

Univerzita Karlova v Praze

Filozofická fakulta

Ústav obecné lingvistiky

Obecná lingvistika

Zuzana B o d n á r o v á

Vend Romani:

**a Grammatical Description and Sociolinguistic
Situation of the so-called Vend dialects of Romani**

Vendština:

**gramatický popis a sociolingvistická situace tzv. vendských
dialektů romštiny**

Disertační práce

vedoucí práce – Mgr. Viktor Elšík, Ph.D.

2015

Acknowledgements

I would like to first thank all my consultants – especially Miska, Hajni and Mónika from Kisbajom – for their patience and assistance during the data collection. I also wish to thank Jakob Wiedner for endless consultations on several chapters of this thesis, Michael Beníšek for assisting me with some etymologies, Eliška Perevorská for revising the English of the thesis, and Viktor Elšík for encouraging me to initiate research on Vend Romani, supervising my studies and giving valuable suggestions. I am also grateful to Kimmo Granqvist for financial support during the writing of my thesis. Without him this work would have never been completed.

I dedicate this thesis to the memory of my mother, Eva Bodnárová (December 1958 – September 2012).

Prohlašuji, že jsem disertační práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne

podpis

Abstract

The thesis provides a detailed grammatical and lexical description of Vend Romani, an under-described dialect of Romani spoken in the Transdanubian region of Hungary, and describes its current sociolinguistic situation. The linguistic data are based on recordings of spontaneous narratives, semi-structured interviews, and linguistic elicitation by means of standardized dialectological questionnaires acquired during linguistic field research.

The thesis is structured into six main chapters: The first chapter deals with the sources of data and methods. The second chapter examines the factors that influence the sociolinguistic vitality of the dialect. The following three chapters are devoted to the grammatical description. The sixth chapter analyses the layers of borrowings in Vend Romani. The thesis also includes the basic vocabulary of Vend Romani translated to English.

Keywords Romani, Vend Romani, grammatical description, sociolinguistic vitality

Abstrakt

Disertační práce je podrobná gramatická a lexikální deskripce maďarské vendštiny, dosud málo popsaného dialektu romštiny zadunajské oblasti Maďarska, a základní popis její současné sociolingvistické situace. Jazyková data vychází jak z nahrávek spontánních narativů a polostandardizovaných rozhovorů, tak z elicitace za použití standardních dialektologických dotazníků pořízených v rámci terénního lingvistického výzkumu.

Práce je rozdělena na šest částí. První část se zabývá zdrojem dat a metodologií. Druhá část zkoumá faktory, které ovlivňují sociolingvistickou vitalitu vendštiny. Následující tři části jsou věnovány gramatickému popisu. Šestá část analyzuje vrstvy přejímek ve vendštině. Součástí práce je také základní slovník vendštiny v překladu do angličtiny.

Klíčová slova romština, vendština, gramatický popis, sociolingvistická vitalita

Table of contents

Abbreviations.....	13
1 Introduction.....	14
1.1 Structure of the thesis.....	14
1.2 Terminology.....	14
Ethnonyms with reference to the own group	15
1.4 Ethnonyms with reference to others.....	21
1.4.1 Boyash	22
1.4.2 Vlax Roma.....	22
1.4.3 Hungarian Roma.....	23
1.4.4 Beggars	23
1.4.5 Sinti.....	25
1.4.6 Romani and Hungarian ethnonyms.....	26
1.5 Geographical distribution.....	28
1.6 Dialect classification	30
1.7 Source of data and methodology.....	31
1.7.1 First stage of research	31
1.7.1.1 Earlier written sources	32
1.7.1.2 Previous field research	32
1.7.1.3 Researchers dealing with Roma in the region	32
1.7.1.4 Population census data	32
1.7.1.5 Online sources	33
1.7.2 Second stage of research.....	33
1.7.2.1 Field research and linguistic questionnaire	33
1.7.2.2 Method of data collection.....	34
1.7.2.3 Amount of data.....	35
1.7.2.4 Selection of consultants.....	38
1.7.3 Third stage of research.....	39
1.7.4 Fourth stage of research.....	39
2 Sociolinguistic vitality.....	40
2.1 Intergenerational language transmission.....	41
2.2 Absolute number of speakers.....	43
2.3 Proportion of speakers within the total population	43
2.4 Trends in existing language domains.....	44
2.5 Response to new domains and media.....	51
2.6 Materials for language education and literacy	52
2.7 Governmental and institutional language attitudes and policies.....	53
2.8 Community members' attitudes toward their own language	53
2.9 Amount and quality of documentation.....	56
2.9.1 Earliest sources on Central Romani in Transdanubia.....	56
2.9.1.1 János Szmodics.....	56
2.9.1.2 József Habsburg.....	59
2.9.2 Documentation of Vend Romani in Hungary	61
2.9.2.1 József Vekerdi	61
2.9.2.2 Ursula Glaeser et al.	62
2.9.2.3 Melinda Rézműves and Károly Bari	62
2.9.2.4 Charles University in Prague.....	62

2.9.3	Documentation of Vend Romani beyond Hungary	63
3	Phonology	64
3.1	Consonants	64
3.1.1	Consonant inventory and graphemes	64
3.1.2	Aspirates and aspiration	67
3.1.3	Consonant clusters	69
3.1.3.1	Word-initial clusters	69
3.1.3.2	Word-medial clusters.....	70
3.1.3.3	Word-final clusters	74
3.1.4	Geminates and gemination.....	76
3.1.5	Voice.....	78
3.1.6	Assimilation and dissimilation.....	80
3.1.7	Palatals and palatalization.....	81
3.1.8	Contraction and related sound changes	84
3.1.9	Apocope	86
3.1.10	Other sound changes	88
3.2	Vowels.....	89
3.2.1	Vowel inventory and graphemes	89
3.2.2	Vowel adaptation	93
3.2.3	Sequences of vowels	94
3.2.4	Diphthongs.....	95
3.2.5	Vowel length.....	96
3.2.5.1	Vowel quantity and processes of vowel lengthening	96
3.2.5.2	Length distribution	99
3.2.5.3	Analogical extension of vowel length in personal pronouns.....	100
3.2.5.4	Analogical extension of vowel length in the imperfective suffix.....	104
3.2.6	Apocope	106
3.3	Stress	108
4	Morphology	110
4.1	Nouns	110
4.1.1	Noun integration	110
4.1.1.1	Nouns borrowed from Slavic.....	110
4.1.1.2	Nouns borrowed from German.....	111
4.1.1.3	Nouns borrowed from Hungarian as masculines.....	112
4.1.1.4	Nouns borrowed from Hungarian as feminines.....	115
4.1.2	Noun formation.....	117
4.1.2.1	Abstract nouns and other minor derivations.....	118
4.1.2.2	Tree names.....	119
4.1.2.3	Occupation and profession names	120
4.1.2.4	Units of measure	121
4.1.2.5	Diminutives	121
4.1.2.6	Feminine formations.....	122
4.1.2.7	Compounding	123
4.1.3	Noun inflection	125
4.1.3.1	Inflection of oikoclitic masculine nouns	126
4.1.3.2	Inflection of oikoclitic feminine nouns	128
4.1.3.3	Inflection of xenoclitic masculine nouns.....	129

4.1.3.4	Inflection of xenoclitic feminine nouns.....	134
4.1.3.5	Pluralia tantum.....	136
4.1.4	Names of localities.....	141
4.2	Adjectives.....	142
4.2.1	Integration of adjectives.....	142
4.2.2	Derivation of adjectives.....	144
4.2.3	Inflection of adjectives.....	147
4.2.4	Comparison.....	149
4.3	Adverbs.....	151
4.3.1	Local adverbs.....	151
4.3.2	Temporal adverbs.....	154
4.3.3	De-verbal adverbs.....	157
4.3.4	Manner adverbs.....	158
4.3.5	Anaphoric adverbs.....	160
4.4	Numerals and quantifiers.....	161
4.4.1	Basic numerals.....	161
4.4.2	Formation of numerals.....	162
4.4.3	Inflection of numerals.....	163
4.4.4	Ordinals.....	164
4.4.5	Multiplicatives.....	165
4.4.6	Quantifiers.....	166
4.5	Pronouns.....	166
4.5.1	Personal and reflexive pronouns.....	167
4.5.2	Reciprocal pronouns.....	169
4.5.3	Demonstratives.....	170
4.5.4	Expletives.....	173
4.5.5	Interrogatives.....	175
4.5.6	Indefinites.....	176
4.6	Articles.....	179
4.7	Verbs.....	181
4.7.1	Adaptation of verbs.....	181
4.7.2	Verb formation.....	181
4.7.2.1	Imported derivational markers.....	181
4.7.2.2	Causatives.....	183
4.7.2.3	Factitives.....	184
4.7.2.4	De-nominal derivations.....	185
4.7.2.5	Iteratives.....	185
4.7.2.6	Middle verbs.....	187
4.7.2.7	Analytic constructions.....	188
4.7.2.8	Compounds and collocations.....	189
4.7.3	Verb inflection.....	191
4.7.3.1	Present stem.....	193
4.7.3.2	Present tense.....	194
4.7.3.3	Future.....	196
4.7.3.4	Imperfect.....	198
4.7.3.5	Imperative.....	200
4.7.3.6	Perfective stem.....	201

4.7.3.7	Preterite.....	204
4.7.3.8	Conditional irrealis	205
4.7.3.9	Uninflected forms.....	207
4.7.4	Copula inflection.....	210
4.7.5	Non-finite forms	215
4.7.5.1	Participles	216
4.7.5.2	Infinitive	216
4.7.6	Verbal particles	217
5	Syntax	227
5.1	Case marking.....	227
5.1.1	Basic syntactic structures.....	227
5.1.2	Possessive constructions	230
5.1.2.1	Predicative possession	231
5.1.2.2	Adnominal possession	232
5.1.2.3	External Possession	233
5.1.3	Comparative constructions	234
5.1.4	Local adverbials.....	235
5.1.5	Temporal adverbials	239
5.1.6	Other adverbials.....	242
5.2	Particles and interjections	250
5.3	Complex syntactic structures	255
5.3.1	Modal constructions.....	255
5.3.2	Complement clauses	256
5.3.3	Relative clauses.....	258
5.3.4	Adverbial clauses.....	260
5.3.4.1	Temporal clauses	260
5.3.4.2	Conditional clauses.....	262
5.3.4.3	Other adverbial clauses.....	264
5.3.5	Coordination	265
5.4	Word order	270
5.4.1	Noun phrase	270
5.4.2	Adpositional phrase	270
5.4.3	Verb phrase	271
5.4.4	Subordinated and interrogative clauses	275
6	Lexicon	277
6.1	Layers of lexical borrowings.....	277
6.2	Semantic domains	281
7	Summary.....	287
	References.....	290
	Appendix: Romani – English vocabulary	298

List of Tables

Table 1 Feminine and masculine forms of ethnonyms	28
Table 2 Documented varieties of Vend Romani.....	36
Table 3 Orthography used for Vend Romani.....	52
Table 4 Comparison of some selected inflective forms of the copula and verb.....	60
Table 5 Consonant phonemes	64
Table 6 Grapheme system.....	65
Table 7 Pre-Hungarian initial <i>CC</i> -clusters in lexical morphemes	70
Table 8 Pre-Hungarian medial <i>CC</i> -clusters in lexical morphemes.....	72
Table 9 Pre-Hungarian medial <i>CCC</i> -clusters in lexical morphemes	73
Table 10 Pre-Hungarian final <i>CC</i> -clusters in lexical morphemes	75
Table 11 Initial sound change <i>j > d'/dž</i>	84
Table 12 Vowel inventory	89
Table 13 Grapheme system.....	91
Table 14 Vowel phonemes in medial, initial and final position	92
Table 15 Long vowel in stressed and unstressed syllables.....	97
Table 16 Pattern of length distribution	99
Table 17 Genitive forms of personal pronouns	100
Table 18 Innovative vowel lengthening in the monosyllabic personal pronouns.....	102
Table 19 Development of vowel length in personal pronouns	102
Table 20 Inflectional markers of imperfect and irrealis	104
Table 21 Development of vowel length in the irrealis forms	105
Table 22 Adaptation of <i>C</i> -final nouns as masculines	113
Table 23 Integration of Hungarian <i>V</i> -final nouns as masculines.....	114
Table 24 Integration of Hungarian <i>V</i> -final nouns as feminines.....	116
Table 25 Inflection of oikoclitic masculine nouns.....	127
Table 26 Inflection of oikoclitic feminine nouns.....	128
Table 27 Inflection of xenoclitic masculine nouns I.....	129
Table 28 Inflection of xenoclitic masculine nouns II.	130
Table 29 Development of the inflectional class of recently borrowed Hungarian nouns.....	132
Table 30 Inflection of Hungarian-borrowed <i>C</i> -final nouns integrated as masculines	133
Table 31 Inflection of xenoclitic feminine nouns	134
Table 32 Inflection of the Hungarian Vend Romani <i>o</i> -masculines	137
Table 33 Inflection of the Burgenland Romani <i>o</i> -masculines	138
Table 34 Inflection of borrowed <i>C</i> -final nouns integrated as masculines	140
Table 35 Inflection of borrowed <i>C</i> -final nouns integrated as feminines	141
Table 36 Adaptation of Hungarian adjectives	143
Table 37 Derivational markers of adjectives	145
Table 38 Inflectional classes of adjectives	147
Table 39 Inflection of adjectives of the Class 1.....	148
Table 40 Inflection of the irregular adjective <i>as-</i> ‘such’	149
Table 41 Comparative and superlative forms of adjectives.....	150
Table 42 Localization and orientation of KR local adverbs	151
Table 43 Marking of local adverbs	152
Table 44 Further local adverbs	153

Table 45 Days of the week	156
Table 46 Compound numerals	163
Table 47 Personal pronouns.....	167
Table 48 Genitive form of personal pronouns	168
Table 49 Demonstrative pronouns	170
Table 50 Deictics	171
Table 51 Expletives	173
Table 52 Specific, free-choice and negative indefinite pronouns and adverbs	177
Table 53 Universal and ‘other’ indefinite adverbs	177
Table 54 Universal and ‘other’ indefinite pronouns.....	178
Table 55 Universal and ‘other’ determiners	178
Table 56 Definite article	179
Table 57 Iterative markers of consonantal verbs	186
Table 58 Tense-aspect-mood marking of lexical verbs	192
Table 59 The functions of indicative tense-aspect-mood values;.....	192
Table 60 Stem allomorphy of contracted verbs	194
Table 61 Present tense inflection of <i>V</i> -verbs, <i>C</i> -verbs and <i>MID</i> -verbs	194
Table 62 Present tense inflection of contracted verbs	195
Table 63 Future tense of <i>C</i> -verbs, <i>V</i> -verbs and <i>MID</i> -verbs.....	196
Table 64 Future tense of contracted verbs	197
Table 65 Inflectional markers in imperfect.....	198
Table 66 Inflectional markers of contracted verbs in imperfect	199
Table 67 Second-person imperative forms	200
Table 68 Second-person imperative forms of contracted verbs	201
Table 69 Perfective markers	202
Table 70 Preterite forms	204
Table 71 Conditional irrealis marker	206
Table 72 Tense-aspect-mood marking of the copula.....	210
Table 73 Present and preterite Copula forms.....	211
Table 74 Copula negation in the present and the preterite	212
Table 75 Present subjunctive, future and (conditional) imperfect forms of the copula.....	213
Table 76 Inflection of the copula.....	213
Table 77 Present copula form	214
Table 78 Past copula form	215
Table 79 Participial markers	216
Table 80 German-borrowed verbal particles	219
Table 81 Hungarian-borrowed verbal particles	221
Table 82 Calqued verbal particles	223
Table 83 Verbal particles in varieties of Vënd Romani.....	225
Table 84 Basic syntactic structures.....	228
Table 85 Case marking in possessive constructions	230
Table 86 Marking of local adverbials	236
Table 87 Marking of names of localities	237
Table 88 Marking of the non-simultaneous temporal relations	240
Table 89 Inflectional marking of other adverbials.....	243
Table 90 Adpositional marking of other adverbials	245
Table 91 Synthetic case marking: Form to function.....	249

Table 92 Complement clauses	257
Table 93 Temporal adverbial clauses	260
Table 94 Coordinating conjunctions.....	266
Table 95 Reconstructed terms denoting human beings	283
Table 96 Terms denoting human beings.....	283

List of Figures

Figure 1 Romani groups according to the Vend Roma	21
Figure 2 Romani ethnonyms used in Zala and Somogy by the Vend Roma	26
Figure 3 Hungarian ethnonyms used in Zala and Somogy by the Vend Roma.....	27
Figure 4 Ethnonyms used in northwestern Hungary by the Vend Roma	27
Figure 5 Dialect classification of Vend Romani.....	30
Figure 6 Amount of data by counties.....	35
Figure 7 Number of consultants and sentence items by gender	38
Figure 8 Number of consultants by age and gender	38
Figure 9 Detail from the manuscript of Szmodis (1827).....	58
Figure 10 Token frequency of geminates	76
Figure 11 Proportion of in/animated nouns in relation to the oblique suffixes <i>-es/-os-</i>	138
Figure 12 Inherited lexicon and layers of lexical borrowings	278
Figure 13 Inherited and borrowed lexicon with regard to some selected semantic domains .	281

List of Maps

Map 1 Southern, Western and Central Transdanubia	29
Map 2 Counties in western Hungary	29
Map 3 Central Romani.....	31
Map 4 Data points.....	37

Abbreviations

0	zero marked	LQCR	elicited language data
1	first person	M 1, 2,...	male consultant 1, 2,...
2	second person	M	masculine
3	third person	MID-verbs	middle verbs
ABL	ablative	^{NAR}	spontaneous language data
ACC	accusative	NEG	negator
AM	adaptation marker	NOM	nominative
^{B94}	Boretzky and Igla (1994: 311–338)	OBL	oblique
C	consonant	PFV	perfective marker/stem
CAUS	causative	PL	plural
COMP	complementizer	PRF	prefix
COND	conditional	PRS	present
COP	copula	PRT	preterite
C-verbs	consonantal verbs	PTC	participle
DAT	dative	Q1, 2,...	quotation 1, 2,...
DEF	definite article	REFL	reflexive pronoun
dial.	dialect	^{RM}	Rézműves (2006)
DIM	diminutive	S	Slavic
EXPL	expletive	SG	singular
F 1, 2,...	female consultant 1, 2,...	SUBJ	subjunctive
F	feminine	V	vowel
FUT	future	VOC	vocative
G	German	VP	verbal particle
GEN	genitive	V-verbs	vocalic verbs
H	Hungarian	ZB	Zuzana Bodnárová
IMP	imperative	[]	Phonetic transcription
IMPF	imperfect	//	Phoneme
INF	infinitive	< >	Grapheme
INS	instrumental	*	Reconstructed or unattested form
IPA	International Phonetic Alphabet	?	Unattested form
IRR	irrealis	~	Alternation of forms
ITER	iterative		
KR	Kisbajom Romani		
LOC	locative		

1 Introduction

1.1 Structure of the thesis

The thesis is structured as follows: An introduction to the ethnonyms used by the Vend Romani speakers, the dialect classification of Vend Romani and the sources of data and methods are presented in the first chapter. The same chapter discusses the reasons for selecting the Romani variety of Kisbajom as a sample representative of Vend Romani, in the locality of which a more intensive research has been conducted. The second chapter deals with the factors influencing the sociolinguistic vitality of the dialect, using the UNESCO (2003) framework elaborated by an expert group. The following three chapters are devoted to the grammatical description: First, the vowel and consonant inventories, the stress pattern and some morpho-phonological processes typical for Vend Romani are introduced. Second, the strategy of loanword incorporation as well as the derivational and inflectional morphology is examined. The third part of the grammatical description deals with the word order and the basic and complex syntactic constructions. The grammatical description is based on the variety of Kisbajom Romani, which is being compared to other Vend Romani varieties at the end of most sections. In the sixth chapter, the basic lexicon of Vend Romani is analysed with a focus on the post-Greek borrowings. The thesis also includes the basic vocabulary of Vend Romani translated to English, which can be found in the Appendix.

1.2 Terminology

The terminology used in this thesis is in large part drawn from Elšík *et al.* (1999), Matras (2002) and Elšík and Matras (2006).

In the thesis I use the newly established terms ‘oikoclitic’ and ‘xenoclitic’ (Elšík & Matras 2006: 324), in place of the earlier terms ‘thematic’ and ‘athematic’ noun classes (e.g. Elšík *et al.* 1999; Matras 2002), respectively. The oikoclitic class in general consists of native (Indo-Aryan) nouns, pre-Greek borrowings and some early Greek borrowings. These borrowings were integrated into the inherited (Indo-Aryan) inflectional classes. The nouns of the xenoclitic inflectional classes, on the other hand, are marked mainly by Greek-origin inflectional suffixes, irrespective of the identity of the source language.

I use the notion ‘inherited’ with reference to the Early Romani features (Matras 2002: 19–20), as well as to features which were inherited into Early Romani, i.e. borrowed from Greek or earlier contact languages. In other words, the pre-Slavic vocabulary and grammar is referred to as ‘inherited’ here.

I have adapted the terms ‘imported’ and ‘extracted affixes’ from Elšík (2007). The former deals with affixes identified within loanwords, while the latter represents those imported affixes that extend to inherited lexemes as well.

The superscript ^{B94} is used to mark the etymologies and reconstructed forms given in the worldlist of Boretzky and Igla (1994: 311–338). The glossed examples are also marked by superscripts, indicating the source of the presented data (see 1.7): ^{NAR} spontaneous narrative, ^{LQCR} elicited data, and ^{RM} Rézműves (2006). Finally, I use the term ‘Rom’ in singular and ‘Roma’ in plural as a noun, and ‘Romani’ as an adjective as well as the name of the language.

1.3 Ethnonyms with reference to the own group¹

A significant finding of the research was that the Roma speaking Vend Romani in Hungary are generally not aware of the designation ‘Vend Roma’ which is used in the Romani literature (e.g. Vekerdi 1984). The speakers call themselves by the professionym *köszörüs* ‘Grinder’² or less commonly *drótos* ‘Tinker’ in Hungarian, and *rom* (PL *róma*) ‘Rom’ in Romani. The consultant of Rézműves (2000: 24–25) used *šlajferitiko rom* or *šlajferi* (G *Schleifer*) ‘Grinder’ when specifying his group belonging. My consultants never referred to the group by this German-origin term, though they were acquainted with it. Once I heard a consultant using the corresponding Hungarian-borrowed *keseřiši* (PL *keseřišta*) ‘Grinder’: *amen eredeti keseřišta sam* ‘we are real Grinders’. On the other hand, Rézműves (2000: 24–25) considers the Grinders to be a subgroup of Vend Roma: ‘Have you ever heard about the Grinders? The Roma that work as grinders or pot-repairmen have this name among the Vend Roma’. This differs from my observation that the members of this group call themselves Grinders even if the traditional occupation of the family was not (only) grinding (Q1).

¹ The following two sections (1.3, 1.4) summarize the author’s earlier papers (2013a) and (2013b).

² Grinding is claimed to be the most common traditional occupation of the group.

(Q1)³

ZB: And what about the *bucsus* Roma?⁴

F 1: The *bucsus* Roma are also Beggars⁵, but they had shooting galleries and they did more of this entertainment stuff.

ZB: But your grandparents did the same (...)

F 2: Yeah, but they were **Roma, Grinders**.

ZB: So they were not Beggars?

F 2: They were not, because they were **Roma**.

F 1: It is only that they also had shooting galleries, but the others did more of those things, you know?

ZB: Who?

F 1: The Beggars, of course. When they stopped making bricks, as it became outdated, they took up the entertaining business.

Only some speakers were familiar with the term ‘Vend’ or ‘Vendel’, considering it to be either an exonym (Q2, 3) or a subdivision within the group (Q4).

(Q2)

F: *O vend sármazášu róma amen sam végülíš; o vend sármazášu róma amen sam. Amen óthar sármozind’am. Upr’ amende ragastinde odá, hod’ kösörüšta, vaš odá mer amare őšök, amare nad’sülök taj még mange te mró dad kösörüši síne. De čak lengeri sakma sin kösörüšök.*

M: *Hivatalošan amaro ánav ‘vend cigány’.*

ZB: *De káj hi odá písím?*

M: *O dél žánel te le ňilvántartínel...*

F: Eventually we are the Roma of Vend origin; we’re the Roma of Vend origin. We originate from this group. They ‘stuck’ to us the name Grinder, because our ancestors, our grandparents, and even my father was a Grinder. But it was just their profession.

M: Our name is officially ‘Vend Roma’.

ZB: But where is it written?

M: God knows if it is recorded somewhere...

³ Henceforth the consultants’ statements in Hungarian are translated to English by the author. The statements in Romani are quoted, and also translated to English. All quotations are marked by Q and numbered consecutively. The abbreviations used in quotations are M for a male consultant, F for a female consultant, and ZB for the author (Zuzana Bodnárová).

⁴ The Romani group called *bucsus* in Hungarian, or *bučuši* in Romani was traditionally providing entertainment, especially on saint’s feast days.

⁵ The Romani group called Beggars is characterized mainly by the lack of Romani language competence (see 1.4).

F: *Ha phúčen amendar, hod' saj sármozáši sal, hát romani žuvli sum, vad' romani (sic!) mánuš sum, vad' romani čhaj sum, vad' faťu sum, vad' romani čalád sam. Ud'hod' na phénas amen odá, hod' vend sam, vad' kösörüšťa sam, de adá kösörüš megjedzész áčhino čak afka upr' amende.*

F: When somebody asks us about our origin, I'm a Romani woman, Romani person, or Romani girl, or Romani boy, or we are a Romani family. So we don't say that we're Vend or Grinders, but this Grinder nickname somehow 'stuck' to us.

(Q3)

F: Mink vendek vagyunk, vend cigányok...

F: We are the Vends, the Vend Roma...

ZB: *De ko phénel upre tumende adá, kaj 'vendek'?*

ZB: But who calls you by this name?

M: *Okola róma.*

M: The other Roma.

ZB: *O kopanášťa?*

ZB: The Boyash?

M: *O kopanášťa, t' o lahój.*

M: The Boyash and the Vlax Roma.

ZB: *Mer tumen so phénen?*

ZB: And how do you call yourself?

F: *Drótošťa, keserišťa.*

F: Tinkers, Grinders.

(Q4)

ZB: Have you heard about the term 'Vend', 'Vend Roma'?

F: Well, we are those. (...) Long ago it was divided. So there were tribes within the Grinders. But I don't know how it was. Because, for example, there were these *tóckó* Roma. So the *tóckó*, *vendel*, or how it is called, *vendel*, or how? And the *zsuklás*, *patavás* Roma, I do not even know what that is, but it is also a kind of tribe among the Grinders. So *patavás*, *vendel*, *tóckó*, *zsuklás*. God knows how many of these exist!

ZB: So they used to marry just between themselves?

F: Not that much... What to say, the one was the same Grinder as the other, but there was a difference between those tribes. Well, they were not happy when a *zsuklás* married a *tóckó*, or the other way around, but it was not that bad as it would have been with a Boyash... It was not possible in the past for a Grinder to marry a Boyash. Now nobody cares.

Others reported to have heard it from the elder members of the group (Q5). It is therefore possible that it has come out of use just recently.

(Q5)

ZB: Who are the Vends? Because I was told that the Grinders are the Vends (...)

F: The older ones are, you know, my grandparents' parents.

ZB: But who did they call Vend? Did they call themselves Vend?

F: Do you know why they were Vendels (!), I can tell you even this. Because in the past the Roma tinkers were called Vendel. Have you heard that they made pots (...)? They used to call them Vendel Roma, or Tinker Roma, they called them many ways.

It appears that most of the group members primarily use the endonym *róma* 'Roma' in Romani, whilst using the professionym Grinders, or less commonly Tinkers, when they define their own group against other Romani groups. The ethnonym Vend or Vendel is not well-known by the group members, though it could have been more dominant in the past.

The Hungarian term *vend cigány* 'Vend Roma' was introduced into the Romani literature by the Hungarian linguist Vekerdi (1984, 1985, 2000). In addition, Vekerdi mentions the Vlach Romani name *vendicko rom* (Vekerdi 1984: 65), while Rézmvés (2000: 24–25) a similar designation *vendetiko rom*. In Hungary the nomenclature 'Vend' is known as the ethnonym of ethnic Slovenes living in the region of *Vendvidék* 'the land of Vends'. This region is situated in southwestern Hungary, near the borders of Hungary with Slovenia and Austria. According to some historical records (Kozár M. 1999), there was a migration of Slovene families of Prekmurje (today, part of Slovenia) and Vendvidék to the Zala and Somogy counties in the 17th and 18th centuries. This Slovene ethnic group has been almost entirely assimilated in language to the Hungarian-speaking majority by now. Nevertheless, it is questionable whether the term 'Vend' indicates that these Roma used to live in Vendvidék, or it has been transmitted to the Roma from the surrounding population of ethnic Slovenes only in Somogy.

I encountered a different use of ethnonyms in the Vas and Veszprém counties. In the former, the term 'Grinders' is not used to designate group identity, as it is considered there to be a derogatory word with the approximate meaning 'vagabond'. The Vend Romani speakers of Vas call themselves *muzsikus cigány* 'Musician Rom' or *magyar cigány* 'Hungarian Rom' in Hungarian, and simply *rom* 'Rom' in Romani. On the other hand, the Vend Romani speakers of Veszprém call themselves *Sinti*⁶; while they are called by some Sinti (i.e.

⁶ In Hungary, the group traditionally called Sinti have spoken the North-western dialect of Romani at least until recently (Mészáros 1980).

Northwestern) Romani speakers *hinsznári*⁷. These Vend Romani speakers consider themselves belonging to the same ethnic group as the Northwestern Romani speakers. Moreover, they perceive the two Romani dialects, Central and Northwestern Romani, to be identical with the only difference that the latter contains more German loanwords.

Rézműves (2000: 24–25) mentions that the Vend Roma are divided into several smaller kin groups, but she does not specify these groups. My consultants confirmed the existence of such subdivision within the group, referring to it as *fajta* ‘kin’, *banda* ‘group’, *faj* ‘race’ or *nemzet* ‘nation’ in Hungarian (Q6).

(Q6)

ZB: What kind of Roma live here?

F: *Patavás'ta, žukláš'ta*, there are many kin groups, the *tóckó* group, there are many groups, *tócke* and so on.

ZB: And the Vendels? Where do they live?

F: We are those (...) *Amen sam odóla* {We are those}, Vendels.

F: It is more than possible that we are the *žukláš'ta*, but we were also called Vendels. So that's why I am telling you that we are also Vendels. Who the hell knows, how it is. I just heard in the past that [we are] such Vendels.

In Romani, the consultants used the Hungarian-borrowed terms *banda* ‘group’ and *fajta* ‘kin’ (Q7), less commonly *nemzečég* ‘kin’ (Q7).

(Q7)

ZB: *Žukláš'ta, prahoš'ta, so h' odá?*

M 1: *Bareder nemzečég. Sanaséje hi. Jék fajta aso hi, jék fajta aso hi.*

ZB: *Žukláš'ta, prahoš'ta*, what is that?

M 1: A bigger kin. You have it everywhere. One kin is this, the other one is that.

⁷ Mészáros (1980: 43–44) claim the term *hinsznári* is used to designate the Vend Romani speakers of Ozora (Tolna) and Városlód (Veszprém). According to my consultants, *hinsznári* is a name given to them by the Sinti Roma of Szentkirályszabadja (Veszprém) and Mohács (Baranya). *Hinsznári* is based on the ethnonym *hienc* which is a term for the Western Transdanubian Swabians, a German-speaking ethnic group in Hungary. Interestingly, Vekerdi (1984: 67) claims that Vend Romani is ‘especially close to the South-Western *Hiencnári* dialect of Romungro’. Thus, it seems that Vekerdi perceived the Vend Romani variety of Veszprém to be a transitional variety between Vend Romani and Romungro Romani (i.e. the northern varieties of South Central Romani, see 1.6).

F: *Na, na. Adá serintem o phure róma upre jékekhávreste afka upre...*

M 2: (...) *upre phende, hod' adá adá hi, oká meg oká hi.*

F: *Áchíno hod' žukláši sal, vad' boboši sal, vad' (...)*

F: No, no. I think the elder Roma started to call each other...

M 2: (...) that the one is this, and the other is that.

F: So the name stayed on you, this *žukláši*, or *boboši*, or (...)

Such kin groups or *fajta* are the *boboši* (PL *bobošťa*, < *bobo* 'bean'), *prahoši* (PL *prahošťa*, < *praho* 'dust'), *žukláši* (PL *žuklášťa*, < *žúkel* 'dog'), *pataváši* (PL *patavášťa*, < *patavo* 'foot-rag'), *feňó* (PL *feňój*) 'pine tree', *tócko* (PL *tócke*, < Hungarian *tót*⁸), *lagaló* (PL *lagalój* < Hungarian dial. *lăgălő*) 'meadow', *čeró* (PL *čerój*, unknown origin) (Rézműves 2000: 24–25). Most of these terms are derived from nouns by the Hungarian-borrowed adjectival suffix *-Vš* (see 4.2.2), while adapted as a noun by *-i* (see 4.1.1). Thus, the literal translations of these sub-groups could be for instance 'bean-like', 'dusty' or 'doggy, dog-like'. The majority of my consultants were aware of their own sub-ethnic identity (Q8), but often considering it as a name given from the outside (Q9).

(Q8)

M: *Mró dad odá kada álo, odoleske odá phennahi boboši. Mro papu meg odá prahošno žukláši sin. T' akor ud'e mró dad lija mrá da. And' adá má amen bobošno žuklášťa sam. Mer má adá keverék hi. Taj afka hi adala, hod' com sármazind'am.*

M: When my father came, they called him *boboši*. And my grandfather was mixed: *prahoši* and *žukláši*. Then my father married my mother. That's why we are mixed: *boboši* and *žukláši*. Because we are already mixed. So this is how we have intermarried.

(Q9)

F 1: What does *zsuklás* mean, or how it's called?

M: Ah, it was invented by the Roma.

F 2: They just invented it.

M: As if I called the one 'dog-headed', and the other 'mongrel'. [laughing] So better not even mention it.

F 2: Or *babosok*, and so on (...)

M: Ah... *babosok*.

⁸ The term *tót* used to be the exonym of the Slovenes, and later of the Slovaks in Hungary.

ZB: Well, which kind are you?

M: They call us *babosok*, we are the *babosok*. *Bobošťa*. What to say, it is true, we like beans... with tasty trotters [laughing].

It is obvious from the quotation above that this system has lost its importance at least in the recent years. Only one of my consultants expressed a negative attribute towards other kin group (Q10).

(Q10)

F: So there were these groups in the past, and the hate remained, or maybe the word hate is too strong... Well, we differentiated between each other. My father used to say that the *tóckó* Roma are double-faced. Once they are eating and drinking with you, next time they stab you. And it is really like that. They are showing the good face, but...

1.4 Ethnonyms with reference to others

In this chapter I will introduce the ethno-classification models encountered in the field. These models show in a schematic way how the Vend Roma define other Romani groups. In other words: who are ‘they’ as compared to ‘us’. According to the most widespread model, found in south-western Hungary, the Vend Roma perceive the existence of five other Romani groups (Figure 1).

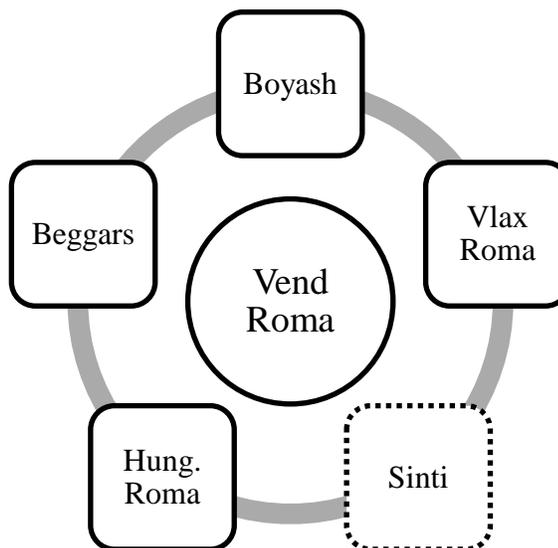


Figure 1 Romani groups according to the Vend Roma

There are a number of Romani and Hungarian designations referring to each of these groups, which I attempted to unify under the terms Boyash, Vlax Roma, Sinti, Hungarian Roma and Beggars. The ‘group’ of *bučuš* Roma ‘saint’s feast day Roma’, mentioned often by the consultants, are not considered to be an individual group entity here. As reported, this group seems to be rather a professional group that includes various Romani groups (Q1, Q11).

(Q11)

F: That’s what I am telling: there are many kinds of Roma. As for example the *bučuš* Roma, they are mixed. They are Roma but they don’t speak Romani, unlike me. There are many nations. There you have the *zsuklás* Roma, this race, that race, there are also many races. Who knows how it is, it was grouped into small tribes before. Vend Romani was the strongest for a time, now it is Vlax Romani.

1.4.1 Boyash

Most Boyash, who traditionally speak a dialect of Romanian, live in Southern Transdanubia (see e.g. Borbély 2001). The members of this group are called by the professionym *kopanáši* (< *kopana* ‘trough’) in Somogy and *koritári* (< *korito* ‘trough’) ‘throughmakers’ in Zala. The respective Hungarian designations are *beás* ‘Boyash’, *oláh* ~ *oláj* ‘Romanian’⁹ or *teknős* ~ *teknővájó* ‘trough maker’. The nicknames *čikno* ‘greasy’ and *šititno* (< H *sötét*) ‘dark’ are also widely used.

1.4.2 Vlax Roma

The members of the group named Vlax in the Romani literature are called *lahó* in Somogy, *lácko rom* in Zala, *olácko* in Veszprém, and *vlahó* in Vas. These designations originate from the Slavic *vlah* or Hungarian *oláh*, referring to ethnic Romanians. The corresponding Hungarian term is *kolompár* (< H *kolomp* ‘cowbell’) in Southern Transdanubia, and *oláh* elsewhere.

⁹ In Southern Transdanubia the Boyash are called *oláh* ~ *oláj* ‘Vlach’.

1.4.3 Hungarian Roma

According to my consultants, there are two Romani groups whose members have been monolingual in Hungarian for at least the last several generations: the Hungarian Roma and the Beggars (see 1.4.4). A Hungarian Rom is called *frakoši* (< H *frakk*) ‘tail-coated’ or, rarely, *rumungro* in Romani, and *magyar cigány* ‘Hungarian Rom’, *muzsikus cigány* ‘Musician Rom’, *úri cigány* ‘Noble Rom’ or *rumungró* (< *rom+ungro*) ‘lit. Rom-Hungarian’ in Hungarian. The latter term has the ‘funny’ counterpart *rúdugró* ‘pole-vaulter’, since it rhymes with it (Q12).

(Q12)

ZB: Are there any Hungarian Roma?

F 1: Of course!

F 2: You mean Vlax Roma, right?

F 1: No, not Vlax Roma! The Hungarian Roma, the *rúdugró* Roma. Of course! They are called *rúdugró* Roma (= pole-vaulters) because of the name *rumungró*. So we named them *rúdugró*. They used to marry between themselves, those *rúdugró* Roma, the Hungarian Roma.

The main characteristics ascribed to the Hungarian Roma are lack of Romani competence and performing music in the past.

1.4.4 Beggars

The other Hungarian monolingual group is referred to with the odd term *kóduši* ~ *kúduši* ‘beggar’, which is generally translated to Hungarian as the dialectal *kódis* (cf. standard H *koldus* ‘beggar’).¹⁰ This ethnonym may also be pejorative in some localities. Interestingly, a few consultants were not even aware of the etymology of the Hungarian dialect word *kódis* (Q13).

¹⁰ Interestingly, the Romani term *kóduš-i* ~ *kúduš-i* is derived from the Hungarian dialectal *kóduš* ~ *kúduš* ‘beggar’, while the respective Hungarian term *kódis* represents the Hungarian dialectal form used in the Austro-Hungarian border region (e.g. H dial. *kuodis* ‘beggar; very poor person’ in Oberwart; Imre 1973: 101), i.e. in the earlier contact language of Vend Romani.

