically by characteristic endings and several derivational tenses, formed on the basis of the basic tenses with certain invariable particles.

As shown above, a small group of lexemes are somewhat anomalous, for the decline morphologically according the noun type of *bakro* or the "long" adjective type *baro*, however their A-Form complete a lacuna in the system of verbal conjugation, chiefly in the Balkan branch, by appearing as the 3rd person singular and plural of the past tense of a reflexive verb, or, in some dialects, an intransitive verb. In fact these lexemes constitute a bridge between the verb class and the noun class. Another bridge between these two classes is made from verbal nouns formed with the endings: *-pe*, *-pa*, *-mos*, etc., from verbal or nominal roots. These are declined as nouns, but at the same time, they exhibit the category of voice (or diathesis). Occasionally they are considered a kind of infinitive. Besides these exceptions the Romani verbal system displays a great deal of consistency.

Except for the imperative, which in all of the words languages constitutes a special subsystem within the verbal system, we can say that from the purely morphological standpoint, as far as the relevant categories are concerned, the Romani verb has the same system of persons as does Albanian, two voices (active and reflexive/passive), and two basic tenses (present and past), while other forms are variants of these, which are expressed by means of various native and foreign particles, which are placed before or after the verb. A short description of these particles is given below, but their treatment will be handled in the section on syntax.

### CLASSIFICATION OF THE Romani VERB

The Romani verb can be divided into two groups according to the final sound of the stem (or theme), which can be the root itself or a root extended with a suffix, e.g. the causative. Verbs, the stem of which ending in an stressed *-a-* are called "thematic" or "group A", whereas all others are called "athematic" or "group E." Their respective conjugations differ greatly, especially in the Balkan branch, but also in many, many dialects of the northern group, since thematic verbs follow a deponent conjugation, as will be shown below. The treatment of verbs is determined according to the characteristic vowel in the ending of the 2nd and 3rd persons of the present tense, in both the singular and the plural. In athematic verbs, the characteristic vowel is *-e-*, which intervenes between the root and the consonant ending, whereas in thematic verbs, the vowel is

-a-, immediately connected with the consonant ending, as shown in the examples below:

|         | Personal Ending | Thematic verb<br>Stem: <b>prasta-</b> «to run» | Athematic verb<br>Stem: <b>phir-</b> «to walk» |
|---------|-----------------|--|--|
| 1st sg. | <b>-av</b>      | <b>prasta-v</b> (< <b>prasta-av</b> )          | <b>phir-av</b>                                 |
| 2nd sg. | <b>-s</b>       | <b>prasta-s</b>                                | <b>phir-e-s</b>                                |
| 3rd sg. | <b>-l</b>       | <b>prasta-l</b>                                | <b>phir-e-l</b>                                |
| 1st pl. | <b>-as</b>      | <b>prasta-s</b> (< <b>prasta-as</b> )          | <b>phir-as</b>                                 |
| 2nd pl. | <b>-n</b>       | <b>prasta-n</b>                                | <b>phir-e-n</b>                                |
| 3rd pl. | <b>-n</b>       | <b>prasta-n</b>                                | <b>phir-e-n</b>                                |

Verbs of Group E (athematic) are divided into three subgroups according to the characteristic formation of the past tense, which is formed by adding the temporal affix (or formant) **d**, **l**, or, more rarely, **t** to the stem before the personal morphemes. The choice of the affix depends on the final sound to the stem, e.g. after a stem ending in **-r-** (as in **phir-** «to walk», the affix is **-d-**: **phir-d-**, whereas after a stem ending in **-kh-** (as in **dikh-** «to see»), the affix is **-l-**: **dikh-l-**. The affix **-t-** is found in only a few verbs (the rules for the choice of past-tense affix will be given below).

#### THE CITATION FORM: 3rd person singular present tense

In languages which possess an infinitive, this usually serves as the citation form. It is said, for example "the Latin verb **videre**, the Serbo-Croatian verb **videti**, the English verb **to see**, the Hindi verb **dekhnā**," etc. In languages which have no infinitive another characteristic form is used, e.g. Albanian cites the verb in the 1st person singular of the present tense: we say "the verb **shoh**" «I see». Modern Greek also cites this same form. In Romani the 3rd person singular of the present tense is accepted as the citation form, for the reason that it indicates the affiliation of the verb, whether it is thematic or athematic, and from it all other conjugated forms can be deduced. Consequently, the function of the citation form in Albanian is assumed by a corresponding Albanian form, i.e. by the 1st person singular present. Thus we say "the Romani verb **dikhel** 'shoh, shikoj' «I see», however the actual meaning is «he sees», but this has nothing to do with the meaning of the verb form, but with the citation form of the verb.

GENERAL INFORMATION ON VERB CONJUGATIONS  
(based on examples of athematic verbs)

In general, verbal endings do not differ markedly among various dialect. On the table below we give the basic forms which characterize the conjugation of the athematic Romani verb: in the two basic tenses (i.e. present and past), in both voices according to the two main patterns (with the affixes **-d-** and **-l-**) of the active verb. For the 1st person singular of the past tense the table below gives the active ending with its variant **-öm**, which is characteristic of the dialects of Bulgaria, but it should not be forgotten that among various dialects the following forms are also encountered: **-om** (e.g. in Mahazeri, Kabuži, etc.) **-üm** (in Arli, Mahazeri, Tharegoni, etc) **-ëm**, e.g. in the Kalderaś dialect of Warsaw), as well as **-em**, which is very widespread, especially among the Gurbeti and Kalderaś.

|               | Subgroup with affix <b>-d-</b><br><b>pherel</b> «to fill» |                             | Subgroup with affix <b>-l-</b><br><b>dikhel</b> «to see» |                             |
|---------------|---|-----------------------------|--|-----------------------------|
|               | active voice  | reflexive/<br>passive voice | active voice   | reflexive/<br>passive voice |
| Present tense |   |                             |  |                             |
| 1st sg.       | <b>pherav</b>   | <b>pherdövav</b>            | <b>dikhav</b>  | <b>diklövav</b>             |
| 2nd sg.       | <b>pheres</b>   | <b>pherdö(ve)s</b>          | <b>dihkes</b>  | <b>dikhlö(ve)s</b>          |
| 3rd sg.       | <b>pherel</b>   | <b>pherdö(ve)l</b>          | <b>dikhel</b>  | <b>dikhlö(ve)l</b>          |
| 1st pl.       | <b>phas</b>   | <b>pherdövas</b>            | <b>dikhas</b>  | <b>dikhlövas</b>            |
| 2nd pl.       | <b>pheren</b>   | <b>pherdö(ve)n</b>          | <b>dikhen</b>  | <b>dikhlö(ve)n</b>          |
| 3rd pl.       | <b>pheren</b>   | <b>pherdö(ve)n</b>          | <b>dikhen</b>  | <b>dikhlö(ve)n</b>          |
| Past tense    |   |                             |  |                             |
| 1st sg.       | <b>pherdöm</b>  | <b>pherdilöm</b>            | <b>dikhlöm</b>   | <b>dikhlilöm</b>            |
| 2nd sg.       | <b>pherdän</b>  | <b>pherdilän</b>            | <b>dikhlän</b>   | <b>dikhlilän</b>            |
| 3rd sg.       | <b>pherdä(s)</b>  | <b>pherdil/o, -i</b>        | <b>dikhlän</b>   | <b>dikhlilän</b>            |
| 1st pl.       | <b>pherdäm</b>  | <b>pherdiläm</b>            | <b>dikhläm</b>   | <b>dikhliläm</b>            |
| 2nd pl.       | <b>pherdën</b>  | <b>pherdilën</b>            | <b>dikhlën</b>   | <b>dikhlilën</b>            |
| 3rd pl.       | <b>pherdë</b>   | <b>pherdile</b>             | <b>dikhlë</b>  | <b>dikhlile</b>             |

As can be seen, the 2nd and 3rd persons of the reflexive/passive present have both a long form and a short form, e.g. **pherdös** vs. **pherdöves**. The full forms of the 3rd person are quite rare, in

both the singular and the plural, and their use is not recommended in the Common Language.

A difference can be noticed in the writing of the 3rd person plural past tense ending: **-lě** in the reflexive/passive mood and **-le** in the active mood of the type with the affix **-l-**, although there is not difference in pronunciation. This is due to the fact that the active mood ending **-lě** (in **dikhlě**), as well as the 2nd person plural ending **-lěn**, is formed from the past tense affix **-l-**, to which the active morphemes **-ě** or **-ěn**, which are characteristic of the 3rd and 2nd persons of the past tense, are added. On the other hand, the ending **-le** of the reflexive/passive forms **pherdile** and **dikhlile** is the ending of a participle with a verbal function, as we have seen above. This corresponds to the forms with **-o** for the masculine singular and **-i** for the feminine singular. As a participle there is no reason to become palatalized since it follows the declension of the "long" adjectives (the type **baro/o**, **-i**, **-e**).

For the sake of ease, it has been proposed that the jotization sign be abandoned in all other forms of 2nd and 3rd persons plural with **-l-**: **pherdilen**, **dikhlen**, and **dikhle**, but also **pherdilem**, **dikhlem**, **dikhlilem** in Gurbeti. Nevertheless, we can recommend another approach, i.e. to write this form according to its nature: verbal forms with jotization and participial forms without jotization. This is proposed not only on grammatical grounds, but also because in a portion of the northern dialects, especially in Poland, the realization **l-** before the vowel **e**, is of two kind, as in other dialects before the vowels **a**, **o**, and **u**. It is usually greatly velarized into **[w]**, except in those cases where it is pre-jotized, when it pronounced as a middle **l** **[l]**.

|            |                               |                           |                         |                           |
|------------|-------------------------------|---------------------------|-------------------------|---------------------------|
| 2nd p. pl. | <b>pherďěn</b><br>[djen, ěen] | <b>pherdilěn</b><br>[len] | <b>dikhlěn</b><br>[len] | <b>dikhlilěn</b><br>[len] |
| in Poland  | [ěen]                         | [len]                     | [len]                   | [len]                     |
| 3rd p. pl. | <b>pherďě</b><br>[dje, ěe]    | <b>pherdile</b><br>[le]   | <b>dikhlě</b><br>[le]   | <b>dikhlile</b><br>[le]   |
| in Poland  | [ěe]                          | [we]                      | [le]                    | [we]                      |

In the above table we have illustrated separately the pronunciation in the dialects of Poland, in which there is a differentiation made between the two kinds of **l**, according to whether or not it is palatalized by the following **e**.

When we analyze the system of simple verbal forms, we arrive at the following summary:

| voice         | active voice   | reflexive/passive    |
|---------------|----------------|----------------------|
| present tense | R + PR         | R + -d/-l + -öv + PR |
| participle    | R + -d/-l- + W |                      |
| past tense    | R + -d/l- + PS | R + -d/l + -il- + PS |

Explanation of the symbols:

R = the verbal stem (**pher-** «to fill», **dikh-** «to see», **per-av-** «to fall», **kal-är-** «to blacken», etc.;

-d/l- = the affix characteristic of the past tense;

W = the vowel characteristic of the noun class: -o, -i, and -e;

PR = the totality of personal present tense ending;

PS = the totality of personal past tense endings;

-öv- = the affix -öv- takes the form -iv- in most Gurbeti dialects, and the form -üv- in most of the others, as in Kalderaś.

As can be clearly seen on the above table, the backbone of the verbal system is the participle, i.e. a lexeme which belongs to the noun class.

## PERSONAL ENDINGS

### I) Present tense personal endings

The present tense personal endings of the Romani verb are as follows:

|        | singular | plural |
|--------|----------|--------|
| 1st p. | -av      | -as    |
| 2nd p. | -s       | -n     |
| 3rd p. | -l       | -n     |

These endings are added immediately to the stem in thematic verbs, i.e. when the stem ends in a stressed -a-, as seen on the table of the verb *prastal* «to run».