(Q13)

F: The Beggars (= *kódis*) is a race. They are similar to the *kolompár*, you know, the Beggars. The word 'beggar' (= *koldus*) means 'beggar'. It is different from the Beggars (= *kódis*), which is a race, you know. Not like Vlax Roma, only similar to the Vlax Roma, but not exactly like them.

ZB: Do they call themselves Beggars (= *kódis*), or it is only you calling them by this name? Do they say, we are Beggars (= *kódis*)...?

F: Maybe. I don't know because I have never heard it from them, but I can say for sure that we call them like this, Beggars (= *kódis*).

The Beggars are characterised by the lack of Romani language competence (Q14–16) and the traditional profession of brick-making (Q14, 15) which is reported to be supplemented by begging and occasional stealing (Q16).

(Q14)

N: There you have such Beggars.

ZB: Beggars?

N: Yeah, Beggars, they are like the Hungarian Roma, those two [groups] do not speak Romani. Those are such Roma.

ZB: So the Beggars do not speak Romani?

N: No, neither do the Hungarian Roma. You have these two types of Roma. So the Hungarian Roma, or *rumungró* Roma, do not speak at all. Neither do the Beggars, only if they mix themselves with others. I mean if they marry somebody, or I don't know how to say it, that the one is from this group, and the other is from that group.

ZB: But why do you call them 'beggars'?

N: Listen, the Beggars, how should I explain it to you... They were such brick-makers and things like that. So they were the Beggars, they were called by this name.

(Q15)

ZB: The Hungarian Roma are called also Beggars, am I right?

F 1: No, that's yet again something different.

F 2: The one who speaks just Hungarian is a Beggar.

F 1: Wait, I am also a Beggar, but I speak both Hungarian and Romani.

F 2: True, but there are those who do not.

F 1: Yeah, there are, there are.

ZB: So you are a Beggar?!

F 1: Yeah, I am. My father was like that, such brick-maker, so he was almost like the Hungarians, or how should I explain it to you... Not Hungarian, but similar...

M: (...) He was mixed.

ZB: But only your father?

F 1: Yes, my mother was a 'normal' Grinder. So that's why I am telling you that I am mixed.

ZB: But your father did not speak Romani?

F 1: Of course he didn't!

ZB: Neither his grandfather, nor his grandmother?

F 1: Not even a bit!

ZB: I am asking because many people mentioned the Beggars, but I did not know who they were exactly.

F 1: Better not even speak with them.

(Q16)

M: *Upro kódušťa meg vaš odá phénen odá, hod' kódušťa (...) maj phukávav tuke. O bučušťa taj ed'ik mášik, ko ed'átalán nisar román na žánel te vakérel, de cigáňok hi.*

ZB: *Ha de ni o bučušťa na žánen román te vakérel?*

M: *Naaa. Averfajta róma hi sar amen (...) Amenge amaro papu odá kóduši sin, mer šoha na žanlahi román te vakérel, čak čórel pekamlahi. Odá kóduši sin.*

M: I tell you why the Beggars are called 'beggars'. So for example the *bučuš* Roma, and the other groups which don't speak Romani at all, but they are still Roma.

ZB: So neither the *bučuš* Roma can speak Romani?

M: No. They are a different kind of Roma. (...) Our grandfather was a Beggar, because he never spoke Romani, he only liked stealing... He was a Beggar.

The number of Hungarian Roma and Beggars is relatively small in Southern Transdanubia. They mainly live in urban areas, and only exceptionally are married to Vend Roma.

1.4.5 Sinti

There are only very few Sinti Roma living in the same area as the Vend Roma, rather individuals than families or groups. That is why this Romani group is marked by a broken line in Figure 1. Some of my consultants reported to have met Sinti Roma in the saint's feast days, as some entertainments were provided by them there. It turned out that they are known under the same professionym 'Grinder', a fact which sometimes complicated my research.

Moreover, some Vend Roma perceived the Sinti as part of their own ethnic group (see 1.3). The term *ninčko rom* (PL *ninčke róma*), and the respective Hungarian *német cigány* ‘German Rom’, also appeared in reference to the Sinti.

1.4.6 Romani and Hungarian ethnonyms

Figure 2 summarizes the most common Romani ethnonyms used by the Vend Roma of Zala and Somogy, while the corresponding Hungarian ethnonyms are featured in Figure 3. They are listed in plural in both figures. The feminine and masculine singular counterparts of ethnonyms found in Figure 2 are given in Table 1.

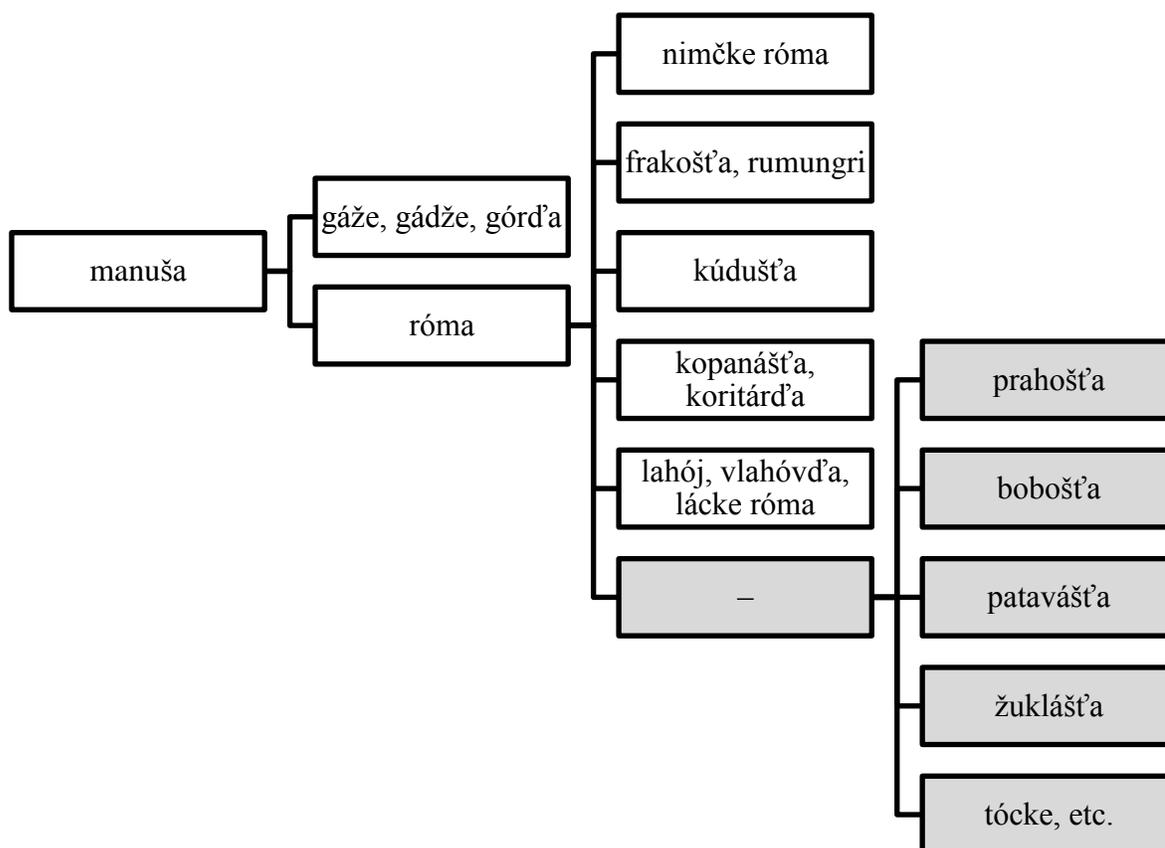


Figure 2 Romani ethnonyms used in Zala and Somogy by the Vend Roma

As modelled in Figure 2, the opposition of Roma and non-Roma is expressed by the terms *róma* ‘Roma’ and *gáže/gádže* or *górd’a* ‘non-Roma’, which are subsumed by the *manuša* ‘people’. The meaning of *gáže* is sometimes narrowed down to ‘Hungarians’, especially in the young generation.

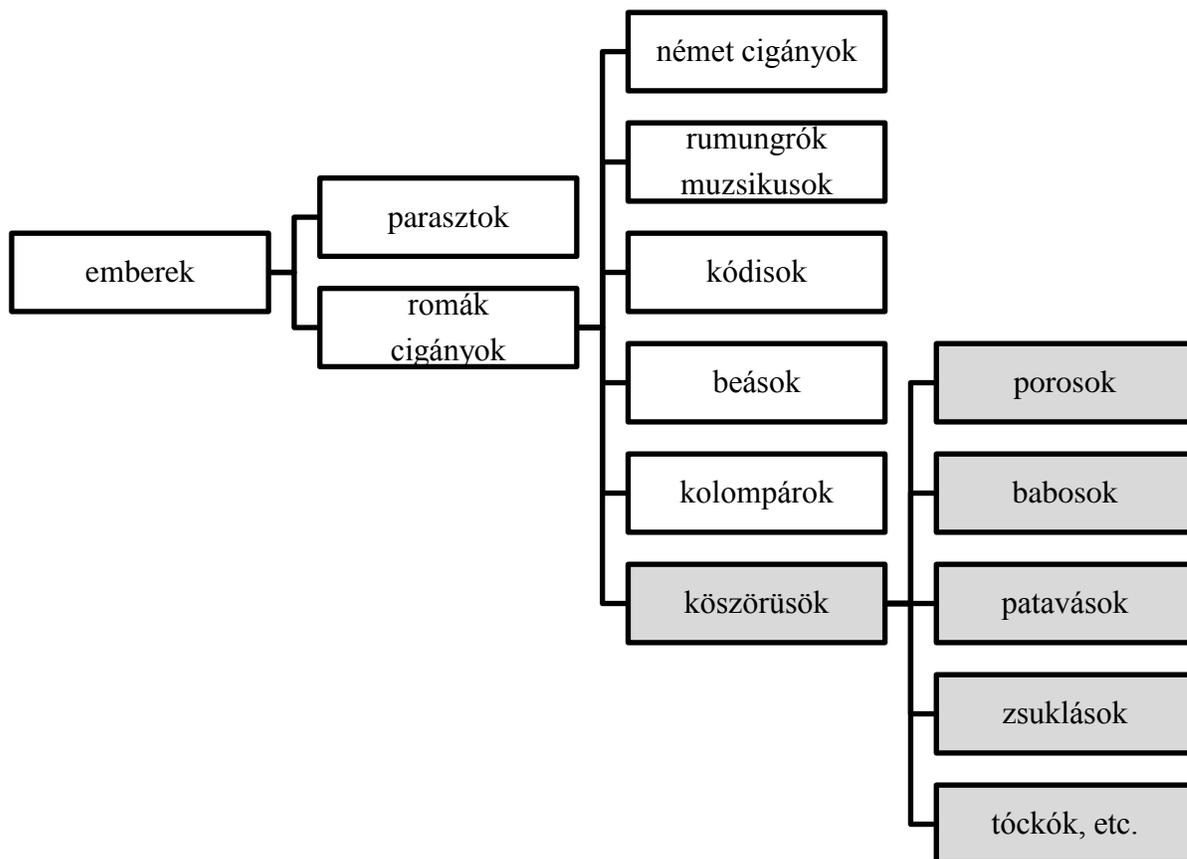


Figure 3 Hungarian ethnonyms used in Zala and Somogy by the Vend Roma

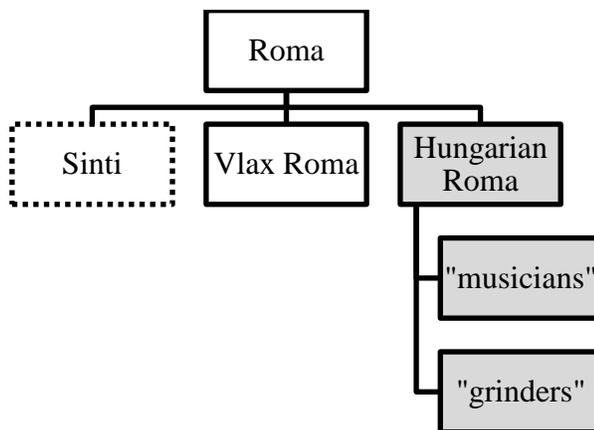


Figure 4 Ethnonyms used in northwestern Hungary by the Vend Roma

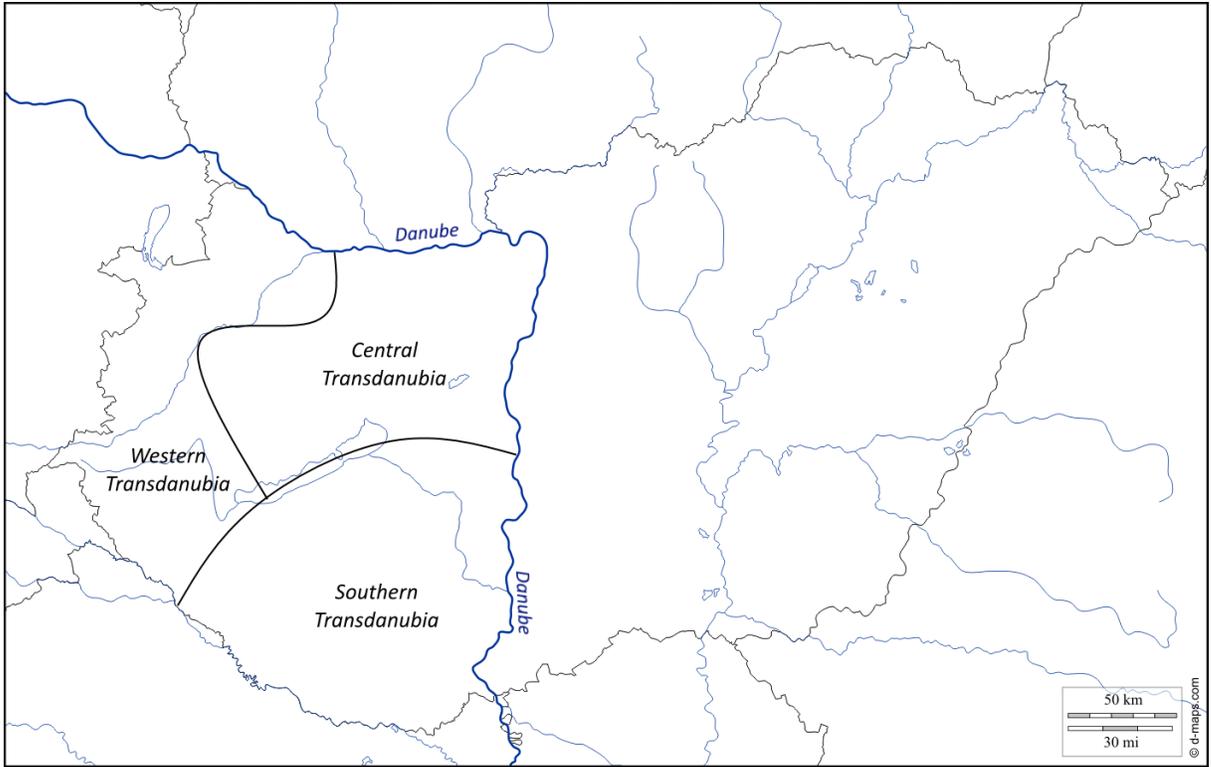
	M	F
Vlax Roma	<i>lahó</i>	<i>lahófkíña</i>
	<i>vlahó</i>	<i>vlahkiña</i>
	<i>lácko rom</i>	<i>lácki romni</i>
Boyash	<i>kopanáši</i>	<i>kopanáškiña</i>
	<i>koritári</i>	<i>koritárkiña</i>
	<i>(čikno, šititno)</i>	<i>(čikni, šititni)</i>
Beggars	<i>kúduši</i>	<i>kúduškiña</i>
Hungarian Roma	<i>frakoši</i>	<i>rumungica</i>
Non-Roma	<i>gážo/gádžo</i>	<i>gáži/gádži</i>
	<i>góri</i>	<i>górkiña</i>

Table 1 Feminine and masculine forms of ethnonyms

I encountered a slightly different classification model in the Győr-Moson-Sopron county, shown in Figure 4. According to this model, the speakers of Vend Romani refer to their own group as Hungarian Roma. The Roma who shifted to Hungarian centuries ago are considered to belong to the same group as well. Nonetheless, there is quite a widespread opinion at least in Transdanubia that the Hungarian or Musician Roma can only be Hungarian monolinguals, a conception which the Vend Romani speakers of Győr-Moson-Sopron would most directly oppose. Here the notions ‘musician Roma’ and ‘grinder Roma’ are treated merely as names of traditional professions. Finally, the Vend Romani speakers distinguish the Vlax and, in case they are aware of them, the Sinti Roma.

1.5 Geographical distribution

The Vend Romani speakers live in Western Transdanubia, and in the western part of Central and Southern Transdanubia (Map 1). The distribution of speakers is, however, quite uneven in this large area. The vast majority of Vend Roma live in the neighbouring counties of Somogy and Zala, including a nearby variety of Baranya. On the other hand, there are only a few isolated varieties in Vas, Veszprém and Győr-Moson-Sopron (Map 2, see also Map 4).



Map 1 Southern, Western and Central Transdanubia; map adapted from http://d-maps.com/carte.php?num_car=3563&lang=en (accessed February 2, 2015)



Map 2 Counties in western Hungary; map adapted from http://d-maps.com/carte.php?num_car=3569&lang=en (accessed February 2, 2015)

Most speakers of the relatively densely populated Somogy are not aware of the speakers in Veszprém, Vas and Győr-Moson-Sopron, often not even of the near-by Zala speakers. Based on my field research, I estimate the total number of localities with Vend Romani speakers at around 75.

1.6 Dialect classification

Figure 5 features the position of the Vend Romani dialect in Romani, according to the most established models of dialect classification (e.g. Boretzky 1999; Matras 2002).

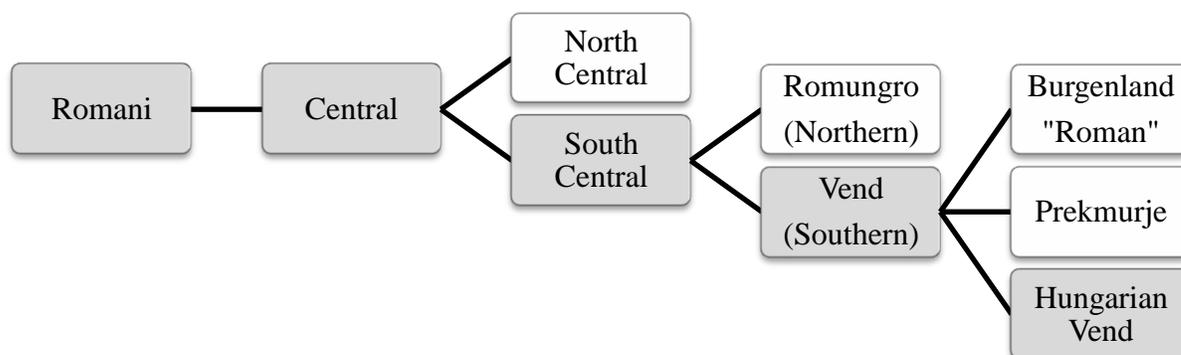
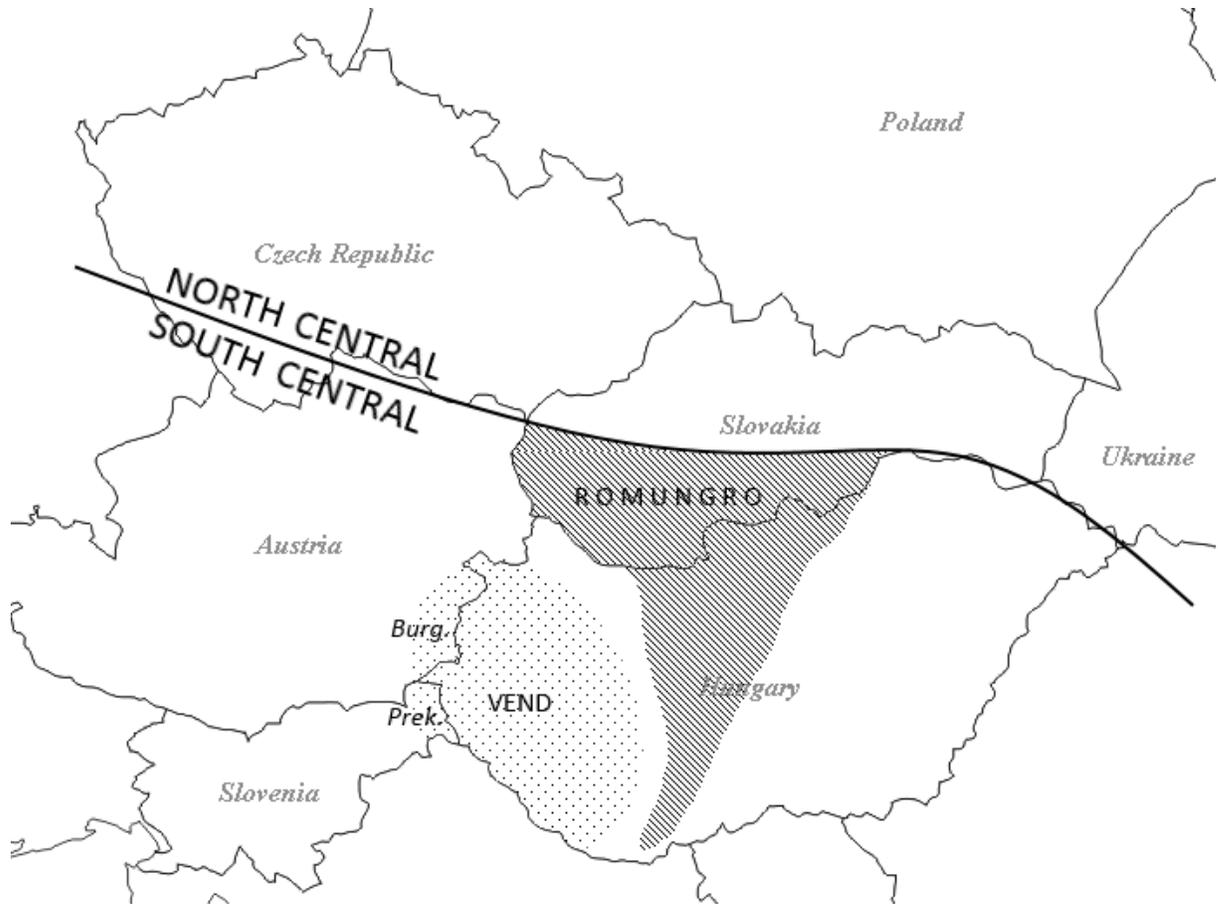


Figure 5 Dialect classification of Vend Romani

The varieties of Vend Romani spoken in Hungary belong to the Vend subgroup of the South Central dialect group. The most closely related varieties of Vend Romani are the Burgenland Romani varieties – also called *Roman* – spoken in eastern Austria (cf. Halwachs 2002), and the varieties spoken in the region of Prekmurje in northern Slovenia (cf. Štrukelj 1980). Other closely related varieties are found in southern Slovakia and northern Hungary, which are termed by the exonym *Romungro* (e.g. Elšík *et al.* 1999: 279). In the thesis I will refer to them as the northern varieties of South Central Romani, as termed in Elšík *et al.* (1999). The less closely related varieties, i.e. varieties of the North Central dialect group, are found in Slovakia and to a lesser extent in Ukraine, Poland and the Czech Republic (Map 3).



Map 3 Central Romani; map adapted from <http://www.worldatlas.com/> (accessed February 2, 2015)

1.7 Source of data and methodology

My research was comprised of four stages with the aim to collect and analyse sociolinguistic and language data on Vend Romani.

1.7.1 First stage of research

The first stage focused on the compilation of a list of localities with possible speakers. I drew on information gained from the following sources, ranked by importance for the research:

- Earlier written sources
- Previous field research
- Researchers dealing with Roma in the region
- Population census data

- Online sources, including forums, social networks, blogs, etc.

1.7.1.1 Earlier written sources

Among the earlier sources mentioning localities with Vend Romani speakers are Glaeser *et al.* (1999), Rézműves (2006), Vekerdi (1984), and Bari (1999) (For more details, see 2.9.2). The storyteller of Glaeser *et al.* originated from Rábahídvég (Vas), while Rézműves recorded several Vend Romani tales in Kisbajom (Somogy). Vekerdi lists the Somogy localities of Nikla, Lengyeltóti, Mesztegyő, Öreglak, Somogyszentpál, Bize-Kelevíz, Újvárfalva and Táská, and some localities with alleged speakers as Devecser, Ajka and Pápasalamon in Veszprém, or Németújfalu in Baranya. Bari interviewed a Vend Romani speaker in Büssü (Somogy), and recorded music in Kaposhomok and Kaposkeresztúr (both Somogy), denominating these samples as ‘Sinto’.

1.7.1.2 Previous field research

During my previous field research on a related variety of Vend Romani in southern Hungary (2008–2009) I came to know about a nearby locality with Vend Romani speakers, Vásárosdombó, where I collected some linguistic data later on. Vásárosdombó is situated in the northern edge of the Baranya county, and it is the only known Vend Romani speaking locality of the county so far.

1.7.1.3 Researchers dealing with Roma in the region

Another source of information on possible localities was researchers and students of the Department of Romani Studies at the University of Pécs in south-western Hungary. I obtained useful information on speakers from the Somogy county, especially from a former student of the department, a native speaker of Vend Romani.

1.7.1.4 Population census data

The population census data of the years 1960, 1990, 2001,¹¹ containing the number of Roma and the Romani language spoken as mother tongue, did not appear to be useful in my research.

¹¹ Hungarian population censuses 1960, 1990 and 2001. http://gis.geox.hu/nkfp/terinf/logon_form.asp (accessed January 29, 2015).

The quality of the social environment, which in case of the Roma in Hungary is rather negative, seemed to influence the results of the census to a large extent. As comparing the census data of 1960 with 2001, a significant difference is found in the number of Roma in some localities. For instance, the official number of Roma in Görgeteg (Somogy) was 80 in 1960, 31 in 1990 and 104 in 2001. Similarly, there were 29 residents declaring Romani identity in 1960, 8 in 1990 and 44 in 2001 in the municipality of Nemesapáti (Zala). Such difference in number is found in many other municipalities, especially the decreased rate in 1990, which cannot be interpreted only by migration or reproductive behaviour. Furthermore, the area of research is inhabited by other Romani groups as well, which makes the results of these censuses difficult to interpret.

1.7.1.5 Online sources

The less effective method in the preliminary part of the research was the use of online sources to gather information on speakers, since I had not been acquainted with the extended use of the Hungarian endonym *kőszörüs* ‘Grinder’ at that time (see 1.3). Nonetheless, in the course of the research, an online article by the Hungarian sociologist Kanczler (2009: 111–122) provided me with crucial information. Kanczler conducted his research on the identity of Roma in Győr-Moson-Sopron, a county in Western Transdanubia. He claims that Fertőrákos, a municipality of this county, is inhabited by approximately two hundred ‘Grinder Roma’. According to Kanczler, the number of passive speakers of Romani in Fertőrákos (see Map 4) is at about 20–30, while there are supposed to be the last one or two active speakers there. This information proved to be of great importance later, as otherwise it would have been difficult to obtain any contact information on the small number of Vend Roma living in Western Transdanubia.

1.7.2 Second stage of research

The second stage dealt with the data collection in the area of research.

1.7.2.1 Field research and linguistic questionnaire

The linguistic field research in the Transdanubian region, on which this thesis is based, was carried out as a part of the Charles University’s project *Linguistic Atlas of Central Romani*

funded by the Czech Science Foundation (Project number: P406/11/0818). The Vend Romani varieties have been documented by the means of the Hungarian-language version of the *Linguistic Questionnaire for the Documentation of Central European Romani*, designed by Elšík (2008–b).¹² It is an elicitation questionnaire that has been constructed for documenting the cross-dialectal variability of Central Romani. The questionnaire consists of 1500 sentence items grouped into several semantic fields such as food and eating, animals, weather, agriculture or modern life in order to facilitate the translation for the speakers. It includes the most important grammatical structures and basic vocabulary of Romani, which are expected to be translated into the local Romani variety.

During the field research, I contacted speakers of Vend Romani in their homes where the elicitation sessions generally took place. During these sessions the speakers orally translated the Hungarian items into Romani, which took about 5–8 hours per session. A session was usually split into two or three days. I also recorded unstructured interviews concerning sociolinguistic matters such as language acquisition, domains of language use, or attitudes towards Vend Romani of the speakers and their family members.

1.7.2.2 Method of data collection

In the absence of any reliable data on the overall population of Vend Roma, I used the research method known as ‘snowball sampling’. This method is especially convenient for locating hidden populations, since the initial consultant is expected to introduce the researcher to additional consultants, likewise these new consultants are then expected to assist in recruiting yet another consultants (Morgan L. David. 2008: 815-816). The name of the method refers to the analogy with the snowball, which increases in size rolling downhill, as the researcher gradually comes into contact with more and more consultants. This method however requires a starting point, an initial consultant. In my research, this consultant was represented by a speaker from Vásárosdombó, where I had recorded some language data earlier (see 1.7.1.2). In case some consultants were not able to refer me to any new

¹² Within the project *Linguistic Atlas of Central Romani*, the German and Slovenian language-versions of the questionnaire were used to collect data in Burgenland and Prekmurje, respectively. These data are analysed in comparison to the Hungarian Vend Romani varieties at the end of several sections in the chapters of Phonology, Morphology and Syntax.

consultants, the list of possible localities identified within the preliminary stage of research was used as guidance. The speakers were for the most part contacted without previous arrangement, by taking advantage of their contact information obtained from earlier consultants. Occasionally, I was directly seeking speakers of Vend Romani, or contacting for help local Roma representatives who were officially elected in several municipalities of Hungary. As for the remote localities where I had strong doubts about the existence of speakers, I made phone calls to the Roma representatives. The elicitation of some sentences through these calls could in all cases erase my doubts.

1.7.2.3 Amount of data

I carried out three research trips to Transdanubia with a total duration of 70 days in 2011: 32 days during March and April, 24 days in August and 14 days in October. During the field research I documented 27 varieties equal to 47 idiolects with regard to representative geographical coverage (Table 2, Map 4).

The amount of collected data (Figure 6) reflects the disproportion in the geographical distribution of speakers, according to which most Vend Roma live in Somogy and Zala, whereas there are only few speakers in the counties of Baranya, Vas, Veszprém and Győr-Moson-Sopron (see 1.5).

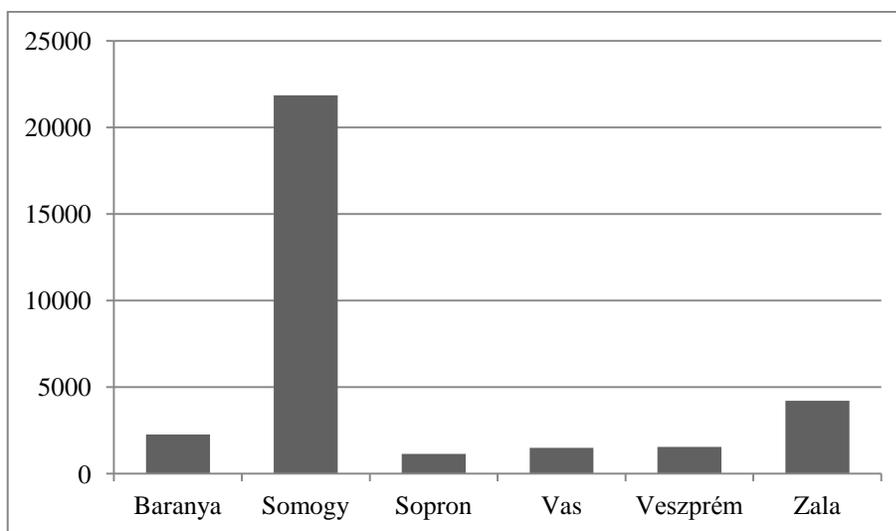


Figure 6 Amount of data (= number of sentence items) by counties

COUNTY	VARIETY	RECORDED IN	IDIOLECT
Baranya	Vásárosdombó	Vásárosdombó & Zimány	5
Somogy	Baté	Baté	1
Somogy	Büssü	Kazsok & Vásárosdombó	3
Somogy	Csokonyavisonta	Heresznye	1
Somogy	Görgeteg	Görgeteg	1
Somogy	Homokszentgyörgy	Homokszentgyörgy	1
Somogy	Kálmánca	Kálmánca	1
Somogy	Kaposkeresztúr	Kaposkeresztúr	1
Somogy	Kaposmérő	Kaposmérő	3
Somogy	Kazsok	Kazsok	1
Somogy	Kisbajom	Kisbajom	4
Somogy	Lengyeltóti	Lengyeltóti	3
Somogy	Mesztegyő	Nikla	1
Somogy	Nikla	Nikla	2
Somogy	Öreglak	Kazsok	1
Somogy	Rinyaújlak	Heresznye	1
Somogy	Tarany	Tarany	1
Somogy	Táska	Táska & Nikla	4
Somogy	Zimány	Zimány	1
Győr-Moson-Sopron	Fertőrákos	Fertőrákos	2
Győr-Moson-Sopron	Kapuvár	Kapuvár	1
Vas	Szakonyfalu	Szakonyfalu	1
Veszprém	Városlőd	Városlőd	2
Zala	Nagykanizsa	Nagykanizsa	1
Zala	Kustánszeg	Kustánszeg	1
Zala	Nemesapáti	Nemesapáti	1
Zala	Németfalu	Kustánszeg & Nagykanizsa	2

Table 2 Documented varieties of Vend Romani



Map 4 Data points [1 Vásárosdombó, 2 Kaposkeresztúr, 3 Baté, 4 Zimány, 5 Büssü, 6 Kazsok, 7 Kálmánca, 8 Csokonyavisonta, 9 Rinyaújlak, 10 Homokszentgyörgy, 11 Görgeteg, 12 Tarany, **13 Kisbajom**, 14 Kaposmérő, 15 Mesztegnyő, 16 Nikla, 17 Táska, 18 Öreglak, 19 Lengyeltóti, 20 Nagykanizsa, 21 Kustánszeg, 22 Németfalu, 23 Nemesapáti, 24 Szakonyfalu, 25 Rábahídvég (Glaeser *et al.* 1999), 26 Városlőd, 27 Kapuvár, 28 Fertőrákos]; map adapted http://d-maps.com/carte.php?num_car=3569&lang=en (accessed February 2, 2015)

1.7.2.4 Selection of consultants

The gender distribution of consultants, although it was not considered as a criterion for selection, is unexpectedly balanced (Figure 7). Moreover, the number of items translated by men and women is almost equal (Figure 7).

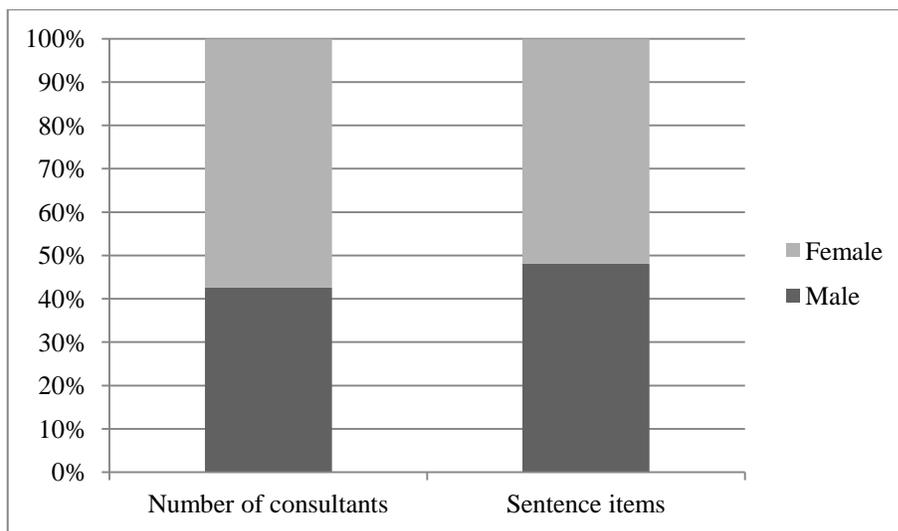


Figure 7 Number of consultants and sentence items by gender

Most of the consultants were middle-aged or older, and only rarely younger than forty years old (Figure 8). The young generation (15–30) is represented by a single consultant with limited proficiency in Romani, by whom not more than six sentence items were translated.

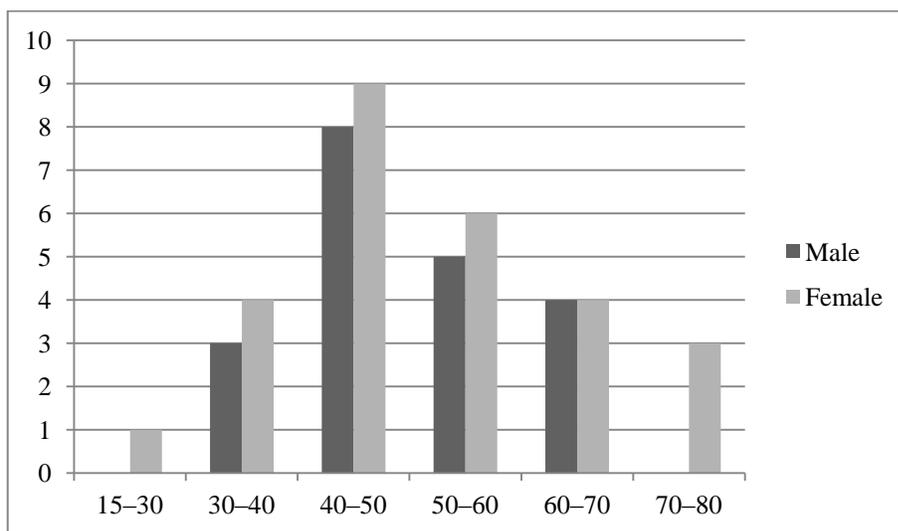


Figure 8 Number of consultants by age and gender

1.7.3 Third stage of research

The third stage aimed to deepen the research in one selected locality, on the variety of which the thesis would be based, taking into consideration the following criteria:

- Variety of relatively high vitality compared to other Vend Romani varieties
- Variety of a region densely populated by Vend Roma (i.e. situated in Somogy)
- Welcoming environment for conducting research

All these conditions were fulfilled in Kisbajom, a locality situated in the central part of Somogy (see Map 4). Kisbajom has in total approximately 450 residents, of which 79, according to the Hungarian Population Census of 2001,¹³ claimed Roma ethnic identity, and 46 considered Romani as their mother tongue. The census data, on the one hand, include the Romanian-speaking Boyash, and on the other hand, the Vend Roma, since both Romani groups co-exist in Kisbajom. Kisbajom Romani (hence KR) is the language of several dozens of Romani residents of Kisbajom born before 1984. In this village I recorded three elicitation questionnaires and some short stories. In addition, some tales were collected in Kisbajom by Rézműves, published in 2006. My attempt to acquire further natural language data by providing some speakers with an audio-recorder was refused by the speakers for various reasons.

1.7.4 Fourth stage of research

The fourth stage comprised the transcription of audio-recordings obtained during the field research, as well as the transcription of existing published sources (Vekerdi 1984, 1985; Bari 1999; Glaeser *et al.* 1999; Rézműves 2000, 2006). The Romani tales from Rézműves (2006) were transcribed from the audio CDs attached to the book, as the printed version is standardized and stylistically revised (e.g. the vowel length is not marked). The analysis of the data was carried out by means of the offline *Linguistic Database for the Documentation of Central European Romani* developed by Elšík (2008–a), and a concordance program which was programmed and personalized by Jakob Wiedner for my specific needs.

¹³ See <http://www.nepszamlalas2001.hu/>. The results published in the Hungarian Population Census from 2001 do not differentiate between varying Romani groups.

2 Sociolinguistic vitality

In order to describe systematically the sociolinguistic vitality of Vend Romani I have used the UNESCO's (2003) Language Vitality and Endangerment evaluation guideline prepared by an expert group. The evaluative factors determining the viability of language are identified in the framework as follows:

1. Intergenerational language transmission
2. Absolute number of speakers
3. Proportion of speakers within the total population
4. Trends in existing language domains
5. Response to new domains and media
6. Materials for language education and literacy
7. Governmental and institutional language attitudes and policies
8. Community members' attitudes toward their own language
9. Amount and quality of documentation

The first six factors dealing with language transmission, number of speakers and domains of language use have the highest importance as they directly verify the language's vitality and state of endangerment. On the other hand, the last three factors related to language attitudes as well as extent of documentation are less crucial in terms of language vitality. The nine factors together aim to characterise the overall sociolinguistic situation of the examined language. Most of the factors are evaluated on a scale ranging from zero to five, where zero represents extinct or highly endangered status, while five stands for safe status. According to these factors, the following degrees of endangerment may be assigned to Vend Romani:

Factor 1.	2	<i>Severely endangered</i> : The language is used mostly by the grandparental generation and upwards.
Factor 2.	–	A few hundred
Factor 3.	2	<i>Severely endangered</i> : A minority speaks the language.
Factor 4.	2	<i>Limited or formal domains</i> : The language is used in

		limited social domains.
Factor 5.	0	<i>Inactive</i> : The language is not used in any new domains.
Factor 6.	0	No orthography is available to the community.
Factor 7.	5	<i>Equal support</i> : Romani and other officially recognized minority languages are protected in Hungary.
Factor 8.	0	The speakers are indifferent with regard to language loss; all prefer to use Hungarian, the dominant language.
Factor 9.	2	<i>Fragmentary</i> : There are some grammatical sketches, word-lists, and texts useful for limited linguistic research but with inadequate coverage. Audio and video recordings exist in varying quality.

The following chapters deal in detail with each factor presented in the guideline.

2.1 Intergenerational language transmission

Probably the most crucial factor regarding language vitality is whether the language is transmitted from the older generation to the younger one or not. According to this scale, most local varieties of Vend Romani are severely endangered because the youngest speakers are of grandparental¹⁴ and older generations. It is rare to find even passive speakers under the age of 20. It follows that the children no longer learn Romani as their mother tongue, since Hungarian has become the dominant or the only language of everyday interactions most of the Vend Roma even in their homes. In addition, Vend Romani is critically endangered especially in Vas, Veszprém and Győr-Moson-Sopron. The youngest speakers of this region are generally of great-grandparental generation.

Based on the typology of speakers presented in Grinevald and Bert (2011: 49–52), my consultants were mainly fluent or semi-speakers, there was a single terminal speaker and probably some ghost speakers (see below). Note that the wide range of speakers with different language skills is not surprising in an ongoing process of language shift. It is certainly difficult

¹⁴ Since teen marriages are quite frequent among the Vend Roma, it is not rare to become grandparent in the person's late thirties.

to estimate the overall competence of speakers after a couple of days spent in the localities. Thus, the following notes on the competence of speakers are mainly based on my observation and impression. A number of my consultants were semi-speakers, as they were relatively fluent in casual conversations, but having varying levels of productive skills. Many of them, for instance, were able to hold conversation without difficulties on the one hand, and having difficulties translating more complex sentences on the other hand. I also interviewed several fluent speakers, mainly of the oldest generations. In Kapuvár, I could contact only a so-called terminal or partial speaker who had been raised by his grandparents. This speaker had only a basic knowledge of the language but mastered a number of fixed expressions. One should also count with the existence of the so-called ghost speakers who deny their knowledge of the language due to the strong negative attitude towards it. It was indeed not rare to meet Vend Roma who first denied being speakers and later started to speak Romani, especially after being addressed in Romani by me.

Mixed marriages have become prevalent just recently, which is another factor influencing language transmission. The high number of intergroup marriages could be explained by the small in number Vend Roma scattered on a relatively large area of Transdanubia. As the consultants reported, such marriages were not tolerated in former times (Q17).

(Q17)

W: Some time ago the Boyash were not allowed to 'enter' our group. It was long ago, everything was different, yeah. Neither the Hungarians could. And today... the Hungarians are also mixing with the Roma and so on.

Today the most widespread mixed marriage is Vend–Boyash. Vekerdi (1984: 65) also notes that the Vend Roma maintain friendly relations with the Boyash. It is not surprising since Boyash is the most numerous Romani group in the region. Marriages between the Vend Roma and Hungarians are also on the increase. In addition, I met some mixed couples where the partner was Vlax, less commonly Beggar or Sinti. The children are generally not brought up bilingually in mixed marriages. The parents reported to opt for Hungarian as the first and only language transmitted to the children. These children are exposed to Romani only indirectly, rarely acquiring passive competence in Romani.

2.2 Absolute number of speakers

It is difficult to estimate the accurate number of speakers in the absence of any official estimation targeting directly the Vend Roma. Relying on my field data, I agree with Vekerdi who estimates the number to be around a few hundred (1984: 65; 2000: 14).

The national census of 2001 first distinguished between the languages spoken by the Roma. The inquired person could claim *cigány / roma* ‘Gypsy / Romani language’, *romani* ‘Romani language’ or *beás* ‘Boyash’ as their mother tongue (Hungarian Population Census 2001). However, the only difference between the dialects entitled as *cigány / roma* and *romani* is that the latter is the adjectival form of the former, which is an apparent mistake of the official authorities. Moreover, the published census data shows only the sum of these three ‘languages’, which makes it impossible to distinguish the number of Romani dialect speakers from the Romanian dialect speaking Boyash. The following Hungarian Population Census (2011) offers the option *cigány (romani, beás)* ‘Gypsy (Romani, Boyash)’ with no further possibility to specify the dialect or language of the speakers. As the Vend Roma coexist with the Boyash in several localities, the census data cannot be used to estimate the number of speakers of any of these groups.

2.3 Proportion of speakers within the total population

Estimating the total number of Vend Roma is even more difficult than estimating the number of speakers. In the majority of localities I could contact at least a few passive speakers, while in some other localities, reportedly, the ‘last’ speaker had passed away recently. This indicates that the language shift towards Hungarian is a relatively recent development. The speakers were also aware of the fact that Vend Romani is spoken by relatively few speakers compared, for instance, to Boyash (Q18).

(Q18)

W: The Grinders are dying away, there are not many left. You have the Boyash everywhere, many of them... You can see them everywhere. So the Grinders are just in these few villages, nowhere else. But the Boyash! There is no village where they would not have been settled down. They procreated a lot, and we are gonna die away, that’s it!

According to the UNESCO (2003) scale, the proportion of speakers within the total number of Vend Romani population could be ranked as severely endangered because only the minority of the population speaks the language.

2.4 Trends in existing language domains

On the scale for this factor, Vend Romani ranks at grade 2 termed as limited domains, because the non-dominant language, Romani, is used only in homes where grandparents or other older family members reside. The dominant language, Hungarian, exclusively occupies the public domains, such as the media, public offices, educational and religious institutions. At the same time, Hungarian is becoming dominant also in the private domains due to the generational gap and lower proficiency of many middle-aged speakers. On the other hand, Romani may also be heard in public spaces where the elders meet or gather, such as in the streets, local markets and shops, festivals or ceremonies. However, it is mainly used in the absence of non-Romani bystanders. The consultants expressed their preference to speak Hungarian among themselves in the presence of non-Roma when using public transport, being in hospitals or in other public spaces with the dominance of non-Roma. It follows that Romani seems to be more vital in residences with Roma-dominant population, and diminishing when the Roma live scattered among the non-Roma.

The switch between Hungarian and Romani is quite common on the word, phrase and sentence levels. Especially the switch from Romani to Hungarian in order to help out with a phrase or idiom is quite widespread in my data (the Hungarian phrase is underlined):

(1)^{NAR} *Adá meg čoro, az az igaššák, na likellah' ánd' oko sabáj.*
And the deceased guy, that's the truth, did not observe the rules.

(2)^{NAR} *Aja akán töb mind valosinü na odolestar hi, hanem ék ár fai'ústar.*
And her father is most probably not this guy, but rather another guy.

(3)^{NAR} *T' adála maj pomožinna kérel le ed'két sóbu.*
And they will help to make it right away (lit. from one-two words).

Furthermore, it has been observed that the speakers often switch to Hungarian when addressing children and babies (Q19), animals (Q20, 21), and in storytelling when quoting people who do not speak Romani (Q22)¹⁵.

(Q19)

W [to the baby]: Yeah, fall asleep! [To the husband]: *Lakeri čánga cili... tél oja hólev. Cide la téle, mer šudre hi lakere čánga!* {Her legs are totally... the trousers are down. Pull them back up, because her legs are cold!} [To the baby]: Your little pussy is frozen! What do you want? What? A horse? [pointing to a plastic horse] That is crappy, that is stinky. It is from the dump, I will give you another horse.

(Q20)

W [to me]: *Há so kéren?* {What are they doing?} [To the children]: What's goin' on? My God! [To herself]: *So hi, so hi? Hát má gondulind'um, hod' valasaj baja hi, no.* {What's goin' on, what's goin' on? I have already thought that there is some problem, ah.} [To the dog]: Go out, go out, come on! Ah, my God! [To the husband]: *Čhiv cuj odá vúdar, mer o žúkel...* {Close that door, because the dog...}

(Q21)

W [to the cat]: And now why aren't you eating? [To herself]: *Hát so h' adla mačkaha, hod' na hal?* {What is wrong with this cat, why is it not eating?} [To the cat]: Eat!

(Q22)

M: *Si duj lumňa vaj trin, ko phénel* {There are two or three women and they say}: Now you, and you, and you come to eat! *Te phend'a* {If he answers} 'I accept, žav' {I go}... *Te phend'a* {If he answers} 'no', then the other should come, who would accept the invitation.

As for the teasing of infants, my data, small in number, does not allow to make any conclusions. However, it may be noted that these data seem to confirm the findings of Réger (1999), according to which the direct sexual teasing is particularly important in the Romani children's early linguistic socialization. My data would then suggest that the Romani pattern

¹⁵ In the following examples, the Hungarian parts are translated to English, while the English translation of the Romani parts follows directly the Romani utterances in curly brackets.

of baby-talk is transferred into Hungarian. Consider, for instance, the utterance ‘Your little pussy is frozen!’ in (Q19), or calling the baby ‘gypsy bitch’ in (Q23):

(Q23)

M [to the baby]: *Čumide man, čumide man, čumide man!* {Kiss me, kiss me, kiss me!} You’ll get a horse, you’ll get a horse, you’ll get a horse, you’ll get a horse, you’ll get a horse. I do not love you, you bastard, Boyash.¹⁶ I’m telling you, I do not love you. A hat? You want it on your head? I should put it on your head, there you go! Op-op-op-op-op, ej-ej-ej! Now you cannot see (because of the hat), you, gypsy bitch, you cannot see, gypsy bitch, gypsy bitch. Ah, here is the horse! [Hands over a plastic horse]: Take it, monkey! A real horse-rider! (...)

Further insults in my data are, for instance, calling the baby ‘monkey’ in (Q23) and ‘ugliness’ in (Q24):

(Q24)

W [to the baby]: Eat it!

M: She does not want.

W [to the baby]: You don’t want it?! Then spit it out, ah! Who are you? (...) Ugliness.

Apart from an insult, the teasing act may also start with mock challenging or threat (e.g. Eisenberg 1986: 183–184). The mock challenge ‘I do not love you’ is encountered in (Q23), while the threat of being beaten in (Q25).

(Q25)

W [to me]: You see how big this girl (= the daughter) is? Almost like you. She is fifteen. But she does not help me with anything; she does not want to help me, just nothing.

ZB: *Lini hi li.* {She is lazy.}

W: *Lini, lini! Močárne line kurvi!* {Lazy, lazy! Ugly lazy bitches!} [To the baby]: Isn’t it true?

[The children are playing with the baby]

W [to the children]: *Na ugrálin!* {Stop jumping!} Stop it! [To me] *Tél la čhiden fejtetőre, ár pharád’ol i men!* {They will knock her down, she’ll break her neck.}

¹⁶ The baby was born to a Boyash-Vend mixed couple. However, the term ‘Boyash’ seems to be used in a derogatory sense in this utterance.

M [to the baby]: Stay there! Otherwise there will be *dáda*¹⁷!

The limited use of Romani by the middle-aged speakers includes greetings and some fixed phrases in order to start a conversation, but often switching to Hungarian for continuation. These speakers excused it by the fact that they learnt Romani simultaneously with Hungarian, and therefore acquired the same proficiency in both languages. Thus, it can be summarized that Romani is generally used in daily life conversations between the members of the elder generation, while the younger generation is addressed almost exclusively in Hungarian.

There is an interesting discourse that emerged in the field regarding the speakers' self-reported perceptions of their own proficiency in Romani, which in many cases contradicts my observation. That is, the speakers often underestimated their own competence in Romani, which they subsequently confronted with the believed higher Romani competence of others.

It is also noteworthy that the consultants often used an 'inclusive strategy' to refer to the Romani language skills, a discourse which may be summarized as 'some non-/Roma speak better than we do'. For instance, the consultants used to overestimate the Romani proficiency of some local non-Roma who learnt Romani to some extent, as well as of the non-Romani researchers interested in the language (Q26).

(Q26) Teasing a woman (W2) because of her supposedly low competence in Romani

W1: *Tu még but site síkjós román.*

W1: You still have a lot to learn in Romani.

W2: *Hát de tu kiűi site síkjohahi manuűni?*

W2: And how much more would you need to learn?

M1 [about me]: *Aja má feder źanel román sar tu.*

M1 [about me]: She speaks Romani better than you do.

M2 [about W2]: *Ha nem tudta, hogy mi a birka, hát bákro hi odá!*

M2 [about W2]: She did not know what the word for sheep is, it is *bákro*!

W2: *Bákro, hát!*

W2: *Bákro*, of course!

M1 [to me]: *His tu źánes, his akor soske?*

M1 [to me]: But you already speak [Romani], so why (are you interested)?

W2: *Sótár kamlahi te kérel.*

W2: She wants to make a dictionary.

ZB: *Mer si but űelvjáráűi.*

ZB: Because there are many dialects.

¹⁷ The baby-talk word *dáda* here means 'beating', i.e. 'you will get a beating'.

W1: *Feder vakérel román sar amen.*

ZB: *Féder na.*

W1: *Dehodnem.*

W1: She speaks Romani better than we do.

ZB: Not better.

W1: Of course better.

The strategy to ascribe higher language skills to the non-Roma, especially to the non-Romani researchers, probably lies in the fact that the general discourse is based on the idea that language competence is generally evaluated through education. If we take into consideration that Vend Romani is not taught in schools, and it does not have a standard form, then it is not surprising that the speakers are less confident about their own proficiency, while the competence of the ones who are believed to have learnt Romani in the school, such the Romani Studies researchers, is perceived by them to be much higher.

On the other hand, the discourse of ‘other Roma speak better than we do’ surfaced in the form of stories about children who speak only Romani (Q27), about Roma communities in poverty or living in segregated localities who speak better Romani than Hungarian (Q27), or about the elders who spoke the ‘pure’ or ‘real’ Romani language, meaning with less Hungarian influence (Q28, 29).

(Q27)

M: *Odoj hi róma, ma dara! Odoj uze kaste žas konkrétan?*

ZB: *Uzi (name).*

M: *Jáj, de! Odóla žánen megen. De odoj mind sako faťú žánen, odoj šaj... odoj perfekten. De odoj sin jék faťú, maj odolendar phuč hod' savo odá, savo čak román žánel te vakérel.*

ZB: *Jaj, láche.*

M: *Mer odá čak román žanlahi. Akébor lo sin sar ák adá [pointing to a little child], taj and' iškola le bičháde taj kova síne: O jék faťú site phučlah' okolestar, hod' savo, so phénel le tanáriske, t' afka vakellahi. Mer na žanlahi te*

M: There are Roma, don't be afraid! To whom exactly are you going there?

ZB: To (name).

M: Ah! They speak well. There all the children speak, perfectly.¹⁸ There was a boy; you can ask them which one it is, who speaks only Romani.

ZB: Oh, that's good.

M: Because that one speaks only Romani. He was of his size [pointing to a little child], and they sent him to school. And there, one boy had to ask another what the teacher said, so they spoke. Because he could not speak

¹⁸ Later, I have visited the locality we are speaking about, where it turned out that no children spoke Romani. The locality is geographically segregated, and inhabited almost exclusively by Roma.

*vakérel. Hát sar afka upre te nevelínel fatún,
hod' te na kova ól!*

The daughter of M: You should also have raised me like that, so that I would not speak any Hungarian...

M: You idiot! (...)

[Hungarian]. How is it possible to raise a child like this!

The daughter of M: You should also have raised me like that, so that I would not speak any Hungarian...

M: You idiot! (...)

(Q28)

M: Neither we can speak a thousand per cent Romani. Because, let me tell you, the old ones spoke it perfectly. But we are already such descendant generations... So it is slowly becoming worn-out. Just like the jeans.

(Q29)

W: There was a basic rule among the elders, a long time ago. I mean not now, maybe among my great-grandparents. It was that the boy or the girl could marry only such person who belonged to them, who was a Grinder. Or more precisely, who belonged to his or her group. Then they mixed together, and that 'very old language' fell apart.

ZB: And they also called themselves Grinders?

W: Of course! Only that they spoke differently, in a way different from how we speak now. We are also carrying the traditions, the Romani language, but they spoke it differently. And there are some who still remember it, you know, and there are those who do not.

The strategy of ascribing Romani language competence to others may also serve for the speakers to minimize the importance of, and their contribution to, the fact that Romani is on the verge of extinction. Nevertheless, such statements could have also been influenced by my presence.

The opposite strategy often taken by the speakers was an 'exclusive strategy', which was manifested in the teasing of children or adults who do not speak (sufficient) Romani. What is interesting is that in most of these acts of teasing, the teased person inevitably loses his/her face, as s/he cannot give an appropriate reply on account of his/her low or zero competence in Romani. The fact of teasing is however evident from the intonation of the teaser. It means that from such a teasing act the teaser (i.e. the Romani speaker) is always the one who comes out as the 'winner'. Consider, for instance, the teasing of the son by his

parents (Q30), the sister who spent her childhood in an orphanage and thus has not learnt Romani (Q31), or the Boyash boy who speaks only Hungarian (Q32).

(Q30)

Father: *De vaker laha valaso!* {Well, say something (in Romani) to her!}

Son: I do not speak.

Father: *Román vaker!* {Speak Romani!}

Step-mother: *Hát de román na žánes ništ?* {But don't you know anything in Romani?}

Father: *De!* {Come on!}

ZB: And your daughter?

Son [to his sister]: Do you speak? Because I don't.

(Q31)

Teaser M: *Uštadi mindž hi aja, mer hárne hi taj kurádi hi, mer mindig kurlahi (...) Sakone čéderi upre peste muklahi.* {She is a walking pussy, because she is short and a bitch, because she was always having sex (...) She let all the 'stallions' on her.}

Teased W: She told me that this little child...

Teaser M [interrupting the W]: Now... *Akán ma ungrika vaker, román vaker kuradi kurva!* {Now do not speak Hungarian, speak Romani, fucking bitch!}

(Q32)

Teaser M: Make also one [a pancake] for this one, to this Boyash bastard. [To the boy]: *Kopanási, saláhi te čórel korháni kopanási?* {Have you been out stealing corn, Boyash?}

Boy: This is a sign language.

Teaser M: Shut up! *Šúne, kopanási, saláhi te čórel korháni?* {Listen, Boyash, have you been out to steal corn?}

The speakers also used to tease each other by challenging each other's Romani proficiency, as it has been illustrated in (Q26). The Romani ethnic identity may also be questioned by teasing, as for instance when one of my consultants addressed another with the words: 'What a Rom you are when you do not speak Romani?'

So far I have dealt with the functions and domains of Vend Romani and Hungarian. Nonetheless, some of my consultants residing in the Austro-Hungarian border region also

reported using German and/or Prekmurje Slovene¹⁹ actively. Unfortunately, I do not have data on the distribution of domains in these language settings. Regarding other Romani dialects, I often witnessed Vend Roma singing in the Vlax dialect of Romani. The Vlax Romani music and dance became popular not only among Vend Roma, but also among other Romani groups in Hungary. It is well illustrated by the fact that the song collection of Bari (1999) also includes a Vlax Romani song performed, with great probability, by a Vend Rom.

2.5 Response to new domains and media

Hungarian is the only language used in new domains such as schools, working place, broadcast media and internet, which places Vend Romani to the last grade named *inactive* in the UNESCO (2003) scale. Vend Romani is not taught at any level of the education system, and the knowledge of the dialect is not required for new working environments. It may theoretically occur in the public television's programme *Roma magazin*²⁰ dealing with the culture and everyday life of Hungarian Roma, broadcasted once a week for half an hour. Vend Romani or its speakers have not received any special attention from the radio or newspapers so far, which was also strongly perceived by the speakers (Q33).

(Q33)

F: Otherwise the language of Grinders wasn't in the newspapers, and there are no books. You have some in Vlax Romani and Boyash, and nothing else. And in Hungary there are the fewest Grinders, the Grinders are the minority. Unlike the Boyash or the Vlax Roma! In our language there wasn't any book in the past, nothing.

The communication through short text messages and emails is mostly used by the younger generation who lacks the necessary Romani language competence. The various social network sites on the internet have become very popular across generations, but a network group promoting the use of Vend Romani is absent. As I am connected on a social network with many of my consultants, I have noted that they never post messages in Romani, except

¹⁹ Prekmurje Slovene is a regional variety of Slovene spoken in the Prekmurje region of Slovenia and in the Vas county of Hungary. The Roma from the Vas county distinguish between Prekmurje Slovene and Slovene.

²⁰ Roma Magazin. <http://www.mediaklikk.hu/musor/roma-magazin/> (accessed June 16, 2014).

for the occasional use of Hungarian slang words with Romani origin such as *dévla* ‘God’, *csávó* ‘guy’, *csóró* ‘poor’, or *dicsak* ‘look’. The multilingual dictionary by Vekerdi (2000) including some hundred Vend Romani words is available online, but none of my consultants were aware of it.

2.6 Materials for language education and literacy

Vend Romani is a strictly oral language with no established orthography. There are no educational materials and programs at any level of the education system, neither institutions nor individuals promoting the standardization and language teaching of Vend Romani. Although there are only a few published texts in this Romani dialect, they were written in several different writing systems. Vekerdi (1984, 1985, 2000) and Glaeser *et al.* (1999) used an orthography common in Romani linguistics inspired by the alphabet of Slavic languages in Europe, while Rézmúves (2000, 2006) was a pioneer in Hungary for using the orthography designed for Romani in general by Marcel Courthiade. Bari (1999) transcribed the Vend Romani songs by means of the writing system based on English and Hungarian which became standardized for the Vlax dialect of Romani in Hungary.

In contrast, I use an alphabet developed to write Czech and Slovak Romani, which differs from the alphabet favoured in Romani linguistics in that the palatal dentals are marked here by caron. In addition, I indicate the long vowels with acute accent. Table 3 compares some selected graphemes from the writing systems promoted by authors which have contributed to the documentation of Vend Romani.

Rézmúves	ś	ž	č	ḍ	ṭ	ṇ, ñ
Bari	sh	zh	ch	gy	ty	ny
Vekerdi	š	ž	č	dj	tj	nj
Glaeser <i>et al.</i>						
Bodnárová	š	ž	č	d'	t'	ň

Table 3 Orthography used for Vend Romani

The revitalization of Vend Romani in Austria and Slovenia resulted in a number of educational materials which are not available to the speakers in Hungary. Burgenland Romani

is transcribed by means of German orthography (e.g. *sch* /š/, *tsch* /č/), while the writing system of Prekmurje Romani is based on Slovenian, including several dialect-specific graphemes such as the palatalized *gj* and *kj* or the diphthongs *au*, *ou*, and *ej*.

My consultants were not familiar either with the few published texts in Hungarian Vend Romani, or with the alphabet used in these texts. Interestingly, none of these alphabets is based exclusively on Hungarian which is the primary language of education of Vend Roma in Hungary. The Vend Romani speakers are thus not literate in the language they speak.

2.7 Governmental and institutional language attitudes and policies

In accordance with the Act LXXVII of 1993 on the Rights of National and Ethnic Minorities, the languages spoken by the Roma and/or Gypsies, namely Romani and Boyash, are officially recognized minority languages of Hungary. The state is obliged to support their teaching within educational institutions when requested by the parents or legal representatives of at least eight students belonging to the same minority group. Teaching of Romani, as well the related standardization and codification efforts are however based on a single Vlax Romani variety called Lovári. Lovári is one of the most vital Romani dialects in Hungary, and there are several activists among the Vlax Roma who actively promote Romani language use and maintenance. The Vlax Romani varieties are however mainly unintelligible to the South Central Romani-speaking Vend Roma.

Thus, Romani as a minority language, irrespective of its dialects, is explicitly protected by the Hungarian government. However, the Roma speaking other Romani dialects than Vlax have not been involved in the revitalization movement which resulted in that standard Romani in Hungary is exclusively based on the Vlax dialect.

2.8 Community members' attitudes toward their own language

My consultants generally expressed a negative attitude toward the Romani dialect they speak. As follows, almost all speakers are indifferent whether Romani is getting lost because they

prefer to use Hungarian in their everyday life. During my field research, I did not meet any speaker who would actively support the maintenance of Romani. Only few of them expressed regret for its decline. This could also be caused by my presence, although I tried to avoid taking any view.

On the other hand, the secret function of Romani was reported by a few consultants as an advantage of Romani speakers (Q34).

(Q34)

W: What to say, there you have my older sister's children, there are three of them. So, my sister, her husband and their family. What do you think about the fact that they cannot say a word [in Romani]! They cannot speak their own language! Because my sister shacked up with a Boyash man and they do not speak Boyash, neither the language of Grinders, only Hungarian. And here is the result: God forbid their child should get somewhere... How good it would be if s/he spoke [Romani]! [W to me]: Well, you are a Hungarian, don't be angry please, I'm telling the truth, you are a Hungarian. But if you go by bus and someone starts to speak Romani, the Grinder language, you can understand every word, what they are saying about you, what they want to do with you, am I right? Well, that is also the case of my daughter. Wherever she goes, because she is already old enough, and if the others would speak about her, she would understand what they want.

The claim from the above quotation 'They cannot speak their own language!' is particularly interesting, since the speaker refers to the Romani language as a property/inheritance exclusively owned by the ethnic Roma. Indeed, being a speaker of Romani, my non-Romani identity and ethnic belonging was challenged by my consultants several times, although I visually do not fit in the stereotypical picture of a Romani woman.

Sometimes the speakers claimed not to be competent in Romani since they associate the language with poverty, lack of education, backwardness or primitivism. Thus, Romani became the symbol of 'gypsiness' in a negative, stereotypical sense in the eyes of some Roma. For instance, one day a Vend Romani couple advised me to go to the Red Cross to get some clothes for free. While the husband held the opinion that I should speak some Slavic language in order to achieve the desired goal more easily, the wife strongly disagreed by stating that I should rather speak Romani (Q35).

(Q35)

W: *Odá žánes, hod' te tut na úle a'i fôti, ande soste te ójs edej Maďarország, hod' edej hi Kapoš, odoj asaj Vöröskereset, káj šaj tuke ár kédes čak afka fôti?*

M: *Karitas.*

W: *O karitas hi, mer odoj den tut. Akárso sükség hi tut upre valasoste, den tut odoj fôti, so tuke való...*

M: *Taj t' ánde gejal, ma ungrika, taj ni román ma vaker! Sar žánes: vaj čejitike vaj lend'elike, akor féder tut den.* [laugh]

ZB: *De te tumen phiren odoj?*

M: *Na, (...)*

W: *Inkáb odoj román te vakérel!*

W: You know, if you do not have enough clothes to wear here in Hungary, there you have Nagykapos, and there is such Red Cross, where you can get clothes just like that.

M: Charity.

W: It is the charity, because they give you stuff. Whatever you need, they give you clothes, in your size...

M: And when you go there, do not speak Hungarian, nor Romani! Speak the languages you know: either Czech or Polish, they will give you stuff more easily. [laugh]

ZB: Do you also go there?

M: No, (...)

W: Would be better for her to speak Romani there!

The quotation above illustrates well how competence in Romani is associated with poverty, at least in the eyes of the wife. There were also speakers who expressed their doubts about the notion that the Hungarian-monolingual Roma (i.e. the Musicians or Beggars) would never speak Romani, as for instance my consultant said: *Mert adá inkáb asaj kódušen, taj lážen i romani čhip.* 'Because they are such Beggars, and they are ashamed of the Romani language.' When I was searching for speakers in the field, it was not unusual to get the answer 'we are not such Gypsies', where 'such' stands for the attributes described above. It also seems that the younger generation may view Romani even more negatively than their parents. It is well illustrated by the situation when I was speaking in Romani with an old couple and their son entered the room addressing them in Hungarian: *Most mit cigánykodtok itt?* 'Why are you "gypsy" here (i.e. behaving like Gypsies)?' Some parents were also reported not to teach their children Romani because the children laugh at them when they speak the language. Others were of the opinion that today's Romani is not 'pure' or 'original' mainly because it is 'contaminated' with a high number of Hungarian loanwords. That's why a few of the speakers were not even sure of their own competence in Romani.

To summarize, it seems that the Vend Roma passively assimilate to Hungarian, while some of the speakers are not even aware about the consequences of such choice, such as that the language will be definitely lost.

2.9 Amount and quality of documentation

The documentation of Vend Romani is inadequate as there is only a brief grammatical sketch, a single word list, and some short fragmentary texts, mainly with no available audio recordings. Vend Romani is therefore only fragmentarily documented, placing it at level 1 on the scale for this factor.

2.9.1 Earliest sources on Central Romani in Transdanubia

2.9.1.1 János Szmodics²¹

The oldest source on South Central Romani spoken in Transdanubia is Szmodics' manuscript *Czigány Grammatika* from 1827, and its shorter version from 1836. The manuscript deals with phonology, morphology and syntax of Romani, without specifying which Romani dialect is the grammar based on. The final part of the manuscript contains several unauthentic letters and conversations in Romani with Hungarian translation, constructed most probably by the author, e.g.:

ROMANI²²

*Lošan, mangav tut, lošan romano národona!
Káj tu akáneč čáčune áthenca šaj dikhes odá
čirla užardo najgereder kam, andro saveste
uštal tuke tra čibakero šukaripnaskeri rátutni*

TRANSLATION

Be felicitous, I beg you, be felicitous Gypsy nation! Because you can see with your own eyes the desired Day when the morning star of your language's improvement will

²¹ The manuscript is archived under the name *Szmódi*. On the other hand, Habsburg (1890) spells his name as *Szmodis*. All other sources introduce this author as *Szmodics* (Vekerdi 1982; Szmodics 1836; Marics 2010).

²² Henceforth, the text is transcribed by means of the orthography used in the thesis. The original text: *Loshán! Mángáv tut loshán románó národona! Káj tu akáneč csácsune átyhencá scháj dikhesz, odá csirlá uzsárdó náj géreder kám! Ándro száveszte ushtyal tuke, trá csibákero schukáripnászkeri rátuni cserhen! The pásche ável, ádgýá báchtáli óri! hávi tut pále ávri lelá ándral odá, pál ádgýá but schel berschengero tchádgyoviben, terdó csernyipen, the túke jék csácsi, jék ákháribnáhá rodini páрни dúd kerelá; pro ucseder szikláríben trá ruminda dumákero!*

*čerhen! The páše ável ad'd'a baxtali óri, havi
tut pále avri lela andral odá pal ad'd'a but šel
beršengero čád'oviben, terdo čerňipen, the
tuke jek čáči, jek akharibnaha rodini pární
dud kerela pro učeder sikláríben tra ruminda
dumakero!*

resurrect. The happy hour is coming, which
will raise you from the darkness of those
hundreds of years, and it will bring you the
real, sought, bright lightness by means of the
honorable teaching of your decayed language.

There is however no doubt that Szmodics' work is for the most part plagiarized from the Czech linguist Puchmayer's grammar (1821) where the Romani varieties spoken in Czechia are described. Habsburg (1888: 304–307) and Vekardi (1982: 2) also expressed their doubt about the origin of his work, owing to a number of Slavisms occurring in the manuscript. The features characteristic to North Central Romani, or some of its varieties include for instance the existence of the voiceless velar fricative /x/ as in *baxtal-o* 'happy', the prothetic /j/ in third person pronouns (*jov* 'he', *joj* 'she', *jon* 'they'), the derivational suffix of abstract nouns in final /n/ (*-iben*, *-ipen*), the innovative final /n/ in *zumi-n* 'soup', the final /s/ in accusative, the imperfective suffix *-as*, the Czech-borrowed question marker *-li*, the interrogative *kaha* 'with whom' (see Figure 9), the relative pronouns *hav-o* 'which' and *har* (alongside *sar*) 'how', the superlative prefix *naj-*, the adposition *mamuj* 'in front of', the nouns *lovina* 'beer' or *lurdo* 'soldier', the adjective *džungál-o* 'ugly', the North Slavisms *musin-* 'need', *národos* 'nation', *divin-* 'to wonder', and many other. These features are, on the other hand, supplemented by some Hungarian loanwords and a few features typical for South Central Romani (Elšík *et al.* 1999), such as the imperfective suffix *-ahi* (alongside *-as*), the interrogative *ki'ti* 'how much, many', the quantifiers *ati* 'so much, many' and *zaloga* 'few', the demonstrative *áthar* 'from here', or the adposition *uze* 'to' (as it is shown on Figure 9)²³.

On the other hand, I have not encountered a single feature in Szmodics' grammar that would be undoubtedly of Vend Romani origin. For instance the contraction of *ave*, *ive*, *ove* and *uve* (see 3.1.8), one of the most typical innovations of Vend Romani, is not attested in his work. Furthermore, some of the words found in his grammar are typical for other varieties of South Central Romani such as the indefinite quantifier *zaloga* 'a few', which has been attested

²³ The Romani text in the right column: *Mángáv tut šukáre áv záloga uze mánde! Soszke? Táha rátyáhá ámen xálaha pre foroszte. Káhá? Read as: Mangav tut šukáre, av zaloga uze mande! Soske? Táha ra'taha amen džalaha pre foroste. Kaha?*

in the easternmost (non-Vend) varieties of Transdanubia and in the historical Nógrád county spoken nowadays (Elšík 2014). The interrogative *kana* ‘when’ and the noun *jilo* ‘heart’ is found in most varieties of Central Romani, but not in Vend Romani, cf. *kada* ‘when’ and *vod'i* ‘heart’, respectively. Thus, Szmodics’ work is a compilation of various sources, including at least Czech Romani and South Central Romani varieties other than Vend.

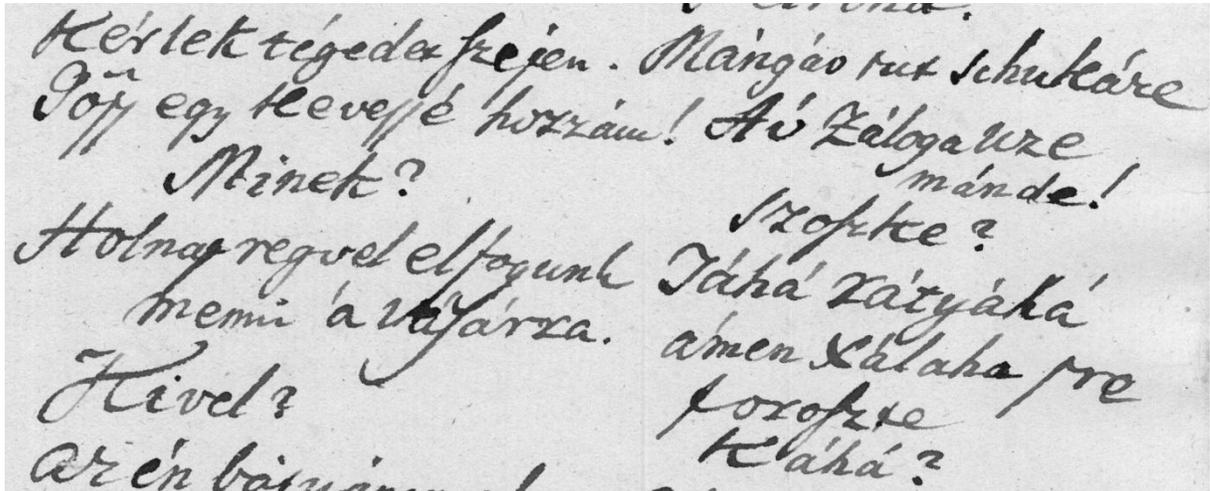


Figure 9 Detail from the manuscript of Szmodis (1827)

Thus, the question arises as to where the South Central Romani data of Szmodics come from. All sources on Szmodics (Habsburg 1888; Vekerdi 1982; Marics 2010) mention that he spoke Romani well, and even preached to his congregation in Romani. There is a legend telling that once Szmodics intended to preach in Romani in Siklós²⁴ (Baranya), but the local priest was against it. He then decided to preach outdoors. The Roma seemed to understand it, but at the same time expressed their disfavour of the fact that the Lords are learning Romani. At the end Szmodics had to escape from the angry crowd (Habsburg 1888: 307).

According to the description of the archived manuscript, Szmodics served as a priest in Gelse (Zala) from 1826 until his death in 1846. Marics (2010: 4) mentions that he was also a priest of Kutas (Somogy), the neighbouring village of Kisbajom. The manuscript was finished in 1827, just a year after he moved to Gelse. Therefore it is more likely that the small in number South Central Romani data were acquired either from the speakers of Kutas, or speakers of his home village Nemespátró (Zala). Today Vend Romani and Boyash is dominant

²⁴ There are no speakers of South Central Romani left in Siklós today.

in the counties of Somogy and Zala where Szmodics spent his life. Nonetheless, the area is also inhabited by some Hungarian monolingual Roma, distinguished as the group of Musicians and Beggars by the Vend Roma (see 1.4). It is therefore more than possible that Szmodics' data originate from the speakers of these groups. As it was already mentioned, the presence of Vend Roma at least in Somogy is most probably explained by their recent migration from what is now the border region of Austria and Hungary. It is conceivable that before their migration, a non-Vend variety of South Central Romani had been spoken in the area, whose remnants can be identified in Szmodics' grammar. However, in a number of instances it is almost impossible to determine whether the form originates from South Central Romani or it was overtaken from other sources.

2.9.1.2 József Habsburg

The following documentation on South Central Romani in the Pannonian Basin, including Transdanubia, comprises the grammar of Habsburg (1888) and the letters addressed to him written by Roma (1890). Habsburg (1888) distinguishes three Romani dialects spoken in the then Austro-Hungarian Empire: Transdanubian and Carpathian Romani, and the dialect spoken by wandering groups. However, none of these dialect groups as they are described in the grammar agrees with the present South Central Romani dialects. Table 4 shows some selected features related to the copula and verb inflection in the three dialects.

The Transdanubian dialect, which could be related to Vend Romani, is without doubt a South Central dialect, though we find some archaic features as the final /s/ in the third person singular preterite form, e.g. *márd'as* (Table 4). After analyzing Habsburg's data on 'Transdanubian Romani', I strongly suppose that his consultants spoke a South Central variety other than Vend, as I could not find any feature or innovation characteristic for Vend Romani. It may also be confirmed by the fact that Habsburg entitled the dialect of two Romani texts 'Transdanubian', which were written by speakers from Vál (Fejér) and Pécs (Baranya; see below). Both localities are situated in the easternmost part of Transdanubia where Vend Romani has most likely not been spoken traditionally.