On the other hand, when dealing with an athematic verb, i.e., when the stem ends in a consonant or an unstressed vowel, then an e intervenes in the 2nd and 3rd persons singular as well as plural, as seen in the verb *phirel* «to walk» in the table below:

|         | Personal Ending | Thematic verb<br>Stem: <i>prasta-</i> «to run» | Athematic verb<br>Stem: <i>phir-</i> «to walk» |
|---------|-----------------|--|--|
| 1st sg. | -av             | <i>prasta-v</i> (< <i>prasta-av</i> )          | <i>phir-av</i>                                 |
| 2nd sg. | -s              | <i>prasta-s</i>                                | <i>phir-e-s</i>                                |
| 3rd sg. | -l              | <i>prasta-l</i>                                | <i>phir-e-l</i>                                |
| 1st pl. | -as             | <i>prasta-s</i> (< <i>prasta-as</i> )          | <i>phir-as</i>                                 |
| 2nd pl. | -n              | <i>prasta-n</i>                                | <i>phir-e-n</i>                                |
| 3rd pl. | -n              | <i>prasta-n</i>                                | <i>phir-e-n</i>                                |

When the stem of the athematic verb ends in a consonant the present tense form presents no difficulty. When it ends in a vowel (unstressed), a connective *v* is inserted, as shown in the examples below:

|         | Personal Ending | Athematic verbs      |                       |
|---------|-----------------|----------------------|-----------------------|
|         |                 | Stem: so- «to sleep» | Stem: sika- «to show» |
| 1st sg. | -av             | so-v-av              | sika-v-av             |
| 2nd sg. | -s              | so-v-e-s             | sika-v-e-s            |
| 3rd sg. | -l              | so-v-e-l             | sika-v-e-l            |
| 1st pl. | -as             | so-v-e-s             | sika-v-as             |
| 2nd pl. | -n              | so-v-e-n             | sika-v-e-n            |
| 3rd pl. | -n              | so-v-e-n             | sika-v-e-n            |

Frequently one cannot distinguish, in the synchrony of the verb, a stem ending in **-v-**, e.g. in **siv-** «to sew», from a verb, the stem of which ends in an unstressed vowel, to which the epenthetic **v** is added throughout the present tense, e.g. in **čí-** «to lay». Consequently, in practice, many describers have treated the endings of the present tense as having two forms:

Endings  $PR_A$ : **-av**, **-as**, **-al**, **-as**, **-an**, and **-an** by inserting the stressed vowel **-a-** to the stem: **prast-av**, **prast-as**, **prast-al**, etc.

Endings  $PR_E$ : **-av**, **-es**, **-el**, **-as**, **-en** and **-en** by adding the epenthetic vowel **-e-**: **phir-av**, **phir-es**, **phir-el**, etc.

These consider that the connective **-v-** is an integral part of the stem: **sov-**, **sikav-**, etc. Such a simplification is better for giving a scholarly description of the conjugation for a specific dialect, but it limits the possibility of studying the verbal system in relationship to all the phenomena among its dialects and in diachrony.

Apparently, the **-v-** in the suffixes **-öv-** (reflexive/passive) and **-av-** (causative) is simply a connective **v**.

## II) Past tense personal endings

The past tense personal endings of the Romani verb are as follows:

|        | singular      | plural |
|--------|---------------|--------|
| 1st p. | -öm, -üm, -em | -äs    |
| 2nd p. | -än           | -ěn    |
| 3rd p. | -ă(s)         | -ě     |



As we have seen above, the 1st person singular past tense exhibits variant forms: -ōm as, e.g. in Erli of Bulgaria, -om, as, e.g. in Mečkari of Albania, -ūm in most dialects of Skopje and Prishtina, -ēm, as, e.g. the Kalderaś dialect of Warsaw, and -em, which is the most widespread, as which is encountered especially in Gurbeti and Kalderaś. Likewise, a number of dialects are characterized by the loss of jotization by this morpheme in all persons, so that we have: -om, -em, -an, -a(s), etc. Most Gurbeti and Kalderaś dialects exhibit this phenomenon. In addition, many other dialects do not jotize the 3rd person plural: phirde «they walked», although they do retain it in the other persons. This is done, for example, in the Mahazeri of Prishtina.

The totality of these endings is symbolized by PS<sub>a</sub>, according to the vowel in the 3rd person singular.

As we have shown, the 3rd person singular and plural of reflexive/passive verbs and certain other verbs, which are intransitive, are lacking and are replaced with a particle, which is determined by the gender and number of the subject:

|       | singular | plural |
|-------|----------|--------|
| masc. | -o       | -e     |
| fem.  | -i       |        |

The totality of these endings are symbolized by PS<sub>o</sub>, according to the vowel of the 3rd person singular masculine.

## FORMS OF THE PARTICIPLE AND THE PAST TENSE

### I) In verbs with a stem ending in a consonant

In cases of athematic verbs ending in a consonant the participle is frequently formed by adding the affix **-d-** to the stem when it ends in an **-r, -l, -n** or **-v** (e.g. in **mar-el** «to beat», we have the participle **mar-d-o** «beaten») and affix **-l** when it ends in **-s, -ś, -ć, -k, -kh, -(n)g** or **-m** (e.g. **piś-el** «to milk; to grind» we have the participle **piś-l-o** «milked; ground»). A limited number of verbs take the affix **-t**. In some dialects there is a vacillation between one form or the other: besides **kamlo** «(be)loved» we also encounter **kamdo, khosto** vs. **khoslo** «wiped», and **phando** vs. **phanlo** «to bind». The forms which conform to the rules given at the beginning of this paragraph, which are the most widespread, are the preferred ones in the Common Language. The past tense is formed by replacing the adjectival endings of the participle **-o, -i, and -e** with endings of the type **PS<sub>a</sub>**, e.g. 3rd person singular **mar-d-ă** «he/she beat», **piś-l-ă(s)** «he/she milked; she/she ground»

The verb **l-el** «to take» has the participial form **lilo**, from which the past tense is regularly formed with the endings **PS<sub>a</sub>**, e.g. **lilăm** «we took». In the past tense a variant on the basis of **li-PS<sub>a</sub>** is also encountered: **liem** «I took», The verb **d-el** «to give» has a special form: **-in-**, to which the affix **-d-** is sometimes added, yielding the participle **dino, dindo**, while the past tense is always formed with the endings **PS<sub>a</sub>**: **dindüm**. Besides this form, we also have **di-PS<sub>a</sub>** in the past tense in many dialects: **diem** «I gave», as shown in the table below:

|         | lel «to take» |                 | del «to give» |                    |
|---------|---------------|-----------------|---------------|--------------------|
|         | present       | past            | present       | past               |
| 1st sg. | lav           | lilom, liem     | dav           | din(d)om, diem     |
| 2nd sg. | les           | lilan, lian     | des           | din(d)an, dian     |
| 3rd sg. | lel           | lila(s), lia(s) | del           | din(d)a(s), dia(s) |
| 1st pl. | las           | lilam, liam     | das           | din(d)am, diam     |
| 2nd pl  | len           | lilen, lien     | den           | din(d)en, dien     |
| 3rd pl. | len           | lile, lie, line | den           | din(d)e, die, dine |

The forms **line** «they took» and **dine** «they gave» in the past tense are characteristic of the Kalderaś dialect.

We have two athematic verbs which end in a consonant but which have a special conjugation:

- **merel** (stem **mer-**) has a completely irregular participle: **mulo** «dead», on the basis of which the past tense is formed with the endings **PS<sub>o</sub>**: **mulām** «I died», **mulo** «he died», etc.

- **perel** (stem **per-**) also has an irregular participle: **pelo** «fallen». In this case the past tense is also formed with the endings **PS<sub>o</sub>**: **pelān** «you fell».

The verb **ušteľ** «to rise» also has a reflexive/passive form in the past tense: **ušteľo**, with dialectal variants **ušteľindilo**, **uxtinlo**, etc., besides the active form **ušteľindā(s)**.

Such verbs, which, when not jotized, follow an active conjugation in the present tense (i.e. without the affix **-ōv-**) and the reflexive/passive in the past tense, are called deponents, i.e. the active conjugation is "left aside" (in Latin **deponent**) in the past tense in order to use the reflexive/passive conjugation. The deponent is characteristic of the Balkan branch, but is also found in several dialects of the northern branch. However, in the northern branch the formation of the past tense with the endings **PS<sub>a</sub>** is more widespread in all dialects: **mulā(s)** «he/she died», **pelā(s)** «he/she fell».

II) In verbs with a stem ending in a consonant

a) If we are dealing with an athematic verb with a final stem vowel, the temporal affix **-d-** can be added to this either immediately or after an epenthetic **v**: **sikado** or **sikav-d-o** «taught».

The verb **čhivel** or **čhuvel** «lay» is a special case, since besides the regular participial form with **-d-**, it can also form the participle in many dialects with **-t-**, so that we have the following forms: **čhido**, **čhivdo**, **čhudo**, **čhuvdo**, **čhito**, **čhuto**. In the past tense we have the following corresponding variants: **čhid-PS<sub>a</sub>**, **čhivd-PS<sub>a</sub>**, **čhud-PS<sub>a</sub>**, **čhuvd-PS<sub>a</sub>**, **čhit-PS<sub>a</sub>**, **čhut-PS<sub>a</sub>**. This verb has another variant, formed with the semi-auxiliary verb **del**: **čhùdel**. See below for the conjugation of this variant.

b) In the case of thematic verbs, i.e. with stems ending in stressed **-a-**, the participle is hypothetical, i.e. it is not found as an existent form in the language, but can only be inferred artificially as a step to the formation of the past tense. Consequently we will discuss the formation of the past tense forthwith.

These verbs form the past tense in two different ways according to dialect, either with the affix **-j-**, which is added to the endings **PS<sub>a</sub>**, e.g. **prasta-** : **prasta-j-a(s)** «he/she ran», or with the affix **-d-**, extended into **-nd-**, and the affix **-il-**, characteristic of the reflexive/passive, to which the endings **PS<sub>o</sub>** are added, e.g. **prasta-** : **prasta-nd-ilo** «he ran».

c) Some irregular verbs have a special participle:

**xal** «to eat» (stem **ha-**) and **piel** «to drink» (stem **pi-**) form the participle with the affix **-l-**: **xalo** «eaten», **pilo** «drunk». The past tense is formed with the endings **PS<sub>a</sub>**: **xalām** «I ate», **pilān** «you (sg.) drank»; **zał** «to go» (stem **za-**) has a completely irregular participle: **gelo**, from which is formed the past tense: **gelem**, **gelōm**, **gelūm**, etc. «I went».

Certain athematic verbs with stems ending in a vowel also belong here:

- **avel** «to come» (stem **a-**) has two different participles, depending on dialect: **a-v-il-o** or **al-o**, as well as two variants in the past tense: **avil-PS<sub>o</sub>** and **al-PS<sub>o</sub>**: **avilen** or **alen** «you (pl.) came»; **ovel** «to become» (stem **o-**) also has two variants in the past tense: **ondil-PS<sub>o</sub>** and **ul-PS<sub>o</sub>**. The participle is hypothetical; **sovel** «to sleep» (stem **so-**) takes the affix **-t-**, which is added immediately to the stem without the epenthetic **v**, while the **-o-** of the root becomes fronted into **u**: **suto**. The past tense is formed with the endings **PS<sub>a</sub>**: **sutās** «he/she slept»; **rovel** «to cry» (stem **ro-**) also fronts the **-o-** in many dialects into **-u-** and takes the affix **-d-** extended into **-nd-**: **rundo**, **rondo** (other dialectal forms with **-l-** also occur). The past tense is formed with the endings **PS<sub>a</sub>**.

In many dialects, especially Gurbeti and the northern branch, the past tense is formed with the affix **-j-**: **rujem** «I cried».

| Summary Table |                                 |   |
|---------------|---------------------------------|---|
| Verb          | Participle                      | Past tense  |
| <b>a-v-</b>   | <b>av-il-o</b> or <b>al-o</b>   | <b>av-il-PS<sub>a</sub></b>                               |
| <b>xa-</b>    | <b>xal-o</b>                    | <b>xal-PS<sub>a</sub></b>                                 |
| <b>l-</b>     | <b>lil-o</b>                    | <b>lil-PS<sub>a</sub></b>                                 |
| <b>mer-</b>   | <b>mul-o</b>                    | <b>mul-PS<sub>o</sub></b>                                 |
| <b>o-v-</b>   | <b>o-nd-il-o</b> or <b>ul-o</b> | <b>o-nd-il-PS<sub>o</sub></b> or <b>ul-PS<sub>o</sub></b> |
| <b>per-</b>   | <b>pel-o</b>                    | <b>pel-PS<sub>o</sub></b>                                 |
| <b>pi-</b>    | <b>pil-o</b>                    | <b>pil-PS<sub>a</sub></b>                                 |
| <b>ro-v-</b>  | <b>ru-nd-o</b>                  | <b>rund-PS<sub>a</sub></b> or <b>ru-j-PS<sub>a</sub></b>  |
| <b>so-v-</b>  | <b>sut-o</b>                    | <b>su-t-PS<sub>a</sub></b>                                |
| <b>za-</b>    | <b>gel-o</b>                    | <b>gel-PS<sub>o</sub></b>                                 |

We have noted these past tense endings according to the system of deponent verbs, since in dialects which have no deponent verbs, all forms take the endings PS<sub>a</sub>. As we have mentioned, deponents are very rare among the athematic verbs, and the vast majority are formed from thematic verbs.

#### THEMATIC VERBS

Thematic verbs, i.e. with the stem ending in a stressed -a-, must be devoted some special attention. They number about twenty. As we have emphasized before most of these form the past tense in the Balkan branch with the endings PS<sub>o</sub>, i.e. they are deponents.:

R + nd + il + PS<sub>o</sub> in the Balkan branch (the past tense affix -d- appears in the form -nd- after the stressed vowel);

R + j(l) + PS<sub>o</sub> in many Central European dialects (the affix -nd- is lacking, while the affix -il- becomes -jl- or -j-);

R + j + PS<sub>a</sub> in dialects without deponents.

In other words, the Balkan dialect usually has the participial endings masc. -o, fem. -i, and pl. -e in the 3rd person, whereas in Central Europe the 3rd person has either the above-mentioned participles or the endings -a(s) in the singular and -e in the plural.

| Present Tense | Past Tense   |                 |                           |             |
|---------------|--------------|-----------------|---------------------------|-------------|
|               |              | Balkan dialects | Central European dialects |             |
| daral         | «to fear»    | daràndilo       | daràjlo                   | daraja(s)   |
| pačal         | «to believe» | pačàndilo       | pačàjlo                   | pačaja(s)   |
| prastal       | «to run»     | prastàndilo     | prastàjlo                 | prastaja(s) |
| dukhal        | «to ache»    | dukhàndilo      | dukhàjlo                  | dukhaja(s)  |

Other verbs of this group are:

|                         |                                     |
|-------------------------|-------------------------------------|
| daral «to laugh»        | sungal «to smell»                   |
| azbal «to touch»        | tromal «to dare»                    |
| ćulal «to leak»         | tuzal «to feel a burning sensation» |
| xasal «to cough»        |                                     |
| izdral «to shudder»     |                                     |
| langal «to limp»        |                                     |
| lažal «to feel ashamed» |                                     |
| prasal «to mock»        | puzgal «to slip»                    |

umblal «to hook»  
urāl «to fly»  
zungal «to feel nausea»

Besides the forms **prasal** and **aprasal** «to mock» with an epenthetic **a-**, there is also the athematic variant **prastavel**.

In some Balkan dialects the verb **azbal** «to touch» exhibits a form with the affix **-d-** with the endings **PS<sub>a</sub>**: **azbandă** «he/she touched» alongside the regular **azbândilo**.

The verb **gilabal** or **gilabel** «to sing», depending on dialects, belong to either the thematic or the athematic group. In addition, it has some other variants: with **-b-** or **-v-** at the end of the root (**gilab-** or **gilav-**), with or without a pre-jotizing **-a-** (**gilăv-** or **gilav-**), with the past tense in **-d-PS<sub>a</sub>** or in **-j-PS<sub>a</sub>**, etc. All the variants are acceptable in the Common Language, but the form **gilăvel**, which corresponds more closely with its etymology (**gili + a + T**), is preferred.

The verb **bagel**, which replaces **gilăvel** in some dialects of Poland, is also acceptable, for it is advised that it only be used for the purpose of poetry reminiscent of the works of the poet Papuša. This verb is apparently related to the Indic **bāg denā** «to crow»

#### FOREIGN WORDS

We must mention here the very important affix, the suffix **-sar**, which is very widespread in dialects of the Gurbeti and Vlah type, and whereby foreign words can be accommodated into the Romani conjugation system. In general, the stem of the foreign verb ends in **-i-** when formed on the basis of a verb in the source language, and an **-o-** when formed on the basis of a noun.

Verbs formed in the above manner are conjugated according to the athematic type, but can be abbreviated but omitting the affix **-sar-** in the present tense and taking only the endings, in which case the ending of the 1st person singular and plural lose the vowel **a**:

|         | From verbal root:<br>traï(sare)l «to reside» |        | From nominal root:<br>praxo(sare)l «to bury» |         |
|---------|--|--------|--|---------|
|         | Present tense                                |        |  |         |
| 1st sg. | traï-sar-av                                  | traï-v | praxo-sar-av                                 | praxo-v |
| 2nd sg. | traï-sar-es                                  | traï-s | praxo-sar-es                                 | praxo-s |
| 3rd sg. | traï-sar-el                                  | traï-l | praxo-sar-el                                 | praxo-l |
| 1st pl. | traï-sar-as                                  | traï-s | praxo-sar-as                                 | praxo-s |
| 2nd pl. | traï-sar-en                                  | traï-n | praxo-sar-en                                 | praxo-n |
| 3rd pl. | traï-sar-en                                  | traï-n | praxo-sar-en                                 | praxo-n |
|         | Past tense                                   |        |  |         |
| 1st sg. | traï-sar-dem                                 |        | praxo-sar-dem                                |         |
| 2nd sg. | traï-sar-dan                                 |        | praxo-sar-dan                                |         |
| 2rd sg. | traï-sar-da(s)                               |        | praxo-sar-da(s)                              |         |
| 1st pl. | traï-sar-dam                                 |        | praxo-sar-dam                                |         |
| 2nd pl. | traï-sar-den                                 |        | traï-sar-den                                 |         |
| 3rd pl. | traï-sar-de                                  |        | traï-sar-de                                  |         |

The verb **traï(sar)el** comes from the Romanian **a traï** «to reside» and has extensively displaced the Romani word **zível**, especially in Gurbeti and Kalderaš, whereas **praxo(sare)l** «to bury» is based on a word the Slavic origin, **pràxo** «dust, ash». The Romani word with this meaning is **paruvel**.

In Montenegro and in some dialects of Bosnia the endings are lost in the 2nd and 3rd persons singular and 1st person plural, taking the form below:

|         |        |         |
|---------|--------|---------|
| 1st sg. | traï-v | praxo-v |
| 2nd sg. | traï   | praxo   |
| 3rd sg. | traï   | praxo   |
| 1st pl. | traï   | praxo   |
| 2nd pl. | traï-n | praxo-n |
| 3rd pl. | traï-n | praxo-n |

Such a simplification, which creates homonyms among personal forms is not acceptable in the Common Language.

As is clear from the above table, the past tense of these verbs is formed regularly with the affix -d- and the endings PS<sub>a</sub>: **trai-sar-d-a(s)**, **praxo-sar-d-a(s)**, **informi-sar-d-a(s)**, etc. Most verbs which follow this conjugation have non-jotizing endings in the past tense.

The formation of the reflexive/passive voice follows a special pattern with the affix -sa- + the connecting -v- in the present tense, and with the same affix in the past tense, which is added to the characteristic affix of the past reflexive/passive -il-. The elements -sa- and -il- are usually written together as -sàil-, -sàjl- or -sal- with a complete loss of the i, and which can be treated as one affix -sàjl-:

| voice         | active voice                  | reflexive/passive               |
|---------------|-------------------------------|---------------------------------|
| present tense | R-i + (sar) + PR              | R-i + sa-v + PR                 |
| participle    | R-i + sar + d + W             |                                 |
| past tense    | R-i + sar + d+PS <sub>a</sub> | R-i + -sàjl + PS <sub>o</sub> . |

In the Balkans these type of foreign words have the corresponding affix -in- (or less common -iz-, e.g. in the Bugurzi dialect of Obilić and the Ursari of Giurgiu in Romania), which form the characteristic features in a completely regular manner, as shown below:

| voice         | active voice               | reflexive/passive                    |
|---------------|----------------------------|--------------------------------------|
| present tense | R + in + PR                | R + in + d + òv + PR                 |
| participle    | R + im + W                 | (or invariable -e)                   |
| past tense    | R + in + d+PS <sub>a</sub> | R + -isà(j)l + PS <sub>o</sub> .     |
|               |                            | or R + in + d + il + PS <sub>a</sub> |

This system can be illustrated with the paradigm of the verb **ofsetisarel** (in the Northern branch) / **ofsetinel** (in the Balkan branch) «to duplicate by off-set»:

In the Northern dialects:

| voice         | active voice         | reflexive/passive |
|---------------|----------------------|-------------------|
| present tense | ofseti(sar)el        | ofsetisavel       |
| participle    | ofsetisard/o, -i, -e |                   |
| past tense    | ofsetisarda(s)       | ofsetisàjl/o, -i  |



In the Balkan dialects:

| voice         | active voice                    | reflexive/passive |
|---------------|---------------------------------|-------------------|
| present tense | ofsetinel                       | ofsetindol*       |
| participle    | ofsetim/o, -i, -e or (ofsetime) |                   |
| past tense    | ofsetindă(s)                    | ofsetisàjl/o, -i  |
|               | or ofsetindil/o, -i             |                   |

Several neologisms that are constructed of purely Romani material are also conjugated according to this model, e.g. **agorisarel** «to end» from **agor** «end», **godisarel** «to think» from **godi** «mind», etc.

The use of the both the forms with the affix **-sar-** and **-in-** is advised in the Common Language in order to break the monotony that would arise from using just one form. They can also be used for euphonic reasons, i.e. to avoid having two similar syllables next to each other, e.g. in the forms **interesinel** «to interest» is preferred over **interesisarel** since it results in the repetition of the sound **s**. On the other hand, the form **filmisarel** «to film» is preferred over **filminel** for similar reasons.

#### VERBS WITH A STRESSED ROOT WITH A STEM FINAL -d-

Verbal roots of this type end in **-d-**, and in the present tense the stress falls on the syllable immediately before this **d** in many dialects, while in other dialects it shifts to the ending, analogously with other verbs. Nevertheless, the penultimate stress is recommended in the Common Language. According to one theory, verbs which follow this pattern were formed long ago from a fusion of a non-verbal root with the semi-auxiliary verb **del** «to give», which explains the exceptional stress and the morphological similarity to the past tense of the verb **del**. In some cases, the meaning of the verb supports this etymology, e.g. in **kàndel** «to listen, to obey», which can be analyzed into **kan** «ear» and **del** «to give». In fact, in many dialects, this verb does not differ from other verbs in its conjugation. This verb forms the past tense as in the verb **del**, e.g. from **kàndel** «to listen» we have the past tense **kandină(s)** or **kandias** «he/she heard». A form with a reduplication of the past tense affix is also encountered: **kandindă** (e.g. in Tharegoni and Mahazeri).

The most common verbs of this class are:

|                                 |                            |
|---------------------------------|----------------------------|
| (am)bòldel I «to wet»           | éadel, éàrdel «to vomit»   |
| (am)bòldel II «to turn»         | éhùdel, éhùrdel «to throw» |
| cirdel, càrdel, càdel «to pull» | xàndel «to scratch»        |
| éumidel, éumindel «to kiss»     | xìndel «to defect»         |
|                                 | kàndel «to listen»         |

kìdel «to gather»  
khàndel «to stink»  
pàldel «to dismiss» (dial. to  
drive)  
phùdel, phùrdel «to blow»

ròdel «to seek»  
tràdel «to send» (dial. «to  
drive»)  
thàvdel «to flow»  
vàzdel «to raise»

The participle of these verbs is usually formed with the ending -mo, which is declined according to the "long" adjective type, or with the invariable ending -me as in foreign words: phurdim/o, -i, -e or phurdime (invariable) «bloated», vazdim/o -i, -e or vazdime (invariable) «raised», etc.

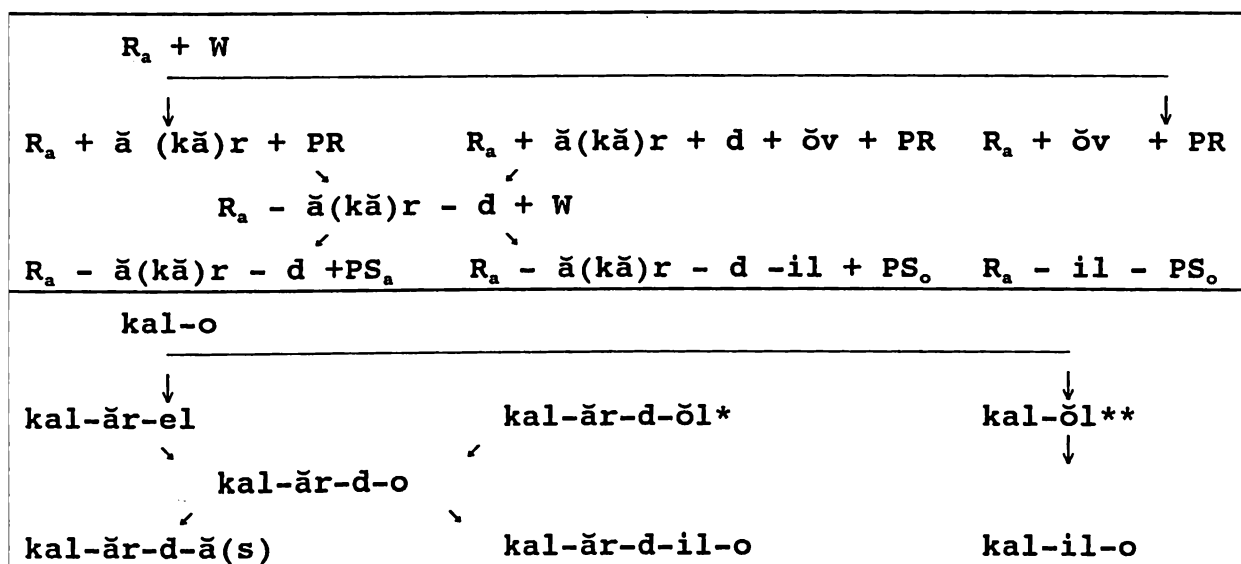
In the past tense, vâzdel and cúmindel have several variants among the dialects: vazdindă, vazlă, vazdlă, vaimda, vazdia, vazăs and cúmindindă, cúmindias, cúmida, cúmizas, etc. The preferred forms in the Common Language are vazdinăs and cúmin(din)dăs for the Balkan variant, and vazla and cúmidla for the other variant.