	Transdanubian	Carpathian ²⁵	Wandering ²⁶
COP.PRS.1SG	<i>síňom</i>	<i>s'om, som</i>	<i>som</i>
COP.PRS.3SG	<i>síňe ~ síňa</i>	<i>s'le ~ s'la, si</i>	<i>si, hi</i>
COP.PRS.3PL	<i>síňe ~ síňan</i>	<i>s'le ~ s'an, si</i>	<i>si, his</i>
COP.PRT.1SG	<i>s(i)ňomahi</i>	<i>s(l)omahi</i>	<i>somas</i>
COP.PRT.3SG	<i>s(i)ňahi</i>	<i>s'ahi ~ sa</i>	<i>has</i>
COP.PRT.3PL	<i>s(i)nehi</i>	<i>s'ehi ~ sa</i>	<i>has</i>
beat.IMPF.1SG	<i>máravahi</i>	<i>maravas</i>	<i>marávas</i>
beat.IMPF.3SG	<i>márelahi</i>	<i>marlas</i>	<i>marélas</i>
beat.IMPF.3PL	<i>márenahi</i>	<i>marnas</i>	<i>marénas</i>
beat.PRT.1SG	<i>márd'om</i>	<i>mard'om</i>	<i>mardom</i>
beat.PRT.3SG	<i>márd'as</i>	<i>marda ~ marde</i>	<i>mardas</i>
beat.PRT.3PL	<i>márde</i>	<i>marde</i>	<i>marde</i>
beat.FUT.1SG	<i>márava</i>	<i>marava</i>	<i>maráva</i>
beat.FUT.2SG	<i>máresa</i>	<i>mareha</i>	<i>marésa</i>

Table 4 Comparison of some selected inflective forms of the copula and verb according to Habsburg (1888: 70–74)

Letter addressed to Habsburg on 30th of May 1890 in Pécs (Habsburg 1890: 757)

ROMANI²⁷

Uprono Mro rom Hercego!

TRANSLATION

My worshipful Lord Prince!

²⁵ The dialect designated as Carpathian seems to consist of features typical for both South and North Central Romani, e.g. the imperfective suffix of verbs is the older *-as*, typical for the North Central varieties, while the copula take the innovative suffix *-ahi* characteristic for the South Central varieties.

²⁶ The copula and verb inflection paradigm of the ‘wandering’ group is the closest to the North Central varieties, except for the unpalatitized stem in preterite and the retention of intervocalic /s/ in the second person singular future form.

²⁷ The original text: *Uprono Mro rom Herczego! – Mé, sukáré Mangáhitut, te oveszáhi aszavo lácsó, te suneszáhi ámen, még na színyatut baszt, te sunen, ola Pécs-szkra, Romane Banda, mé ándé kámáhi tuke té szikáven mra Pécs-szkra Angluno Románe Banda, szopálál amen szinyam téle csittó upro Balatonfüredate té czídel, mé sukáre mangavtut Uprono Mro-rom té oveszáhi aszavo lácsó, te suneszáhi ámen. suzsipeha ácsóvav. – Farkas Sándor, angluno primási.*

Me šukáre Mangáhi tut, te ovesáhi asavo láčo, te šunesáhi amen. Még na síňa tut bast te šunen ola Péč(i)skra Romane Banda. Me ánde kamáhi tuke te sikáven mra Péč(i)skra Angluno Romane Banda, so pálal amen siňam téle čitto upro Balatonfüredate te cídel. Me šukáre mangav tut Uprono Mro rom, te ovesáhi asavo láčo, te šunesáhi amen.

Šužipeha áčovav.

Farkas Sándor

angluno primáši

I would like to beg you nicely to be so kind as to listen to us. You have not had so far the luck to listen to the Romani Band of Pécs. I would like to present you my Prominent Romani Band of Pécs, which used to play down there in Balatonfüred. I nicely beg you my worshipful Lord to be so kind as to listen to us.

Yours Sincerely,

Farkas Sándor

first violinist

Transdanubian Romani as described by Habsburg shares a few features with Vend Romani as for instance the verb *čiv-* (PTC *čitt-*) with the meaning ‘to put’, which is marked by bold letters in the sample above. These dialects may be better seen as transitional varieties between Vend and Romungro (see 1.6), though the Romungro features clearly prevail. Thus, Habsburg, similarly to Szmodics, named as Transdanubian the Romungro varieties (with some Vend features) spoken in the Transdanubian region, most probably in its eastern part. On the other hand, Vend Romani most probably developed in the western part of Transdanubia or beyond, i.e. in the periphery of the South Central Romani area. This means that the present geographical location of speakers may only be explained by recent migration (see 1.5).

2.9.2 Documentation of Vend Romani in Hungary

2.9.2.1 József Vekerdi

The first reliable documentation of Vend Romani in Hungary dates back to the second half of the 20th century. It was carried out by the Hungarian linguist József Vekerdi. It includes the very first – brief in extent – grammatical sketch of Vend Romani published in 1984, where the most specific features of Vend Romani are compared to other South Central varieties. It is not clear whether his data are based only on the Vend Romani text recorded in 1981 which follows the grammatical part. This text was elicited by a native speaker from Nikla (Somogy). In addition, the paper includes a short word list of those words which in most parts differ from other Romani dialects, with indication of the Romungro form, and eventually the German or South-Slavic origin of the form.

The collection of tales and stories in different Romani dialects by Vekerdi (1985) contains the same tale as in Vekerdi (1984: 75–86), as well as an elicited text by a speaker of Öreglak (Somogy). There is no audio-recording available with the texts.

In 2000, Vekerdi published a multidialectal dictionary with indication of the dialectal affiliation of words, where several hundred words are marked as being of Vend Romani origin. The lexical items are translated to Hungarian and English. It is a solid source, though with some inaccuracies such as the occasional indication of the archaic non-contracted form of verbs in *áve*, *óve*, *úve* and *íve*, e.g. **l-áve-l* (Vekerdi 2000: 71), cf. *l-á-l* ‘to comb’.

2.9.2.2 *Ursula Glaeser et al.*

Four Vend Romani tales recorded in Ozora (Tolna) were published by Glaeser, Halwachs and Heinschink (1999), which document the local dialect of Rábahídvég (Vas) where the storyteller comes from. The paper in addition compares some selected grammatical features of Vend Romani with a closely related variety spoken in Austria. The vowel length, a distinctive phonological feature of Vend Romani, is not marked. The respective audio recordings are archived at the University of Graz.

2.9.2.3 *Melinda Rézműves and Károly Bari*

A brief narrative on recent history and traditions of Vend Roma was published by the Hungarian ethnologist Rézműves (2000). The narrative, as well as some Romani wisdoms, comes from a speaker of Somogy. In addition, the collection of Romani tales published by Rézműves in 2006 contains six Vend Romani tales from Kisbajom (Somogy). The book includes a CD-ROM on which the Romani and Hungarian versions of the tales are recorded. Bari (1999) recorded some several-sentence long songs and texts in Somogy which are available on the CD-ROM attached.

2.9.2.4 *Charles University in Prague*

The first attempt to extensively document Vend Romani – among other Romani varieties – and its intra-dialectal variation was made within the frame of the project *Linguistic Atlas of Central Romani* by a team based at the Charles University in Prague (see 1.7.2.1), which I was a member of. The three-year project (2011–2013) with the aim to document and analyse

several Central Romani varieties resulted in several hundred dialectological maps representing the geographical distribution of various linguistic features. As it has been already mentioned, the data analysed in the present thesis in large part come from this project.

Furthermore, the most recent documentation and description of Vend Romani includes three short Vend Romani stories and a song (Bodnárová 2013c), the ethnic and kin terminology of the Vend Roma (Bodnárová 2013b), and the process of loanword integration in Vend Romani (Bodnárová 2014).

2.9.3 Documentation of Vend Romani beyond Hungary

The first documentation of Vend Romani varieties spoken in Austria dates back to 1953, when Knobloch published several Vend Romani texts from Burgenland. Knobloch's texts can be considered the oldest source on Vend Romani in general. On the other hand, the first documentation on the neighbouring varieties in Slovenia was carried out by the ethnologist Štrukelj in 1980, followed by the first documentation of Vend Romani in Hungary by Vekerdi in 1984 (see above). The thorough documentation of Vend Romani spoken in Slovenia and Austria started only in the mid-nineties, as result of the revitalization processes. Although a large number of publications, including educational materials, have been produced in/on Prekmurje and Burgenland Romani since then, the Vend Romani speakers of Hungary are not familiar with these sources.

3 Phonology

3.1 Consonants

3.1.1 Consonant inventory and graphemes

The consonant phonemes of KR are indicated in Table 5.

	Labial	Alveolar	Post- alveolar	Palatal	Velar	Glottal
Nasal	m	n		ɲ		
Stop	p	b	t	d	c	ɟ k g
Aspirated	p ^h	t ^h			c ^h	k ^h
Affricate		ts	dʒ	tʃ	dʒ	
Aspirated				tʃ ^h		
Fricative	f	v	s	z	ʃ	ʒ h
Trill		r				
Approximant		l			j	

Table 5 Consonant phonemes

KR has preserved all Early Romani phonemes including the aspirates (cf. Matras 2002: 56), but not the voiceless velar fricative /x/. This sound was replaced by the glottal fricative /h/ due to South Slavic or/and Hungarian influence (Elšík *et al.* 1999: 295–297), e.g. *ha-* < *xa-* ‘to eat’. Furthermore, the voiced palatal affricate /dʒ/ merged with the post-alveolar fricative /ʒ/ in initial and intervocalic positions (e.g. **dža*^{B94} > *ža-* ‘to go’, **ladžatar*^{B94} > *lažatar* ‘shame.ABL’), while it is pronounced as voiceless /tʃ/ in word-final position, e.g. *mindž* /mintʃ/ ‘vagina’, *pándž* /pa:ntʃ/ ‘five’, *ládž* /la:tʃ/ ‘shame’. The sound /dʒ/ is realized, on the other hand, in medial position when preceded by the nasal /n/, such as in *prindžár-* /prindʒa:r/ ‘to know’, *pándžvarděš* /pa:ndʒvarde:ʃ/ ‘fifty’ or *mindža* /mindʒa/ ‘vagina.PL’. Unlike in Early Romani (Matras 2002: 56), the palatal dentals /c/ and /ɟ/ are distinctive phonemes in KR, as well as the palatalized counterpart of the nasal /n/. The phoneme /n/ has the allophones [ɲ] and [n] which are in complementary distribution. The former occurs before velar stops (e.g. *sun g-* ‘to smell’), while the latter is found elsewhere.

The German velar fricative /x/ is generally omitted when borrowed to KR, e.g. *lajt* (< G leicht) ‘light’, *gá* (< G dial. [ga:x])²⁸ ‘fast’. On the other hand, it is transmitted to KR in the form of the velar stop *k* regarding the loanword *óbok* (< G *Obacht*) ‘attention’. Unlike in the present dialects of German spoken in the Austro-Hungarian border region, the word-final trill is pronounced in the German loanwords of KR, e.g. in *níder* < G [ni:dɐ]. On the other hand, the trill became vocalized when followed by a consonant, e.g. *fút* (< G dial. [fuət]) ‘away’. The adaptation of German loanwords containing a lenis is dealt with in section 3.1.5, and a fortis stop /k/ in section 3.1.2.

The Hungarian phonemes are transferred by means of borrowings to KR without a major interference, as the consonant inventory of the two languages in contact largely coincide. Hungarian seems to have brought only the affricate /dʒ/ into KR.²⁹

The grapheme system used in the present thesis is shown in Table 6.

Grapheme	<i>b</i>	<i>c</i>	<i>č</i>	<i>čh</i>	<i>d</i>	<i>d'</i>	<i>dz</i>	<i>dž</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>j</i>	<i>k</i>	<i>kh</i>	<i>l</i>
IPA	/b/	/ts/	/tʃ/	/tʃʰ/	/d/	/j/	/dʒ/	/dʒ/	/f/	/g/	/h/	/j/	/k/	/kʰ/	/l/
Grapheme	<i>m</i>	<i>n</i>	<i>ň</i>	<i>p</i>	<i>ph</i>	<i>r</i>	<i>s</i>	<i>š</i>	<i>t</i>	<i>t'</i>	<i>th</i>	<i>t'h</i>	<i>v</i>	<i>z</i>	<i>ž</i>
IPA	/m/	/n/	/ɲ/	/p/	/pʰ/	/r/	/s/	/ʃ/	/t/	/c/	/tʰ/	/cʰ/	/v/	/z/	/ʒ/

Table 6 Grapheme system

This writing system roughly agrees with the one developed for Slovak and Czech Romani, which is based on the Slovak/Czech alphabet (Hübschmannová 1995: 197). It is characterised by the marking of the palatalized dentals /c ɟ ɲ/, the affricate /tʃ/ and the fricatives /ʃ ʒ/ by placing a caron over the respective graphemes. The aspiration is denoted by a digraph consisting of the aspirated sound together with the grapheme <h>. Additionally, the grapheme <t'h> has been introduced for the phoneme /cʰ/ which is absent from the orthography of Slovak/Czech Romani. The phoneme /dʒ/ proper to Hungarian loanwords is also marked by a digraph, namely by <dz>.

²⁸ Henceforth the German dialectal forms will be transcribed phonetically, since there are considerable differences in pronunciation between standard German and the German dialects of Eastern Austria.

²⁹ The consonant /dʒ/ in the original form **dzar* ‘body hair’ was replaced by /z/ in KR, i.e. *zár*.

All KR consonants may occur in initial position, except of the palatal dental t^{30} and its aspirated counterpart t^h ³¹. The palatal dental \check{n} in initial position is admissible only in Hungarian loanwords, e.g. *ňomo* ‘foot-print’, *ňujtín-* ‘to fasten’.

As for medial position, the only consonant which has not been attested in the data is the aspirated labial stop ph , except for in some compound words or superlative formations, e.g. *unoka-phen* ‘niece’, *leg-phureder* ‘the oldest’. The phoneme $/dʒ/$ has been only encountered in the intervocalic position, more precisely in the Hungarian-borrowed verbs *jedzín-* ‘to engage’ and *hedzín-* ‘to point’.

Most KR consonants are found in final position. Exceptions are the fricatives h ³² and \check{z} , and the aspirates ph , th and t^h . The word-final voiced consonants b , d , g , d' , v and z are generally realized as [p], [t], [k], [c], [f] and [s], respectively, due to devoicing rules (see 3.1.5). The consonants t' and f are also absent in word-final position, unless they substitute their voiced counterparts. The aspirated velar stop kh and palatal affricate $\check{c}h$ are realized at the end of the word as $/k/$ and $/tʃ/$ respectively, according to the rule of deaspiration (see 3.1.2). The consonants c , \check{n} , z and d' in word-final position are only attested in recent loanwords, e.g. *gonc* < G dial. [gɒnts] ‘quite’, *kormáň* < H *kormány* ‘government’, *doboz* < H *doboz* /*dobos* ‘box’, *vad'* /*vac*/ < H *vagy* ‘or’. Nevertheless, the palatal nasal may occur in some inherited words as well, due to the optional elision of final i , e.g. *páň* ~ *páňi* ‘water’.

Consonant inventory and graphemes in other varieties of Vend Romani

A significant difference is found between the consonant inventories of the northwestern (Sopron³³ and Veszprém Romani) and the southern varieties (Somogy, Zala and Vas Romani). That is, the former group prefers the affricates \check{c} , $d\check{z}$ and dz in place of the latter group’s palatal stops t' , d'/\check{z} and the fricative z , respectively. Since the latter group’s consonant system roughly

³⁰ Although the initial t' is permissible in Hungarian, there has not been attested such a loanword in KR.

³¹ The original nouns in initial t^h such as **thil* ‘butter’ and **thilav* ‘plum’ were replaced in KR by the borrowed *vaj* (< H *vaj*) and *slíva* (< S *šljiva*), respectively.

³² Like in Hungarian (Kenesei *et al.* 1998: 386), the glottal h is not pronounced at the end of the word in KR, e.g. *če* < H *cseh* /tʃɛ/ ‘Czech’, *ša* < **šah* ‘cabbage’ (cf. OBL *šah-en-*), *gá* < G dial. [gax] ‘fast’ (cf. COMP *gáh-éder*). Nonetheless, it is preserved in the inflectional forms of these words as well as in adapted loanwords (e.g. *pleh-o* < H *pléh* /ple:/ ‘tin’).

³³ Henceforth the Romani variety of the Győr-Moson-Sopron county will be called ‘Sopron Romani’.

agrees with the system described above on the example of KR, here I will deal only with the differences found in the former group as compared to KR.

The Sopron-Veszprém group is more archaic in having preserved the sound *dž* in the initial position in contrast to the latter group, e.g. *dža-* ‘to go’, cf. KR *ža-*. In addition, the number of occurrences of *dž* is increased in these varieties by the sound changes **ž > dž* and **gi > dži*, e.g. **žuto > džuto* (cf. S *žut*) ‘yellow’, **gili*^{B94} *> džili* ‘song’. The palatal *d* at the morphological boundary also results in *dž* (e.g. *kerdž-um* ‘do.PRT-1SG; I did’ < *kerd-*), while it occurs only sporadically alongside *dž*, e.g. *ódža ~ ód’a* ‘(to) there’. Similarly, the affricate *č* occupies the place of the palatal *t*, e.g. *kiči*, cf. KR *kiči* ‘how much/many’; *súč-um* ‘sleep.PRT-1SG; I slept’ < PTC *sút-*, cf. KR *sút’-um*. The approximant *j* is generally replaced either by *dž* or *č* when follows a consonant (see 3.1.7). In Burgenland and in some varieties of Prekmurje, the palatals *d*’ and *t*’ have been entirely replaced by *dž* and *č*, respectively. Thus, with regard to the consonant inventory, Sopron and Veszprém Romani is similar to Burgenland Romani and to some peripheral varieties of Prekmurje Romani.

Another shared feature of Sopron and Veszprém Romani is that the sound *dz* is not limited only to Hungarian borrowings (as it is in KR), but it also appears as an optional variant of an initial *z*, e.g. *dzumi ~ zumi* ‘soup’, *dzija ~ zija* ‘duvet’, *dzorál-o ~ zorál-o* ‘strong’.

3.1.2 Aspirates and aspiration

The aspirated sounds are phonologically distinctive compared to their non-aspirated counterparts, as it is for instance illustrated by the minimal pair *ker* ‘make!’ vs. *kher* ‘house’ (the same example also in Halwachs 2002: 4). KR has preserved the Early Romani aspirated stops *kh*, *ph*, *th*, *tʰ* and the affricate *čh* (cf. Matras 2002: 54). The aspirate *tʰ* developed through various processes: In the nouns *mátʰin* ‘fly’ and (**)mortʰin* ‘leather’ (attested only as an adjective *mortʰuno* ‘leather’), the root-internal **khi* was substituted by *tʰi* (cf. Elšík *et al.* 1999: 291). By contrast, the aspirate *tʰ* resulted from palatalization in *átʰa* ‘eye.PL’ (< **akhja* < **jakha*), *dítʰ-o(v)-* (< **dikh-jov-*) ‘to appear, seem’ and *naʰh-ov-fer* (< **nakh-jov-*)³⁴ ‘to pass, elapse’ (cf. *ibid*: 318, 366).

³⁴ Only the middle form of the inherited verb *nakh*^{B94} ‘to pass’ has been attested in the data, i.e. *naʰhov-* ‘to pass, elapse’.

The aspirated velar stop *kh* is not limited in its occurrence to the inherited lexicon, since it also appears in some German loanwords of South Bavarian³⁵ origin (e.g. *khafě* < G *Kaffee* [k^hafɛ:] ‘coffee’, *khirin-* < G dial. [k^hiɾŋ] ‘to shout, scream’, *khuglina* < G *Kugel* [k^huɡl] ‘bullet, shot’), and even in the Hungarian loanword *khártá*³⁶ (< H *kártya*) ‘card’. By contrast, the velar aspirate is not preserved when followed by a consonant, e.g. *klát* < G dial. [k^hladl] ‘dress’.

The aspirated sounds generally occur in initial and intervocalic positions, while the aspiration is lost in word-final position, e.g. *likh* /lik/ ‘nit’ vs. *likhenca* /li:k^hentsa/ ‘nit.INS’. The loss of final aspiration is not reflected in the spelling if the paradigm of the word indicates the existence of an aspirated sound, as it is shown in the previous instance. The aspiration is only exceptionally retained when followed by a word in initial vowel, e.g. *Lakh* /lak^h/ *ár so h’ odá!* ‘Guess what it is!’, cf. *Dikh* /dik/ *adá fe pášal!* ‘Look at it closely!’ The aspirates are usually omitted also before consonants at the morphological boundary, e.g. *dik-jam* (< **dikh-jam*) ‘see-PRT.1PL’, *áč-lahi* (< **áčh-lahi*) ‘stay-3SG.IMPF’. This morphophonological process is on the other hand reflected in the orthography. The aspiration is preserved before a consonant only in case of the word-initial consonant cluster /phr/, e.g. *phral* ‘brother’, *phrál* ‘to open’.

Aspirates and aspiration in other varieties of Vend Romani

In place of the KR aspirate *tʰ* the affricate *č(h)* appears in Veszprém, Sopron and Burgenland, and in most varieties of Prekmurje Romani. The aspirate *čh* has only occasionally been attested in Sopron and Veszprém, which may also be caused by the lower language competence of the speakers. The aspiration is also audible in the reciprocals *jékekhráve* and *kh ráve* ‘each other, one another’ in Vásárosdombó (Baranya) and Csokonyavisonta (Somogy), even though it is followed by a consonant. The aspirate *th* emerged in the Zala and Veszprém Romani loanword *frajthov* (< G compound *Fried-hof* ‘lit. peace-yard’) ‘graveyard’ through the coalescence of the German sounds *t* and *h* at the morphological boundary.

³⁵ In contrast to standard German, only the fortis stop /k/ is aspirated in the South Bavarian varieties, while the other stops /d t b p g k/ are realized as voiceless lenis.

³⁶ It is possible that this form has been contaminated by the corresponding German form *Karte* /k^harte/) ‘card’.

3.1.3 Consonant clusters

Consonant clusters may occur in every position of the word in KR: word-initially, word-medially and word-finally. The most widespread are the clusters consisting of two consonants, while those with three consonants occur mostly in medial, and rarely in initial position. The number of medial *CC*-clusters occurring in inflected and/or derived words exceeds the number of clusters found in lexical morphemes. Here I will introduce only the pre-Hungarian consonant clusters other than geminates. For the geminates consult section 3.1.4, and for the translation of the lexemes in Table 7–10 consult the vocabulary in the Appendix.

3.1.3.1 Word-initial clusters

The initial clusters of inherited words consist of obstruent³⁷-liquid /br dr gr p^hr pr sr tr/, obstruent-obstruent /ʃt sv/ and nasal-liquid /mr/ combinations. The latter cluster is represented only by the possessive pronoun *mr-o* ‘my’ which developed through the contraction **mir-o* > *mr-o*. The initial clusters /kl kr pl sl ʃp st/ entered KR through Slavic, while the clusters /ts fl/ through German borrowings (see Table 7).

I	S	G		I: inherited, S: Slavic, G: German, (?): unknown
✓			/dr/	I: <i>drákhi, drab, drom</i>
✓			/p ^h r/	I: <i>phrál, phral</i>
✓			/ʃt/	I: <i>štár</i>
✓			/mr/	I: <i>mro</i>
✓	✓		/gr/	I: <i>gra</i> ; S: <i>gráblálinel, grobo</i>
✓	✓		/br/	I: <i>bríšind</i> ; S: <i>bríga</i>
✓	✓		/sr/	I: <i>srasta</i> ; S: <i>srída</i>
✓	✓	✓	/tr/	I: <i>tro, trádel, tranda, trómal, trin, truš</i> ; S: <i>trézvisajol, tresánel, trašilo</i> ; G: <i>tráninel</i>
✓	✓	✓	/sv/	I: <i>svíri</i> ; S: <i>sveci</i> ; G: <i>svituri</i>
✓	✓	✓	/pr/	I: <i>pro, prengéro, prastál, prindžárel, prásal</i> ; S: <i>praho, prik, prosto, primínel</i> ; G: <i>práni</i> ; (?) <i>pruťinel</i>

³⁷ The obstruents of KR include the stops /b d g p p^h t t^h k k^h/, the fricatives /v z ʒ f s ʃ h/ and the affricates /dʒ dʒ ts tʃ tʃ^h/.

I	S	G	I: inherited, S: Slavic, G: German, (?): unknown
✓		/pl/	S: <i>plasta, plán, plajci</i> ; (?) <i>plín</i>
✓		/sl/	S: <i>slívi</i>
✓		/st/	S: <i>staklo</i>
✓		/kl/	S: <i>klédalo, klačimo, klinci, klúčo, klínind'um</i> ; also I (e) <i>k-lik</i>
✓	✓	/kr/	S: <i>kruška</i> ; G: <i>kráksni, krécuno, krót, krumpin</i>
✓	✓	/ʃp/	S <i>špita</i> ; G: <i>špajterutno, špírínel, špót, špajz, špórulínel</i>
	✓	/tʃv/	G <i>cvituri</i>
	✓	/fl/	G <i>flohi</i>

Table 7 Pre-Hungarian initial CC-clusters in lexical morphemes

Initial three-consonant cluster is found only in the loanwords *straka* ‘magpie’, *štrajtín-* ‘to quarrel’ and *štrimfi* ‘stockings’. The first has undoubtedly Slavic-origin, while the latter two are borrowed from either German or the local dialect of Hungarian.

3.1.3.2 Word-medial clusters

There are a wide number of clusters in medial position (see Table 8–9). The most common combinations are the obstruent-obstruent /fk ft ʃc sv vʃ zd ʃk st ʃt/, obstruent-liquid /dl kr tl vl vr dr kl pr tr/, liquid-obstruent /lv rts rd rh rʰ lh rʃ rk rv/, obstruent-nasal /kn kɲ ʃn sn/ and nasal-obstruent /nd ndʒ ng nh/. The liquid-nasal /rm rɲ rn/, nasal-nasal /mn mɲ/, nasal-liquid /ml/ and obstruent-glide /vj/ types of clusters are less frequent.

I	S	G	I: inherited, S: Slavic, G: German, (?): unknown
✓		/dl/	I: <i>adla</i>
✓		/fk/	I: <i>afka, d'efkar</i>
✓		/ft/	I: <i>ófto, éfta</i>
✓		/kn/	I: <i>čikno, tikno, phukni, biknel</i>
✓		/kɲ/	I: <i>cukňúdi</i>
✓		/kr/	I: <i>bákro</i>
✓		/lv/	I: <i>balvas</i>
✓		/ml/	I: <i>umlál</i>

I	S	G	I: inherited, S: Slavic, G: German, (?): unknown
✓		/mn/	I: <i>somnak, lumni, khámni</i>
✓		/mp/	I: <i>čumňik</i>
✓		/nd/	I: <i>tranda, khándel, ando</i>
✓		/ndʒ/	I: <i>prindžárel</i>
✓		/ng/	I: <i>nango, bango, šinga, angušto, angar, angáli, khangéri, žangál, sungel, mángel, čhungárel</i>
✓		/nh/	I: <i>kanhi</i>
✓		/rts/	I: <i>vurci</i>
✓		/rd/	I: <i>hurdo, verda, órde, murdárel</i>
✓		/rh/	I: <i>kirhaj, čerhéni, cerha</i>
✓		/rm/	I: <i>kirmo, ermáňa, ármi, kermúso</i>
✓		/rɲ/	I: <i>burňik, hurňa</i>
✓		/rɕʰ/	I: <i>mortʰin</i>
✓		/fn/	I: <i>bášno, hušnel</i>
✓		/fɕ/	I: <i>parašt'uva, ušt'el</i>
✓		/sv/	I: <i>nasvalo, ásvin</i>
✓		/tl/	I: <i>šutlo</i>
✓		/vɟ/	I: <i>karavd'in, avd'in</i>
✓		/vj/	I: <i>bávjal</i>
✓		/vl/	I: <i>žuvli, kóvlo</i>
✓		/vr/	I: <i>avral, pašávro, čhavri</i>
✓		/zd/	I: <i>azdel, rezdal, pizdi</i>
✓	✓	/dr/	I: <i>vodro, šudro, pedro</i> ; S: <i>modro</i>
✓	✓	/kl/	I: <i>ráklo, miriklo, čirikli</i> ; S: <i>rokla, staklo, dokle</i>
✓	✓	/lh/	I: <i>bulho</i> ; S: <i>mulhi</i>
✓	✓	/pr/	I: <i>upro</i> ; S: <i>opruja</i>
✓	✓	/rɟ/	I: <i>urd'el, pherd'as, hurd'aléko</i> (< <i>hurdo</i>); S: <i>erd'avo</i>
✓	✓	/rk/	I: <i>kurko, harkum, korkóro, kerko</i> ; S: <i>morkoňi</i> ; (?) <i>cirki</i>
✓	✓	/rn/	I: <i>pherno, parno, terno, kirno, harno</i> ; S: <i>mirno</i> ; H: <i>varno</i>
✓	✓	/rv/	I: <i>parvárel, barválo</i> ; S: <i>garvano</i>

I	S	G		I: inherited, S: Slavic, G: German, (?): unknown
✓	✓		/ʃk/	I: <i>maškar</i> ; S <i>kruška, bubreško</i>
✓	✓		/sn/	I: <i>khosno</i> ; S <i>lasno</i>
✓	✓		/tr/	I: <i>žamutro, pátrin, mutrel, patrāja</i> ; S <i>petrežilo</i>
✓	✓	✓	/st/	I: <i>anrusti, srasta, sósten, sásto, prastál, pobistérel, astárel</i> , OBL <i>grastes-</i> < <i>gra</i> , OBL <i>vastes-</i> < <i>va</i> ; S: <i>ostolo, prosto, plasta, dosta, misto</i> ; G: <i>lajsti, fajst</i>
✓	✓	✓	/ʃt/	I: <i>angušto, uštál, uštídel, náštig</i> ; S: <i>ništa</i> ; G: <i>ajšti</i>
		✓	/nts/	S: <i>klinco, ninco</i>
		✓	/sk/	S: <i>vusko</i>
		✓	/ʃɲ/	S: <i>češňáko</i>
		✓	/vd/	S: <i>évda</i>
		✓	/zv/	S: <i>trézvisajol</i>
✓	✓		/mb/	S: <i>žamba</i> ; G: <i>bimbi</i>
✓	✓		/nk/	S: <i>šunka</i> ; G: <i>henkínel, lanko</i>
		✓	/fn/	G: <i>háfni</i>
		✓	/gl/	G: <i>khuglini</i>
		✓	/mp/	G: <i>ampós, krumpin</i>
		✓	/nt/	G: <i>runto</i>
		✓	/tn/	G: <i>sajtni</i>

Table 8 Pre-Hungarian medial CC-clusters in lexical morphemes

Several clusters have only recently arisen through processes such as the vowel elision (e.g. *adla* < *adala* ‘these’, *balvas* < **balevas*^{B94} ‘bacon’), devoicing (e.g. *afka* < **avka* ‘so’), metathesis (e.g. *d’efkar* < **d’ekfar* ‘once’), palatalization (e.g. *burňik* < **burnik*^{B94} ‘handful’), nasalization (e.g. *čumňik* < **čupni*^{B94} ‘whip’), simplification of consonant cluster (e.g. *umlav-* < **umblav-*^{B94} ‘to hang oneself’), or compounding (e.g. *khán-d-* ‘to stink’, cf. *khan* ‘smell’, *d-* ‘to give’). The medial CC-clusters /mb nts nk sk ʃɲ vd zv/ were introduced to KR from Slavic, while the clusters /fn gl mp nt tn/ from German. The CCC-clusters /ndr ngl ngr ʃkr str/ are encountered in inherited words, while there are only three loan-nouns comprising CCC-clusters, namely the Slavic-origin *ninčko* ‘German, Germany’, and the German-origin *kráksni* ‘tool bag’ and *bumtni* ‘wound’.

I	S	G	I: inherited, S: Slavic, G: German
✓		/ndr/	I: <i>lindra, andro</i>
✓		/ngl/	I: <i>kángli, anglo</i>
✓		/ngr/	I: <i>angrusti, ungriko</i>
✓		/ʃkr/	I: <i>maškro</i>
✓		/str/	I: <i>sástro</i>
	✓	/ntʃk/	S: <i>ninčko</i>
		✓ /ksn/	G: <i>kráksni</i>
		✓ /mtn/	G: <i>bumtni</i>

Table 9 Pre-Hungarian medial CCC-clusters in lexical morphemes

The vast majority of medial CC-clusters resulted from inflection or derivation, i.e., at the morphological boundary. Below I have listed the most common types of such clusters as well as the morphological context they emerge from:

- /C-l/ in third-person singular future and imperfect, e.g. *áč-la* ‘stay-3SG.FUT’, *cid-la* ‘pull-3SG.FUT’, *phág-lahi* ‘break-3SG.IMPF’, *dik-lahi* ‘watch-3SG.IMPF’,
- /C-n/ in second-person and third-person plural future and imperfect, e.g. *áč-na* ‘stay-2/3PL.FUT’, *cid-na* ‘pull-2/3PL.FUT’, *phág-nahi* ‘break-2/3PL.IMPF’, *dik-nahi* ‘watch-2/3PL.IMPF’,
- /C-ɟ C-c C-j/ especially in preterite (except of the third-person plural), and in several other inflected and derived forms, e.g. *háb-d’a* ‘food-PL’, *phág-d’a* ‘break-PRT.3SG’, *kováč-t’a* ‘blacksmith-PL’,
- /C-d C-t/ in participle and third-person plural preterite forms, e.g. *khel-de* ‘dance-PRT.3PL’, *ker-do* ‘made-PTC’, *běš-te* ‘sit-PRT.3PL’,
- /s-t/ in the third-person masculine singular forms of ablative and locative, e.g. *les-te* ‘he-LOC’, *les-tar* ‘he-DAT’,
- /s-k/ in the masculine singular of genitive and dative, e.g. *les-ker-o* ‘he-GEN-M.SG’, *les-ke* ‘he-DAT’,
- /n-g/ in genitive plural (e.g. *len-ger-o* ‘they-GEN-M.SG’) as well as in iterative derivations such as *čhin-gér-* ‘cut-ITER-’,

- /n-d/ in locative and ablative plural, e.g. *len-de* ‘they-LOC’, *len-dar* ‘they-ABL’,
- /n-ts/ in the instrumental plural, e.g. *len-ca* ‘they-INS’,
- /n-k/ in iterative derivations, e.g. *sikavin-kér-* ‘show-ITER-’,
- /C-t/ in derivations with the ordinal *-to*, e.g. *štárto* ‘fourth’, *šófto* ‘sixth’,
- /C-v C-f/ in derivations with the multiplicative *-val* ~ *-far*, e.g. *štárval* ‘four times’, *butfar* ‘many times’,
- /C-k/ in feminine derivation of the *-kiňa* type, e.g. *pádárkiňa* ‘female doctor’, *bótoškiňa* ‘saleswoman’, *igazgatófskiňa* ‘directress’,
- /g-C k-C/ in formations with the superlative *leg-* ‘the most’, e.g. *leglasnéder* ‘the cheapest’, *legféder* ‘the best’ (*gf* > *kf*),
- and /r-C/ in derivations with the indefinite *akár-* ‘any’ and *aver-* ‘other’, e.g. *akárso* ‘anything’, *akárko* ‘anybody’, *averthán* ‘somewhere else’.

There are only few three-consonant clusters found in other than lexical morphemes. The most common are the /skr/ and /ngr/, which originated from the elision of /e/ in the genitive suffixes *-ker-* SG, *-ger-* PL, e.g. *les-kr-o* < *les-ker-o* ‘he-GEN-M.SG’, *len-gr-o* < *len-ger-o* ‘they-GEN-M.SG’. The cluster /ndl/ is in free variation with /nd/ in the preterite form of third-person plural verbs in stem-final /n/, e.g. *phen-dle* ~ *phen-de* ‘tell-PRT.3PL’. The CCC-clusters /ngl/ and /ngn/ appear in certain imperfective and future forms of the verbs *máng-* ‘to beg’ and *sung-* ‘to smell’, e.g. *máng-la* ‘beg-3SG.FUT’, *sung-nahi* ‘smell-2/3PL.IMPf’.

3.1.3.3 Word-final clusters

The final consonant clusters are the least in number (Table 10). They are composed of two consonants, while the second consonant is generally voiceless (see 3.1.5). In inherited words we find the nasal-obstruent /nk ntʃ nt/, liquid-obstruent /rʃ rt/ and obstruent-obstruent /st ʃt/ pairs of clusters. The latter cluster is also encountered in the apocopated form of the Slavic-origin word *ništ* < *ništa* ‘nothing’. The clusters /ft mp nts/ were brought into KR via German borrowings.

I	S	G		I: inherited, S: Slavic, G: German
✓			/ntʃ/	I: <i>pándž, mindž</i>
✓			/rʃ/	I: <i>murš, berš</i>
✓			/rt/	I: <i>čhungard, phurt</i>
✓			/st/	I: <i>bast</i>
✓	✓		/ʃt/	I: <i>kašt, vóšt</i> ; S: <i>ništ</i>
✓		✓	/nk/	I: <i>beng, čang</i> ; G: <i>ánk, link</i>
✓		✓	/nt/	I: <i>brišind, dand, d'énd</i> ; G: <i>feront, núrunt</i>
			✓ /ft/	G: <i>luft</i>
			✓ /mp/	G: <i>phumb</i>
			✓ /nts/	G <i>gonc</i>

Table 10 Pre-Hungarian final CC-clusters in lexical morphemes

Consonant clusters in other varieties of Vend Romani

In this section I will focus only on the initial and final clusters of other Vend Romani varieties which are absent in KR. In other varieties of Vend Romani, the initial CC-clusters /k^{hr}/ and /sk/ have been also encountered in the inherited lexicon. The former occurs in the reciprocal *khráve* ‘each other, one another’ in Csokonyavisonta (Somogy), and the latter in the noun *skámi* ‘chair’ found in several Vend Romani varieties. Further initial CC-clusters entered especially Szakonyfalu (Vas) and Zala Romani through Slavic contact. These are the /gl/ found in *glédalo* (cf. KR *klédalo*) ‘mirror’, /ʃk/ in *škola* ‘school’ (Vas and Zala), /vl/ in *vlahó* ‘Vlax’ and /vr/ in *vrištán-* ‘to scream’. The cluster /gj/ in *gjonó* (< **gnojo*, cf. S *gnój*) ‘dung’, which is attested in Zala Romani, resulted from the metathesis of the sounds *n* and *j*. The German-contact has in addition brought several initial clusters especially to Vas, Sopron and Veszprém Romani, such as /ʃl/ in *šlekt* ‘bad’ and *šlajferi* ‘grinder’, /ʃr/ in *šrajferi* ‘grinder’, /ʃn/ in *šnǒ* ‘fast’, /ʃt/ in *štüj* ‘quiet’, and /fr/ in *frajli* ‘of course’, *frajnézi* ‘nettle’, *frajthov* ‘graveyard’, *frančoft* ‘relative’, *fró* ‘happy’, *frogastica* ‘butterfly’.

As regards the final CC-clusters, the sequence /vr/ is found in the Zala Romani syncopated pronoun *ávr* (~ *ár*, < **aver*^{B94}) ‘other’, and the sequence /ht/ in the Vas Romani nouns *liht* ‘light’ and *óboht* ‘attention’, both borrowed from German.

3.1.4 Gemimates and gemination

All KR consonants have their long counterparts, except for the fricatives *f h v ž* and the nasal *m*. The existence of consonant length in KR is most probably triggered by the prolonged contact with Hungarian, that is, with a language with distinctive consonant length. The only minimal pairs attested in the data are *idegen-o* (noun) vs. *idegen-no* (adjective) (cf. H *idegen*) ‘foreign’, *sen-o* ‘coal’ vs. *sen-no* < **sent-no* (cf. H *szent*) ‘saint, holy’, and *ola* DEF.OBL.F vs. *olla* ‘those’ (see below). The two former pairs are constituted each by two Hungarian-borrowed items, while the latter comprises two inherited items. Like in Hungarian (Kenesei *et al.* 1998: 386), long consonants generally occur in word-medial position, while they are not allowed in initial position. Gemimates are rather rare in final position, as they become degeminated when adapted into KR, e.g. *mijelöt* (< H *mielőtt*) ‘before’, *ked* ~ *ket* (< H *kedd*) ‘Tuesday’.

I have counted overall 214 occurrences of gemimates in my Vend Romani text corpus comprising 300 thousand characters (ca. 80 thousand word tokens). The token frequency of the individual gemimates is illustrated in Figure 10.

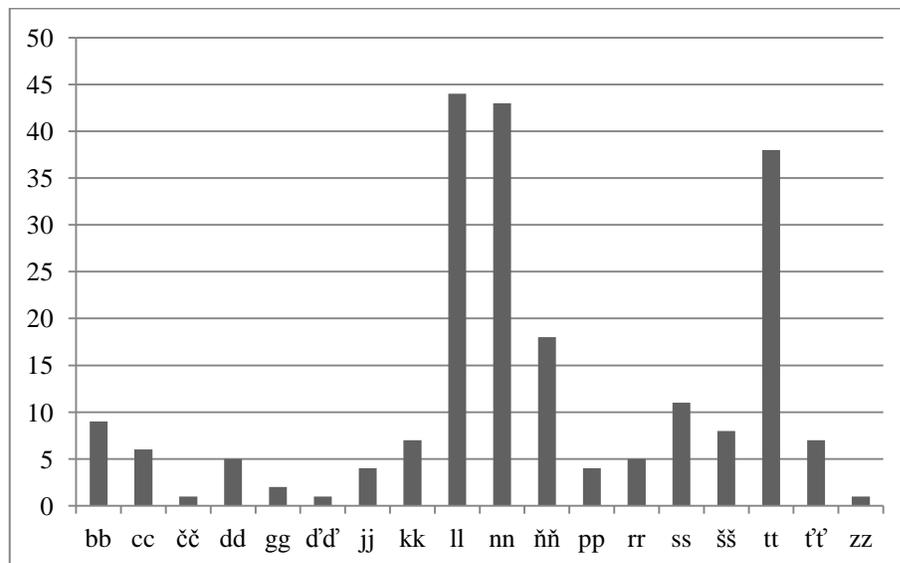


Figure 10 Token frequency of gemimates

The most common gemimates in KR are *ll*, *nn* and *tt*. The high number of *ll* and *nn* is caused by the syncope of *e* in the third-person singular and second/third-person plural personal markers (before the future or imperfective marker), respectively. This change only affects the

verbs in stem-final *n* or *l*, e.g. *gén-nahi* < **gen-en-ahi* ‘read.2/3PL.IMPF’, *gén-la* < **gen-el-a* ‘read.3SG.FUT’, *khel-lahi* < **khel-el-ahi* ‘read.3SG.IMPF’. Furthermore, Hungarian adjectives in final *n* are adapted into KR by *-n-*, which also gives rise to the geminate *nn*, e.g. *hütlen-n-o* (< H *hűtlen*) ‘unfaithful’, *išmeretlen-n-o* (< H *ismeretlen*) ‘stranger’, *semtelen-n-o* (< H *szemtelen*) ‘rude’. The number of *ll* geminates is increased by the assimilative sound change *rl* > *ll* which surfaces in third-person singular future and imperfect forms of verbs in stem-final *r*, e.g. *pel-la* ~ *per-la* ‘fall-3SG.FUT’ (cf. *pér-* ‘to fall’), *kel-lahi* ~ *ker-lahi* ‘make-3SG.IMPF’ (cf. *kér-* ‘to make’).

The large number of *tt* geminates entered into KR through the Hungarian dialect-borrowed factitive verbs in *-itt* (cf. standard H *-ít*), e.g. *épittín-* (< H dial. *építt*) ‘to build’, *sorittín-* (< H dial. *szorítt*) ‘to press’, *takarittín-* (< H dial. *takarítt*) ‘to tidy’. In general, intervocalic consonants tend to lengthen in KR, as well as in the local Hungarian dialect (Király 2005: 27), e.g. *hüttő* (< H *hűtő*) ‘fridge’, *koppín-* (< H *kap*) ‘to catch’, *alacsoňňan* (< H *alacsony*) ‘low’, or *hangoššan* (< H *hangosan*) ‘loudly’. On the other hand, the quantity of the consonant is not necessarily preserved in Hungarian loanwords, e.g. *tavassal* ~ *tavasal* (< H *tavasszal*) ‘in spring’, *sempilla* ~ *sempila* (< H *szempilla*) ‘eyelashes’, *akkor* ~ *akor* (< H *akkor*) ‘then’, *d’orš-abb-an* ~ *d’orš-ab-an* (< H *gyors-abb-an*) ‘faster’.

Assimilation of consonant clusters also led to the emergence of geminates, for instance, in *gullo* < **gudlo*^{B94} ‘sweet’, *kello* < **kerlo* ‘throat’, *čilla* < **čirla* ‘long ago’, *fitti* < G *fertig* ‘ready’, or *rittín- com* ‘make the bed’ < G *richten* ‘prepare, mend’. The intervocalic consonant in **trito* ‘three’ became spontaneously geminated in Vend Romani, giving rise to the form *tritto*. The words *adla* ‘these’, *odla* ‘those’ and *pášjov-* ‘to lie’ may optionally be pronounced with a geminate, i.e. *alla*, *olla*, *páššov-*.

The consonants *d t n d’ t’ ň* may become long before *j*, conforming the lengthening rule of Hungarian (Kenesei *et al.* 1998: 440), e.g. *bút’ťa* < **bútja* ‘works’, *rat’ťaha* ~ *rat’aha* < **ratjaha* ‘morning’, maybe also *eňňa* ~ *eňa* < Greek *έννιά*³⁸ ‘nine’. The origin of the geminate *jj* is found in the optional elision of *i* in the perfective marker of middle verbs in other than third-person plural forms, e.g. *erd’avisaj-j-um* < *erd’avisaj-ij-um* ‘worsen-PFV-1SG; I worsened’, *parvardisaj-j-al* < *parvardisaj-ij-al* ‘grow_up-PFV-2SG; you grew up’.

³⁸ The geminate *ňň* was most probably reintroduced rather than preserved in this case.

Geminates and gemination in other varieties of Vend Romani

For the same reason as in KR, the geminates *ll* and *nn* are overrepresented in other Vend Romani varieties as well. The geminate *tt* occurs only sporadically in Vas and Sopron Romani, which is most probably conditioned by the local Hungarian dialect.

3.1.5 Voice

The voiced and voiceless pairs of consonants in KR are:

voiced *b d g dž d' dz v z ž*
voiceless *p t k č t' c f s š*

The aspirates *čh kh ph th t'h* and the phoneme *h* do not have voiced counterparts; and, on the other hand, the voiceless counterparts of sonorants *l m n r j* and palatal nasal *ň* are absent. The voice opposition is phonologically distinctive, illustrated for instance by the minimal pairs *bal* 'hair' vs. *pal* 'behind', or *daj* 'mother' vs. *taj* 'and'. Both voiced and voiceless consonants may occur in initial and medial position. In word-final position, the voiced consonants undergo devoicing when the word is pronounced on its own, or when the following word begins with a voiceless consonant (4). The final consonant retains its voice quality when the adjacent word begins with a vowel (5) or with a consonant having the same voice quality (6).

(4)^{RM} *hod' tu /hoc tu/ nad'on lačhe vódiskro mánuš sal.*
COMP 2SG very good hearted man COP.2SG
(...) that you are a very kind person.

(5)^{LQCR} *hod' adá /hoj ada:/ fer fogineha te phukál.*
COMP this.M VP will.FUT.2SG COMP say.INF
(...) that you will say it.

(6)^{NAR} *hod' bári /hoj ba:ri/ vaj tikni.*
COMP big or small
(...) whether it is big or small.

Like in Hungarian (see e.g. Bárkányi & Kiss 2010), the postvocalic voiced fricative *v* in word-final position generally loses its voicing while the frication is preserved (e.g. *žánav* [ʒa:nɔv] ‘I know’). Moreover, the final voiceless consonant became lexicalized in some words, such as in *íč* < **ídž*^{B94} ‘yesterday’ (cf. *íčutno* ‘yesterday’s’) or *čhip* < **čhib*^{B94} ‘language’ (cf. *čhip’ta* ‘languages’). The voice alternates in case suffixes (a) as well as in some palatalized inflectional and derivational markers (b) (cf. Matras 2002: 53–54). The voiced variants of these suffixes are attached to stems in final voiced consonants, while the voiceless variants are employed elsewhere. The iterative derivational marker has also voiced and voiceless variants, i.e. *-in-gér-* ~ *-in-kér-*. The distribution of these suffixes is conditioned by the syllabic structure of the stem and the origin of the verb (see 4.7.2.5).

- | | | | |
|----|-------------------------------------|----|------------------------------|
| a. | GEN <i>-k(e)r-</i> ~ <i>-g(e)r-</i> | b. | PL <i>-d’a</i> ~ <i>-t’a</i> |
| | DAT <i>-ke</i> ~ <i>-ge</i> | | PTC <i>-d-</i> ~ <i>-t-</i> |
| | ABL <i>-tar</i> ~ <i>-dar</i> | | |
| | LOC <i>-te</i> ~ <i>-de</i> | | |

In the Slavic loanword *opruja* (< S *obrva*) ‘eyebrows’, the voiced bilabial stop /b/ changed to the voiceless /p/. German words that contain lenis are often adapted into KR with voiceless consonants, such as *pon* < G dial. [bɔʊn] ‘train’, *práni* < G dial. [brɔʊn] ‘brown’, *ampós* < G [ambɔs] ‘anvil’, *kráksni* < G dial. [graksn] ‘tool bag’, *krumpa* < G dial. [grɔmpan] ‘potato’, *krót* < G dial. [grɔt] ‘straight’. On the other hand, the lenis of German-borrowed words became voiced in *gá* < G dial. [gax] ‘fast’, *gonc* < G dial. [gɔʊnts] ‘quite’, *níder* < G [ni:ðɐ] ‘low’ and *éza* < G [e:zɪ] ‘donkey’. Hungarian words that end with a geminate are degeminated and optionally pronounced voiceless when borrowed to KR, e.g. *inkáb* ~ *inkáp* (< H *inkább*) ‘rather’, *legaláb* ~ *legaláp* (< H *legalább*) ‘at least’.

Voice in other varieties of Vend Romani

In Sopron, the velar *v* tends to vocalize into the approximant /v/ in the intervocalic position as well as in the cluster *VvC*, e.g. *garuvav* /garuvav ~ garuvav/ ‘I hide’, *dživdžár-* /dʒivdʒa:res/ ‘to fire’. However, the presence of devoiced *v* in *afka* (/afka/ < **avka* /avka/) ‘so’ suggests that the realization of postvocalic *v* through devoicing is older than its realization through vocalization.

The inflectional forms of the nouns *ič* (< **idž*) ‘yesterday’ and *čhip* (< **čhib*) ‘language’ show that the original voiced word-final consonants were devoiced in Somogy, Vas and Zala Romani, too.

3.1.6 Assimilation and dissimilation

All obstruent consonants of KR assimilate in voicing to the following consonant, thereby following the assimilation rule of Hungarian (Kenesei *et al.* 1998: 441–2). For instance, the first consonant of the cluster *gf* in *legföder* /lekfe:der/ ‘the best’ becomes voiceless, while the first consonant of the cluster *šb* in *kišbajum* /kizbajum/ ‘Kisbajom’ undergoes voicing. In contrast to colloquial Hungarian, the fricative /v/ may trigger voicing on the preceding consonant (e.g. *hétvége* /he:dve:ge/ ‘weekend’). The origin of this feature can be found in the Hungarian dialect spoken in Somogy (Király 2005: 28), the present contact language of KR. Interestingly, the fricative /v/ may become devoiced in the second position of certain words, which as well replicates the pattern found in the local Hungarian dialect (*ibid.*), e.g. *borotfa* < H *borotva* (H dial. *borotfa*) ‘razor’, *ötfen* < H *ötven* (H dial. *ötfen*) ‘fifty’. The devoicing of *v* has been attested also in some inherited words (e.g. *butfar* < **butvar* ‘many times’, *ratfál-o* < **ratvál-o* ‘bloody’), while the voicing triggered by *v* is reserved only to Hungarian loanwords. Thus, the voiced variants **budvar* and **radval-o* seem not to be possible.

The complete assimilation of consonants is found especially in Hungarian borrowings, since Hungarian exhibits a broad variety of assimilatory changes (Kenesei *et al.* 1998: 436–46). These borrowings are transcribed phonologically in the present thesis, e.g. *teccínel* /tetstsi:nel/ < H *tetszik* /tetstsik/ ‘likes’, *barra* /barra/ < H *balra* /barra/ ‘to the left’. Complete assimilation may also be found at the morphological boundary, such as in *herceg-kiňa* /hertsek-kija/ ‘princess’ (cf. *hercego* ‘prince’). For the complete assimilation of consonant clusters found in inherited words refer to section 3.1.4.

Diachronically, the labial obstruent changed to nasal consonant through assimilation with the following nasal, such as in the feminine nouns *gurumni* < **guruvni* ‘cow’, *somnak* < **sovnak*^{B94} ‘gold’, *khamni* < **khabni*^{B94} ‘pregnant’, *lumni* < **lubni*^{B94} ‘woman’, *čumňik* < **čupni*^{B94} ‘whip’, or in the adaptation marker of borrowed adjectives in final vowel *-mn-*, e.g. *utolšó-mn-o* (< **utolšó-vn-o*, cf. H *utolsó*) ‘last’, but not in *moštó-vn-o* (cf. H *mostoha*) ‘step-’. An example of a distance dissimilation is found in *bávjal* (< **bavlal* < **balval*^{B94}) ‘wind’ and

ángjal (< **anglal*^{B94}) ‘in front’, while a long-distance assimilation of the lateral *l* to the palatal approximant *j* in *khujája* ‘toilet.PL’ (< **khulája*, cf. SG *khuláli*), and in the derived forms of *mejajár-* (< **melajár-*) ‘to make dirty’ and *mejajov-* (< **melajov-*) ‘to become dirty’. The assimilation of adjacent consonants is also frequent between individual words. Consider the neighbouring consonant *r* and *l* in the following example:

- (7)^{NAR} mer le /mel le/ bálen még na din te hal.
 because DEF.OBL pig.ACC.PL yet NEG give.PRT.3PL COMP eat.INF
 (...) because they haven’t fed the pigs yet’.

Assimilation in other varieties of Vend Romani

The sound *v* may trigger voicing or become devoiced in Somogy and Zala Romani (e.g. *butfar* ‘many times’), but not in Sopron, Vas and Veszprém Romani (e.g. *butvar* ‘many times’). The derivational suffix *-vn-* does not undergo assimilation in Vas and Zala Romani, e.g. *židóvno* ‘Jewish’. While the dissimilated forms *bávjal* ‘wind’ and *ángjal* ‘in front’ appear in Somogy, Zala and Vas Romani, the original forms *bávlal* and *ánglal* are found in Sopron and Veszprém Romani.

3.1.7 Palatals and palatalization

The palatals of KR include the stops *d’*, *t’* and *tʃh* and the nasal *ň*. Palatal consonants may emerge before the vowel *i* and the palatal approximant (i.e. the yod). In KR, the sound *i* only occasionally triggered palatalization of the preceding dental (a) or velar stops (b) in the root.

- | | |
|--------------------------------------------------------------|------------------------------------------------------|
| a. <i>karavd’in</i> < * <i>karavdi</i> ^{B94} ‘crab’ | b. <i>avd’in</i> (~ <i>avgin</i>) ‘honey’ |
| <i>ód’a</i> < * <i>odija</i> ‘(to) there’ | <i>d’ili</i> < * <i>gili</i> ^{B94} ‘song’ |
| <i>at’i</i> < * <i>ati</i> ‘so much/many’ | <i>d’iva</i> < * <i>giv</i> ^{B94} ‘wheat’ |
| <i>kit’i</i> < * <i>keti</i> ^{B94} ‘how much/many’ | <i>vód’i</i> < * <i>ogi</i> ^{B94} ‘heart’ |
| | <i>pot’in-</i> < * <i>pokin-</i> ‘to pay’ |
| | <i>mát’hin</i> < * <i>makhi</i> ^{B94} ‘fly’ |

The unpalatalized sequence *gi* is preserved in *avgin* ‘honey’ (alongside the palatalized *avd’in*), the sequence *ki* in *kín-* ‘to buy’, *kíral* ‘cottage cheese’, *kirivo* ‘godfather’, *kirmo* ‘worm’ and *kiťi* ‘how much/many’, the sequence *ti* in *tikn-o* ‘small’, *uštíd-* ‘to get’, *náštig* ‘cannot’, *kijaráti* ‘evening’, *búti* ‘work’ and *angrusti* ‘ring’, and the sequence *khi* in *khin-o* ‘tired’ (cf. Elšík *et al.* 1999: 288–295).

The nasal changed to palatal when followed by *i* only in *burňik* < **burnik*^{B94} ‘handful’, *čerhěni* (~ *čerhěni*) ‘star’, *čumňik* < **čupni*^{B94} ‘whip’, *háňig* < **xanig* (cf. **xaning*^{B94}) ‘well’, *páňi* < **pani*^{B94} ‘water’, *ráňik* < **ránik* (cf. **Ran*^{B94}) ‘wand’ and *sapuňi* < **sapuni* (cf. **sapni*^{B94}) ‘soap’. By contrast, the unpalatalized cluster *ni* has been retained in feminine nouns such as in *khúni* ‘elbow’, *čoháni* ‘witch’, *manušni* ‘woman’, *piráni* ‘fiancée’, *ráni* ‘lady’, or *phukni* ‘blister’. The inflectional stem of verbs is palatalized before the perfective marker of middle verbs *il-* ~ *-ín-*, e.g. *khiň-ín-* ‘tired-PFV-’ < *khin-o* ‘tired’.

A yod may also trigger palatalization, resulting in the following outcomes: 1) the preceding dentals *d t n* become palatals (i.e. *d’ t’* and *ň*), 2) the preceding velars *g k j* become palatals (i.e. *d’ t’ th*), 3) the dental *l* is delateralized into the approximant *j* (Elšík *et al.* 1999: 294), while 4) other preceding sounds (except of the glottal *h*, see below) are accompanied by the palatal glide *j*, or in postconsonantal position by the palatals *d’* and *t’*. The two latter palatals are in complementary distribution: *d’* is employed after voiced consonants, and *t’* after voiceless consonants. The realization of *j* in postconsonantal position as *d’* ~ *t’* is also typical to several Hungarian dialects of Transdanubia (Király 2005: 26–28), which points to the fact that we are dealing with a contact-induced change in KR.

Palatalization triggered by the yod is found before the nominative plural marker in the nouns *angušto* (PL *angušt’a*) ‘finger’ and *kermúso* (PL *kermúst’a* < **kermúsja*) ‘mouse’ (see 4.1.3.1), as well as in the non-base forms of feminine nouns, e.g. *gurumňa* ‘cow.PL’ < *gurumni* ‘cow’. Furthermore, the perfective stem of verbs becomes palatalized before the personal concord marker in other than third person plural forms, e.g. *phučť-a* ‘ask.PRT-3SG; s/he asked’ < PFV stem *phučť-*. The derivational morphemes that induce palatalization include the marker of middle verbs *-(j)ov-*³⁹ (e.g. *čaj-ov-* ‘to eat oneself full’ < *čál-o* ‘full’), causatives *-(j)ár-* (e.g. *khamň-ár-* ‘to make pregnant’ < *khámn-i* ‘pregnant’), and the adjectival markers -

³⁹ Henceforth, when the type of the palatalization is not specified the palatal approximant *j* will be used to mark the palatalization.

(j)án- (e.g. *gurumň-án-o* ‘of cow’ < *gurumn-i* ‘cow’), -(j)ikán- (e.g. *lumň-ikán-o* < *lumn-i* ‘woman’) and -(j)ál- (e.g. *hev-d’ál-o* ‘leaky’ < *hév* ‘hole’). An irregular pattern of palatalization is found in *kaňh-a* ‘chicken.PL’ (< **kanhj-a*, cf. SG *kanh-i*) and *bujh-ár-* ‘to set the table’ (< **bulhj-ár-*, cf. *bulh-o* ‘wide’), since the penultimate consonants of the inflectional stems have turned into a palatal or a yod. This rule seems to be applied for words with stem-final *h*.

An on-going development in KR is that the initial palatal approximant *j* tends to change to *d’*, which affects merely the inherited words:

<i>j</i>	<i>j ~ d’</i>	<i>d’</i>
<i>jakh</i> ‘eye’	<i>jékh ~ d’ékh</i> ‘one’	<i>d’áro</i> < * <i>járo</i> ‘flour’
<i>jag</i> ‘fire’	<i>jefkar ~ d’efkar</i> ‘once’	<i>d’alo</i> < * <i>jalo</i> ‘raw’
<i>jagakéro</i> ‘police’		<i>d’iv</i> < * <i>jiv</i> ‘snow’
		<i>d’énd</i> < * <i>jénd</i> ‘winter’

The change *j > d’* is contact-induced, since the occurrence of *d’* or *dž* in place of the initial *j* is typical for several Hungarian dialects of Transdanubia (Imre 1971: 9, 50). It is interesting, however, that today none of the Hungarian loanwords display this change, e.g. *jego* < H *jég* ‘ice’, *jágeri* < H *jáger* ‘hunter’ (cf. G dial. [jaga]). Thus, the origin of initial *d’* in the inherited lexicon was induced by the former Hungarian dialect KR was in contact with, while the disappearance of initial *d’* in Hungarian loanwords is triggered by the present Hungarian dialect which, most probably, does not display the sound change *j > d’*.

Palatals and palatalization in other varieties of Vend Romani

The realization of the initial *j* as *d’/dž* is the most widespread in Zala and Vas Romani, while in Veszprém (and probably also in Sopron) Romani the change is reserved for the nouns *dživ* ‘snow’ and *džénd* ‘winter’ (Table 11). In Zala and Vas Romani, the change *j > d’* affected also the inherited noun *d’erni* < **jerni* (< **erni*^{B94}) ‘file’ as well as several loanwords, e.g. *d’upa* < S *jopa* ‘coat’, *d’ágeri* < H/G *jáger/Jäger* ‘hunter’, or *d’ego* < H *jég* ‘ice’.

	KR	VESZPRÉM	ZALA/VAS
* <i>jakh</i> ‘eye’	<i>j</i>	<i>j</i>	<i>d'</i>
* <i>jag</i> ‘fire’	<i>j</i>	<i>j</i>	<i>d'</i>
* <i>jekh</i> ‘one’	<i>j ~ d'</i>	<i>j</i>	<i>d' (j)</i>
* <i>járo</i> ‘flour’	<i>d'</i>	<i>j</i>	<i>d'</i>
* <i>jálo</i> ‘raw’	<i>d'</i>	<i>j</i>	<i>d'</i>
* <i>jiv</i> ‘snow’	<i>d'</i>	<i>dž</i>	<i>d'</i>
* <i>jénd</i> ‘winter’	<i>d'</i>	<i>dž</i>	<i>d'</i>

Table 11 Initial sound change $j > d'/dž$

In Veszprém and Sopron Romani the original palatal lateral **l'* was depalatalized into the dental *l* (cf. Elšík *et al.* 1999: 294), e.g. *siklol* (< **sik'lol*, cf. KR *sikjol*) ‘learns’, *kamlom* (< **kaml'om*, cf. KR *kamjom*) ‘I wanted’, *molaha* (< **mol'aha*, cf. KR *mojaha*) ‘with wine’, *džuvla* (< **džuv'la*, cf. KR *žuvja*) ‘women’. This also applies to Hungarian loanwords with the historical palatal lateral **l'*, such as H *petrezselyem* > KR *petrežilo* ‘parsley’ or H *király* < KR *királi* ‘king’.⁴⁰ On the other hand, the palatal lateral **l'* was delateralized into the palatal approximant *j* in the perfective marker **-il'*- of d-verbs (e.g. *rod-ij-* ‘searched’), irregular verbs such as *uštij-* ‘woke up’, *urdžij-* ‘dressed’, *áčhij-* ‘stayed’, *lij-* ‘took’ and *dij-* ‘gave’, and the copula *új-* ‘was/were’.

In the same varieties, the original palatal approximant **j* is realized as *dž* (cf. KR *d'*) after voiced and *č* (cf. KR *t'*) after voiceless consonants, e.g. *fačuvdža* (**fačuvja*) ‘children’, *skámdža* (**skámlla*) ‘tables’, *zenésča* (**zenésja*) ‘musicians’.

3.1.8 Contraction and related sound changes

A distinctive feature of Vend Romani (including KR) in comparison to the northern varieties of South Central Romani is the existence of a wide variety of sound changes caused by contraction. The source of it may be found in the German and/or Hungarian dialect spoken in the Austro-Hungarian border-region, since contraction is very common in both of these

⁴⁰ Note that the Hungarian dialects in Western Transdanubia display the same development (Király 2005: 26; Imre 1971: 51).

contact languages (Imre 1971: 11, 55). The most striking example of it is the omission of intervocalic *v* in the cluster *Vve* (namely in *ave*, *ove*, *uve* and *ive* sequences), which probably first resulted in hiatus and later in various diphthongs depending on the adjacent vowel, e.g. **tavel* > **tael* ‘cooks’. The same development is attested in the Hungarian dialect spoken in Oberwart (Austria), which was triggered by the German contact, e.g. H standard **hova* > Oberwart Hungarian *hoa* ‘to where’ (Imre 1971: 55). In KR the diphthongs were presumably replaced by a long vowel after the loss of German contact, i.e. *ae* > *á*, *oe* > *ó*, *ue* > *ú* and *ie* > *í*.

The contraction *Vve* > *V*: that has developed through the monophthongization of diphthongs is typical in the present third person singular and second/third-person plural forms of verbs with stem-final *v*, e.g. **tael* > *tál* ‘cooks’, **ziel* > *žil* ‘lives’, **thoen* > *thón* ‘you/they wash’, **garuen* > *garún* ‘you/they hide’. In the second person singular, the cluster *Vve* is reduced to *Vj*, e.g. *garujs* < **garuves* ‘you hide’, *hajojs* < **hajoves* ‘you understand’.⁴¹ The following words are also contracted: *god’ár* < **god’aver* ‘smart’, *árto* < **averto* ‘the next’ and *dí* < **dives* ‘day’ (but not *divése* ‘during the day’), as well as the second and third-person causatives and middle verbs, such as *hajol* < **hajovel* ‘understands’ or *kerál* < **keravel* ‘makes so. do’. The sequence *Vve* is preserved, on the other hand, in the second-person plural imperative forms (e.g. *phukaven* ‘tell.2PL.IMP’, cf. **phukán*), in the case-inflected forms of nouns in stem-final *v* (e.g. *thaveha* ‘thread.INS’, *alaveske* ‘word.DAT’, *gaveskero* ‘village.GEN’), and in several Hungarian loanwords, such as in *haveri* (< H *haver*) ‘friend’.

The cluster *iva* (> *ija*) was shortened to *a* in the first-person forms of *čhiv-* ‘to put’, i.e. *čhav* (< **čhijav* < **čhivav*) ‘I put’, *čas* (< **čhijas* < **čhivas*) ‘we put’. This type of contraction has not affected the first-person forms of the verbs *živ-* ‘to live’ and *siv-* ‘to sew’, where a glide is inserted to break the hiatus after the loss of *v* i.e. *žijav* (< **živav*) ‘I live’, *žijas* (< **živas*) ‘we live’, *sijav* (< **sivav*) ‘I sew’, *sijas* (< **sivas*) ‘we sew’.⁴² The glide *j* also replaced the fricative *v* in the borrowed noun *opruja* (< **opruva* < S *obrva*) ‘eyebrows’, in the multiplicative *trijal* (< **trival*) ‘three times’ and in the derived abstract nouns in stem-final *v*, such as in *rojibe* (< **rovibe*) ‘cry’, *tájibe* (< **távibe*) ‘cooking’ or *uštajibe* (< **uštavibe*) ‘step’.

⁴¹ It is possible that the change *Vve* > *Vj* previously occurred also in the third person singular and second/third-person plural forms, while the glide *j* has been only preserved before the sibilant *s*.

⁴² The cluster *iva* has been preserved only in the forms *kitivar* ‘how many times’ and *aivar* ‘so many times’ in KR.

The contraction *ije* > *í* is found in the second and third-person forms of *pij-* ‘to drink’ (e.g. *píl* < **pijel* ‘drinks’), while the cluster *ame* is replaced by *á* in *pekál* < **pekamel* ‘need.3SG’. The contraction *eve* > *é* affected the nouns *d’énd* < **d’evend* ‘winter’ (but not *d’evénde* ‘in winter’) and *dél* < **devel* ‘God’ (but not the inflected form *dévl-*).

The sound change *ve* > *j* is typical at the end of the word such as in the plural forms *saj* (< **save*) ‘what kind of’, *asaj* (< **asave*) ‘such’ and *lój* (< **lóve*) ‘money’. The change *vi* > *j* occurs in the plural forms of xenoclitic nouns (e.g. *patkój* < **patkóvi* ‘horseshoe.PL’), and in the feminine singular forms *saj* (< **savi*) ‘what kind of’ and *asaj* (< **asavi*) ‘such’. As a result of these sound changes, the two latter forms have homonymous forms for various distinct functions, e.g. *asaj* < **asave* ‘such-PL; such-M.SG.OBL; such-PL.OBL’ and *asaj* < **asavi* ‘such-F.SG’. In the masculine singular, the contraction *avo* > *ó* is attested, i.e. *asó* < **asavo*. Furthermore, the contraction **ava* > *á* is found in the first-person singular future and imperfect, while the contraction **ahahi* > *áhi* in the third-person singular irrealis forms.

A rather uncommon reduction of stem is displayed in the first-person perfective forms *phom* < **phend’om* ‘I told’ and *žom* < **žand’om* ‘I knew’.

Contraction in other varieties of Vend Romani

In addition to the sound changes mentioned above, the sound change *ove* > *ó* affected the Vas Romani noun *tór* < **tover* ‘axe’ (cf. KR *hokono* ‘axe’), while the change of the final *ve* into *j* occurs in the accusative form of the Veszprém Romani noun *čhá* ‘boy, son’, i.e. *čháj* < **čháve* ‘boy/son.ACC’. The contraction *avo* > *á* in *sári žéne* (< **savore žene*) is attested only in Szakonyfalu (Vas). In the same variety, the final cluster *avo* is contracted to the diphthong *au* in *čhau* < **čhavo* ‘boy, son’ and *asau* < **asavo* ‘such’. In contrast to KR, the second/third-person forms of *pij-* ‘to drink’ are not contracted in Zala and Prekmurje, e.g. *pijel* ‘drinks’, cf. KR *píl*. Finally, the reduced forms *žom* ‘I told’ and *phom* ‘I said’ have developed only in Somogy Romani.

3.1.9 Apocope

A shared feature of Vend Romani (including KR) and other South Central varieties is the loss of word-final *s* and *n* (cf. Elšík *et al.* 1999: 297–300). In KR, the final *s* is omitted in the

adverbs *dí* < **dives* ‘day’ and *adí* < **adadives* ‘today’, in the nominative singular of xenoclitic nouns (e.g. *combo* < **combos* ‘thigh’, *daráži* < **darázis* ‘bee; wasp’), accusative singular of masculine nouns (e.g. *gáže* < **gážes* ‘non-Roma.OBL’), in the masculine third-person pronoun *le* < **les* and reflexive pronoun *pe* < **pes*, in the third-person singular preterite forms (e.g. *márd’a* < **márd’as* ‘beat.PRT.3SG’) and adverbs (e.g. *čorikán* < **čorikánes* ‘poorly’). By contrast, the word-final *s* has been preserved in *balvas* ‘bacon’, *mas* ‘meat’, *khas* ‘hay’, in the accusative form of the interrogative *ko* ‘who’ (i.e. *kas*), and in the present indicative forms of second-person singular and first-person plural (cf. *ibid.*), e.g. *máres* ‘beat.2SG’, *máras* ‘beat.1PL’.

The word-final *n* was deleted in *verda* < **verdan* < **vordon*^{B94} ‘cart, car’ and *hábe* < **xaben* ‘food’, in the derivational suffixes of abstract nouns (i.e. *-ibe* < **-iben*, *-ipe* < **-ipen*), and in the xenoclitic participial suffix (i.e. *-im* < **-ime* < **-imen*). It has been retained, on the other hand, in *ásvin* ‘tear’, *avd’in* ‘honey’, *trin* ‘three’, *kólin* ‘chest’, *pápin* ‘goose’, *pátrin* ‘leaf’, and in the derivations of the names of trees as in *phabalin* ‘apple tree’ or *kruškulin* ‘pear tree’ (cf. Elšík *et al.* 1999: 297–300). Furthermore, the loss of the final cluster *st* in *va* < **vast*^{B94} ‘hand’ and *gra* < **grast*^{B94} ‘horse’ is a shared feature of many South Central varieties including KR. This sound change has not affected the Slavic-borrowed *mist* (< **misto* < S *mesto*) ‘because of’, neither the inherited noun *bast* ‘luck’ nor *srasta* < **srast* ‘iron’. As pointed out by Elšík *et al.* (*ibid.*), this sound change took place before the loss of the uvular in **baxt* (> *bast*) and before the metathesis in **saster* (> **srast* > *srasta*).

In the flow of the speech, the borrowed conjunction *vad’* ~ *vaj* ‘or’ may occasionally be reduced to *v* when precedes a word with an initial vowel, e.g.:

(8)^{NAR} *na kopanáši sin v’ osó.*
 NEG Boyash COP.PRT.3 or such
 He was not a Boyash or something like that.

(9)^{NAR} *phendle hod’ i rák tut v’ adá v’ odá te hal.*
 tell.PRT.3PL COMP DEF cancer you.ACC or this.M or that.M COMP eat.INF
 They said that you may get cancer, or this or that.

Apocope in other varieties of Vend Romani

In some other varieties of Somogy, the auxiliary *šaj* ‘can’ is optionally reduced to *š* when the adjacent word begins with a vowel (cf. KR *vaj* ~ *v* ‘or’), e.g. *saki perc edej š’ól* ‘s/he can be here any minute’.

3.1.10 Other sound changes

A distinctive feature of the South Central Romani varieties (including KR) is the debuccalisation $s > h$ (Elšík *et al.* 1999: 300–301). In KR, the original intervocalic s was replaced by h in the instrumental singular (e.g. *leha* < **leša* ‘with him’), in the second person singular and the first person plural forms in future (e.g. *kereha* < **kereša* ‘you.SG will do’, *keraha* < **kerasa* ‘we will do’, but not in *hasa* ‘you.SG will eat’, see *ibid.*) and imperfect (e.g. *kerehahi* < **kerešasi* ‘you.SG were doing’, *kerahahi* < **kerasasi* ‘we were doing’, but not in *hasahi* ‘you.SG/we were eating’, see *ibid.*), in the imperfective suffix *-ahi* < **-asi* (see the previous examples), in the irrealis suffix *-áhi* < **-asi* (e.g. *kerd’omáhi* < **kerd’omasi* ‘I would have done’), and in the noun *táha* < **tasja*^{B94} ‘tomorrow’.

The original initial s has been preserved in KR such as in the interrogatives *so* ‘what’, *sar* ‘how’, *savo* ‘what (kind of)’, *soske* ‘why’ and in the determiners *sa* ‘all’ and *sako* ‘every’. The present third-person copula forms *hi* and *si* have been inherited into KR from Early Romani (see Matras 2002: 69).

Similarly to other South Central dialects (see Elšík *et al.* 1999), the voiceless velar fricative /x/ was replaced by /s/ in *bast* < **baxt*^{B94} ‘luck’ and /f/ in *ofto* < **oxto*^{B94} ‘eight’, the trill /r/ is lost in *cíd-* < **cird-*^{B94} ‘to pull’ and *phúd-* < **phurd-*^{B94} ‘to blow’, and the nasal /n/ is deleted in *máro* < **manRo*^{B94} ‘bread’, *mro* < **minRo* ‘my’ and *pro* < **pinRo*^{B94} ‘foot’. A prothetic /j/ is found in *d’áro* < **aRo*^{B94} ‘flour’, *d’iv* < **iv*^{B94} ‘snow’, *d’énd* < **ivend*^{B94} ‘winter’, *d’ékh* < **ekh*^{B94} ‘one’, *d’efkar* < **ekhvar*^{B94} ‘one’. The latter two forms are also permissible with the prothesis of /j/, i.e. *jékh*, *jefkar*. The prothetic /v/ is attested in the inherited *vóšt* (< **ošt*^{B94}) ‘lip’, and in the borrowed *vodro* (< Old Church Slavonic **odrŭ*; Elšík 2009: 270) ‘bed’ and *vusko* (< S *uzak*) ‘narrow’. The consonant /n/ is added to the end of the word in *karavd’in* < **karavdi*^{B94} ‘crab’ and *mát’hin* < **makhi*^{B94} ‘fly’.

In KR, the glottal fricative *h* tends to move after the nasal or liquid, as for instance in *kanhi* (< **kahni*) ‘chicken’, *bulho* (< **buhlo*^{B94}) ‘wide’ and *mulhi* (< **muhli*^{B94}) ‘fog’. On the other hand, the labial fricative *v* seems to be placed before the obstruent or liquid as it is found in *bávjal* (< **balval*^{B94}) ‘wind’, *jevkar* (< **jevkar* < **jekvar*) ‘wind’ and *évda* (S < **jedva*) ‘hardly’. The latter metathesised form could also have been directly borrowed from a local dialect of South Slavic. The inherited nouns **nilaj* ‘summer’ and **saster* ‘iron’ underwent long-distance metathesis to *linaj* and *srasta*, respectively.

Other sound changes in other varieties of Vend Romani

The prothetic /j/ in *d’ilo* < **ilo*^{B94} has been attested only in Vásárosdombó (Baranya) and Csokonyavisonta (Somogy). The older form *nilaj* ‘summer’ has been preserved in Vas and Zala as well as in some peripheral varieties of Somogy.

3.2 Vowels

3.2.1 Vowel inventory and graphemes

The vowel system of KR comprises fourteen vowel phonemes, seven of which are long (Table 12).

	Front		Back					
	unrounded		rounded		unrounded		rounded	
	short	long	short	long	short	long	short	long
Close	i	i:	y	y:			u	u:
Mid	ε	e:	ø	ø:			o	o:
Open					a	a:		

Table 12 Vowel inventory

The front rounded short and long vowel pairs are reserved for the recent Hungarian loanwords, such as *bölcő* < H *bölcső* ‘cradle’ and *čütörtök* < H *csütörtök* ‘thursday’ (see also 3.2.2).

The phoneme /a/ may become realized as [a] and [ɒ], while the phoneme /a:/ as [a:] and [a:]. The most common short variant is the back unrounded [a], found in pre-Hungarian words as well as in the morphologically adapted (i.e. older) loanwords from Hungarian, e.g. *barát-o* [bara:to] < H *barát* [bɒra:t] ‘friend’. The slightly rounded variant [ɒ], which agrees with the pronunciation of the phoneme in standard and colloquial Hungarian, is reserved for the recently borrowed Hungarian items. That is, for those Hungarian-origin words which are morphologically not adapted into KR, such as *balaton* [bɒlɒton] < H *Balaton* [bɒlɒton] ‘Lake Balaton’. The rounded variant is also often used in recently borrowed verbs, although these verbs are morphologically adapted. To give an example, the borrowed stem of the preterite form *takarittind'am* [tɒkɒrittind'am] ‘we saved’ (< H *megtakaritt* [mɛgtɒkɒritt] ‘to save money’) is pronounced according to the Hungarian pattern, while the vowel of the perfective marker *-am* follows the general Romani pattern of pronunciation. Similarly to the short /a/, the allophones of the long phoneme /a:/ are distributed complementarily: The sound [a:] is generally realized in native words, while the sound [a:] occurs in recent loanwords, e.g. *d'áro* [ɟa:ro] ‘flour’, *hijába* [hija:bɒ] < H *hijába* [hi⁽ⁱ⁾a:bɒ] ‘in vain’. The distribution of the two allophones is not as straightforward as it has been described, since the pronunciation of recently borrowed items may vary from speaker to speaker or even in the speech of a single person. In addition it seems that the complementary distribution of the long allophones [a: a:] is less striking than that of the short allophones [a ɒ] since there is a tendency of /a:/ to expand to the recently borrowed items as well.