Some irregularities in conjugation, which can be observed in various dialects, are not recommended in the Common Language, e.g. the peculiar forms of vâzdel «to raise» and cúmindel «to kiss», given above, or the formation of the past tense without the affixes -d- or -l-, characteristic of this tense, which is the case in the Drindari of Bulgaria: kër-ăs «he/she did», mar-ăs «he/she beat» (vs. ker-d-ăs, mar-d-ăs in the Common Language. Likewise the ending -al in the 2nd person singular of the past tense is also to be avoided. This ending is used instead of -an in some dialects of Hungary and Slovakia. In Romania and Ukraine many dialects pronounce the ending -av of the 1st person plural of the present tense as [-aʋ], but write -av like everyone else.

VERBS DERIVED FROM ADJECTIVES, AND CAUSATIVE VERBS

It is an interesting fact that, just as reflexive/passive verbs are formed from participle, a verb of the reflexive/passive voice can also be formed from any adjective. Verbs acquiring this form demonstrate a process leading to its formation. For example from the adjective *kalo* is formed *kal-öl* «to become black», and the past tense *kal-il-o*. Compare this to the reflexive/passive voice formation, e.g. from *mar-el* «to beat», participle *mar-d-o* «beaten», reflexive/passive *mar-d-öl* «is beaten», and the past tense *mard-il-o*.

At the same time, a adjectival causative verb can also be formed by means of the affix *-är-* (with its variants *-kär-* and *-ker-*) + PR. These causatives have their own participle from which the reflexive/passive is formed, as is shown on the table below using as a example the adjective *kalo* «black» ( $R_a$  = adjectival root):



\*) 1st p. sg. *kal-är-d-öv-av*

\*\*) 1st p. sg. *kal-öv-av*

The meanings of *kal-är-d-öl* and *kal-öl* are very close, but there is nevertheless a difference. The form formed directly from the adjective indicates a process of development from the individual himself, whereas the reflexive/passive and causative indicate that this process is being developed by another actor.

Besides the causatives formed from adjectival roots with the affix *-ä(kä)r-*, there is also a causative formed from verbs (usually intransitive) with the affix *-av-* (or *-v-* when a deponent verb is taken as the base), e.g. from *nakh-el* «to pass» we form *nakh-av-el* «to accompany», from *dara-l* «to be afraid» we form *dara-v-el* «to frighten», etc.

Causatives with -av- are formed as follows:

a) in verbs of the E group

|  |   |   |  |
|--|---|---|--|
| R + PR<br>\<br>R + d/l + W<br>/<br>R + d/l + PS <sub>o</sub> | → | R + av + PR<br>\<br>R + a(v)d + W<br>/<br>R + a(v)d + PS <sub>a</sub> | R + a(v)d + öv + PR<br>\<br>R + a(v)d + W<br>\<br>R + a(v)d + il + PS <sub>o</sub> |
| nakh-el<br>\<br>nakh-l-o<br>/<br>nakh-l-o                    | → | nakh-av-el<br>\<br>nakh-a(v)d-o<br>/<br>nakh-a(v)d-ã                  | nakh-a(v)d-öl*<br>\<br>nakh-a(v)d-o<br>\<br>nakh-a(v)d-il-o                        |

b) in verbs of the A group (deponents)

|  |   |  |  |
|--|---|--|--|
| R + PR<br>\<br>*R + nd<br>/<br>R + nd+ il- + PS <sub>o</sub> | → | R + av + PR<br>\<br>R + a(v)d + PS <sub>a</sub>    | R + a(v)d + öv + PR<br>\<br>R + a(v)d + W<br>\<br>R + a(v)d + il + PS <sub>o</sub> |
| dara-l<br>\<br>*dara-nd<br>/<br>dara-nd-il-o                 | → | dar-av-el<br>\<br>dara-(v)d-o<br>/<br>dara-a(v)d-ã | dar-(v)d-öl*<br>\<br>dara-(v)d-o<br>\<br>dara-(v)d-il-o                            |

\*) 1st person singular: dara-(v)d-öv-av

In the present state of the language reflexive/passive causatives are more virtual than effective. However, they are occasionally used in the Common Language when meaning demands.

The derivational system of the causatives is similar to the system of formation of verbal pair *čar-av-al* «to graze» (transitive) and *čar-öl* «to graze» (intransitive). It has not quite clear whether these derive from *čar-el* «to lick» or *bar* «grass».

|             |   |                               |                                     |
|-------------|---|-------------------------------|-------------------------------------|
| čar-öl<br>↓ | → | čar-av-el<br>\<br>čar-a(v)d-ã | čar-a(v)d-öl<br>\<br>čar-a(v)d-il-o |
| čar-il-o    |   | čar-a(v)d-ã                   | čar-a(v)d-il-o                      |

Three pair of similar verbs of this type can be listed, as given on the complete table of existing forms below:

a) *ikal-el / ikl-öl* «to take out / to go out»

| Transitive                   | Reflexive/<br>Passive           | Intransitive                  |                                  |
|------------------------------|---------------------------------|-------------------------------|----------------------------------|
|                              |                                 | short form                    | long form                        |
| <i>ikal-PR</i>               | <i>ikal-d-öv-PR</i>             | <i>ikl-PR</i>                 | <i>ikl-öv-PR</i>                 |
|                              | <i>ikal-d-W</i>                 | ↓                             | ↓                                |
| <i>ikal-d-PS<sub>a</sub></i> | <i>ikal-d-il-PS<sub>a</sub></i> | <i>ikl-ind-PS<sub>a</sub></i> | <i>ikl-ist-il-PS<sub>o</sub></i> |

Note: Besides the form *ikl-ist-il-PS<sub>o</sub>*, there are also other variants, among them, *ikl-il-PS<sub>o</sub>*, *ikl-isl-PS<sub>o</sub>* and *ikl-ist-PS<sub>o</sub>*. In some dialects an initial *n-* is added: *nikal-*, etc.

b) *xuläv-el / xul-öl* «to lower / to descend»

| Transitive                      | Reflexive/<br>Passive              | Intransitive                  |                                  |
|---------------------------------|------------------------------------|-------------------------------|----------------------------------|
|                                 |                                    | short form                    | long form                        |
| <i>xuläv-PR</i>                 | <i>xulä(v)-d-öv-PR</i>             | <i>xul-PR</i>                 | <i>xul-öv-PR</i>                 |
|                                 | <i>xulä(v)-d-W</i>                 | ↓                             | ↓                                |
| <i>xulä(v)-d-PS<sub>a</sub></i> | <i>xulä(v)-d-il-PS<sub>a</sub></i> | <i>xul-ind-PS<sub>a</sub></i> | <i>xul-ist-il-PS<sub>o</sub></i> |

Note: Besides the form *xul-ist-il-PS<sub>o</sub>*, there are also other variants, among them, *xul-il-PS<sub>o</sub>*, *xul-isl-PS<sub>o</sub>* and *xul-ist-PS<sub>o</sub>*, as well as a form with *PS<sub>o</sub>* in the Northern branch.

c) *ukläv-el / ukl-öl* «to mount / to climb up»

| Transitive                      | Reflexive/<br>Passive              | Intransitive                  |                                  |
|---------------------------------|------------------------------------|-------------------------------|----------------------------------|
|                                 |                                    | short form                    | long form                        |
| <i>ukläv-PR</i>                 | <i>uklä(v)-d-öv-PR</i>             | <i>ukl-PR</i>                 | <i>ukl-öv-PR</i>                 |
|                                 | <i>uklä(v)-d-W</i>                 | ↓                             | ↓                                |
| <i>uklä(v)-d-PS<sub>a</sub></i> | <i>uklä(v)-d-il-PS<sub>a</sub></i> | <i>ukl-ind-PS<sub>a</sub></i> | <i>ukl-ist-il-PS<sub>o</sub></i> |

Note: Besides the form *xul-ist-il-PS<sub>o</sub>*, there are also other variants, among them, *xul-il-PS<sub>o</sub>*, *ukl-isl-PS<sub>o</sub>* and *xul-ist-PS<sub>o</sub>*, as well as a form with *PS<sub>o</sub>* in the Northern branch.

The various dialects usually choose a preferential form for the intransitive series from the short form (with endings of the

active type) or the long (with endings of the reflexive/passive type). The preference is not always consistent for all three verbs. For example, the Mahazeri of Prishtina usually uses the long form for the verb *ikl-* and the short form for the other two in the present tense, whereas there is no preference in the past tense. In the Common Language the complete table of forms is observed. Nuances of meaning can develop; the short forms emphasize dynamism of the action.

#### IMPERATIVES

The imperative is a verbal form which has the particularity that it exists only in the present tense and in the 2nd person singular and plural. The plural imperative form corresponds to the present tense, as in Albanian: *pheren* «fill!», *dikh* «look!», *xa* «eat!», etc. The imperative is rare in the reflexive/passive conjugation, but when it does appear, e.g. with an ironic value or in a narrative function (see syntax), it is only the short form with *-ön* is used, rather than *-öven*.

In the singular the imperative usually correspond to the root without an ending: *pher* «fill!», *dikh* «look!», *xa* «eat». Roots which acquire a connecting *v* in the present reflexive/passive retain this *v* in the imperative: *sikav manqe!* «show me!» (sg.). There are cases where the root is extended with a vowel, mainly *e*, but also *i* and *a*. The vowel *e* is added in the monosyllabic verbs *del* «to give» - imperative *de*, and *lel* «to take» - imperative *le*, as well as in verbs formed with *-del*: *vàzdel* «to lift» - imperative *vazde* «lift!», *kìdel* «to gather» - imperative *kide* «gather!», etc. This vowel can also be added in verbs with two root-final consonants, e.g. *bistrel* «to forget» - imperative *bistre* or *bister*. Some verbs ass the vowel *i* (*xuli* «lower!», *ukli* «mount!», *uští* «rise!») or the vowel *a* (*ava* «come!», *da* «give!» - variant of *de*).

The imperative has other variants in the various dialects, but, considering the strong emotional character of the imperative and the fact that the differences do not cause any confusion, all forms are acceptable in the Common Language.

The stressed particle *ta* is often added to the singular imperative: *dikh-ta* «just look!», or *vaker-ta* «just tell!»

The imperative serves to express not only a command, but can also have a narrative function: *o éhavo ròde, ròde, i angrusti na arakhlă* «The boy searched and searched, but didn't find the ring».

## FORMATION OF TENSES DERIVED MORPHOLOGICALLY

A quite large number of particles are used throughout the different dialects to express various modalities. Occasionally the same particle will have different values depend on the dialect. The most important of these particles in the Common Language are the following:

### 1) Prepositive particles

**ka** - this particle is used to form the future tense in the Balkan branch of dialects: **ka dikhas** «we will see». This can be elided into **k** before a vowel: **k aven** or **ka aven** «they will come». There are also the variants **kam** and **kan**, which are not recommended in the Common Language, as they are too regional. The particle **ka(m)** derives from the verb **kamel** «to want», while the particle **do** in Albanian and **ða** in Greek are innovations.

**më** - This is synonymous with **ka** in some dialects of Bulgaria: **më aven** «they will come». It is not used in the Common Language, as it is too regional.

**ne** - «let», exists in many dialects in Yugoslavia and probably derives from the Serbian word with the same meaning **neka**. It is not recommended in the Common Language, and is replaced **mek**.

**mek** - «let», is very widespread in Central Europe. It probably derives from the verb **mekhel** or **mukhel** «to let», in the same way that **ka(m)** derives from **kamel**: **mek aven** «let them come».

**te** - In the present tense, this can be synonymous with the particle **mek** (**te aven!** «let them come»), or it can function as a conjunction (**mangav te siklövav** «I want to learn», whereas in other tenses it frequently has a conditional meaning (**te phirden sig, šaj reses len** «if you walk fast, you will catch up with them»). The combination of this particle with a verb is often treated as a special mood, called the subjunctive. Its use is briefly explained in the section on syntax.

### 2) Postpositive particles

**a** - This is very important, and has two different meanings depending on the dialect group. In most dialects of the Balkan group and in some archaic forms of Central Europe **a** expresses the present tense: **avèna** «they come, they are coming», whereas in other dialects it expresses the future tense: **avèna** «they will come». In the Common Language the use of the present tense verb without the particle **a** is recommended (**aven** «they come, they are coming», and in the future tense, the use of **a** or **ka**, depending on the variant used, is recommended (**ka aven = avèna** «they will come»).

- **sas, sine (hine)** - Depending on dialect, this particle expresses either the continuity of action in the past: **avènsas** «they were coming», **sutās sine** «they were sleeping», or the antecedence of an action in comparison to the contextual tense (i.e. the pluperfect): **avile sine** «they had come». In the Mahazeri dialect of Prishtina the combination of the particles -a ... **hine** is used with a perfective meaning: **ma ker akaja buti, me kerdūma la hine** «Don't do this work, I have already done it.»

- **ne, na:** Combines with the copula to express a continuous state in the past.

- **tano:** This particle, which declines according to the "long" pattern of adjectives, fem. **tani**, pl. **tane**, usually expresses continuity of action: **phirel tano** (Mahazeri dialect) «he is walking», **sutā tano** «he was sleeping», etc. It can also function as a copula: **but butārni tani akaja čhaj** «this girl is very hard-working».

- **bi:** This particle of a Slavic origin, which has a conditional meaning, occasionally also comes before the verb. It can be used in conjunction with the particle **sine**.

These particles are written as part of the verb when no other element can be inserted between the particle and the verb, e.g. before the particles -a, -sas and -ne (-na): **avèna, sovènsas, našlìsas, hiúmne**, etc. Other particles are written separately, for another element may be inserted between the particle and the verb, e.g. **dikhlo man hine** «he had seen me». In the case of the particles such as **ka, ne**, etc., the possibility of another element's being inserted between the particle and the verb depends on dialect. Thus, these particles are written separately, and each follows the syntax of its own dialect as regards intervening elements.

These particles combine with each other in some dialects in order to express fine nuances, e.g, in Mahazeri of Prishtina, which presents a extraordinary wealth in expressing fine nuances by combining particles. The system of particles in each dialect definitely needs to be studied in detail to collect this data in a monographic form.

#### NOMINAL LEXEMES FORMED FROM VERBS

Certain lexemes which decline like nouns or are invariable are formed from the verbal root:

##### a) Participles

We have examined participles above. They are declined according to the "long" pattern of adjectives.



## b) Verbal adverbs

Gerunds are formed by adding the ending **-ndoj** (or one of its variants **-ndor** or **-ndo**) to the verbal root, e.g. **prasta-ndoj** «(while) running». When the root ends in a consonant, a connecting **-i-** is inserted: **phir-i-ndoj** «(while) walking». When the root ends in an unstressed vowel and take an epenthetic **-v-**, the **-v-** is retained and the connective **i-** is added to it, yielding the construct: **R-v-i-ndoj**, e.g. **si-ka-v-i-ndoj** «(while) showing», **ro-v-i-ndoj** «(while) crying». The gerund is invariable.

In practice, the gerund formed from reflexive/passive verbs is not encountered. Notwithstanding, the enrichment of the Common Language by means of these virtual forms can be proposed: **siklōndoj** or **siklōvindoj** «(while) learning», **kalōndoj** or **kalōvindoj** «pulling out».

## c) Verbal nouns

We have mentioned the ending **-pe** and its variants **-pa**, **-pen**, **-be**, **-ba**, **-ben**, **-mos**, **-mo**. This ending forms the verbal noun from a verbal root, e.g. **prasta-pe** «running», **paća-pe** «belief», etc. When the root ends in a consonant a connective **i** is inserted: **phir-i-pe** «stroll», **naś-i-mos** «escape», etc. When the root ends in an unstressed vowel and it takes an epenthetic **-v-**, the **v** can be retained or omitted: **ro-v-i-pe** or **ro-i-pe** «crying», **sika-v-i-pe** or **sik-a-i-pe** «instruction». Verbal nouns from type A verbs must be carefully distinguished from verbal nouns made from their causatives, e.g. **dara-pe** «fright» (vs. **dar** «fear») **dara-l** «to fear», but **dara-i-pe** or **dara-v-i-pe** «(act of) frightening» from **dara-v-el** «to frighten».

Verbal nouns follow a special declension:

- The singular B-Form is **-mas-**: **roimas-** from **roipe** «crying». The B-Form links with postpositions like any other noun, e.g. with the postposition **-qe**: **roimasqe** «in order to cry», with **-ça**: **roimaça** «with crying», as well with the others. In some dialects in which the verbal adverb is lacking, the verbal noun with the postposition **-ça** substitutes for this, e.g. **avili roimaça** for **avili roindoj** «she came crying». The B-Form also has the dialectal variants **-mnas-** and **-mnās-**, which are used less often in the Common Language.

- The plural of the verbal noun is rare: the A-Form is characterized by the affix **-mata**, while the B-Form is **-maten-**, as we have seen in the chapter on the morphology of the masculine noun. A small number of dialects, e.g. the Kabuži of Korçë, use the Form-B only in the plural instead of the singular: **ali roimnānça** «she came crying». This tendency is not recommended in the Common Language.

The verbal noun also expresses the category of voice, since it is also formed from the reflexive/passive verb, e.g. *siklōpe* «learning, study» from *siklōl* «to learn», or *buxlōpe* «widening» (passive process) from *buxlōl* «to be widened», compared to *buxlipe* «wideness», etc. Many dialects have lost the ability to form verbal nouns from reflexive/passive verbs, however this ability needs to be widely developed in the Common Language, since it is a source of natural enrichment.

Verbal nouns with *-pe* (or its variants) are sometimes considered a kind of infinitive, and do indeed fulfil that function. In addition, verbal nouns express the category of voice and sometimes can take an object, and from this aspect it resembles the infinitive, however, on the other hand, abstract nouns with *-pe* are formed not only from the verbal root, but also from nominal bases (e.g. *boripe* «bridehood» from *bori* «bride», or adjectival bases (e.g. *ternipe* «youth» from *terno* «young», and for that reason it is difficult to consider this form a simple infinitive. In fact, it is a transition between the verbal class and the nominal class.

#### THE COPULA

The copula «to be» in Romani exists only in the present tense, and is lacking completely in some dialects. Two series of forms are acceptable in the Common Language:

|        | Series with i     |              | Series without i |            |
|--------|-------------------|--------------|------------------|------------|
|        | singular          | plural       | singular         | plural     |
| 1st p. | <i>sium, siom</i> | <i>siam</i>  | <i>sem</i>       | <i>sam</i> |
| 2nd p. | <i>sian</i>       | <i>sien</i>  | <i>san</i>       | <i>sen</i> |
| 3rd p. | <i>(i)si</i>      | <i>(i)si</i> | <i>si</i>        | <i>si</i>  |

The series with *i* is preferred in the Balkan variant of the Common language. This variant is also characterized by the endings  $PS_a$  of the type *-ūm* or *-ōm*, *-ān*, *-ā(s)*, etc. in the past tense.

The series without *i* is preferred in the northern variant of the Common Language. This variant is also characterized by the endings  $PS_a$  on the type *-em*, *-an*, *-a(s)*, etc. in the past tense.

Both variants have a variant of the copula with initial *h-* and *s-*: *hium*, *hian*, etc., or *hem*, *han* (or *hal* in Czechoslovakia). Some dialects, especially in Greece, but also in Albania, prefix *i* to the forms of the copula: *isininom*, *isininan*, etc., as the presence of an initial *i* in the 3rd person (*isi*) is very widespread. In addition the interposition of an *[-n-]* or *[-nj-]* after between the element *si-* and the copula's endings is also

seen in the southern dialects, e.g. in the dialects of the Mečkari of Muzeqe: (i)sinom, (i)sinan, etc., or of the Erlikes of Sofia: sinõm, sinãn, etc. All these forms are acceptable in the Common Language, as long as a comprehension is not hindered.

The semantic and syntactic relationship between the copula and the verbs ovel «to become» and avel «to come», as well as adjectival types such as tano, fem. (tani, pl. tane) or talo, fem. tali, pl. tale is analyzed in the section on syntax. It will suffice to say here that in the past tense the copula has no form comparable with these verbs. The meaning in the past tense is usually provided by the modal particle -ne, -sas, -àsa, etc. depending on dialect, e.g. «I was» is expressed in the Tharegoni dialect as hiùmne, the Mahazeri dialect as hiùmna, and in the Mečkari dialect as sinòmsa or sinomàsa, etc. For the 3rd person, besides the form sine, the particle sas is also found, which we have seen in the formation of the imperfect past tense and which is very widespread, especially in the dialect of Central Europe and in the Gurbeti dialects. All these forms have an imperfect past tense meaning.

Romani has a special invariable copula to express price: mol, e.g. akava kher mol X miliònã «This house costs X million» It can be combined with the verb ovel with a syntactic nuance: o kher mol ovel X miliònã «The price of the house amounts to X million». In the past tense mol is used together with the respective form of the verb ovel or the copula: o kher mol sine X miolònã «The house cost X million».

## INVARIABLE LEXEMES

Sections on morphology dedicated to invariable lexemes is always shorter than others, for the reason that its study is really more appropriate under lexicography. Let us mention here the most important elements.

### ADVERBS

#### 1) Adverbs formed from old case forms:

- from the locative: **gave** «in the country», **khere** «homeward», **averčhane** «otherwise», **jekthane** «together», etc.

- from the ablative: **dural** «from afar», **kheral** «from home», **andral** «from inside», etc. (a complete list of these adverbs can be found on page 48 in the section dedicated to the declension of Romani nouns).

As we have already seen, this class of adverbs can be enriched with international words with the ending **-al** and having a the connotation of an adverb of manner: **normal**, **natural**, **original**, etc., as well as the original creation: **ačaral** «usually, as a rule», from **ačar** «rule», or **automatikal**, etc.

#### 2) Adverbs formed from adjectives:

- with the same form as the adjective when it ends in a consonant (the "short" type): **šukar** «good», **kuć** «dear», etc., as well as with comparative forms with **-dar**.

- with the ending **-es** in adjectives of the "long" type: **rromanes** «in Romani; in a well-behaved manner», **čaćes** «in a truthful manner», **lačhes** «well», **lokhes** «easily», etc.

#### 3) Adverbs formed on the basis of a demonstrative: Most of these indicate location or movement to or form a place:

- Place where something is located: The suffix **-the** is added to the stem of demonstrative adjectives: **odothe**, **gothe**, **othe**, etc. «there». When the stem contains a **k**, the suffix the unaspirated variant **t** emerges: **akate** «here», **okote** «there», etc. In some dialects the adverbial suffix, for location, is **-j**: **akaj**, **odoj**, etc.

- Moving away from a place: In a similar manner adverbs are formed which indicate the moving away from a place. The suffix **-thar** is added to the stem of the demonstrative adjective (**-tar** when the stem ends in **k**), e.g. **odothar**, **othar** (Mečkari), **okotar** «from there, thence», **akatar** «from here, hence», etc.

- with motion towards something: The suffix **-rik** or the variant **-ringa** is added to these same stems: **adarik** «up to here», **odorik** «to there», etc. This suffix is probably related to the Romani word **rig** «side».
- motion along: This uses the suffix **-nă** or **-nia**: **akănă**, **akănia** «this way», etc.

Another adverb with temporal or locative meaning is formed on the base of a demonstrative, i.e. the word **akana** «now». A series of adverbs with temporal meanings can be formed in an analogous manner, e.g. **okona**, **adana**, etc. Such adverbs, especially **okona** «then, at that time», are very useful, since Romani usually uses the foreign words **ònda**, **ather**, **atuné**, etc., to express this concept. The word **atòska** «then» is also found, especially in Gurbeti. However, this word usually has a relative meaning, i.e. «therefore» rather than a purely temporal one. It is good to retain this distinction in the Common Language: **okona** «then, in that case» besides **atòska** «therefore».

4) Adverbs formed from a prefix and a nouns denoting units of time: This prefix, which can be **a-**, **e-**, or **ada-** appears to derive from an old demonstrative. We have the following forms:

- from **dive(s)** «day»: **avdive(s)**, **avdie**, **aděš**, etc. «today» ;
- from **rat** «night»: **erati**, **irat**, **arăt**, **arati**, etc. «this evening» (also sometimes «tonight»)
- from **bersé** «year»: **adabersé** «this year».