It is interesting to point out that the quality of the short-long phoneme pair /a a:/ realized in inherited and older loanwords is roughly reversed as compared to its realization /ɒ a:/ in the recent loanwords. This may be explained by the recent phenomenon found in the local Hungarian dialect, where the Hungarian dialectal pronunciation is gradually retreating in favour of the more prestigious, colloquial Hungarian, pronunciation. This development is found also in other Hungarian dialects (Imre 1972: 93). Thus, KR conserved the Hungarian dialectal pronunciation in the inherited lexicon and older loanwords, while the newly borrowed items reflect the colloquial Hungarian pronunciation, which has become popular among the local Hungarian speakers.

The vowel phoneme /ɛ/ has three variants. These are the mid [e], open-mid [ɛ] and the near-open [æ]. The latter two variants are in free variation, while being in complementary distribution with the first variant, e.g. *perse* [pæɛse] ~ [pɛɛse] < H dial. *persze* [pæɛse] ~

[pɛrʃe], cf. standard H *persze* [pɛrʃɛ] ‘of course’. The distinction between the open-mid [ɛ] (~ [æ]) and mid [e] sounds is taken over from the Hungarian contact dialect, where the two sounds are treated as individual phonemes (e.g. Kenesei *et al.* 1998: 385). Both sounds may occur in borrowed as well as in inherited words, e.g. *edej* [ɛdej] ~ [ædej] ‘here’, *de* [dɛ] ~ [dæ] < H *de* ‘but’. The distribution of these allophones in the Hungarian lexicon is for the most part unpredictable, as it is related to the origin of the word. However, in the pre-Hungarian lexicon there is a tendency for the mid variant to occur in the final position, while the open-mid variant is found elsewhere. The open front variant [æ] is also present in the Hungarian dialect that KR is in contact with (Király 2005: 35). In KR, this sound tends to occur in stressed syllables, such as in the first syllable in the borrowed *vesprim* [væsprim] ~ [vɛsprim] < H *Veszprém* (town in Hungary). In addition, it also occurs in the diphthong *aj* [æɪ], which is found in several German loanwords, as for instance in *rajn* [ræɪn] (< G dial. [ræ:ɪn]) ‘clean’ or *cajt* [tsæɪt] (< G dial. [tsæ:ɪt]) ‘time’. As it may be observed on the two latter examples, the KR pronunciation of the diphthong agrees with its pronunciation in the German dialect spoken in Eastern Austria (see Vollmann & Moosmüller 2001). The diphthong *aj* may also be realized by the more back [aɪ], e.g. *rajn* [raɪn] ‘clean’.

The vowel phonemes of KR are represented by the graphemes listed in Table 13.

Grapheme	<i>a</i>	<i>á</i>	<i>e</i>	<i>é</i>	<i>i</i>	<i>í</i>	<i>o</i>	<i>ó</i>	<i>ö</i>	<i>ő</i>	<i>u</i>	<i>ú</i>	<i>ü</i>	<i>ű</i>
IPA	/a/	/a:/	/ɛ/	/ɛ:/	/i/	/i:/	/o/	/o:/	/ø/	/ø:/	/u/	/u:/	/y/	/y:/

Table 13 Grapheme system

The long vowels <á é í ó ú> are marked with a single acute, the Hungarian-borrowed long vowels <ő ü> with a double acute. The front rounded short vowels are written with an umlaut, i.e. <ö ü>.

As regards the position of vowels, all fourteen vowel phonemes are allowed in medial position, while only thirteen have been found in initial or final position Table 14. The long /y:/ seems to be excluded in the initial position, and the short /ø/ in final position. As it has been mentioned above, front rounded vowels appear only in Hungarian words. In addition, the final /u:/ is also allowed just in Hungarian loanwords (e.g. *búčú* < H *búcsú* ‘saint’s day’), while the final /e:/ is reserved for German and Hungarian loanwords (e.g. *té* < G *Tee* ‘tea’, *tévé* < H *tévé* ‘TV’). The phonemes *ú* and *ű* are rare in the final position, as they often become shortened

already in the Hungarian contact dialect, e.g. *faťú* ~ *faťu* ‘boy’ < H dial. *fattyu*, cf. H *fattyú* ‘bastard’.

	MEDIAL	INITIAL	FINAL
/a/	<i>kařuko</i> ‘deaf’	<i>akán</i> ‘now’	<i>afka</i> ‘such’
/a:/	<i>zorálo</i> ‘strong’	<i>ánav</i> ‘name’	<i>vakerá</i> ‘I will speak’
/ɛ/	<i>zeleno</i> ‘green’	<i>eňa</i> ‘nine’	<i>me</i> ‘me’
/e:/	<i>věř</i> ‘forest’	<i>éfta</i> ‘seven’	<i>khafé</i> ‘coffee’
/i/	<i>cipa</i> ‘skin’	<i>irín-</i> ‘to write’	<i>ári</i> ‘out’
/i:/	<i>sír</i> ‘garlic’	<i>ič</i> ‘yesterday’	<i>dí</i> ‘day’
/o/	<i>vodro</i> ‘bed’	<i>odá</i> ‘that’	<i>ko</i> ‘what’
/o:/	<i>vóra</i> ‘yard’	<i>ój</i> ‘she’	<i>asó</i> ‘such.M’
/ø/	<i>kölčön</i> ‘loan’	<i>öltöň</i> ‘suit’	–
/ø:/	<i>sőňeg</i> ‘carpet’	<i>ős</i> ‘autumn’	<i>égő</i> ‘bulb’
/u/	<i>žuvli</i> ‘woman’	<i>upral</i> ‘above’	<i>papu</i> ‘grandfather’
/u:/	<i>žúkel</i> ‘dog’	<i>účo</i> ‘tall’	<i>faťú</i> ‘child, son’
/y/	<i>büske</i> ‘proud’	<i>üd’ešn-o</i> ‘skillful’	<i>ed’bü</i> ‘immediately’
/y:/	<i>hűtő</i> ‘fridge’	–	<i>körű</i> ‘around’

Table 14 Vowel phonemes in medial, initial and final position

Vowel inventory and graphemes in other varieties of Vend Romani

Pronunciation of vowels in individual Vend Romani varieties seems to coincide with their pronunciation in the respective local Hungarian dialects. The front rounded vowels may also be found in German loanwords in the western varieties of Hungarian Vend Romani, such as in *cvüllen* (< G dial. [tʃvʏlɪŋ]) ‘twins’ and *sőma* (< G dial. [sœ(ɪ)vœ]) ‘from itself’ in Vas Romani, or in *šnő* (< G dial. [ʃnœ(ɪ)]) ‘fast, quickly’ and *štüj* (< G dial. [ʃtʏ(ɪ)]) ‘quietly’ in Sopron Romani. The near-open [æ] represents a distributional variant of /e/ in Zala Romani, being optionally realized in the diphthong *ej*, e.g. *čhej* [tʃæɪ] ‘girl, daughter’, *dej* [dæɪ] ‘mother’ or *šej* [ʃæɪ] ‘can’.

3.2.2 Vowel adaptation

The articulation of KR vowels agrees with that in its Hungarian contact dialect (i.e. Southern Transdanubian), which facilitates the process of loanword adaptation. This means that most Hungarian sounds are transferred without being adapted into the sound system of KR. By contrast, the Hungarian front labialized vowels are generally adapted by delabialization, which is characteristic also for other South Central varieties in contact with Hungarian (Elšík *et al.* 1999: 309):

ö > *e*: H *kökény* > KR *kekěň-i* ‘blackthorn’

ő > *é*: H *csődör* > KR *čéder-i* ‘stallion’

ü > *i*: H *ügyes* > KR *id’ešn-o* ‘skillful’

ű > *í*: H *hűs* > KR *hiš-o* ‘shadow’

The sound change *ü* > *i* and *ű* > *í* is also found in Somogy Hungarian which is the contact dialect of KR (Király 2005: 36). This may imply that it was already the delabialized, Hungarian dialectal, form which was adapted into KR, e.g. H dial. *szirke* (cf. H *szürke*) > KR *sirk-asto* ‘grey’, H dial. *míanyag* (cf. H *műanyag*) > KR *míaňag-ošno* ‘plastic’. The German front labialized vowels seem to have been adapted by the same process, as it is shown by the KR loanword *él-o* (< G *Öl* [ø:l]) ‘oil’.

Several characteristic features of the Southern Transdanubian dialect of Hungarian had a varying impact on KR. For instance, the sound change *o* > *u* affected the Hungarian derivational marker of inchoatives (i.e. *-ol* > *-ul*) as well as several lexical items of the local Hungarian dialect (e.g. *ibid*: 35; Imre 1971: 24), and so it was brought to KR via borrowings, e.g. *špór-ul-ín-* < H dial. *spór-ul* (cf. H *spór-ol*) ‘to save money’, *gond-ul-ín-* < H dial. *gond-ul* (cf. H *gond-ol*) ‘to think’, *ňum-ín-* < < H dial. *nyum* (cf. H *nyom*) ‘push’, *baluk-n-o* < H dial. *balug* (cf. H archaic *balog*) ‘left’, *čurg-ín-* < H dial. *csurog* (cf. H *csorog*) ‘stream’. The change from *o* to *u* is also found in the Slavic loanwords *gulubica* (< S *golobica/golubica*) ‘dove’ and *puruč-ín-* (< S *poruč-iti*) ‘order’, and in the inherited perfective first-person singular marker *-um* (< *-om). However, these changes most probably happened before the Hungarian contact. Note that a similar sound change is attested in the Romani dialects in current contact with Slovenian, such as in Prekmurje Romani as well as further to the south, in the non-related dialect of Dolenjska Romani (Cech 2006: 2). The sound *o* changes to *u* before

r and *l* also in some German dialects of Austria (Wiesinger 1967). Thus, this sound change may be considered a shared areal feature of the wider region.

Similarly to the Southern Transdanubian dialect of Hungarian (Király 2005: 35), the sound *a* is often replaced by *o* when preceded by a syllable containing long *á*, such as in *lázošn-o* (< H dial. *lázos*, cf. H *lázás*) ‘fevered’ and *láboš-kiňa* (< H dial. *lábos*, cf. H *lábás*) ‘pot’. The elision of the syllable-final *l* causes compensatory lengthening, e.g. *legétet-ín-* (< H dial. *legétet*, cf. H *legeltet*) ‘to graze’. The occurrence of *ö* in Hungarian loanwords has been increased by the sound change *e* > *ö* which underwent in Southern Transdanubian Hungarian (ibid. 28), e.g. *böčületešen* (< H dial. *böcsületesen*, cf. H *becsületesen*) ‘honestly’. Long closed vowels are often pronounced short in KR, so are they in the Hungarian dialect of Central Somogy (ibid: 26), e.g. *alaminijum* (< H dial. *alámínium*, cf. H *aluminium*) ‘aluminum’, *šürü* (< H dial. *sürü*, cf. H *sűrű*) ‘thick’, *bučuz-ín-* (< H dial. *bucsuz-ik*, cf. H *búcsúz-ik*) ‘to say goodbye’. The vowel *é* was replaced by *i* only in some Hungarian loanwords, such as in *nípo* (< H dial. *níp*, cf. H *nép*) ‘people’ and *kípo* (< H dial. *kíp*, cf. H *képo*) ‘people’. This sound change is, however, absent in the Hungarian dialects of Central Somogy where KR is spoken (ibid: 35). This means that KR has fossilized the older, Hungarian dialectal pronunciation of these words.

Vowel adaptation in other varieties of Vend Romani

The sound change *o* > *u* in the first person singular perfective marker is found in the majority of Somogy Romani varieties as well as in Prekmurje Romani. It is, on the other hand, absent in Zala, Vas, Veszprém and Burgenland Romani.

3.2.3 Sequences of vowels

Two consecutive vowels are rare in KR. It may occur in Hungarian loanwords (e.g. *kakaó* < H *kakaó* ‘chocolate milk’, *január* < H *január* ‘January’, *téesi* < H *téesz* ‘collective farm’, *eccerüen* < H *egyszerűen* ‘simply’) and between the constituent elements of compounds (e.g. *priko-íč* ‘before yesterday’, *elő-írás* < H *elő-írás* ‘regulation’). The VV-cluster is exceptionally pronounced due to the drop of the intervocalic *v*, such as in *asao* (~ *asavo* ~ *asó*) ‘such’ and

hegedua (~ *hegeduva*) ‘violin’. Two identical adjacent vowels are found in the inherited negation particle *naa* ‘no(pe)’ in emphatic use:

- (10)^{NAR} *na, phénav: and'al o šélo? phend'a: naa.*
 well say.1SG bring.PRT.2SG DEF rope say.PRT.3SG nope
 Well, I asked him: Have you brought the rope? He said: Nope.

Similarly to Hungarian (Kenesei *et al.* 1998: 413), a /j/ is often inserted between the vowel sequences comprising *i* or *í* to prevent hiatus, e.g. *jú.li.juš* < H *július* /júli^(j)uʃ/ ‘July’, *ka.mi.jon* < H *kamion* /kɔmi^(j)on/ ‘camion’, *in.di.ja* < H *India* /indi^(j)ɔ/ ‘India’.

Sequences of vowels in other varieties of Vend Romani

The glide *j* is occasionally audible between two neighbouring vowels at the word boundary in some Somogy Romani varieties, such as in Zimány:

- (11)^{LQCR} *na ʃ ali khér.*
 NEG come.PRT.3SG.F home
 She has not come home.

3.2.4 Diphthongs

The diphthongs of KR are falling diphthongs, being mostly composed of a vowel in combination with the glide *j*. These are the *aj*, *áj*, *ej*, *oj*, *ój* and *uj*. They are found in indigenous (a) and borrowed stems (b), as well as in various inflectional and derivational forms (c). The diphthongs *aj*, *oj* and *uj* are outstanding in the number of occurrences, since they mostly result from the contraction *Vve > Vj* in the second person (see 3.1.8).

a.	b.	c.
<i>aj kir.haj</i> ‘boots’	<i>cajt</i> (< G) ‘time’	<i>ajs</i> ‘you come’
<i>áj káj</i> ‘where’	<i>bo.čaj.tín-</i> (< H) ‘to forgive’	<i>dáj</i> ‘mother.VOC’
<i>ej e.dej</i> ‘here’	<i>nej.lon</i> (< H) ‘nylon’	–
<i>oj o.doj</i> ‘there’	<i>ši.poj.ka</i> (< H) ‘flute’	<i>sojs</i> ‘you sleep’

a.	b.	c.
<i>ój</i> <i>šój</i> ‘whistle’	–	<i>ójs</i> ‘you become’
<i>uj</i> <i>muj</i> ‘mouth’	<i>šuj.bín-</i> (< G) ‘to push’	<i>duj.to</i> ‘second’

The German diphthongs were adapted as long monophthongs into KR, e.g. *fút* < G dial. [fuɔt] ‘away’, *šír* < G dial. [g̊ʃiɔ] ‘dish’, *móm* < G dial. [muɔm] ‘aunt’, *klát* < G dial. [kʰladl] ‘clothes’, *práni* < G braun ‘brown’, *háfní* < G Haufen ‘heap’. The monophthongization of German diphthongs most probably proceeded under the Hungarian influence, since Hungarian originally does not possess diphthongs. The vowel sequence *au* of Hungarian loanwords – which are also loans in Hungarian – is pronounced either with /aw/ or simply with a long /a:/ in KR, e.g. *autómata* [awto:mata] ~ *átómata* [a:to:mata] < H *automata* [ɒwtomɒtɒ] ‘automata’.

Diphthongs in other varieties of Vend Romani

In Zala Romani, the syllable and word-final sequence *Vv* is articulated with the respective vowel in combination with the glide /w/, e.g. *žav* [zaw] ‘I go’, *suv.d’a.ha* [suwɔaha] ‘with needle’, *hev.d’a.lo* [hewɔa:lo] ‘leaky’, *živ.la.hi* [ziwlahi] ‘s/he lived’. Furthermore, the falling diphthong /ua/ is attested in some recently borrowed German loanwords of Sopron Romani, such as in e.g. *fuat* [fuat] < G dial. [fuɔt] ‘away’, *muam* [muɔm] < G *Muhme*, G dial. [muɔm] ‘aunt’.

3.2.5 Vowel length⁴³

3.2.5.1 Vowel quantity and processes of vowel lengthening

The introduction of vowel length into KR was most probably triggered by prolonged contact with Hungarian, a language that has length opposition. As a result, Vend Romani has ten independent vowel units, five short (a) and five long (b), in addition to the front-rounded short (c) and long (d) vowel pairs that are borrowed from Hungarian:

⁴³ This chapter is a shorter version of the author’s manuscript (Bodnárová & Wiedner 2015b; submitted for publication).

- a. a /a/, e /ɛ/, i /i/, o /o/, u /u/
- b. á /a:/, é /e:/, í /i:/, ó /o:/, ú /u:/
- c. ö /ø/, ü /y/
- d. ő /ø:/, ű /y:/

Vowel length is a distinctive phonological feature in KR that has been established based on minimal pairs such as the following:

ov ‘be.IMP.2SG’ vs. *óv* ‘he’

sapano ‘wet’ vs. *sápano* ‘snake’s, of snake’

phral ‘brother’ vs. *phrál,* ‘s/he opens’

ásal ‘s/he laughs’ vs. *asál* ‘s/he makes so. laugh’

khul ‘excrement’ vs. *khúl* ‘s/he weaves’

so ‘what’ vs. *só* ‘what kind of’

urd’al ‘fly’ vs. *urd’ál* ‘make so. dress’

daj ‘mother’ vs. *dáj* ‘mother.VOC’

šudro ‘cold’ vs. *šudró* (< H dial. *sudró*) ‘rolling pin’

meg (< H *meg*, verbal particle denoting perfective aspect) vs. *még* (< H *még*) ‘still, yet’

Long vowels may occur in any position in a word (e.g. *ícutno* ‘yesterday’s’, *adi* ‘today’, *paramísi* ‘story, tale’), and a word may contain more than one long vowel (e.g. *pékíbe* ‘cake’, *khírínel* ‘s/he shouts’, *lákjáhi* ‘s/he would have found’). The distribution of long vowels is not bound to the position of stress (see 3.3). The fact that long vowels and stress are independent of one another is exemplified in Table 15 where stress is marked in bold.

#SYLLABLES	LONG VOWEL	
	STRESSED SYLLABLE	UNSTRESSED SYLLABLE
1	<i>bár</i> ‘garden’	-
2	<i>mánuš</i> ‘human’	<i>márá</i> ‘I will beat’
3	<i>dandérel</i> ‘s/he bites’	<i>lákjáhi</i> ‘s/he would have found’

Table 15 Long vowel in stressed and unstressed syllables

In KR, as well as in Romani in general, long vowels have been introduced through contraction and compensatory lengthening (see Matras 2002: 60), e.g. *žá* < **žava* ‘I will go’, cf. *žav* ‘I go’; *dí* < **dives* ‘day’. Vowels have also become lengthened in open syllables in pre-tonic position (ibid.), e.g. CVCV **ka.lo* > **ká.lo* ‘black’, cf. **šud.ro* ‘black’. This development represents an earlier stage of the language when the stress fell on the final syllable. More recently, the stress has shifted to the penultimate syllable in Vend Romani, resulting in the coincidence of long vowels with stressed open syllables, e.g. CVCV *tá.ha* ‘tomorrow’ (cf. CVCCV *tík.no* ‘small’), CVCVC *má.nuš* ‘human’ (cf. CVCCVC *čum.ňik* ‘whip’), CVCVCV *mo.mé.li* ‘candle’ (cf. CVCVCCV *či.rík.li* ‘bird’). Vowel length is thus to a large extent predictable in case of the inherited lexicon. Hungarian loan words, on the other hand, are always adapted together with their vowel length, e.g. *té.vé* < Hung. *tévé* ‘television’, *pu.ló.ve.ri* < Hung. *pulóver* ‘pullover’, *pa.lo.ta* < Hung. *palota* ‘palace’. In these loans, the initial syllable is stressed.

Elšík *et al.* (1999: 311) report that vowel length became morphologically relevant in South Central Romani, where the inherited adjectives and the polysyllabic possessive pronouns become shortened when used attributively, e.g. *amaro verda* ‘our car’, cf. predicative adjective use *o verda amáro hi* ‘the car is ours’. They (ibid: 312) also noticed that the vowel length of the base form is generally preserved in derivations as well as throughout the inflectional paradigm of the word, e.g. *pé.kav* ‘I bake’ vs. *pék.řum* ‘I baked’, but not in *phé.nav* ‘I say’ vs. *phen.dřum* ‘I said’, etc. However, a thorough analysis is needed to determine in which environment the vowels become lexicalized.

Analogical extension may also account for the introduction of vowel length into certain KR words and word forms. This process is understood here to be a type of a linguistic change involving a less common form that extended and therefore has become the more common form (see Hock 1986: 238–279; Anttila 2003; Blevins & Blevins 2009: 1–12). I will demonstrate this particular change in the sections 3.2.5.3 and 3.2.5.4 by examining certain personal pronouns and a part of the inflectional verb paradigm. I will also attempt to identify the constraints of vowel lengthening by considering the phonological, morphological and semantic contexts in which this pattern extension arises.

3.2.5.2 Length distribution

The most common pattern of length distribution in the inherited lexicon is that the long vowel occurs in the penultimate open syllable of polysyllabic words (Table 16).

	DISYLLABIC	TRISYLLABIC
noun	<i>á.nav</i> ‘name’	<i>mo.mé.li</i> ‘candle’
adjective	<i>ló.ko</i> ‘slow’	<i>nas.vá.lo</i> ‘sick’
verb	<i>va.ké.r-</i> ‘to speak’	<i>pa.ri.ké.r-</i> ‘to thank’
adverb	<i>tá.ha</i> ‘tomorrow’	<i>kor.kó.ro</i> ‘alone’

Table 16 Pattern of length distribution

Exceptions from this rule are:

a) the polysyllabic lexical items where the penultimate syllable is closed but contains a long vowel, such as the nouns *ár.mi* ‘cabbage’, *ás.vin* ‘tear’, *bák.ro* ‘lamb, sheep’, *báš.no* ‘cock, rooster’, *báv.jal* ‘wind’, *pát.rin* ‘leaf’, *rák.lo* ‘non-Romani boy’, *sást.ro* ‘father-in-law’, *sós.ten* ‘underpants’, *pa.šáv.ro* ‘rib’; the adjectives *bán.go* ‘crooked, curved’, *hár.no* ‘short’, *kóv.lo* ‘soft’, *pár.no* ‘white’, *sás.to* ‘healthy’, *tér.no* ‘young’; the middle verbs *páš.tov-* ‘to lie’, *hán.d’ov-* ‘to itch, tickle’, *tér.d’ov-* ‘to stand’, *bár.d’ov-* ‘to grow’; the local adverbs *án.de* ‘inside’, *án.gjal* ‘in front’, *án.gle* ‘forward’, *áv.ral* ‘outside’, *ór.de* ‘hither’; or the numerals *éf.ta* ‘seven’ and *óf.to* ‘eight’.

b) the polysyllabic lexical items where the penultimate syllable is open but contains a short vowel, such as the nouns *bi.jav* ‘wedding’, *ma.sek* ‘month’, *du.muk* ‘fist’, *zi.jand* ‘pity’, *ba.ba* ‘grandmother’, *ci.pa* ‘skin; leather’, *du.mo* ‘back’, *ko.va* ‘thingummy’, *pa.pu* ‘grandfather’, *zu.mi* ‘soup; juice’, *ba.li.čho* ‘piglet’, *ho.ko.no* ‘axe’, *ki.ri.vo* ‘godfather’, *ko.ka.lo* ‘bone’, *ko.pa.na* ‘trough’, *pa.raš.tu.va* ‘Friday’, *po.si.ta* ‘pocket’, *ra.ta.ha* ‘morning’, *sa.pu.ňi* ‘soap’, *si.la.vo* ‘pincers’, *si.ri.mi* ‘belt’; the adjectives *ci.lo* ‘whole, all’, *čo.ro* ‘poor; deceased’, *ko.ro* ‘blind’, *di.li.no* ‘stupid’, *ka.šu.ko* ‘deaf’; the middle verbs *ha.jov-* ‘to understand’ and *na.thov-fer* ‘to pass, elapse’; the adverbs *e.dej* ‘here’, *ki.ti* ‘how much/many’, *kha.tar* ‘from where’, *o.doj* ‘there’, *ra.ta.ha* ‘in the morning’; or the numerals *a.ti* ‘this/that much’, *ču.lo* ‘few, little’ and *e.ňa* ‘nine’.

As for the monosyllabic lexicon, the vowel became long in the nouns *bár* M ‘stone’, *bár* F ‘garden; fence’, *búr* ‘bush’, *čár* ‘grass’, *čór* ‘thief’, *dár* ‘fear’, *hév* ‘hole’, *íč* ‘yesterday’, *kár* ‘penis’, *láč* ‘shame’, *pér* ‘belly; stomach’, *pór* ‘feather’, *sír* ‘onion’, *šój* ‘whistle’, *věš* ‘forest’, *zár* ‘hair’, *zór* ‘strength, power’ and *vóšt* ‘lip’; in the adverbs *dúr* ‘long, far’, *íč* ‘yesterday’ and *káj* ‘where, to where’; in the numerals *jékh* ‘one’, *pándž* ‘five’, *šóv* ‘six’, *štár* ‘four’ and optionally in *děš* (~ *deš*) ‘ten’. The monosyllabic verbs which keep the long vowel in the various inflected or derivational forms are *áčh-* ‘to dwell; stay’, *běš-* ‘to sit’, *čár-* ‘to lick’, *čór-* ‘to steal, rob’, *gén-* ‘to read; count’, *hán-* ‘to dig’, *khós-* ‘to wipe’, *kúr-* ‘to have sex’, *kúš-* ‘to peel’, *lád-* ‘to drive’, *lákh-* ‘to find’, *mákh-* ‘to oil, grease’, *már-* ‘to beat’, *náš-* ‘to run’, *pék-* ‘to bake’, *píš-* ‘to milk’, *rés-* ‘to fit; reach’, *rúš-* ‘to be angry’ and *trád-* ‘to drive’.

Due to the apocope of the final vowel, the long vowel is sometimes located in the last syllable, e.g. *thán* < **tháne* ‘in place’, *khér* < **khére* ‘at home’.

3.2.5.3 Analogical extension of vowel length in personal pronouns

Table 17 illustrates the genitive forms of personal pronouns that occur in KR, where we can distinguish between the monosyllabic (*mr-*, *tr-*, *pr-*) and the polysyllabic pronouns (*les-kr-* ~ *les-ker-*, *la-kr-* ~ *la-ker-*, *amar-*, *tumar-*, *pumar-*, *len-gr-* ~ *len-ger-*).

1SG	<i>mr-</i>	1PL	<i>amar-</i>
2SG	<i>tr-</i>	2PL	<i>tumar-</i>
3SG.REFL	<i>pr-</i>	3PL.REFL	<i>pumar-</i>
3SG.M	<i>les-kr-</i> ~	3PL	<i>len-gr-</i> ~
3SG.F	<i>les-ker-</i>		<i>len-ger-</i>
	<i>la-kr-</i> ~		
	<i>la-ker-</i>		

Table 17 Genitive forms of personal pronouns

These pronouns take the adjectival endings that are represented by the short vowels M.SG *-o*, F.SG *-i* and PL *-e* in the nominative (a), and M.SG *-e*, F.SG *-a* and PL *-e* in the oblique (b).

a.
mr-o murš ‘my-M.SG husband’
mr-i čhaj ‘my-F.SG daughter’
mr-e khéra ‘my-PL houses’

b.
mr-e muršes- ‘my-OBL.M.SG husband.OBL’
mr-a čha- ‘my-OBL.F.SG daughter.OBL’
mr-e kheren- ‘my-OBL.PL houses.OBL’

In KR most of the singular gender markers of monosyllabic possessive pronouns have developed a long vowel counterpart, which is only employed when the head noun is *daj* ‘mother’ or *dad* ‘father’, e.g. *mr-í daj* (< **mr-i daj*) ‘my-F.SG mother’, *mr-ó dad* (< **mr-o dad*) ‘my.M.SG father’. The lengthening of vowels in the monosyllabic pronouns *mr-V* ‘my’, *tr-V* ‘your’ *pr-V* ‘his/her own’ can most likely be traced to the commonly used phrases of ‘my/your/his/her own mother’ and ‘my/your/his/her own father’, which have become compounds in KR, e.g. **mri+daj* < **mri daj*, **mro+dad* < **mro dad*. This development is particularly interesting because compounding is not a productive means of word-formation neither in KR nor in Romani in general (Matras 2002, 119).

In the newly emerged compounds, the gender markers of pronouns have become located in open pre-tonic syllables and have therefore undergone lengthening, e.g. **mri.daj* < **mri daj*, **mró.dad* < **mro dad*. Following the shift in stress patterns, long vowels are now found in stressed syllables, e.g. *mrí.daj* < **mri.daj*, *mró.dad* < **mró.dad*. While it still inflects for gender, the pronoun constituent of the compound is strongly bound to the head noun, e.g. *mr-í+daj* ‘my-F.SG+mother’ (long vowel), but cf. *mr-i moštóvni daj* ‘my-F.SG step-mother’ (short vowel).

Subsequently, the occurrence of long vowels has become analogically extended to the feminine oblique forms of the monosyllabic pronouns when followed by the nouns *daj* ‘mother’ and *dad* ‘father’ (Table 18).⁴⁴ In the oblique, the stress falls on the oblique suffix of the noun instead of the pronoun, i.e. *d-a-* < **daj-a-* ‘mother-OBL.F.SG-’, *dad-es-* ‘father-OBL.M.SG-’. It has been mentioned in section 3.2.5.1 that vowel length is generally maintained in the inflectional forms of words. According to this, one would expect that the feminine oblique form of the compound *mrídaj* ‘my mother’ will be *mrida-*. This is, however,

⁴⁴ This does not hold for the irregular vocative form *mri dáj*, ‘my.F.SG.NOM mother.VOC’, which consists of the nominative form of the pronoun and the apocopated form of the noun (*dáj* < **dáj-e*, ‘mother.VOC’, cf. *daj*, ‘mother’).

not the case in KR where the pronoun constituent takes the feminine oblique marker *-a*, which becomes lengthened analogically to the nominative forms, e.g. *mr-á da-tar* ‘my.OBL.F.SG mother-ABL; from my mother’, cf. *mr-a čha-tar* ‘my.OBL.F.SG daughter-ABL; from my daughter’. Note that the pronoun element of the compound is directly followed by the head noun *daj* ‘mother’ in cases other than nominative, e.g. long vowel in *mrá+da-ha* ‘my.OBL.F.SG+mother.OBL.F.SG-INS; with my mother’, but short vowel in *mra moštóvna da-ha* ‘my.OBL.F.SG step-mother.OBL.F.SG-INS; with my step-mother’.

	‘my ~ your ~ his/her own mother’	‘my ~ your ~ his/her own father’
NOM	<i>mr-í+daj ~ tr-í+daj ~ pr-í+daj</i>	<i>mr-ó+dad ~ tr-ó+dad ~ pr-ó+dad</i>
OBL	<i>mr-á+da- ~ tr-á+da- ~ pr-á+da-</i>	<i>mr-e+dades- ~ tr-e+dades- ~ pr-e+dades-</i>

Table 18 Innovative vowel lengthening in the monosyllabic personal pronouns

The development of long vowels in the monosyllabic pronouns of Vend Romani is illustrated in Table 19. First, the nominative nouns *daj* ‘mother’ and *dad* ‘father’ have become compounded with the pronouns *mr-* ‘my’, *tr-* ‘your’ and *pr-* ‘his/her own’. Second, the occurrence of long vowels has analogically extended to the feminine oblique forms of these pronouns.

	-MOTHER	-FATHER
	↓	↓
Compounding	my.NOM your.NOM his/her_own.NOM	my.NOM your.NOM his/her_own.NOM
Analogical extension	my.OBL you.OBL his/her_own.OBL	–

Table 19 Development of vowel length in personal pronouns

This analogical change in vowel length seems to have been driven by certain phonological and semantic constraints: The former constraint may account for the absence of a long vowel in the masculine pronoun, e.g. *mr-e*, cf. **mr-é*. The KR data suggest that the sound /é/ occurs

significantly less frequently in the word-final position than the sounds /á, ó, í/ (e.g. *mr-á* ‘my-F.SG.OBL’, *mr-ó* ‘my-M.SG’, *mr-í* ‘my-F.SG’). Moreover, the final /é/ appears exclusively in the German and Hungarian borrowed nouns (e.g. *khafé* < G *Kaffee* ‘coffee’, *tévé* < Hungarian *tévé* ‘television’), while the final /á, ó, í/ sounds are also, or exclusively, encountered in inherited words. According to this, it is highly unlikely that the sound /é/ would occupy the final position, especially in an inherited word, such as the masculine oblique pronoun. The number of syllables seems to also play a decisive role in the introduction of long vowels into personal pronouns. It has been shown that only monosyllabic pronouns are influenced by the innovation, while those with more than one syllable have remained unchanged, e.g. *a.ma.ra da-*, not *a.ma.rá da-* ‘our mother.OBL’. Finally, the introduction of vowel length seems to be semantically constrained by the lexical field comprising the kinship terms ‘mother’ and ‘father’. The fact that only these nouns became compounded with the personal pronouns may be due to a general tendency towards analogical change in the names denoting kinship. Winter (1969: 39ff.), for instance, compares several Indo-European languages and draws the conclusion that analogical change often takes place in the semantically closely related sets of kinship terms, such as mother : father. He provides an example of this type of change found in Tocharian, where the root vowel in *mācer*, *mācar* ‘mother’, is analogously taken over from *pācer*, *pācar* ‘father’. Winter cites another example from Old Church Slavonic, where the accusative form of ‘mother’ was replaced by the genitive form that is in agreement with the inflectional pattern of ‘father’, that is, *otīca* ‘father.GEN/ACC’ \Rightarrow *matere* ‘mother.GEN/ACC’ < *’mother.ACC’. In the northern varieties of South Central Romani, we find similar analogical change in the kinship pair brother : sister. For instance, instead of the inherited terms *phral* ‘brother’ and *phen* ‘sister’, the borrowed *ečč-o* (< Hung. *öcs* ‘younger brother’) ‘brother’, and its feminine derivation, *eč-kiňa* ‘sister’, is used in Dunajská Streda Romani (Slovakia), *tāštív-r-o* (< Hung. *testvér* ‘sibling’) ‘brother’ and its feminine derivation *tāštív-r-kiňa* ‘sister’ in Vlčany Romani (Slovakia), or *endáň-i* (< *endáňi* ‘kin’) ‘brother’ and its feminine derivation *endán-kiňa* ‘sister’ in Mátraverebély Romani (Hungary). Thus, the formal or structural similarity of the kinship terms developed through analogy appears to be rather common, at least in the Indo-European languages.

3.2.5.4 Analogical extension of vowel length in the imperfective suffix

One of the main isoglosses that distinguish the South and North Central dialect groups of Romani is the imperfective suffix, which generally has the form *-ahi* in the former (South) and *-as* in the latter (e.g. Elšík *et al.* 1999: 351). In most varieties of South Central Romani the suffix *-ahi* is realized with a short vowel. In some other varieties, the first vowel of the suffix is either long (*-áhi* or *-áj* in the Žitný ostrov region, Slovakia), or the short and long variants are in free variation (*-ahi* ~ *-áhi* in Pest county, Hungary). KR is exceptional in the sense that the long- and short-vowel forms of the suffix have become functionally distinct. For example, the suffix *-ahi* is applied in the imperfect and in the formally analogous conditional (e.g. *kereh-ahi* ‘you were doing; you would do’), while the marker *-áhi* is used to form irrealis conditional (e.g. *kerd'al-áhi* ‘you would have done’). Compare the imperfective and irrealis forms that are demonstrated by the verb *már-* ‘to beat’ in Table 20.

	IMPERFECT	IRREALIS
1SG	<i>már(-)áhi</i>	<i>márd^ʔ-um-áhi</i>
2SG	<i>már-eh-ahi</i>	<i>márd^ʔ-al-áhi</i>
3SG	<i>már-l-ahi</i>	<i>márd^ʔ(-)-áhi</i> ↓
1PL	<i>már-ah-ahi</i>	<i>márd^ʔ-am-áhi</i>
2/3PL	<i>már-n-ahi</i>	<i>márd^ʔ-en-áhi</i>

Table 20 Inflectional markers of imperfect and irrealis

The imperfect is formed by attaching the personal concord markers (*-a(v)-*, *-eh-*, *-l-*, *-ah-*, *-n-*) as well as the imperfective suffix *-ahi* to the stem. Furthermore, the first-person imperfective form is irregular, as it results from the contraction of the personal concord marker *-av-* and the suffix *-ahi*, i.e. *már(-)áhi* < **már-av-ahi* ‘I was beating’. On the other hand, the irrealis form consists of the perfective form of the verb stem (e.g. *márd-*) followed by the personal concord markers (*-’um-*, *-’al-*, *-’a-*, *-’am-*, *-’en-*) and the long-vowel counterpart of the imperfective suffix, *-áhi*. In this paradigm, the third-person singular form is irregular (indicated in grey in Table 20) due to the contraction of the personal marker *-a* and the suffix *-ahi*, i.e. *márd^ʔ(-)-áhi* < **márd^ʔ-a-ahi* < **márd^ʔ-ah-ahi*. The stress does not interact with the vowel quantity of neither the imperfective nor irrealis suffix. Instead, the stress

generally falls on the personal concord marker, or less commonly, on the preceding syllable due to the elision of /e/ in the third-person singular and the second/third-person plural markers, e.g. *már-l-ahi* < **már-el-ahi*. The first syllable of the imperfective/irrealis suffix, having been merged with the personal marker, became stressed in the contracted form of the imperfect and irrealis, that is, in the first- and third-person singular, respectively. I argue that the latter form was the trigger for the vowel length becoming generalized – or analogically extended – across the whole irrealis paradigm (Table 21).

	PERSON	FORM
Contraction	3SG	*stem- <i>ah-ahi</i> > stem- <i>áhi</i>
	↓	↓
Extension	2SG	* stem- <i>al-ahi</i> > stem- <i>al-áhi</i>
	2/3PL	* stem- <i>en-ahi</i> > stem- <i>en-áhi</i>
	1SG/PL	* stem- <i>um-ahi</i> > stem- <i>um-áhi</i>

Table 21 Development of vowel length in the irrealis forms

On the other hand, the imperfective paradigm seems to be resistant to the ‘possible’ analogical pressure of the first-person form. Nonetheless, let us first consider the paradigm affected by the change. Following the markedness theory (see e.g. Jakobson 1939; Croft 1990; Greenberg 1966), we can determine that the third-person singular is the unmarked member of the paradigm, while the other persons are marked. The term unmarked refers here to the shortest, ‘least marked’ or zero-coded elements of the respective paradigm that occurs more frequently than the marked forms. According to this, the third-person singular form of the irrealis paradigm is the least marked because the morphological boundary between the personal *-a* and the imperfective suffix *-ahi* became blurred. This is also supported by the results of Elšík and Matras (2006: 361–362) based on the sample of various Romani dialects in which the singular and, to a lesser extent, the third-person form can be considered unmarked or, according to their terminology, ‘default values’ in Romani. Returning to our example, the vowel length became the only indicator of the underlying personal marker in the third-person singular. Furthermore, it is commonly acknowledged that analogical change tends to be based on the most frequent, unmarked form with highest frequency, while the least frequent, marked, forms naturally tend to regularize. As follows, the third-person suffix *-áhi*

was re-analysed as an irrealis suffix and, subsequently, extended through the paradigm. In other words, the original imperfective marker *-ahi* split into two distinct markers in KR: into the imperfective *-ahi* and irrealis *-áhi*. As a consequence, the irrealis third-person singular form has become even less marked than before, since the personal marker ceased to be encoded by the vowel length. This form may therefore be analysed as consisting of a perfective stem and an irrealis suffix (in long vowel), while the person is zero-coded, e.g. *márd'-áhi* ‘beat.PFV-0-IRR’ < **márd'(-)áhi* ‘beat.PFV-3SG-IRR’.