5) Adverbs ending in **-ine(a)** in the Kalderaš dialect:

The Kalderaš dialect has two series of adverbs ending in **-ine** or **-one** which are formed from the names of the days of the week: **lûja** «Monday», **luine** «on Monday», **sàvato** «Saturday», **savatone** «on Saturday», etc., whereas for Sunday a genuine locative case is used: **kurke** «on Sunday», from **kurko** «week» and [purano] **kurko** «Sunday» (literally «the old week»).

6) Interrogative adverbs

Interrogative adverbs of place and time are closely connected with those formed on the basis of demonstratives, except that their base is the interrogative element **ka-**. From the semantic point of view, these are quite confused throughout the dialects, however it is recommended that they be differentiated in the Common Language. The interrogative adverbs pertain to the four directions:

- place where something is located: **kaj** «where» (without motion), also **kate?**, e.g. **kaj** (or **kate**) **bešel?** «Where does he live?»
- moving away from a place: **kotar?** and the variant **katar?** «where

from, whence», e.g. **Kotar aven?** «Where do you come from?».

- with motion towards something: **karik?** or **korik?** or the shorter variant **kor?** «where (to), whither?». This adverb is quite rare and is usually replaced by **kaj?**, e.g. **karik gelo** «Where did he go?».

- motion along: **kànă?** «along what/». This is also quite rare and sometimes can indicate motion towards something: **kànă gelăn?** **Kànă phirdăn?** «where/along what did you go?», «Where/along what did you walk?»

Besides the adverb **kotar**, which is well preserved in all dialects, usually one of the other interrogatives has assumed the meaning of the others, e.g. **kor** in the Tharegoni dialect, which has replaced **kană** for "motion along" and often is used in place of **kaj** to express "place where something is located."

From the same base **ka-** the temporal adverbs are also formed: **kana?** «when?», e.g. **Kana avilăn?** «When did you come?»

We must also mention here the interrogative adverb of manner: **sar?** «how?», which is not connected with any system of word formation: **Sar san?** «How are you?».

7) Adverbs which are not related to any other words:

These are: **sig** «quickly», **mišto** «well», and naturally the foreign adverbs. Here belong also the words meaning «here», which vary greatly from dialect to dialect and are generally not connected to a system. The most common are **gäja** in Gurbeti and in many dialects of Central Europe, and **azukha**, **azahar**, etc. in the Balkan dialects.

Although the ending **-ta** is probably related to the postposition **-Ōe** or with the suffix **-ta/-the** of adverbs in Group 3, whereas the ending **-al** is of a very old origin, we cannot link the words **inčäte** or **inčal** «at a distance» with any root **\*inča**, for the reason that no such root exists in modern Romani.

#### PREPOSITIONS AND POSTPOSITIONS

Adverbs of place ending in **-e** and **-al** usually correspond to the prepositions below:

| Adverb      |              | Preposition          |
|-------------|--------------|----------------------|
| <b>pale</b> | <b>palal</b> | <b>pal-</b> «before» |
| <b>opre</b> | <b>opral</b> | <b>opr-</b> «above»  |
| <b>tele</b> | <b>telal</b> | <b>tel-</b> «under»  |
| <b>paše</b> | <b>pašal</b> | <b>paš-</b> «near»   |

The article is connected immediately to the preposition and is pronounced as one word together with it: **pal-o khel** «in front of the house». When the article contains a front vowel, the rules for softening of the **l** are in effect.

These prepositions usually take a connective vowel when they precede a consonant: **paše** (or **paša**) **le rroma** «near the Rom».

We must emphasize here a peculiarity of these prepositions. When they accompany an adjective, the adjective takes the locative postposition **-Ōe**, and we could say that we are dealing with a pre-postposition: **pal-...-Ōe**, **opr-...-Ōe**, etc., e.g. **pala manŌe** «in front of me», **pal amenŌe** «in front of us», etc. It can also be seen that in this case, these take a connective vowel before the initial consonant of the adjective.

In addition to these prepositions which form a series, let us also mention **and-** (or **an-**) «in» and **kotar** «from». Romani also has a pre-postposition with the meaning «without», which we have met above: **bi...-go**, e.g. **bi kheresgo** «homeless». As we have seen, this structure also exists in the languages of India. It can also be connected with the old postposition **-go** with the meaning of manner or position, which has been preserved only in some dialects of Turkey and Albania in expressions such as **koćiengo** «one one's knees», **pirengo vasten go** «on one's hands and feet», **bulāgo** «seated», **peresgo** «on one's stomach», etc.

#### CONJUNCTIONS

The study of conjunctions belongs under syntax.

#### PARTICLES OF NEGATION

The system of negation in Romani deserves a study of its own. Let us mention here only the main points, although his question belongs to syntax more than morphology.

The indicative mood is negated, according to dialect, with the particles **na**, **ni**, or **ći**. Since **na** is the most widespread and is in common with the languages of India, it is preferred in the Common Language, however **ći** is used extensively in the Northern dialect, whereas it is recommended that **ći** be limited to word formation: **nivar** «never», **nisar** «in no way», **nijekh** «no one», etc. A complete list is given in the synoptic table of correlative words. A repeated **ni** is also used with the meaning "neither ... nor," e.g. **na avile ni e phure ni e terne** «neither the old no the young came».

In all forms of the Common Language it is recommended to use **ći** to negate a verb after the interjection **na**. In this way a misunderstanding in response to a question containing a negative element can be avoided, e.g. to the question **Na manges te xas?**

«Don't you want to eat?» it is better to answer **Na, cí mangav** «No, I don't (want)», than **Na, na mangav**, because the second response could be understood as **Na, na, mangav** «No, no, I do (want)».

Other particles of negation are **bi**, which is used in word formation and the pre-postposition **bi...-go** «without», and **ma** «don't!», which is used with imperatives: **ma azba!** «Don't touch!» (in India the corresponding particle is **mat**). This is also used in dependent clauses with **te** (the relative or conditional word): **Dikh śukar ma te peres!** or **Dikh śukar te ma peres!** «Watch that you don't fall!». In some dialects this particle is replaced with **na**, but it is advisable to retain this preserve this in the Common Language with its semantic distinction.



## THE RELATIONSHIP BETWEEN VARIANTS OF THE COMMON LANGUAGE

Throughout this study we have referred several times to the two variants of the Common Language. Like the Common Language itself, these are not two petrified models, but are rather goals, two foci for the convergence of dialectal features. As we have seen, the two dialects which materialize from these variants are a summary of natural elements which are the most favorable for bringing together, union, and unification of the dialects of the Balkans with the dialects of Central Europe. As the same time, thanks rules of reading, which are distinct for each dialect, the realization of their variants remains very close to spontaneous speech, which is the most certain source for supporting the development of any language. More than 400 years ago Martin Luther, who worked to standardize the German language, avowed, in order to know how to speak German, "observe how the mother in the home, the child in the street, and the simple man in the market speak".

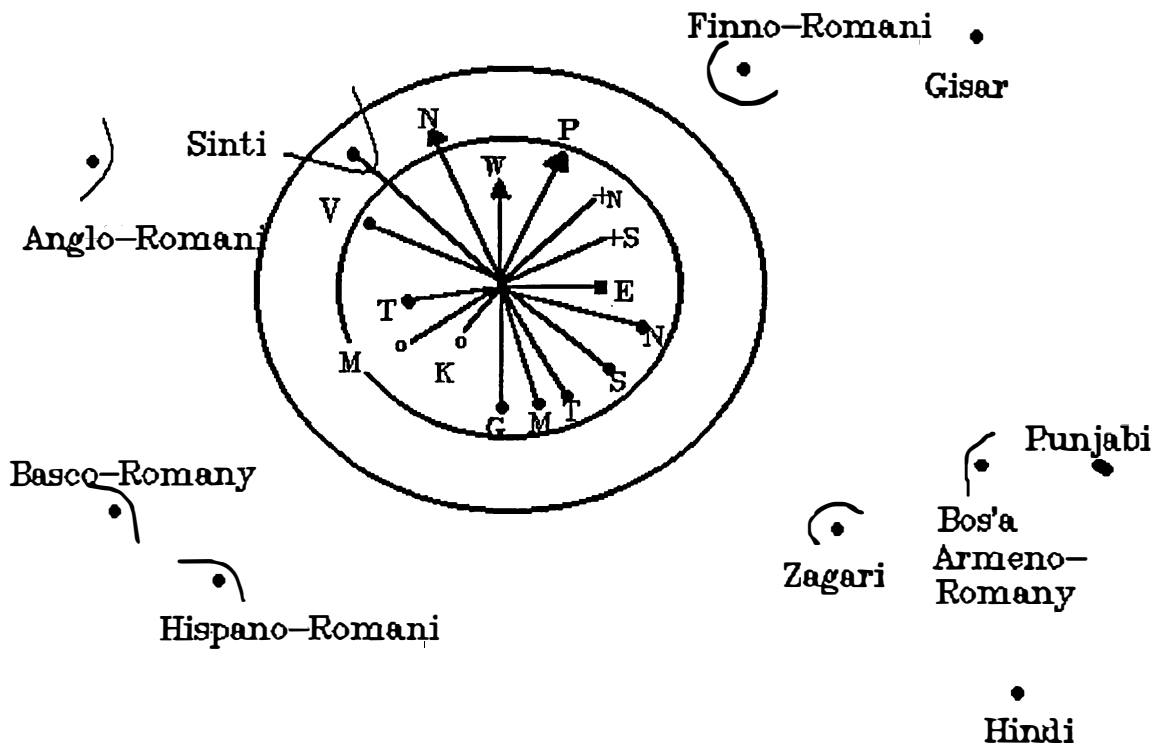
Thus, while relying heavily on actual local dialects and while reinforcing their common elements, Romani can manage to concentrate itself around two variants. Practice will demonstrate and develop this process further, until one distinctive form is created, or we will retain both variants forever side-by-side. Practice will also show the way to develop the pronunciation of these variants, or each dialect will retain its own distinctive features of pronunciation and with its own distinctive reading rules, or the inclination towards a distinct pronunciation will appear, which could be a literal realization of the letters. The common orthography, *maškarunto lekhipe*, accommodates both these possibilities. In this way, this orthography meets the requirements of the 1st Romani Congress, "nane jekh dialèkta po lači le avere dialekten@ar, ama mangas jekh čhib maškarutni (centr@lo) normalisardi, te vakeras la and-e kongrèsã thaj te lekhas la and-e internacion@lo lila," i.e. "No dialect is better than any other, but we require a common normalized language (central) whereby we can speak as congress and to write international papers" (Komisia e čhibãqi / The Linguistic Commission, Romano Kongreso, London April 8-12, 1971), chaired by the Yugoslav Rom poet Slobodan Berberski).

In fact, the two variants which we have defined above represent a kind of convenience, for each is closer to spontaneous dialects than to the single central language, and, at the same time, the differences between the two variants is actually quite small. It is useful here to give these differences on a formal table, by arranging these which are most characteristic. On this table are placed differences which can be distinguished in the variants in question, e.g. the pronunciation of the letter *ž* as [j] or a soft [ž], whereas differences which coexist inside this variant are not reflected. These are, e.g. the realization of [k / k' / q / ç / t'] for k before a front vowel, since different realizations

are encountered in the Balkan variant as well as in the Central European variant. In the area of morphology we can also mention the difference -ă / -a in feminine endings (boră / bora), both the pre-jotizing form -ă and the nonpre-jotizing -a are encountered in the two variants. The table also presents differences which, without being distinctive features between the variants of the Common Language, are encountered so often in one of them that we can say that they are characteristic of it. Thus, for example, the verbal affix -sar-, which is very widespread in Central Europe, but is rather rare in the true Balkan dialects, is reflected.

The table presents only phonetic, morphological, and lexical differences, and not syntactic ones, for the reason that particular differences in syntax directly connected with the variants in question have not been noted.

Diagram #1  
 The Position of Romani Dialects in Comparison  
 to Common Romani, and Hindi and Punjabi



a) True Romani dialects:

Inside the circle are shown the position of 16 Romani dialects from certain European countries in relation to the Common language (which is in the center of the circle):

- dialects of Poland are indicated with a triangle, as follows:

- N = the dialect of the field plains zone
- W = dialect of the mountainous zone;
- P = Kalderaś

- dialects of the Soviet Union are indicated with a cross: N = northern dialects, and S = southern dialects.

- Erli dialect of Sofia, Bulgaria is indicated with a E and a square.

- dialects of Yugoslavia are indicated with a dot, as follows:

- N = Ubit dialect of Serbia;
- S = Thamari of Prokuple.