The question arises, therefore, as to why the contracted first person singular form has not triggered similar vowel lengthening in the imperfective paradigm, though it is the unmarked member in the respective paradigm. It seems that it is more essential to maintain the distinction between the imperfective and irrealis paradigms than the restriction imposed by the person/number value.

Vowel length in other varieties of Vend Romani

The analogical extension of vowel length in the possessive pronouns has been also attested in other Vend Romani varieties in Hungary, as well as in the closely related and geographically neighbouring variety of Versend Romani (Bodnárová 2009: 28). This change has not been detected in the Vend Romani varieties that are spoken beyond Hungary.

3.2.6 Apocope

The apocope of final vowel is very common in KR, especially of final *e*. Since long vowels are most often found in open penultimate syllables (see 3.2.5.2), the frequent apocope of final vowels results in that several words have long vowels in their final syllables. The apocope of final *e* is common before the sonorants *l r m n* and *j*. It occurs in 1) preterite third-person plural of MID-verbs (e.g. *hajin* < **hajin-e* ‘obey.3PL’), 2) preterite third-person plural of the *d*-verbs as well as the verbs *l-* ‘to take’ and *d-* ‘to give’, 3) in adverbs formed by the suffixes *-ón* (e.g. *bokhal-ón* < **bokhal-ón-e* ‘hungry’), *-án* (e.g. *rom-án* < **rom-án-e* ‘in Romani’) and *-ún* (e.g. *parašt'-ún* < **parašt'ún-e* ‘on Friday’), 4) in the adverb *khetán* (< **khe-tán-e* ‘one.OBL-place-on’) ‘together’, 5) in some local adverbs with directive and stative orientation (e.g. *tél* ~ *tél-e* ‘down’, *khér* < **khér-e* ‘at/to home’), 6) in the participial marker *-im* < **-im-e*

of loanwords, *d*-verbs as well as the verbs *l*- ‘to take’ and *d*- ‘to give’ (*dim* < **dim-e* ‘given’), 7) in the preterite third person copula form *sin* (~ *sín-e*) ‘s/he was; they were’, 8) in the negated present third person copula form *nán* (~ *nán-e*) ‘s/he was not; they were not’, and 9) in the fossilized vocative form *dáj* < **daj-e* ‘mother.VOC’.

The final *a* may be optionally dropped in the adverb *akán* ~ *akán-a* ‘now’ as well as in the indefinite pronoun *ništ* ~ *ništ-a* ‘nothing’. The apocope of final *i* may occur in the local adverb *ár* (~ *ár-i*) ‘out’ and in the nouns with the stem-final palatal *ň*, such as in *páň* (~ *páň-i*) ‘water’ or *zubuň* (~ *zubuň-i*) ‘coat’.

Apart from what was described above, the word-final unstressed vowel is sometimes not pronounced before another vowel in the flow of the speech. In KR, it is especially the case of the word-final *i* (12), *a* (13) and *e* (14) (cf. Elšík *et al.* 1999: 313):

- (12)^{LQCR} *káj h' [= hi] oja lumni?*
 where COP.3 that.F woman
 Where is that woman?

- (13)^{LQCR} *mro papu and' [= and'a] ék góno d'áro*
 my grandfather bring.PRT(-3SG) a sack flour
 My grandfather brought a sack of flour.

- (14)^{NAR} *com ameng' [= amenge] ál kécázned'ven ezer forint*
 together 1PL.DAT come.3SG two_hundred_and_forty thousand forint
 We get together two hundred and forty thousand forint.

Apocope in other varieties of Vend Romani

The final *e* is also elided in **korkór-e* (> *korkór-Ø*) ‘alone’ in Sopron Romani and in some Somogy Romani varieties. Furthermore, in some Somogy Romani varieties the apocope of final *o* and long *á* seems to be also allowed, e.g. (Kaposmérő Romani):

- (15)^{LQCR} *ad' [= adá] avrijal iste kéres*
 this(-M) another_way should do.2SG
 You should do it another way.

- (16)^{LQCR} *ék tikno žúkel béšt' [= béšto] uzar o kher*
 a small dog sit.PRT(-3SG.M) at DEF house
 A small dog was sitting at the house.

Exceptionally, the merger of two adjacent vowels at the word boundary results in the lengthening of the second vowel, such as $a+o > ó$ in Baté Romani (Somogy):

- (17)^{LQCR} *n' óvla [= na ovla] má tateder*
 NEG COP.FUT.3SG any_more warmer
 It is not going to get warmer.

3.3 Stress

The stress pattern of KR agrees to a great extent with that described in Halwachs (1998b: 26–29) for Burgenland Romani, according to which the stress tends to fall on the penultimate syllable of polysyllabic words. More precisely, the stress is placed on penult in most adjectives, in the nominative and accusative forms of nouns (e.g. *žuv.li* ‘woman’, *žuv.ja* ‘woman.ACC’), and in the present (e.g. *kér.el* ‘s/he does’), future (except for the first person singular; e.g. *ker.la* ‘s/he will do’), preterite (e.g. *ker.d'a* ‘s/he did’) and irrealis forms of non-contracted verbs (e.g. *ker.d'á.hi* ‘s/he would have done’). In other than nominative and accusative cases, the stress falls on the oblique suffix, e.g. *fú.ró.ha* ‘with drill’, *vé.šu.va.tar* ‘from chisel’. As a result, the stress is positioned on the antepenultimate syllable in case of the non-contracted genitive forms, e.g. *da.des.ke.ro* (~ *da.des.kro*) ‘father’s’. The antepenultimate syllable is stressed, too, in imperfect second person plural and third person forms of verbs, e.g. *már.la.hi* ‘s/he was beating’.

Due to contraction, final stress emerged in the future first person singular (e.g. *márá* < **márava* ‘I will beat’), in other than first person present forms of contracted verbs (e.g. *sikál* < **sikavel* ‘teaches’), in Hungarian loanwords with final long vowels (e.g. *fat'ú* < **fat'úvo* ‘child, son’), in the demonstrative pronouns *adú* < **adava* ‘this’, *odú* < **odova* ‘that’ and *okú* < **okova* ‘that other’, and in the adjectives *asó* < **asavo* ‘such.M’, *asaj* < **asavi* ‘such.F’ and *god'ár* < **god'aver* ‘smart’. The final stress is further found in the adjective *šukár* ‘beautiful’, in the local adverbs *edej* ‘here’ and *odoj* ‘there’, in the personal pronouns *amen* ‘we’ and

tumen ‘you.PL’, in the third-person plural reflexive pronoun *pumen* ‘themselves’ and, due to the apocope of final vowel (see 3.1.9), in several other words and word forms.

The initial syllable is stressed in nominative forms of unadapted Hungarian items. This means that these loans are taken over together with the Hungarian stress pattern, e.g. *felvonuláš* < H *felvonulás* ‘procession’. The penultimate stressed syllables and the final syllables of apocopated forms often contain long vowels (see 3.1.9).

4 Morphology

4.1 Nouns

4.1.1 Noun integration⁴⁵

KR contains a number of nouns originating from South Slavic, German and Hungarian. These nouns are either adapted by means of the Greek-origin adaptation suffixes, or integrated but morphologically unadapted into the masculine or feminine xenoclitic gender classes. The following sections discuss the patterns of noun integration separately for the individual contact languages, since they display some differences. I will pay special attention to the features absent or peripheral in the northern varieties of South Central Romani, such as the fact that the recently borrowed *C*-final nouns do not require adaptation suffixes, or that these nouns may optionally become feminines in Romani.

4.1.1.1 Nouns borrowed from Slavic

The *C*-final nouns that have been borrowed from South Slavic are adapted by the Greek-origin suffix *-o* as masculine nouns,⁴⁶ e.g. *grob-o* < S *grob* ‘grave’, *noj-o* < S *gnoj* ‘dung’, *prah-o* < S *prah* ‘dust; ash’. An exception is the noun *vór-a* (< S *dvor*) ‘yard’ which was adapted by *-a* into xenoclitic feminine class. The Slavic nouns in final *-o* were integrated into xenoclitic masculine class (e.g. *trašil-o* < S *strašil-o* ‘scarecrow’, *klédal-o* < S *ogledal-o* ‘mirror’), while the Slavic nouns in final *-a* became feminines in KR (*slív-a* < S *sliv-a* ‘plum’, *subot-a* < S *sobot-a* ‘Saturday’, *motik-a* < S *motik-a* ‘hoe’). The *C*-final Slavic nouns *plán* (< *plande*; Vekardi 1984: 74) ‘noon’ and *pómod’* (< S *pomoč*) ‘help’ were borrowed without any adaptation suffixes. The noun *plán* became masculine in KR, while the noun *pómod’* received feminine gender value. These nouns were probably borrowed in the time of, or after the loss of, the German contact, since the process of unadapting the *C*-final nouns is clearly a German-induced development in KR (see 4.1.1.2).

⁴⁵ This chapter is partly derived from the author’s earlier paper (Bodnárová 2014).

⁴⁶ No masculine adaptation by means of the Greek-origin suffix *-i* has been found in the data.

4.1.1.2 Nouns borrowed from German

German nouns that have been borrowed into KR have retained their original gender value. This means that German masculine nouns were subsumed in the class of KR masculines, while German feminine nouns became also feminine nouns in KR. Either masculine or feminine gender may have been assigned to German neutral nouns:

neutral → masculine: G *Öl* > *él-o* ‘oil’

neutral → feminine: G *Reh* > *ré* ‘deer’

In this regard, it is interesting how Romani dealt with the borrowing of the *C*-final German nouns. These nouns were partly adapted into xenoclitic classes with the regular adaptation suffixes, i.e. with *-o* or *-i* into masculine classes (a) and with *-a* or *-i* into feminine classes (b), and partly integrated but unadapted into xenoclitic classes (c).

- a. *cvitur-i* < (G *Zwitter* ‘hermaphrodite’) ‘homosexual’
- b. *éz-a* < G *Esel* ‘donkey’
- c. *mašin* < G dial. *Maschin* ‘machine’
ampós < G *Amboß* ‘anvil’

Since a number of German *C*-final nouns are feminines in German, and thus are integrated as feminines into KR, the number of feminine loanwords in KR has considerably increased (see also 4.1.1.4).⁴⁷

The German nouns in final vowel are morphologically not adapted either, e.g. G *Reh* /re:/ (> *ré*) ‘deer’. The intensive German contact has also brought some uncountable nouns into KR. These are the masculine *khafě* < G *Kaffee* ‘coffee’, *té* < G *Tee* ‘tea’, *šír* < G *Geschirr* ‘dish’, *rajs* < G *Reis* ‘rice’ and *klát* < G *Kleid* ‘dress’. The adjective *rajn* (< G *rein*) ‘clean’ is also used as a noun meaning ‘cleanliness, tidiness’.

⁴⁷ Before the German contact, the *C*-final nouns had been generally adapted by the Greek-origin adaptation suffixes into xenoclitic classes of Romani masculines.

4.1.1.3 Nouns borrowed from Hungarian as masculines

The strategy of unadapting the *C*-final nouns, which developed during the German contact, has been applied to the Hungarian nouns too. However, given that Hungarian does not distinguish gender, the *C*-final Hungarian nouns were more or less randomly integrated into the KR feminine and masculine classes (see also 4.1.1.4.). The gender value is determined only for some recently unadapted nouns, while other nouns may be assigned either masculine or feminine gender. Nevertheless, it seems that the masculine integration slightly outnumbers the feminine integration.

The adaptation pattern of the recently borrowed *C*-final nouns depends on the following factors:

- Time of borrowing
- Phonetic type of the final consonant in the source word
- Number of syllables in the source word

The first factor deals with the age of the loanwords. It affects the integration pattern exclusively of those nouns which are borrowed from Hungarian. More precisely, Hungarian loanwords that have been – presumably – borrowed earlier are adapted by the Greek-origin adaptation suffixes *-o* or *-i*, such as *világ-o* < H *világ* ‘world’ or *šógor-i* < H *sógor* ‘brother-in-law’. On the other hand, Hungarian loanwords that have been – presumably – borrowed after a prolonged contact with the language are unadapted, i.e. integrated into Vend Romani by a zero suffix, e.g. *leptop-0* < H *leptop* ‘laptop’, *táršašág-0* < H *társaság* ‘company’, *silvester-0* < H *szilveszter* ‘New Year’s Eve’. Moreover, the recently borrowed Hungarian items containing the vowel *a* are pronounced in KR with the Hungarian-specific slightly rounded open back vowel [ɒ], while elsewhere – in pre-Hungarian words as well as in adapted loanwords from Hungarian – the unrounded vowel [a] is used (see 3.2.1). This type of loanword integration is very unusual, as these loanwords are an integral part of the KR lexicon, but they are neither adapted morphologically nor phonologically (for the stress pattern see 3.3).

The latter two factors that influence loanword adaptation have been described by Elšík *et al.* (1999: 322–323) for the northern varieties of South Central Romani. The second factor determines the adaptation suffix of the older loanwords. Accordingly, the Hungarian nouns

ending in the sonorants *m n l*, the plosives *p b t d k g* and the fricative *v* generally take the adaptation suffix *-o* (Table 22). With regard to borrowing of Hungarian nouns ending in any other final consonant, the third factor is relevant, i.e. the number of syllables in the source word.

Thus in the case of Hungarian nouns in the sibilants *c ě s š z ž*, the sonorants *j r*, the palatals *ň d'* and the glottal *h*,⁴⁸ the choice between the adaptation suffixes *-o* or *-i* is determined by the number of syllables of the Hungarian source word. The monosyllabic Hungarian nouns take the *-o* suffix, whereas the polysyllabic nouns take the *-i* suffix, as it is exemplified in Table 22.

	<i>m n l p b t d k g v</i>	<i>c ě s š z ž j r ň d' h</i>
Monosyllabic <i>-o</i>	<i>hang-o</i> < H <i>hang</i> ‘voice’	<i>híš-o</i> < H <i>hús</i> ‘shadow’
Polysyllabic <i>-o</i>	<i>somséd-o</i> < H <i>szomszéd</i> ‘neighbour’	
Polysyllabic <i>-i</i>		<i>šárkáň-i</i> < H <i>sárkány</i> ‘dragon’

Table 22 Adaptation of C-final nouns as masculines

There is in addition a significant variation between the integration of a single loanword with or without the adaptation suffixes *-o* / *-i*, e.g. *határ-i* alongside *határ-o* < H *határ* ‘border’, *koš-o* alongside *koš-o* < H *kos* ‘ram, tup’, or *šajt-o* alongside *šajt-o* < H *sajt* ‘cheese’. This means that the original, overtly adapted, form occasionally occurs alongside the prevalent innovative, unadapted, form. In part, this variation reflects the differences between idiolects of different speakers or correlates with the degree of their competence in Romani.

The adaptation suffix is generally added to the Hungarian base stem, less commonly to inflectional stem, e.g. *lelk-o* < H *lélek* (infl. stem *lelk-*) ‘soul’, *mírg-o* < H *méreg* (infl. stem *mérg-*) ‘anger’, *cukr-o* < H *cukor* (infl. stem *cukr-*) ‘son’, *tev-o* < H *tő* (infl. stem *töv-*) ‘trunk’. Irregularity is found in the stem of the noun *tetev-o* < H *tető* (infl. stem *tetej-*) ‘roof’, which could have arisen through analogy to the form *tev-o* (see above).

⁴⁸ No Hungarian loanwords in the fricative *f* or the palatal *t'* have been attested in the data.

Hungarian nouns in final vowel may also become either masculines or feminines in KR. It seems to be a rule that human nouns of male sex are generally adapted as masculine nouns, independently of the final sound of the source word, such as the masculines *faťú* < H *fattyú* ‘son’ and *átóserelő* < H *autószerelő* ‘car mechanic’. Other examples imply that objects traditionally used by men may be subsumed in the masculine class, such as the masculine *ňakkendő* < H *nyakkendő* ‘tie’. Nonetheless, it is not clear which other factors, if any, play a role in classifying further inanimate nouns into gender classes (see also 4.1.1.4).

Table 23 exemplifies the types of masculine nouns in final vowel found in KR.

IN FINAL	e.g.
a	<i>terorišta</i> < H <i>terrorista</i> ‘terrorist’
ó	<i>koporšó</i> < H <i>koporsó</i> ‘coffin’
ő	<i>lépčő</i> < H <i>lépcső</i> ‘stairs’
ő > ó	<i>felhó</i> < H <i>felhő</i> ‘cloud’
u	<i>kuku</i> < H dial. <i>kuku</i> ‘egg’
ú ~ ú > u	<i>faťú</i> ~ <i>faťu</i> ‘son’ < H <i>fattyú</i> ‘bastard’
ű > ó	<i>betó</i> < H <i>betű</i> ‘letter’
i	<i>bači</i> < H <i>bácsi</i> ‘uncle’

Table 23 Integration of Hungarian V-final nouns as masculines

The vast majority of these Hungarian nouns are borrowed without any phonological changes. The delabialization of the word-final front rounded vowel is typical only in masculine adaptation. Compare, for instance, the delabialized masculine *d’epló* (< H *gyeplő*) ‘rein’ with the labialized feminine noun *hüttő* (< H *hűtő*) ‘fridge’. It is further interesting that the quality of the final vowel changes beside its delabialization, such as in *betó* < H *betű* ‘letter’. Furthermore, the word-final long *ú* may shorten, which is a common feature of colloquial Hungarian. The nouns *lah-ó* (< H *oláh*, or S *vlah*) ‘Vlax Romani man’ and *hábor-i* (< H *hábor-ú*) ‘war’ are adapted irregularly, as the regular forms would be **laho* and **háború* ~ **háboru*.

Also noteworthy is the seeming adaptation by the suffix *-k-* in the masculine *burđu-k-o* (< H dial. *borgyu*) ‘calf’ and feminine *čipá-k-a* (< H *csipa*) ‘eye’s sand’, pointed out also by Elšík *et al.* (1999: 324–325) in some lexemes of the northern varieties of South Central

Romani. Similarly to Elšík *et al.* (ibid.), I suppose that the consonant *k* could be a part of a dialect or archaic form of the respective lexeme, in the same way as for instance the noun *teržek-o* ‘trunk’ has been borrowed from the archaic Hungarian *törzsök* (cf. standard H *törzs*), or the noun *metelik-i* from the dialect Hungarian *metélk-e* ‘noodles’, cf. standard Hungarian *metélt*. An alternative explanation can be that the plural forms were borrowed into KR, which is marked by *-k* in Hungarian, i.e. H dial. *borgyu-k* ‘calf-PL’ and H *csipá-k* ‘eye’s sand-PL’.

4.1.1.4 Nouns borrowed from Hungarian as feminines

In the case of *C*-final feminine adaptation, only the first factor dealing with the time of borrowing is relevant (see 4.1.1.3). According to it, the older loanwords in final consonant in the source language are adapted by the suffix *-a* (e.g. *keňv-a* < H *könyv* ‘book’), while the recent loanwords are unexpectedly marked by a zero-suffix, e.g. *bus-0* < H *busz* ‘bus’, *ed’etem-0* < H *egyetem* ‘university’, *bál-0* < H *bál* ‘ball’. The reason for unadapting Hungarian *C*-final nouns as feminines may be found in that, before the Hungarian contact, KR had already possessed a number of *C*-final feminine nouns borrowed from German (see 4.1.1.2). The German-induced integration pattern of *C*-final nouns was thus replicated for the Hungarian nouns, too. Given that in the ancestor varieties of KR the *C*-final nouns used to be adapted exclusively as masculines and the *V*-final nouns as feminines, the innovative integration pattern considerably raised the number of feminine nouns in KR.

Since gender does not exist in Hungarian, there are some tendencies to integrate Hungarian nouns in final consonant to the feminine gender class (cf. Elšík *et al.* 1999: 320–327): First, animate nouns generally reflect the sex of the referent, e.g. the feminine *aňoš* < H *anyós* ‘mother-in-law’. Second, inanimate nouns referring to some objects used by women are integrated as feminines, such as *harišňanadrág* < H *harisnyanadrág* ‘tights’, *vajlinga* < H dial. *vajling* ‘basin-like pot’, and *sandál* < H *szandál* ‘sandal’. Third, the gender of a replaced word may be preserved in the loanword, such as the feminine gender in *sarvaš* < H *szarvas* ‘deer’ as compared to the older feminine noun *ré(j)* borrowed from G *Reh* ‘deer’. However, there are still many instances of feminine loanwords, where the reason for this type of integration remains unclear, e.g. the feminines *kološtor-0* < H *kolostor* ‘monastery’, *d’ár-0* < H *gyár* ‘factory’, *sekrěň-0* < H *szekrény* ‘wardrobe’, *naranč-0* < H *narancs* ‘orange’, *forgáč-a* < H *forgács* ‘chip’, etc. Furthermore, the feminine and masculine gender of the same loanword

may also vary, e.g. *dél* (F ~ M) < H *dél* ‘South’, *vár* (F ~ M) < H *vár* ‘castle’, *boritték* (F ~ M) < H *boriték* ‘envelope’. It thus seems that the Hungarian C-final nouns are in some cases admissible to adapt into both gender classes.

Hungarian nouns in a final vowel are mostly integrated into the class of feminine nouns (Table 24).

IN FINAL	e.g.
a	<i>buborka</i> < H <i>uborka</i> ‘cucumber’
e ~ e > a	<i>kečk-e</i> ~ <i>kečk-a</i> < H <i>kecske</i> ‘goat’
é	<i>tévé</i> < H <i>tévé</i> ‘TV’
i	<i>buli</i> < H <i>buli</i> ‘party’
ó	<i>rádijó</i> < H <i>rádió</i> ‘radio’
ó > ova	<i>hintó-v-a</i> < H <i>hintó</i> ‘spring carriage’
ő	<i>terittő</i> < H dial. <i>terítő</i> ‘tablecloth’
ő > ova	<i>mézo-v-a</i> < H <i>mező</i> ‘meadow’
ű > uva	<i>hegedu-v-a</i> < H <i>hegedű</i> ‘violin’

Table 24 Integration of Hungarian V-final nouns as feminines

In addition, the nouns in final *e* may also be adapted by the marker *-a*. The adaptation marker *-v-a* of Hungarian nouns in final long vowel has been preserved only in a few earlier loanwords. Like in masculine nouns, the final labial vowel becomes delabialized. Furthermore, the form *kuňu-v-a* < H *kunyhó* ‘hovel’ resulted from the assimilation of the final vowel to the preceding vowel. The form *šerpeň-a* from the Hungarian *serpenyő* ‘pan’ is irregular, since it is adapted by *-a*.

When adapting nouns into KR, the adaptation suffix is attached either to the Hungarian base stem or inflectional stem (cf. 4.1.1.3), e.g. *šork-a* < H *sarok* (infl. stem *sark-*) ‘heel’, *d’omr-a* < H *gyomor* (infl. stem *gyomr-*) ‘stomach’. On the other hand, the morphological boundary of Hungarian nouns in final short vowel is reanalysed as being without a final vowel, e.g. the H stem *kecske* is analysed as *kečk-e* ~ *kečk-a* in Romani.

Noun integration in other varieties of Vend Romani

The patterns of loanword integration described above for KR roughly agrees with the patterns found in other Vend Romani varieties of Hungary. As it has been mentioned, some loanwords may have their gender value determined, whereas in case of other nouns the specific gender value seems to be only randomly assigned. However, the KR nouns with ‘fixed’ gender may have different gender value in other varieties of Vend Romani. For instance, the nouns *bus* ‘bus’ and *hüttő* ‘fridge’ are feminines in KR but masculines in Zala Romani, or the noun *khafě* ~ *kafě* ‘coffee’, which is masculine in KR and feminine in Veszprém Romani. The irregularly adapted noun *hábor-i* (< H *háború*) ‘war’ occurs also in other varieties of Somogy, and, surprisingly, in the geographically distant Sopron Romani, while the regular *háború* ~ *hábarú* is attested in other varieties of Vend Romani. Furthermore, the KR irregularly adapted *šerpeň-a* (< H *serpenyő*) ‘pan’ is adapted as *šerpeňo-v-a* in Zala and Vas Romani, while unadapted (i.e. *šerpeňő*) in other varieties of Vend Romani. This noun has generally become feminine in Vend Romani, but not in Veszprém Romani.

The process of unadapting the *C*-final feminine nouns is also common to Burgenland and Prekmurje Romani. For instance, Burgenland Romani adapted a number of *l*-final German nouns by the suffix *-in-* into the class of xenoclitic feminines (NOM.SG *-in-a*, NOM.PL *-in-i*) (Halwachs 1998a: 23). This suffix has been preserved in the noun *khugl-in-a* ~ *kugl-in-a* (< G *Kugel*) in KR and Veszprém Romani, *vajgl-in-a* (< G *Weidling*) ‘metal bowl’ in Zala Romani, in the Hungarian-borrowed noun *kifl-in-a* (< H *kifli*) in Csokonyavisonta (Somogy), and in the plural nouns *ágl-in-i* (< G *Ohrring*) ‘earrings’ in Kálmánca (Somogy) and *štrimf-ini* (< G *Strumpf*) ‘socks’ in Táska (Somogy).

4.1.2 Noun formation

This chapter deals with the various processes involved in noun formation. It will be shown that in KR new nouns are mainly formed from the existing ones by means of suffixation. The genitive derivations, and derivations which can be considered ‘unmarked’, are rather exceptional. Neither compounding is a productive means to create nouns. The relatively large number of compounds found in KR is either borrowed or calqued from Hungarian.

4.1.2.1 Abstract nouns and other minor derivations

The inherited suffixes *-ibe* and *-be* derive abstract nouns from consonantal and vocalic verbs, respectively, e.g. *már-ibe* ‘beating’ < *már-* ‘to beat’, *ásá-be* ‘laugh, smile’ < *ása-* ‘to laugh, smile’. They have been attested only with inherited verbs. Middle verbs take also the suffix *-ibe*, placing it after the participial stem, e.g. *sikl-ibe* ‘learning’ < PTC *sikl-*. The final consonant of the stem in *nand'-ibe* ‘bath’ (< PTC *náng-*) becomes palatalized before the derivational suffix. In contracted verbs, the suffix *-ibe* is attached to the full stem, e.g. *pj-ibe* ‘drink’ < *pi(j)-* ‘to drink’. Some of the derived nouns have a lexicalized meaning such as *astar-ibe* ‘prison’ < *astár-* ‘to catch’ or *pék-ibe* ‘cake’ < *pék-* ‘to bake’.

The voiceless counterpart of the deverbal suffix, *-ípe*, derives abstract nouns from adjectives, e.g. *gull-ípe* ‘candy’ < *gull-o* ‘sweet’, *kuč-ípe* ‘expensiveness’ < *kuč* ‘expensive’. The suffix is added directly to the adjectival stem, e.g. *dilin-ípe* ‘stupidity’ < *dilin-o* ‘stupid’, *sast-ípe* ‘health’ < *sást-o* ‘healthy’, *nasval-ípe* ‘sickness’ < *nasvál-o* ‘sick’. In contrast to *-ibe*, the deadjectival suffix has been also attested with borrowed stems, such as with the German-borrowed *rajn* ‘clean’ > *rajn-ípe* (alongside *rajn*, see below) ‘clearness’, or with the Slavic-borrowed *zelen-o* ‘green’ > *zelen-ípe* ‘greenness’. The meaning ‘grease, fat’ may be expressed by both derived forms *čikn-ípe* and *čikň-ípe*. The inflectional stem of the former is the adjective *čikn-o* ‘greasy, fatty’, while the latter is presumably derived from the middle verb **čikň-ov-* ‘to become fat’, though this verb is absent from my data.

The individual voiced (*-ibe*) and voiceless (*-ípe*) pair of suffixes are strictly reserved for deverbal and deadjectival derivations, respectively. It means that my data show little alternation as regards the interchangeability of these suffixes, e.g. *nand'-ípe* ~ *nand'-ibe* ‘bath’ < *náng-o* ‘naked’, *barval-ípe* ~ *barval-ibe* ‘wealth’ < *barvál-o* ‘rich’. It is also possible that this alternation is caused by the lower language proficiency of some speakers.

The Hungarian derivational suffixes *-ság* ~ *-ség*, which derive abstract nouns from adjectives and nouns, are borrowed into KR together with the Hungarian root, e.g. *bátoršág* < H *bátorság* ‘courage’, cf. *bátor* ‘brave’; *ünnepšég* < H *ünnepség* ‘celebration’, cf. *ünnep* ‘feast’. These loanwords are not adapted in KR (see 4.1.1), except for the noun *kirájság-o* < H *királyság* ‘kingdom’, which takes the adaptation marker *-o*.

The *-šág-* allomorph of the suffix, preceded by a remnant of the adaptation marker *-ín-*, is used to derive nouns from verbs in KR (cf. Elšík *et al.* 1999: 330), i.e. *-i-šág-o*. The formant

-i-šág-o is productive in KR, since it has been attested in the Slavic-origin *pis-i-šág-o* ‘writing’ < *pis-ín-* ‘to write’ and the German-origin *štrajt-i-šág-o* ‘quarrel’ < *štrajt-ín-* ‘to quarrel’. Hungarian employs different suffixes for this function, namely *-ás ~ -és*. Thus, in Hungarian the corresponding forms of the previous examples would be *ír-ás* ‘writing’ < *ír* ‘to write’ and *veszeked-és* ‘quarrel’ < *veszeked-ik* ‘to quarrel’. In Hungarian, nouns are also derived from verbs by the suffixes *-at ~ -et*. Nevertheless, these suffixes are borrowed into KR together with the Hungarian root, e.g. *élet-o* < H *élet* ‘life’, *vašaláš* < H *vasalás* ‘ironing’, *kerittész* < H dial. *kerítettés* ‘fence’.

The German borrowed *khírer-i* (< G dial. *khierer*) ‘scream’ appears to contain the German deverbal suffix *-er-*, cf. *khír-ín-* ‘to scream’. It is not clear whether the same derivational morpheme is found in *síder-i* (maybe the dial. form of the G *Sieb*) ‘sieve’, or the morpheme *-er-* constitutes the part of the root here.

The abstract nouns *rajn* ‘clearness’ (cf. *rajn* ‘clean’) and *pótin* ‘payment’ (cf. *pořín-* ‘to pay’) do not involve any derivational markers. A seemingly similar example is the noun *pómod* ‘help’. However, this form was rather directly borrowed from the South Slavic *pomoč* ‘help’ than internally derived from the verb *pomožín-* ‘to help’. The genitive form of some nouns have a different meaning compared to that of the source noun, e.g. *jaga-kér-o* ‘fire.OBL-GEN-M.SG; safety match’ < *jag* ‘fire’, *moja-kér-o* ‘wine.OBL-GEN-M.SG; pub’ < *mol* ‘wine’, or *pren-gér-o* ‘feet.OBL-GEN-M.SG; policeman’ < *pre* ‘feet’ (cf. see also the genitive form of the numeral *šóv* ‘six’ > *šóves-kér-o* ‘six.OBL-GEN-M.SG; gun’).

4.1.2.2 *Tree names*

The suffixes *-alin ~ -ulin* are used to derive names of trees from the names of their fruits. They are lexically restricted, since the former allomorph is encountered only in *phab-alin* ‘apple tree’ (< *pháb-a* ‘apple’), while the latter only in *krušk-ulin* ‘pear tree’ (< *krušk-a* ‘pear’). Other names of trees are composed of the adjectival form of the fruit’s names and the noun *kašt* ‘tree’, which is a calque from Hungarian, e.g. *čerešnitiko kašt* (cf. H *cseresznye-fa* ‘lit. cherry tree’) ‘cherry tree’, *slivitiko kašt* (cf. H *szilva-fa* ‘lit. plum tree’) ‘plum tree’, *agácitiko kašt* (cf. H *akác-fa* ‘lit. wattle tree’) ‘wattle tree’. The calqued periphrases *phábuno kašt* ‘lit. apple tree’ and *kruškano kašt* ‘lit. pear tree’ may occasionally substitute the derivational forms *phabalin* and *kruškulin*, respectively.

4.1.2.3 Occupation and profession names

The only derivational morpheme of occupation names which shows some productivity is *-áš ~ -oš* borrowed from Hungarian, *-(V)s*. It is found in the ethnonym *kopan-áš-i* (< S *kopan-a* ‘trough’) and in some designations of sub-ethnic groups such as *žukl-áš-i* (< *žúkel* ‘dog’), *patav-áš-i* (< *patav-o* ‘foot-rag’), *bob-oš-i* (< *bob-o* ‘bean’), and *prah-oš-i* (< *prah-o* ‘dust’). Despite of the small in number data, I would argue that the vowel component of the derivational suffix *-Vš-* was contaminated by the corresponding Hungarian forms here, except of the form *kopan-áš-i*, i.e.:

Hungarian		KR
<i>kutyá-s</i> (< <i>kutya</i> ‘dog’)	→ <i>kuty-ás</i>	<i>žukl-áš-i</i>
<i>kapcá-s</i> (< <i>kapca</i> ‘foot-rag’)	→ <i>kapc-ás</i>	<i>patav-áš-i</i>
<i>bab-os</i> (< <i>bab</i> ‘bean’)		<i>bob-oš-i</i>
<i>por-os</i> (< <i>por</i> ‘dust’)		<i>prah-oš-i</i>

The first two examples show that the morphological boundary of the derivational morphemes in the corresponding Hungarian forms was reanalysed as *-ás*, and consequently replicated in the Romani example in the form *-áš*. The latter two examples display the same derivational marker as their Hungarian pairs. The form *kopan-áš-i* does not copy the respective Hungarian form *teknő-s*, it may be thus explained by analogical change in accordance with the forms *žukl-áš-i* and *patav-áš-i*. Other imported suffixes from Hungarian indicating certain occupations are the *-ás ~ -és* (< H *-ász ~ -ész*) and *-išta* (< H *-ista*), e.g. *vad-ás-i* < H *vad-ász* ‘hunter’, *zen-és-i* < H *zen-ész* ‘musician’, *terror-išta* < H *terror-ista* ‘terrorist’. The German denominal suffix *-er* which is found in *šlajf-er-i* < G *Schleif-er* ‘grinder’ and *jág-er-i* < G *Jäg-er* or H dial. *jáger* ‘hunter’, is also imported. The South Slavic derivational suffix *-ár* (< S *-ar*) has been preserved in the loanwords *pék-ár-i* (< S *pek-ar*) ‘baker’ and *pád-ár-i* ‘doctor’ (< Slovenian *pad-ar* ‘quack’). The noun *món-ár-i* ‘miller’, which also contains the South Slavic-origin suffix *-ár*, was most probably borrowed via Hungarian, cf. H *moln-ár*.

4.1.2.4 Units of measure

The extracted adjectival formant *-oš-n-* appears in derivations of banknote names, as it is shown by the only attested example *šel-oš-n-i* ‘one hundred note’. Note that the feminine form of the adjective is used as a noun, i.e. *šelošn-i* not *šelošn-o*. The base form is the inherited numeral *šel* ‘hundred’ here, while the derivational suffix is the allomorphic variant of the Hungarian *-as*, which occurs in the corresponding Hungarian translation *száz-as* ‘one hundred note’ < *száz* ‘hundred’.

4.1.2.5 Diminutives

Diminutives are derived by means of the inherited suffix *-ór-* from nouns (e.g. *rakl-ór-i* < *rákl-i* ‘non-Romani girl’, *pér-ór-o* < *pér* ‘belly’, *kher-ór-o* < *kher* ‘house’), as well as from adjectives that are used as nouns (e.g. *tikn-ór-i* ‘baby’ < *tikn-i* ‘the small’). The suffix *-ór-* is productive. It has been attested in some recent loanwords, such as in *kuňuv-ór-i* < *kuňuv-a* (< H *kunyhó*) ‘hovel’, or *fat’uv-ór-o* < *fat’ú* ‘child, boy’ (cf. H *fattyú* ‘bastard’). The latter example shows that the derivational stem is not contracted in contrast to the base form, i.e. **fat’úv-o* < *fat’ú* ‘child’. The *-ór-* derivation is used also as a respectful way to address God or older persons, e.g. *dél-ór-o* ‘dear God’ < *dél* ‘God’, *phur-ór-i* ‘granny’ < *phúr-i* ‘old-F.SG’.

The adjective *tikn-o* ‘small, little’ placed before the noun is another common means to express diminutivity in KR. For instance, the expressions *tikno fat’ú* ‘little boy’ and *tikni čhaj* ‘little girl’ have been attested several times in the data, while the corresponding derivational form **čhaj-ór-i* is absent, and *fat’uv-ór-o* occurs only rarely. The combination of a diminutive derivation with the adjective *tikn-o* seems to be restricted to the storytelling speech style, e.g. *tikno kher-ór-o* < *kher* ‘house’, *tikni marh-ór-i* < *marha* ‘animal’, *tikni phur-ór-i* < *phúri* ‘old woman’ (for more examples see Rézmúves 2006).

The names of animals’ young ones are only rarely derived by *-ór-*. I have encountered only the derived names *žukl-ór-o* ‘puppy’ < *žúkel* ‘dog’, and *marh-ór-i* ‘little animal’ < *marha* ‘animal’. These forms were, in addition, preceded by the adjective *tikn-o*. Names of the animals’ young ones are mostly composed of the adjective *tikn-o* which is followed by the name of the respective adult animal, e.g. *tikni čirikli* ‘nestling’ < *čirikli* ‘bird’, *tikni ré* ‘doe’ < *ré* ‘deer, roe’, *tikni mačka* ‘kitten’ < *mačka* ‘cat’. Among the non-derived names of animal

babies we can find, for instance, the inherited *baličho* ‘piglet’, *čhavri* ‘chicken’, or the borrowed *čikó* (< H *csikó*) ‘foal’ and *burd’uko* (< H *borjú*) ‘calf’.

4.1.2.6 *Feminine formations*

The inherited suffix *-n-* of animate nouns denoting females is not productive. In KR it is found in *manuš-n-i* ‘woman’ < *mánuš* ‘man’, *rá-n-i* ‘lady’ < *raj* ‘lord’, *gurum-n-i* ‘cow’ < **gúruv* ‘bull’ and *lum-n-i* ‘woman’. The female counterpart of some other inherited nouns denoting males is marked by final *-i*, which results in the change of the inflectional class (Elšík *et al.* 1999: 331–332), e.g. *gáž-i* ‘non-Romani woman’ < *gáž-o* ‘non-Romani man’, *kiriv-i* ‘godmother’ < *kiriv-o* ‘godfather’, *pirán-i* ‘fiancée’ < *pirán-o* ‘fiancé’ and *rákl-i* ‘non-Romani girl’ < *rákl-o* ‘non-Romani boy’. The feminine inflection by *-i* is also found in nominalized adjectives, e.g. *kedvešn-i* ‘girlfriend’ < *kedvešn-o* ‘boyfriend’, *phúr-i* ‘old woman’ < *phúr-o* ‘old man’.

A productive means to derive animate feminine nouns in KR is the South-Slavic suffix *-kiň-* (cf. Elšík *et al.* 1999: 332–333). It is generally applied in place of the Hungarian nouns *nő* or *asszony* meaning ‘woman’, e.g. *šógor-kiň-a* (cf. H *sógor-nő*) ‘sister-in-law’, *somséd-kiň-a* (cf. H *szomszéd-asszony*) ‘neighbour woman’. These Hungarian nouns that refer to the gender are sometimes borrowed into KR together with their head nouns, e.g. *meň-assoň* (< H *meny-asszony*) ‘bride’. Unlike in Hungarian, in KR the suffix *-kiň-* occurs in feminine forms of ethnic nouns (e.g. *židóf-kiň-a* (cf. H *zsidó*) ‘Jewish woman’), and rarely elsewhere (e.g. *bosorkán-kiň-a* (cf. H *boszorkány*) ‘witch’). Other feminine ethnic nouns are expressed by a phrase, e.g. *romani lumni* ‘Romani woman’.