T = Topani of Skopje .  
M = Mahazer of Prishtina  
G = Gurbeti of Obilić  
V = the dialect of Vojvodina  
T = the old dialect of Titograd

- dialects of Albania are indicated with a little circle::

K = Kabuži of Korçë  
M = Meçkari of Tirana

b) Peripheral dialects: Finno-Romani, and Zagari of Iran.

ć) Creolized dialects: the diagram shows Anglo-Romani of Great Britain, Basco-Romany in the Basque regions, Hispano-Romani of Spain, and Boša of Armenia, as well as the semi-creolized Sinti dialect of Italy.

d) We have added to this diagram two languages of India: Hindi and Punjabi, as well as the Gisar dialect of Tajikistan, which are sometimes considered to be related to Romani.

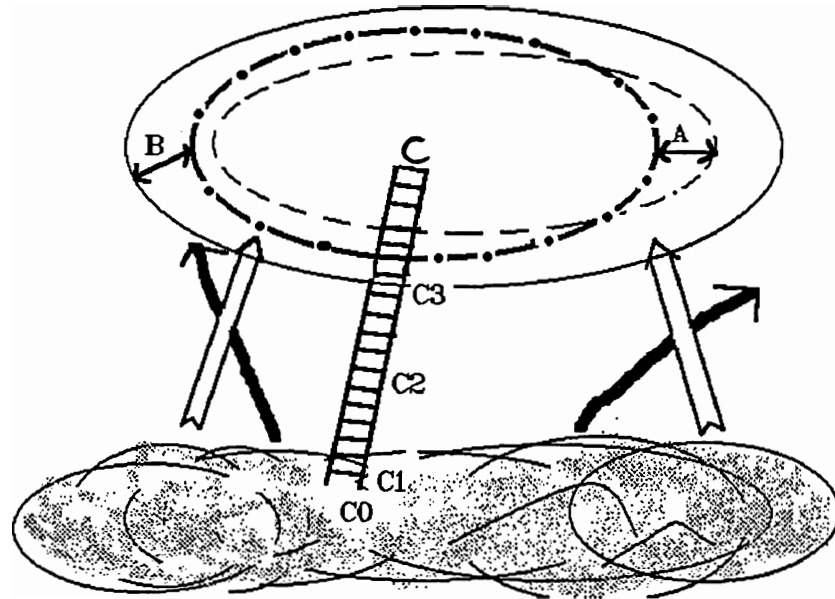
Meaning of the symbols: Those dialects with little dialectal variation (the standard language or very local dialects) are indicated with by their small size, e.g. with a dot, whereas the larger size of the symbol indicates variation in the language/dialect in question.

The material distance from the center of Romani expresses the dialectal difference, measured by lexico-statistical methods (the radius of the large circle encompass the Romani dialects is equal to 1, beyond which the distance between two dialects shows that they are separate languages).

It can be seen that, except for the plains dialect of Poland, all other Romani dialects have a distance of less than one between them.

When one of the para-Romani dialects exhibits a relatively well-preserved grammar, this is indicated with half circle (Finno-Romani and Zagari), whereas when it is creolized it is indicated with a parabola; the more creolized the dialect, the more open the parabola.

Diagram #2  
The Path of Romani Standardization



The above diagram shows Romani in its spontaneous spoken condition, parcelled dialectically. The solid zones represent local foreign words which make up foreign material which is being eliminated or replaced gradually with Romani words. However, international and foreign words that have been assimilated are not distinguished from the native vocabulary.

The white path symbolizes convergence which results from standardization, whereas the black path represents the enrichment and expansion of the language through word formation and acceptable foreign words.

The intersecting ellipses above represent the standard with two centers (the northern variant and the Balkan variant), as is proposed for the written referential language.

The three levels of tolerance are as follows:

A) between the two standard written variants;

B) between the different spoken realizations of the spoken standard, which arise from the written standard through particular lexical rules for each dialect (the ellipsis with dashes);

C) the third level is represented as a ladder which symbolizes the continuity between the spoken forms C1, C2, C3, etc. These levels are used according to custom and spread from the familial idiolect C up to one of the standardized realizations. Naturally the speaker will choose a realization close to his own.

Table #1: Relationship between Nouns and Verbs in Romani

| Verb       |        |         |
|------------|--------|---------|
| Transitive |        |         |
| present    |        |         |
| 3rd p.     | kerel  | dikhel  |
| past       |        |         |
| 1st p.     | kerdüm | dikhlüm |
| 2nd p.     | kerdän | dikhlän |
| 3d p.      | kerdä  | dikhläs |

|   | Verb         |          |       | Adj.  | Noun   | Form      |
|---|--------------|----------|-------|-------|--------|-----------|
|   | Intransitive |          |       |       |        |           |
| present   | nakhel       | bešel    | merel |       |        |           |
| 3rd p.  |              |          |       |       |        |           |
| past  |              |          |       |       |        |           |
| 1st p.  | nakhlüm      | bešlüm   | mulüm |       |        |           |
| 2nd p.  | nakhlän      | bešlän   | mulän |       |        |           |
| masc.   | nakhlo       | bešlo    | mulo  | kalo  | bakro  | A-Form m. |
| 3rd p.  | nakhlä =     | bešli    | muli  | kali  | bakri  | A-Form f. |
| fem.  | (nakhles)    | (bešles) | mules | kales | bakres | B-Form m. |
|   | (nakhlä)     | (bešlä)  | mulä  | kalä  | bakrä  | B-Form f. |
| Forms in parentheses are used with a postposition:<br>mules-ge, kalä-ge, etc. |              |          |       |       |        |           |

Table #2: Synoptic Overview of Plural Forms of Nouns Ending in a Consonant in the Singular

| end-<br>ing | masculine    |              |            | feminine      |                 |                   |            |
|-------------|--------------|--------------|------------|---------------|-----------------|-------------------|------------|
|             |              |              | with<br>-a |               |                 | with<br>-a        | with<br>-ă |
| -b          | drab         | medicine     | draba      | dab           | slap            | daba              |            |
| -p          | sap          | snake        | sapa       | čhib          | tongue          | (čhiba)           | čhibă      |
| -m          | drom         | road         | droma      | pošom<br>čham | wool<br>face    | pošoma<br>(čhama) | čhamă      |
| -d          | dad          | father       | dada       |               |                 |                   |            |
| -t          | grast        | horse        | grasta     | bust          | skewer          |                   | bustă      |
| -n          | kan          | ear          | kana       | phen          | sister          |                   | phenă      |
| -g          | śing         | horn         | śinga      | jag           | fire            | jaga              |            |
| -k          |              |              |            | čhik<br>phak  | mud<br>arm      | čhika<br>phaka    |            |
| -kh         | rukḥ         | tree         | rukha      | jakh          | eye             | jakha             | jakhă      |
| -s          | phus         | straw        | phusa      | kris          | court           |                   | krisă      |
| -z          |              |              |            | diz           | city            | (diza)            | diză       |
| (-z)        |              |              |            |               |                 |                   |            |
| (-ž)        |              |              |            |               |                 |                   |            |
| (-ć)        |              |              |            | koć           | knee            | koća              | koćă       |
| -ž          |              |              |            | laz           | shame           | laža              |            |
| -l          | phral        | brother      | phrala     | phal          | board           | phala             | phală      |
| -j          | muj          | mouth        | muja       | baj           | sleeve          | baja              |            |
| -v          | guruv        | ox           | guruva     | zuv<br>suv    | louse<br>needle | zuva<br>(suva)    | suvă       |
| -x          | tirax        | shoe         | tiraxa     | mamux         | sloe            | mamuxa            |            |
| -r          | angar<br>por | coal<br>pora | angara     | por           | guts            | (pora)            | poră       |

Note: The variants in parentheses are less preferred in the Common Language.

Table #3: Sandhi System

|     |                       | Sandhi I | Sandhi II                          | Sandhi III<br>(with palatalization before<br>i and e)          |
|-----|-----------------------|----------|------------------------------------|--|
| -q- | after n               | /g/      | none                               | [ng] [ng'] [nʒ]<br><br>with assimilation:<br>[ñ] [ň]           |
|     | in other<br>positions | /k/      | - -sk-<br>+ -hk-<br>+ -k-<br>+ -s- | [sk] [sk'] [sč]<br>[hk] [hk'] [hč]<br>[k] [k'] [č]<br>none [s] |
| -ṭ- | after n               | /d/      | none                               | none [nd]<br><br>with assimilation [n]                         |
|     | in other<br>positions | /t/      | - -st-<br>+ -ht-<br>+ -t-          | [st]<br>[ht]<br>[t]  |
| -ç- | after n               | /c/      | none                               | none [c]   |
|     | in other<br>positions | /s/      | - -s-<br>+ -h-<br>+ -j-<br>+ -0-   | none [s]<br>none [h]<br>none [j]<br>none [0]                   |



Table #4: Sandhi System (example of application)

|       |            | Sandhi I | Sandhi II                  | Sandhi III<br>(after i and e)   |
|-------|------------|----------|----------------------------|---|
| -q    |            |          |                            |   |
| lenqo | «their» m. | /lengo/  |                            | [lengo]<br>[leño]   |
| lenqi | «their» f. | /lengi/  |                            | [lengi] [leng'i] [lenzi]<br>[leñi] [leñi] [leñi]  |
| laqo  | «her» m.   | /lako/   |                            | [lako]  |
| laqi  | «her» f.   | /laki/   |                            | [laki] [lak'i] [laći]   |
| lesqo | «his» m.   | /lesko/  | -ske-<br>-h-<br>-k-<br>-s- | [lesko]<br>[lehko]<br>[leko]<br>[leso]  |
| lesqi | «his» f.   | /leski/  | -ske-<br>-h-<br>-k-<br>-s- | [leski] [les'ki] [lesći]<br>[lehki] [lehk'i] [lehći]<br>[leki] [lek'i] [leći]<br>[lesi] |

Naturally, we must also add here the variant with postpositions, e.g. final -ro, -ri, -re: lenqoro, lenqiri, lenqro, lenqere, etc.

| -@     |             |          |                     |          |
|--------|-------------|----------|---------------------|----------|
| len@ar | «from them» | /lendar/ |                     | [lendar] |
| la@ar  | «from her»  | /latar/  |                     | [latar]  |
| les@ar | «from him»  | /lestar/ | -st-<br>-ht-<br>-t- | [lestar] |

Here belongs also the postposition -@e.

| -ŕ      |             |            |                         |   |
|---------|-------------|------------|-------------------------|---|
| lençar  | «with them» | /lanca(r)/ |                         | [lenca]                                     |
| laça(r) | «with her»  | /lasa(r)/  | -s-<br>-h-<br>-j-<br>-0 | [lasa]<br>[laha]<br>[laja]<br>[laa] or [la] |
| leça(r) | «with him»  | /lesa(r)/  | -s-<br>-h-<br>-j-<br>-0 | [lesa]<br>[leha]<br>[leja]<br>[lea]         |

Table #5: Comparison of the Postpositional Systems of Hindi and Romani

- the pronoun is masculine, in the singular, and in the A-

Form:(Hindi -ō / Romani -o)

laRke kā bāl  
e raklesqo bal  
«the boy's hair»

laRki kā bāl  
e raklāqo bal  
«the girl's hair»

- the pronoun is in the plural or in the B-Form

(Hindi -e / Romani -e)

- in the plural:

laRke ke ā'khé  
e raklesqe jakha  
«the boy's eyes»

laRki ke ā'khé  
e raklāqi jakha  
«the girl's eyes»

- in the B-Form:

laRke ke gā'w me  
e raklesqe gaves<sup>0</sup>e  
«in the boy's village»

laRkī ke ā'khé  
e raqlāqe gaves<sup>0</sup>e  
«in the girls village»

Table #6a: Declension of Nouns  
Types bakro m. «ram» and bakri f. «ewe»

|  | bakro m. «ram»               |                              | bakri f. «ewe»  |   |
|--|------------------------------|------------------------------|---|---|
|  | singular                     | plural                       | singular  | plural  |
| A-Form   | bakro                        | bakre                        | bakri   | bakria,<br>bakrā  |
| B-Form   | bakres                       | bakren                       | bakria,<br>bakrā  | *bakrien,<br>bakrēn,<br>•bakrān   |
| with postpositions:  |                              |                              |   |   |
| locative -@e<br>@=[d/t]  | bakres@e                     | bakren@e                     | bakria@e,<br>bakrā@e  | *bakrien@e,<br>bakrēn@e,<br>•bakrān@e   |
| ablative -@ar<br>@=[d/t]   | bakres@ar                    | bakren@ar                    | bakria@ar,<br>bakrā@ar  | *bakrien@ar,<br>bakrēn@ar,<br>•bakrān@ar  |
| dative -qe<br>q=[k/g]  | bakresqe                     | bakrenqe                     | bakriaqe,<br>bakrāqe  | *bakrienqe,<br>bakrēnqe,<br>•bakrānqe   |
| <sup>1</sup> possessives<br>m. -qo(ro)<br>f. -qi(ri)<br>pl. -qe(re)<br>q=[k/g] | bakresqoro<br>-qiri<br>-qere | bakrenqoro<br>-qiri<br>-qere | bakriaqoro<br>-qiri<br>-qere<br><br>bakrāqoro<br>-qiri<br>-qere | *bakrienqoro<br>-qiri<br>-qere<br><br>bakrēnqoro<br>-qiri<br>-qere<br><br>•bakrānqoro<br>-qiri<br>-qere |
| comitative<br>-ça(r)<br>ç=[s/h/j]  | bakreça                      | bakrença(r)                  | bakriaça,<br>bakrāça  | *bakriença(r),<br>bakrēnça(r),<br>•bakrānça(r)  |
| with the pre-postposition bi...-qo   |                              |                              |   |   |
| privative  | bi bakresqo                  | bi bakrenqo                  | bi bakriaqo,<br>bi bakrāqo                                      | *bi bakrienqo,<br>bi bakrēnqo<br>•bi bakrānqo   |

The \* indicates a Balkan form; the • indicates a Central European form.

<sup>1</sup>and their variants: m. -qero, -qro, f. -qeri, gri, pl. -qre.

Table #6b: Declension of Nouns  
Types rrom m. «Rom man» and rromni f. «Rom woman»

|  | rrom m. «Rom man»            |                              | romni f. «Rom woman»  |  |
|--|------------------------------|------------------------------|---|--|
|  | singular                     | plural                       | singular  | plural   |
| A-Form   | rrom                         | rroma                        | rromni  | rromnia,<br>rromnă   |
| B-Form   | rromes                       | rromen                       | rromnia,<br>rromnă  | *rromnien,<br>rromnën,<br>•rromnän   |
| with postpositions:  |                              |                              |   |  |
| locative -@e<br>@=[d/t]  | rromes@e                     | rromen@e                     | rromnia@e,<br>rromnă@e  | *rromnien@e,<br>rromnën@e,<br>•rromnän@e   |
| ablative -@ar<br>@=[d/t]   | rromes@ar                    | rromen@ar                    | rromnia@ar,<br>rromnă@ar  | *rromnien@ar,<br>rromnën@ar,<br>•rromnän@ar  |
| dative -qe<br>q=[k/g]  | rromesqe                     | rromenqe                     | rromniage,<br>rromnäqe  | *rromnienge,<br>rromnënge,<br>•rromnänge   |
| <sup>2</sup> possessives<br>m. -qo(ro)<br>f. -qi(ri)<br>pl. -qe(re)<br>q=[k/g] | rromesqoro<br>-qiri<br>-qere | rromenqoro<br>-qiri<br>-qere | rromniaqoro<br>-qiri<br>-qere<br><br>rromnäqoro<br>-qiri<br>-qere | *rromniengoro<br>-qiri<br>-qere<br><br>rromnëngoro<br>-qiri<br>-qere<br><br>•rromnänqoro<br>-qiri<br>-qere |
| comitative<br>-ça(r)<br>ç=[s/h/j]  | rromeça                      | rromença(r)                  | rromniaçã,<br>rromnăçã  | *rromniençã(r),<br>rromnënçã(r),<br>•rromnänçã(r)  |
| with the pre-postposition bi...-qo   |                              |                              |   |  |
| privative  | bi rromesqo                  | bi rromenqo                  | bi rromniaqo,<br>bi rromnäqo                                      | *bi rromniengo,<br>bi rromnëngo<br>•bi rromnänqo   |

The \* indicates a Balkan form; the • indicates a Central European form.

<sup>2</sup>and their variants: m. -qero, -gro, f. -qeri, qri, pl. -qre.

Table #6c: Declension of Nouns  
Types dad m. «father» and \*daj, •daj f. «mother»

|  | dad m. «father»             |                             | *daj, •dej f. «mother»                                      |  |
|--|-----------------------------|-----------------------------|---|--|
|  | singular                    | plural                      | singular  | plural   |
| A-Form   | dad                         | dada                        | *daj, •dej  | dia, daja, dǎ,<br>dà   |
| B-Form   | dades                       | daden                       | dia, daja,<br>dǎ, dà  | *dien,<br>dajen,<br>•dǎn   |
| with postpositions:  |                             |                             |   |  |
| locative -@e<br>@=[d/t]  | dades@e                     | daden@e                     | dia@e,<br>daja@e  | *dien@e,<br>dajen@e,<br>•dǎn@e   |
| ablative -@ar<br>@=[d/t]   | dades@ar                    | daden@ar                    | dia@ar,<br>daja@ar  | *dien@ar,<br>dajen@ar,<br>•dǎn@ar  |
| dative -qe<br>q=[k/g]  | dadesqe                     | dadenqe                     | diage,<br>dajaqe  | *dienqe,<br>dajenqe,<br>•dǎnqe   |
| <sup>3</sup> possessives<br>m. -qo(ro)<br>f. -qi(ri)<br>pl. -qe(re)<br>q=[k/g] | dadesqoro<br>-qiri<br>-qere | dadenqoro<br>-qiri<br>-qere | diaqoro<br>-qiri<br>-qere<br><br>dajaqoro<br>-qiri<br>-qere | *dienqoro<br>-qiri<br>-qere<br><br>dajenqoro<br>-qiri<br>-qere<br><br>•dǎnqoro<br>-qiri<br>-qere |
| comitative<br>-ça(r)<br>ç=[s/h/j]  | dadeça                      | dadença(r)                  | diaça,<br>dajaça  | *diença(r),<br>dajença(r),<br>•dǎnça(r)  |
| with the pre-postposition bi...-qo   |                             |                             |   |  |
| privative  | bi dadesqo                  | bi dadenqo                  | bi diaqo,<br>bi dajaqo                                      | *bi dienqo,<br>bi dajenqo<br>•bi dǎnqo   |

The \* indicates a Balkan form; the • indicates a Central European form.

<sup>3</sup>and their variants: m. -qero, -qro, f. -qeri, qri, pl. -qre.

Table #6d: Declension of Nouns  
Types marro m. «bread» and \*pani, \*paj m. «water»

|  | marro m. «bread»                                  |                              | *pani, *paj m. «water»   |  |
|--|---|------------------------------|--|--|
|  | singular  | plural                       | singular   | plural   |
| A-Form   | marro   | marre                        | *pani, *paj  | *pană,<br>*paja  |
| B-Form   | These have no free Forms-B, as they are inanimate |                              |  |  |
| with postpositions:  |   |                              |  |  |
| locative -@e<br>@=[d/t]  | marres@e  | marren@e                     | *paněs@e,<br>*pajes@e  | *paněn@e,<br>*pajen@e  |
| ablative -@ar<br>@=[d/t]   | marres@ar   | marren@ar                    | *paněs@ar,<br>*pajes@ar  | *paněn@ar,<br>*pajen@ar  |
| dative -qe<br>q=[k/g]  | marresqe  | marrenqe                     | *paněsqe,<br>*pajesqe  | *paněŋqe,<br>*pajenqe  |
| <sup>4</sup> possessives<br>m. -qo(ro)<br>f. -qi(ri)<br>pl. -qe(re)<br>q=[k/g] | marresqoro<br>-qiri<br>-qere                      | marrenqoro<br>-qiri<br>-qere | *paněsqoro<br>-qiri<br>-qere<br><br>*pajesqoro<br>-qiri<br>-qere | *paněŋqoro<br>-qiri<br>-qere<br><br>*pajenqoro<br>-qiri<br>-qere |
| comitative<br>-ça(r)<br>ç=[s/h/j]  | marreça   | marrença(r)                  | *paněça,<br>*pajesça   | *paněnça(r),<br>*pajença(r)                                      |
| with the pre-postposition bi...-qo   |   |                              |  |  |
| privative  | bi marresqo                                       | bi marrenqo                  | bi paněsqo,<br>bi pajesqo  | *bi paněŋqo,<br>*bi pajenqo                                      |

The \* indicates a Balkan form; the • indicates a Central European form.

<sup>4</sup>and their variants: m. -qero, -qro, f. -qeri, qri, pl. -gre.

Table #6e: Declension of Nouns  
Types muj m. «mouth» and sasuj f. «mother-in-law»

|  | muj m. «mouth»   |  | sasuj f. «mother-in-law»   |   |
|--|--|--|----------------------------|---|
|  | singular   | plural   | singular                   | plural  |
| A-Form   | muj  | muja   | sasuj                      | sasa  |
| B-Form   | This has no free Forms-B,<br>as it is inanimate              |  | sasa                       | *sasen,<br>•sasān   |
| with postpositions:  |  |  |                            |   |
| locative -@e<br>@=[d/t]  | mos@e,<br>mujes@e  | mon@e,<br>mujen@e  | sasa@e,                    | *sasen@e,<br>•sasān@e   |
| ablative -@ar<br>@=[d/t]   | mos@ar,<br>mujes@ar  | mon@ar,<br>mujen@ar  | sasa@ar,                   | *sasen@ar,<br>•sasān@ar   |
| dative -qe<br>q=[k/g]  | mosqe,<br>mujesqe  | monqe,<br>mujenqe  | sasaqe,                    | *sasenqe,<br>•sasānqe   |
| <sup>5</sup> possessives<br>m. -go(ro)<br>f. -qi(ri)<br>pl. -qe(re)<br>q=[k/g] | mosqoro<br>-qiri<br>-qere<br><br>mujesqoro<br>-qiri<br>-qere | monqoro<br>-qiri<br>-qere<br><br>mujenqoro<br>-qiri<br>-qere | sasaqoro<br>-qiri<br>-qere | *sasenqoro<br>-qiri<br>-qere<br><br>sasānqoro<br>-qiri<br>-qere |
| comitative<br>-ça(r)<br>ç=[s/h/j]  | moça,<br>mujeça  | monça<br>mujença(r)  | sasaça                     | *sasença(r),<br>•sasānça(r)                                     |
| with the pre-postposition bi...-go   |  |  |                            |   |
| privative  | bi mujesgo   | bo mongo<br>bi mujenjo                                       | bi sasaço                  | *bi sasenjo,<br>•bi sasānjo                                     |

The \* indicates a Balkan form; the • indicates a Central European form.

<sup>5</sup>and their variants: m. -qero, -gro, f. -qeri, qri, pl. -qre.

Table #7: Correlative Words

| Meaning      | Quest-ion  | Indefinite                                      | Demonstrative                   |                            | General                    | Negation   |
|--------------|--|---|---------------------------------|----------------------------|----------------------------|--|
|              |  |   | near                            | far                        |                            |  |
| locative     | kaj?   | varekaj,<br>kaj god                             | PB-the/-te,<br>-j               | PB-the/-ta<br>-j           | sarkaj,<br>sakaj           | nikaj,<br>nijekthane   |
| ablative     | kotar?   | varekotar,<br>kotar god                         | PB-thar/-<br>tar                | PB-thar/-<br>tar           |                            | nikotar,<br>niekthanes $\Theta$<br>ar,<br>khatinen $\Theta$ ar |
| temporal     | kana?  | varekana,<br>kana god                           | akana                           | odona,<br>okona,<br>atòska | savaxt                     | nivaxt,<br>nivar   |
| personal     | kon?   | varekon,<br>kon god,<br>jekh manuś              | PB-va                           | PB-va                      | sa,<br>sa(vor)e,<br>sarkon | niekh,<br>nikon,<br>na...khonik                                |
| object       | so?  | vareso,<br>so god,<br>khanć,<br>jekh<br>čhumuni | PB-va                           | PB-va                      | sa                         | niso,<br>khanć   |
| causative    | sosqe?<br>sos $\Theta$ ar?<br>sos $\Theta$ e?<br>(dial.) |   | akalaqe,<br>alakesqe            | odolaqe,<br>okolaqe        |                            | nisosqe,<br>nisos $\Theta$ ar                                  |
| quantitative | kobor?<br>kozom?<br>sode?<br>kedi?<br>(dial.)            | cärra,<br>xàca,<br>xàri,<br>harica              |                                 | odobor(om),<br>ponabrakha  | sa                         | nibor,<br>khanć  |
| qualitative  | savo?  | savona,<br>nesavo                               | asavo                           | gasavo                     | sarsavo                    | nisavo   |
| manner       | sar?   | varesar,<br>sar god                             | akajal,<br>ajal,<br>akhal, gäja |                            | sajal                      | nisar  |



1) Distinction in meaning: vare- corresponds to "some-" in English, while god corresponds to "some-... or other"; 2) PB = pronominal base: aka-, oko-, aga-, ga-, etc.; 3) Romani does not distinguish between these forms based on meaning.