The suffix *-kiň-* has been attested in combination with two South Slavic loanwords: *prost-óf-kiň-a* (cf. S *prost-a-kinj-a*) ‘non-Romani woman’ and *pékár-kiň-a* (cf. S *pekar-ka*) ‘female baker’. The latter form appears to be derived internally. The derivation *faťuf-kiň-a* ‘girl’ (< *faťú* ‘boy, child’) has been attested only rarely alongside the inherited noun *čhaj* ‘girl’, as well as the derivation *manuš-kiň-a* ‘woman’ (< *mánuš* ‘man’) alongside the inherited form *manušni* ‘woman’. A rather uncommon feminine derivation from an inanimate object is the noun *láboš-kiň-a* (cf. H dial. *lábos*) ‘metal bowl’. The motivation for it may be found in the fact that the metal bowl is an object used traditionally by women.

Other male-female pairs are the lexical *čhaj* ‘girl’ vs. *fat’ú* ‘boy’, *phen* ‘sister’ vs. *phral* ‘brother’, *baba* ‘grandmother’ vs. *papu* ‘grandfather’, or *sási* ‘mother-in-law’ vs. *sástro* ‘father-in-law’. One of my consultants with lower Romani language skills used systematically the feminine noun *sástr-i* ‘mother-in-law’ as an analogical form to *sástr-o*, and even once the masculine form *sás-o* ‘father-in-law’ as an analogical form to *sás-i*. The nouns *unoka* ‘grandchild’ and *unokateštvér* ‘cousin’ stands for both male and female referents, while the sex is distinguished only in their inflected forms.

Generally, there is a single form to refer to animals irrespective of their sex, e.g. *bákro* ‘sheep, ram, dam’, *ré* ‘deer, buck, roe’, *pápin* ‘goose, gander’, *éza* ‘donkey, jennet, jackass’, or *róka* ‘fox, vixen’. The female counterpart of *žúkel* ‘dog’ is *žukli* ‘bitch’, and of *bálo* ‘pig, boar’ the borrowed *geba* (< H *göbe*) ‘sow’. In both examples, the masculine form is used as the gender-indifferent form. Similarly to the noun *mánuš* ‘human; man’ which has both gender-specific and gender-indifferent meaning (cf. *manušni* ‘woman’), or to the noun *fat’ú* ‘child; boy’ (cf. *čhaj* ‘girl’). In contrast to it, the feminine form is the general term used for the animals ‘cattle’, ‘cat’ and ‘chicken’: *gurumni* ‘cattle, cow’ vs. *bika* ‘bull’, *mačka* ‘cat’ vs. *muř mačka* ‘tomcat’, *kanhi* ‘chicken, hen’ vs. *bášno* ‘rooster’. The gender-indifferent form and both gender-specific forms have been attested only for ‘horse’, i.e. the inherited *gra* ‘horse’ in gender-indifferent, and the borrowed *čéderi* (< H *csödör*) ‘stallion’ and *kanca* (< H *kanca*) ‘mare’ in gender-specific reference.

4.1.2.7 *Compounding*

KR has a great number of compounds which are directly borrowed from Hungarian. These are mainly noun-noun (e.g. *ňak-kendő* < H *nyak-kendő* ‘lit. neck-scarf; tie’), and less commonly adjective-noun compounds (e.g. *háló-soba* < H *háló-szoba* ‘lit. sleeping-room; bedroom’). The number of compounds which comprise also inherited matter is rather small. It includes some preposition-noun compounds, such as *pal-o-plán* ‘lit. after-the-noon; afternoon’, *prik-o-táha* ‘lit. beyond-the-tomorrow; day after tomorrow’ and *kija-ráti* ‘lit. towards-night; evening’, in addition to some noun-noun compounds, such as *muř fat’ú* ‘lit. man child; boy’, *muř mačka* ‘lit. man cat; tom-cat’ or *kašt čóribe* ‘lit. wood stealing; wood-stealing’. The noun-noun compounds generally replicate the respective Hungarian expressions, cf. the above examples with H *fíú-gyermek* ‘lit. boy-child’, H colloquial *fíú-macska* ‘lit. boy-cat’, and H *fa-lopás* ‘lit.

wood-stealing'.⁴⁹ An exception is the compound *gra verda* 'horse carriage', which calques the German noun-noun compound *Pferde-kutsche* 'lit. horse-carriage'. The corresponding Hungarian expression is *lovas-szekér*, which comprises adjective and noun constituents, i.e. *lovas* 'horse (adj.)' < *ló* 'horse' and the noun *szekér* 'carriage'. The pronoun and noun elements are compounded in the meaning 'my/your/his/her own mother' and 'my/your/his/her own father' (see 3.2.5.3).

Other noun phrases of merely inherited matter are modelled on the corresponding Hungarian expressions, e.g. *bari móm* < H *nagy-néni* 'lit. big aunt; aunt', *parne fóti* < H *fehér-nemű* 'lit. white sort; underwear', *teluni hólev* < H *alsó-nadrág* 'lit. under trouser; undershorts', *murdali háñig* < H *dög-kút* 'lit. carcass well; carcass pit', and *žuti ripa* < H *sárga-répa* 'lit. yellow carrot; carrot'. The numeral *epaš* 'half' is compounded with the noun *irat* 'night' in the meaning 'midnight', i.e. *epaš-irat*.

Noun formation in other varieties of Vend Romani

Similar devices are used to form nouns also in Vend Romani varieties other than KR. In Vas Romani, the suffixes *-ípe* and *-íbe* freely alternate when attached to verbs, e.g. *živ-ípe* ~ *živ-íbe* 'life' < *ži(v)-* 'to live'. Unlike in KR, the deverbal derivation suffix of vocalic verbs *-be* is added to the stem extended by *v* in Vas Romani, e.g. *asa-v-íbe* 'laugh' < *ása-* 'to laugh', *dukha-v-íbe* 'pain' < *dukha-* 'to hurt'. The stem extension of vocalic verbs is most probably motivated by the stem alternation found in contracted verbs, cf. the base forms of the vocalic *dukha-* 'to hurt' and the contracted *phuka-* ~ *phukav-* 'to tell'.

In Vas Romani and in the Somogy Romani variety of Kálmánca, the suffix *-íbe* is attached to the root instead of the inflectional stem in *pót-íbe* 'salary' (not *pot'in-íbe*; < *pot'in-* 'to pay'). This form could have been also derived from the noun *pót-in* 'payment, salary', but then the derivational form in *-ípe* would have to be expected. The suffix of abstract nouns denotes quality in the case of the deadjectival nouns *lačh-ípe* 'goodness' (< *lačh-o* 'good') in Zala Romani and *kuč-ípe* 'high prices' (< *kuč* 'expensive') in some varieties of Somogy Romani.

In Vas Romani and in the neighbouring varieties of Baté and Kaposkeresztúr (Somogy), the formant *-i-šág-o* is a productive means to derive abstract nouns from adjectives

⁴⁹ Note that instead of the adjectival *muršikano* 'male' the noun *murš* 'man' is used in the KR compounds.

(e.g. *phur-i-šág-o* ‘old age’ < *phúr-o* ‘old’, *barval-i-šág-o* ‘richness’ < *barvál-o* ‘rich’), nouns (e.g. *nimc-i-šág-o* ‘military service’ < *nimc-o* ‘soldier’), and verbs (e.g. *ás-i-šág-o* ‘laugh’ < *ás-a-* ‘to laugh’, *khír-i-šág-o* ‘roar’ < *khírín-* ‘to scream’).

The inherited diminutive suffix *-ór-* occurs in all varieties of Vend Romani. It seems, however, that the suffixation is progressively taken over by the process of compounding the adjective *tikn-o* ‘small, little’ with the respective noun. In Zala, the diminutive form of *čhej* ‘girl’ is reduced to *čhá-ri* (< **čhaj-óri*, cf.) ‘little girl’, while the diminutive form of *čhá* ‘boy’ (< **čhav-o*) remains the regular *čhav-ór-o* ‘boy’.

In Vásárosdombó (Baranya), the derivation suffix of feminines *-kiň-* freely alternates with the non-palatalized *-kin-*. In the same Romani variety, the inherited noun *phen* ‘sister’ is interchangeable with the derived forms *phen-kiň-a* ~ *phen-kin-a* and *phral-kiň-a* ~ *phral-kin-a* (< *phral* ‘brother’). Moreover, the derived form *manuš-kiň-a* ‘woman’ (< *mánuš* ‘man’) is more common than the older feminine derivation *manuš-n-i*. In this regard, a speaker of Baté Romani (Somogy), an adjacent locality of Vásárosdombó, reported the non-derived Baté Romani form *šlífer-a* ‘laddle’ as compared to the corresponding derived Vásárosdombó Romani form *šlífer-kiň-a* ‘laddle’ to be a distinguishing dialectological feature.

4.1.3 Noun inflection⁵⁰

Nouns in KR are inflected for gender, number and case. Following Matras (2002: 78–80) and Masica (1991: 232ff.), two distinct layers in the nominal case system of KR can be distinguished. The Layer I cases are the nominative and accusative, while there are only remnants of the vocative case. The Layer II cases include the dative (*-ke*, *-ge*), locative (*-te*, *-de*), ablative (*-tar*, *-dar*), instrumental (*-ha*, *-ca*) and genitive (*-k(é)r-*, *-g(é)r-*). The suffixes of the Layer II cases are added to the oblique stem of the noun. The allomorph *-ca* of the instrumental marker is employed after /n/ in the oblique suffix, while the allomorph *-ha* elsewhere. For the distribution of allomorphs of other Layer II cases refer to section 3.1.5. The functions of cases are dealt with in detail in chapter 5.1.

The accusative, which marks the animate direct object (Elšík 2000a: 13), agrees with the oblique stem, while the final *s* is dropped in the masculine singular forms, e.g. *murš-e*

⁵⁰ This chapter is partly derived from the author’s earlier paper (Bodnárová 2014).

‘man.ACC’, cf. *murš-es-* ‘man.OBL-’. The exceptions are the recently integrated Hungarian loanwords in final consonant, having the accusative forms homonymous with the nominative forms (see 4.1.3.3 and 4.1.3.4).

The nominative case is used to address persons (18)–(19), animals or objects, and the direct addressing is often placed at the end of the sentence.

(18)^{NAR} *román vaker, kuradi kurva!*
 Romani speak.IMP.2SG fucking bitch
 Speak in Romani, fucking bitch!

(19)^{RM} *ere šun, čoro rom*
 here listen.IMP.2SG poor Romani_man
 Listen, poor Romani man!

Only the nouns *daj* ‘mother’ and *dad* ‘father’ have different forms when used for addressing. The former has the long-vowel form *dáj* ‘mother.VOC’, which most probably resulted from the compensatory lengthening after the elision of the vocative suffix *-e*: *dáj* < **daj-e* ‘mother-VOC’. The latter form, on the other hand, has preserved the vocative suffix *-e*: *dad-e* ‘father-VOC’. The determiner of the head noun is in nominative, such as in (20) and (21).

(20)^{RM} *lačo dí tuke mr-i dáj!*
 good day 2SG.DAT my-F.SG mother.VOC
 Good afternoon, my mother!

(21)^{RM} *mr-o dade, ma röv!*
 my-M.SG father.VOC NEG.IMP cry.IMP.2SG
 Don’t cry, my father!

The vocative plural suffix **-ale* has been entirely lost in KR.

4.1.3.1 Inflection of oikoclitic masculine nouns

Four classes of oikoclitic masculine nouns are distinguished in KR (Table 25).

SAMPLE	NOM.SG	NOM.PL	OBL.SG	OBL.PL
<i>gav</i> ‘village’	- <i>o</i>	- <i>a</i>	- <i>es-</i>	- <i>en-</i>
<i>bál-o</i> ‘pig’	- <i>o</i>	- <i>e</i>	- <i>es-</i> ~ - <i>os-</i>	- <i>en-</i>
<i>angušt-o</i> ‘finger’	- <i>o</i>	- <i>ja</i> ~ - <i>ʔa</i>	- <i>es-</i>	- <i>en-</i>
<i>há-be</i> ‘food’	- <i>b-e</i>	- <i>b-dʔa</i>	- <i>b-es-</i>	- <i>b-en-</i>
<i>már-íb-e</i> ‘beating’	- <i>íb-e</i>	- <i>íb-dʔa</i>	- <i>íb-es-</i>	- <i>íb-en-</i>
<i>gull-íp-e</i> ‘candy’	- <i>íp-e</i>	- <i>íp-ʔa</i>	- <i>íp-es-</i>	- <i>íp-en-</i>

Table 25 Inflection of oikoclitic masculine nouns

Most of the inherited nouns end in a consonant or the vowel *-o*. There are in addition several derived abstract nouns in *-(i)be* and *-ípe*. The diminutives in *-óro* are inflected according to the paradigm represented by *bálo* in Table 25, i.e. NOM.PL *-ór-e*, OBL.SG *-ór-es-* ~ *-ór-os-*, OBL.PL *-ór-en-*. The class represented in Table 25 by *angušto* ‘finger’ includes furthermore the noun *kermúso* ‘mouse’.

In plural, the *C*-final nouns take the suffix *-a*, and the *o*-final nouns the suffix *-e*. In case of the abstract nouns and the nouns *angušto* and *kermúso* (i.e. *anguštʔ-a* and *kermús-ʔa*), the plural suffix involves in addition the consonants *j* ~ *ʔ* ~ *dʔ*, depending on the quality of the stem-final consonant (see 3.1.7). The marker of the oblique singular is generally *-es-*, whereas the oikoclitic and xenoclitic suffixes (see 4.1.3.3) alternate in *o*-final nouns. It seems therefore that the xenoclitic inflection extends to the oikoclitic one. The oblique plural marker *-en-* is the same across the paradigms.

Inherited nouns that have the same form for nominative singular and plural are the *C*-final *bal* ‘hair’, *dand* ‘tooth, teeth’, *kašt* ‘tree(s)’, *vóšt* ‘lip(s)’, and the *V*-final *dí* ‘day(s)’, *ša* ‘cabbage(s)’, *gra* ‘horse(s)’ and *va* ‘hand(s), arm(s)’. The *C*-final nouns take the regular oblique suffixes: SG *-es-*, PL *-en-*. The noun *dí* has the non-contracted oblique stem SG *dives-es-* and PL *dives-en-*, while the other *V*-final nouns have full, non-apocopated, stems in oblique, i.e. SG *šah-es-*, PL *šah-en-*; SG *grast-es-*, PL *grast-en-*; SG *vast-es-*, PL *vast-en-*. The oblique stem of the noun *žúkel* ‘dog’ is irregular, since the vowel /e/ becomes elided: *žukl-*. The contracted noun *dél* ‘God’ has full stem in plural and oblique: *dévl-* (*dévl-es-* OBL.SG).

4.1.3.2 Inflection of oikoclitic feminine nouns

There are six classes of feminine oikoclitic nouns (Table 26): three in final consonant, two in final *-i*, and one in final *-in*.

SAMPLE	NOM.SG	NOM.PL	OBL.SG	OBL.PL
<i>žuv</i> ‘louse’	<i>-0</i>	<i>-a</i>	<i>-a-</i>	<i>-en-</i>
<i>hév</i> ‘hole’		<i>-ja ~ -t'a ~ -d'a</i>	<i>-ja- ~ -t'a- ~ -d'a-</i>	<i>-jen- ~ -t'en- ~ -d'en-</i>
<i>phabalin</i> ‘apple tree’		<i>-d'a</i>	<i>-d'a-</i>	<i>-d'en-</i>
<i>kúč-i</i> ‘cup’	<i>-i</i>	<i>-a-</i>	<i>-a-</i>	<i>-en-</i>
<i>čhúr-i</i> ‘knife’		<i>-ja ~ -t'a ~ -d'a</i>	<i>-ja- ~ -t'a- ~ -d'a-</i>	<i>-jen- ~ -t'en- ~ -d'en-</i>
<i>ásv-in</i> ‘tear’	<i>-in</i>	<i>-a</i>	<i>-a-</i>	<i>-en-</i>

Table 26 Inflection of oikoclitic feminine nouns

The main difference between these paradigms lies in the nominative singular marking, since the plural and oblique suffixes are identical. The suffixes with *j t' d'* are employed in the nominative plural and the oblique in nouns with final consonants other than *j, k, kh, g, s, š, ž* and *č*, e.g. *meň-a* ‘necks’ < *men* ‘neck’, *hév-d'a* ‘holes’ < *hév* ‘hole’, *moj-a* OBL.SG < *mol* ‘wine’ (see 3.1.7). The following nouns are irregular: *lim* ‘snot’, *píšum* ‘flea’, *avd'in* ‘honey’, *karavd'in* ‘crab’, *žuv* ‘louse’ and *ármi* ‘cabbage’; they take the suffixes without *j t' d'*. The German-origin noun *cajt* ‘time; weather’ is also inflected as the inherited *C*-final nouns: *cajt'-a-* OBL.SG.

The inflection of feminine diminutives agree with the inflection of *i*-feminines, e.g. *phenór-i* ‘sister.DIM-F.SG’, *phenór-d'a* ‘sister.DIM-F.PL’, *phenór-d'a-* ‘sister.DIM-OBL.SG’, *phenór-d'en-* ‘sister.DIM-OBL.PL’. The *in*-class includes only the nouns *mát'hin* ‘fly’, *ásvin* ‘tear’ and *pátrin* ‘leaf’, but not *pápin* (NOM.PL *papiň-a*, OBL.PL *papiň-en-*) ‘goose’, *avd'in* (OBL.SG *avd'in-a*) ‘honey’, *karavd'in* (OBL.PL *karavd'in-en-*) ‘crab’ and *ricin* ‘resin’, neither the derived names of trees such as *kruškulin* ‘pear tree’ and *phabalin* ‘apple tree’. The attested non-base forms of these nouns are indicated in brackets. Note that these nouns are inflected according to the *C*-final oikoclitic feminine nouns (Table 26): *pápin* is inflected as *hév*; *avd'in* and *karavd'in* as *žuv*; while *kruškulin* and *phabalin* constitute a separate subclass by taking the plural and oblique suffixes with *d'* (i.e. *kruškulin-d'a* ‘pear tree.PL’, not *kruškuliň-a*).

The nouns *čár* ‘grass’, *čhip* ‘language; tongue’ and *masek* ‘month’ have the same form for both nominative singular and plural. On the other hand, the number is distinguished in the oblique, e.g. *čhip-ťa-* OBL.SG and *čhip-ťen-* OBL.PL, *čár-d’a-* OBL.SG and *čár-d’en-* OBL.PL, *masek-a-* OBL.SG and *masek-en-* OBL.PL. The noun *jakh* ‘eye’ has the irregular non-base stem *át’h-*, while the nouns *daj* ‘mother’ and *čhaj* ‘daughter’ have the reduced stems *da-* and *čha-*, respectively, in the oblique singular.

4.1.3.3 Inflection of xenoclitic masculine nouns

The inflectional paradigms of xenoclitic masculine nouns are shown in Table 27 (older loanwords) and Table 28 (recent loanwords).

SAMPLE	NOM.SG	NOM.PL	OBL.SG	OBL.PL
<i>barát-o</i> ‘friend’	<i>-o</i>	<i>-i</i>	<i>-os-</i>	<i>-en-</i>
<i>zubuň-(i)</i> ‘coat’	<i>-i ~ -o</i>	<i>-a</i>	<i>-is-</i>	<i>-en-</i>
<i>pádár-i</i> ‘doctor’	<i>-i</i>	<i>-ja ~ -ťa ~ -d’a</i>	<i>-is-</i>	<i>-en-</i>
<i>pap-u</i> ‘grandfather’	<i>-u</i>	<i>-u</i>	<i>-us-</i>	?

Table 27 Inflection of xenoclitic masculine nouns I.

The vast majority of xenoclitic masculine nouns are those adapted by the Greek origin markers *-o* and *-i*. Apart from the recently borrowed nouns, the class of *o*-masculines includes the Greek-origin nouns *fór-o* ‘town’, *kokal-o* ‘bone’ and *silav-o* ‘pincers’, and the class of *i*-masculines the inherited *vód-i* ‘heart’ and *páň-(i)* ‘water’, and the Greek-origin *sapuň-(i)* ‘soap’, *svír-i* ‘hammer’ and *sirim-i* ‘belt’. The final *-i* may be optionally dropped in nouns in stem-final *ň*. In nominative plural, the *o*-masculines take the suffix *-i*, while the *i*-masculines employ the suffix *-a*. The plural suffixes *-ja ~ -ťa ~ -d’a* are reserved for the *i*-masculines in final *s š č c ž* and *r*, e.g. *daráž-i* ‘bee’ > PL *daráž-d’a*, *bohóc-i* ‘clown’ > PL *bohóc-ťa*, *bogár-i* ‘bug’ > PL *bogár-d’a*. The plural suffix does not trigger gemination when attaches to a stem in final palatal, e.g. *kiráj-i* ‘king’ > PL *kiráj-a*, *šárkán-i* ‘dragon’ > PL *šárkán-a*.

The Greek-borrowed *papu* ‘grandfather’ is the only noun which consists of the adaptation marker *-u*. Although the consultants apparently found it difficult to form the

respective plural form, they congruently used the singular form also in plural, e.g. *mre duj papu* ‘my two grandfathers’.

The oblique singular suffix is *-os-* for the *o*-masculines, *-is-* for the *i*-masculines, and *-us-* for the noun *pap-u*. The oblique plural marker is *-en-* for the *o*-masculines and *i*-masculines. The oblique plural form of the noun *papu* has not been attested in the data.

SAMPLE	NOM.SG	NOM.PL	OBL.SG	OBL.PL
<i>bač-i</i> ‘uncle’	<i>-i</i>	<i>-ik</i>	<i>-is-</i>	<i>-en-</i>
<i>vajd-a</i> ‘leader’	<i>-a</i>	<i>-i</i>	<i>-as-</i>	<i>-an-</i>
<i>kuk-u</i> ‘egg’	<i>-u</i>	<i>-i</i>	<i>-us-</i>	<i>-en-</i>
<i>felhó</i> ‘cloud’	<i>-o</i>	<i>-j</i>	<i>-s-</i>	<i>-n-</i>
stem. <i>C</i>	<i>-o</i>	<i>-Vk</i>	<i>-os-</i>	<i>-en-</i>
			<i>-is-</i>	

Table 28 Inflection of xenoclitic masculine nouns II.

The recently borrowed Hungarian nouns are those that end in a short vowel (*-i*, *-a*, *-u*), long vowel (*-ó*, *-ő*, *-ú* ~ *-u*) or in a consonant (Table 28). In plural, the nouns ending in a short vowel take the suffix *-i* (except of the *i*-final nouns), and the nouns ending in a long vowel take the suffix *-j* (< **-vi*). The nominative plural form of the recently borrowed *i*-masculines agrees with the Hungarian plural form, i.e. it is formed by the Hungarian plural suffix *-(V)k*. The oblique singular suffix of *i*-masculines is *-is-*, that of *a*-masculines is *-as-*, that of *u*-masculines is *-us-*, and that of the nouns in final long vowel *-s-*. The oblique plural suffix of the *i*-masculines and *u*-masculines⁵¹ is *-en-*, that of *a*-masculines is *-an-*, and that of the nouns in final long vowel is *-n-*. The oblique forms of the consonant-final nouns are discussed below.

The nominative plural forms of *C*-final xenoclitic nouns equal to the Hungarian plural forms of the borrowed nouns, which end in *-(V)k*, e.g. *pléd-ek* (OBL.SG *pléd-os-*, OBL.PL *pléd-en-*) < H *pléd* ‘blanket’, *telefon-ok* (OBL.SG *telefon-os-*, OBL.PL *telefon-en-*) < H *telefon* ‘phone’, *ellenšég-ek* (OBL.SG *ellenšég-os-*, OBL.PL *ellenšég-en-*) < H *ellenség* ‘enemy’. The Hungarian plural marker *-(V)k* is not productive in KR, as it does not occur in

⁵¹ The noun *kuku* is the only attested representative of the respective paradigm.

other than Hungarian-origin words. The strategy of borrowing a noun together with its plural marker is not new in KR. I have found in the data a limited number of nouns which were borrowed from German together with the German plural suffix *-(e)n*, such as the feminine noun *cigrétl-en* < G *Zigarette-n* ‘cigarette-PL’, cf. G dial. *Zigaretterl* ‘cigarette(s)’ (see 4.1.3.4). One of my consultants used the German-origin plural together with the regular plural marker for the noun *kuk-u*: PL *kuk-i-jen* instead of *kuk-i*. The German plural suffix *-(e)n* also expanded to the masculine and neutral German nouns which in German as a rule take the plural ending *-e*, e.g. *réj-en* < G *Reh-e* ‘deer-PL’, cf. singular *ré(j)* < G *Reh* ‘deer’.

A striking development in KR is that the accusative forms of the *C*-final animate nouns are homonymous with their nominative forms, both in the singular (22)–(23) and the plural (24).

(22)^{NAR} *anáďa_ár* *le* *álatorvoš-0*.
 send_for.PRT.3SG DEF.OBL veterinarian-(ACC)
 He sent (someone) for the veterinarian.

(23)^{LQCR} *phučťum* *le* *polgármešter-0*.
 ask.PRT.1SG DEF.OBL mayor-(ACC)
 I asked the mayor (...)

(24)^{LQCR} *fenon_čhingerďa* *le* *bákren,* *taj t' ole* *juh-ok-0*.
 rip.PRT.3SG DEF.OBL lamb.ACC.PL and also DEF.OBL sheep-PL-(ACC)
 He ripped the lambs and also the sheep.

In other words, the accusative case is markerless and equals to the inflectional stem of the noun. Thus, not even the Hungarian accusative suffix *-Vt* is borrowed.⁵² On the other hand, the accusative plural form of the older animate loanwords takes the inherited oblique plural suffix *-en* (25), e.g. *barát-o* < H *barát* ‘friend’.

(25)^{LQCR} *o* *žúkel danderďa* *mre* *barát-en*.
 DEF dog bite.PRT.3SG my.PL friend-ACC.PL
 The dog bit my friends.

⁵² Cf. the corresponding accusative forms in Hungarian: *álatorvos-t*, *polgármester-t*, and *juh-ok-at* [PL-ACC].

According to my hypothesis, as regards the unadapted loanwords in final consonant, the NOM.SG and NOM.PL forms of Hungarian nouns were borrowed first (indicated in dark colour in Table 29). In the second stage, the ACC.SG form of these loanwords emerged beside the NOM.SG form, as an analogy to the homonymy between the NOM.SG and the ACC.SG forms of the adapted loanwords (indicated by arrows in Table 29). At first glance, the nominative forms of these integrated loanwords seem to be code-switches, as these nouns are not adapted phonologically or morphologically. However, the accusative forms of these nouns disprove this hypothesis, as they are not identical with the Hungarian accusative forms. The most curious, however, is the development of the ACC.PL form. Here, the systematic homonymy between the ACC.PL and the NOM.PL forms must have been based on the homonymy between the ACC.SG and the NOM.SG forms of the unadapted loanwords (indicated by parenthesis in Table 29), as there is no other source for it (the ACC.PL forms of the adapted loanwords are systematically distinct from the respective NOM.PL forms).

	ADAPTED	UNADAPTED
NOM.SG	<i>šógor-i</i> →	<i>polgármešter-0</i>
NOM.PL	<i>šógor-d'a</i>	<i>polgármešter-ek</i>
ACC.SG	<i>šógor-i</i> →	<i>polgármešter-0</i>
ACC.PL	<i>šógor-en</i>	<i>polgármešter-ek</i>

Table 29 Development of the inflectional class of recently borrowed Hungarian nouns

In other than nominative and accusative cases the unadapted loanwords have the same inflectional suffixes as the adapted loanwords. Thus, the oblique singular suffixes *-os-* or *-is-* are used in both of these classes of loanwords, see e.g. *áломášiha* in (26).

- (26)^{LQCR} *žan uzo áломáš-0, (...) odoleha sembe, ole áломáš-i-ha (...)*
 go.IMP.2PL to station that.INS opposite DEF.OBL station-OBL.M-INS
 you go to the station, (...) in front of it, of the station (...)

Like in adapted loanwords, the type of the final consonant and the number of syllables in the source word (as well as in the loanword) determine which of the two oblique suffixes is

added to the stem. The unadapted loanword *álmás* (< H *állomás* ‘station’), for example, ends in the sibilant *š* in both Hungarian and Romani, and it is polysyllabic. Therefore, the oblique singular suffix added is *-is-*, which is based on the underlying (but actually non-occurring) nominative singular adaptation suffix *-i*. The oblique plural suffix *-en-* is found both in the adapted and unadapted loanwords, e.g. see the instrumental form *fotel-en-ca* (27) of the unadapted loanword *fotel* < H *fotel* ‘armchair’.

- (27)^{NAR} *rendešne uri fotel-en-ca*
 real.PL noble armchair-OBL.PL-INS
 with real noble armchairs

Table 30 summarizes the inflectional paradigm of loanwords of *C*-final Hungarian nouns which are integrated into KR as masculines. The adapted loanwords *somséd-o* (< H *szomszéd* ‘neighbour’) and *tanár-i* (< H *tanár* ‘teacher’) represent the two original xenoclitic masculine classes. The unadapted loanwords *serb* (< H *szerb* ‘Serbian male’) and *polgármešter* (< H *polgármester* ‘mayor’) represent, on the other hand, the innovative xenoclitic masculine classes. The former loanword belongs to the class of ‘latent’ *o*-masculines (as it ends in a sonorant), while the latter to the class of ‘latent’ *i*-masculines (as it ends in a vibrant and is polysyllabic). The parts of the paradigms which are distinct from the original xenoclitic ones are indicated in dark colour.

	<i>-o</i>	<i>-i</i>	<i>-o</i>	<i>-o</i>
NOM.SG	<i>somséd-o</i>	<i>tanár-i</i>	<i>serb</i>	<i>polgármešter</i>
NOM.PL	<i>somséd-i</i>	<i>tanár-d’a</i>	<i>serb-ek</i>	<i>polgármešter-ek</i>
ACC.SG	<i>somséd-o</i>	<i>tanár-i</i>	<i>serb</i>	<i>polgármešter</i>
ACC.PL	<i>somséd-en</i>	<i>tanár-en</i>	<i>serb-ek</i>	<i>polgármešter-ek</i>
OBL.SG	<i>somséd-os-</i>	<i>tanár-is-</i>	<i>serb-os-</i>	<i>polgármešter-is-</i>
OBL.PL	<i>somséd-en-</i>	<i>tanár-en-</i>	<i>serb-en-</i>	<i>polgármešter-en-</i>

Table 30 Inflection of Hungarian-borrowed *C*-final nouns integrated as masculines

The adapted and unadapted loanwords show some differences not only in their inflection, but also in their stress pattern (see also 3.3). The location of the stress in Hungarian

loanwords is indicated in bold font in the Table 30. The unadapted loanwords employ the regular stress pattern in the nominative and the accusative, with stress on the penult. In the other cases, the stress falls also on the penult (e.g. *ta.ná.ris.ke* ‘teacher.DAT’); except for the genitive case, where the antepenultimate syllable is stressed (e.g. *ta.ná.ris.ke.ro* ‘teacher.GEN.M’). The unadapted loanwords, however, have the stress on the first syllable in both the nominative and the accusative. These nouns are apparently borrowed together with their stress pattern, since in Hungarian the stress regularly falls on the first syllable. In other than nominative and accusative cases, the penultimate syllable, or in the case of genitive the antepenultimate syllable is stressed. Thus exactly the same part of the inflectional paradigm which differs from the original xenoclitic one with regard to the segmental form of the inflections also exhibits a different stress pattern.

The material masculine nouns *khafé* ‘coffee’ and *té* ‘tea’ have the same form for singular and plural. The respective oblique singular forms are *khafé-s-* and *té-s-*.

4.1.3.4 Inflection of xenoclitic feminine nouns

Table 31 summarizes the inflectional paradigms of xenoclitic feminine nouns.

SAMPLE	NOM.SG	NOM.PL	OBL.SG	OBL.PL
<i>bubork-a</i> ‘cucumber’	<i>-a</i>	<i>-i ~ (-ák)</i>	<i>-a-</i>	<i>-en-</i>
<i>kečk-e ~ kečk-a</i> ‘goat’	<i>-e ~ -a</i>	<i>-i ~ -ék</i>	<i>-a-</i>	<i>-en-</i>
<i>bul-i</i> ‘party’	<i>-i</i>	<i>-ik</i>	?	?
<i>tév-é</i> ‘TV’	<i>-é</i>	<i>-ék</i>	?	?
<i>terittő</i> ‘tablecloth’	<i>-0</i>	<i>-j ~ -j</i>	<i>-0-</i>	<i>-n-</i>
<i>mézo-va</i> ‘meadow’	<i>-va</i>	<i>-j</i>	<i>-va-</i>	<i>-ven-</i>
stem.C	<i>-0</i>	<i>-Vk</i>	<i>-a-</i>	<i>-en-</i>

Table 31 Inflection of xenoclitic feminine nouns

The vast majority of xenoclitic feminine nouns ends in *-a*. This class includes a number of borrowed nouns (e.g. *bót-a* < H *bolt* ‘shop’, *kosa* < S *kosa* ‘scythe’), the feminine derivations by the formant *-kiň-a* (e.g. *fiškároškiňa* ‘female lawyer’), and some inherited nouns such as *šinga* ‘horn’, *drákha* ‘grape’, *čéňa* ‘earring’, *phába* ‘apple’, *lindra* ‘sleep’, *cipa*

‘skin’, or *vrca* ‘hair’. Hungarian nouns in final *e* are subsumed under the same class, e.g. *medv-e* ‘bear’, *pinc-e* ‘cellar’, *lepk-e* ‘butterfly’. These borrowed nouns are optionally adapted by the inherited suffix *-a*, e.g. *medv-a*, *pinc-a*, *lepk-a* (cf. with the examples above). The plural suffix of the *a*-feminines is generally *-i*. However, the recently borrowed *a*-nouns and *e*-nouns may also occur with the Hungarian plural suffix *-Vk*. The oblique stem of these nouns is formed by *-a-* in singular and *-en-* in plural.

There are only few examples of feminine loanwords in final *-i*, such as *bicigli* ‘bicycle’, *bimbi* ‘chamber-pot’, *bud’i* ‘pants’ or *mozi* ‘cinema’, while loanwords that end in *-é* are represented only by the noun *lé* ‘juice’ alongside the sample *tévé* listed in Table 31. The nominative singular and plural forms of these loanwords agree with the respective singular and plural forms found in Hungarian, while their oblique forms have not been attested in the data.

Either feminine or masculine gender may be assigned to the nouns in final long vowel, namely in *-ó* and *-ő*. The difference between the two paradigms lies in the oblique singular form, which is *-ő-* ~ *-ó-* for the feminine nouns (e.g. *törüköz-ő-ha* in 28), and *-ős-* ~ *-ós-* for the masculine nouns.

- (28)^{LQCR} *tél le khósta la törüköz-ő-ha.*
 down 3SG.ACC wipe.PRT.3SG DEF.OBL.F.SG towel-OBL.F.SG-INS
 S/he has dried him with the towel.

Some loanwords in final long vowel in the source language have preserved the older adaptation suffix *-v-* in nominative singular and oblique, such as in *hegedu-v-a* (< H *hegedű*) ‘violin’. Furthermore, the *C*-final feminine loanwords take the Hungarian-borrowed suffix *-(V)k-* in plural, e.g. *bus-ok* < H *busz-ok* ‘bus-PL’, cf. *bus* F < H *busz* ‘bus’. Similarly to the *C*-final masculine loanwords, the accusative and nominative forms of animate feminines are homonymous both in the singular and the plural, e.g. *sarvaš-0* < H *szarvas* ‘deer’ (cf. ACC in 29), *rák-ok* < H *rák-ok* ‘crab-PL’ (cf. ACC in 30).

- (29)^{NAR} *dikjum ék bara sarvaš-0⁵³.*
 see.PRT.1SG a big.OBL.F.SG deer-(ACC)

⁵³ Otherwise, in KR the ‘deer’ is called *ré* F.

I saw a big deer.

- (30)^{LQCR} *rák-ok-0* *tádam.*
crab-PL-(ACC) cook.PRT.1PL
We were cooking crabs.

On the other hand, the oblique suffix of unadapted loanwords is the inherited *-a-* in singular (31), and *-en-* in plural, e.g. *bus-en-ca* < ‘bus-OBL.PL-INS’.

- (31)^{NAR} *bus-a-ha* *áthar* *žánes* *te* *žal.*
bus-OBL.F.SG-INS from_here can.2SG COMP go.INF
You can go from here by bus.

Some German-borrowed nouns take the German-origin plural suffixes *-en* or *-in*, which are treated as irregular here. These are the *krump-a* ‘potato’ < PL *krump-in*, *ré* ‘deer’ < PL *réj-en*, and *cigrétl-i* ‘cigarette’ < PL *cigrétl-en*. In addition, the rather unusual form *móm-er-d’a* has been elicited from a KR speaker as a plural form of the noun *móm* ‘aunt’. The source of the suffix *-er-* is most probably the German plural marker *-er*.

4.1.3.5 *Pluralia tantum*

KR have some nouns that are used only in plural, such as *d’ív-a* (< **d’iv*) ‘wheat’, *sáj-a* (< **saj*) ‘saliva’, *armán-a* (< **arman*) ‘curse’, spit’, *kukur-d’a* (< **kukur*) ‘hailstone(s)’, *pór-d’a* (< **por-i*) ‘intestine(s)’, *lój*⁵⁴ (< **lój*) ‘money’, *svec-i* (< **svec-o*) ‘feast(s)’, *cirk-i* (< **cirk-a*) ‘greaves’. The inherited noun *čhor* ‘moustache’, which is originally the singular masculine form, is used only in plural number (32).

- (32)^{LQCR} *tél* *čhind’a* *pr-e* *čhor.*
down cut.PRT.3SG REFL.GEN-PL moustache
He shaved his moustache.

Analogically, only the plural form of the borrowed **sakál-o* (< H *szakáll*) ‘beard’ is used in KR, i.e. *sakál-i* ‘beard-PL’ (33).

⁵⁴ The stem is also contracted in oblique, i.e. *lój-n-*, cf. **love-n-*.

- (33)^{LQCR} *asa-j bar-e sakáli hi le!*
 such-PL big-PL beard COP.3 3SG.ACC
 He has such a big beard!

Noun inflection in other varieties of Vend Romani

The vocative form *dáj* ‘mother.VOC’ has been attested only in a few varieties of Somogy, as well as in some varieties of Burgenland and Prekmurje Romani. In other Vend Romani varieties, the nominative (short-vowel) form *daj* is used to address the mother. The origin of the Sopron Romani vocative formation *haj-daj* ‘mother.VOC’ may most probably be found in the vocative particle *hej* ‘oh’ used in the Austrian dialect⁵⁵.

In Veszprém Romani, the instrumental marker *-ca* competes with the form *-car*.

One of the most interesting developments of the Hungarian Vend Romani varieties including KR is that the oikoclitic *o*-masculines may inflect for the oblique case as the xenoclitic *o*-masculines (indicated in dark colour in Table 32).

	originally oikoclitic		originally xenoclitic	
	<i>bál-o</i> ‘pig’	<i>khóro</i> ‘jug’	<i>barát-o</i> ‘friend’	<i>bloko</i> ‘window’
NOM.SG	<i>-o</i>	<i>-o</i>	<i>-o</i>	<i>-o</i>
OBL.SG	<i>-es- ~ -os-</i>	<i>-os- ~ -es-</i>	<i>-os-</i>	<i>-os-</i>

Table 32 Inflection of the Hungarian Vend Romani *o*-masculines

A similar development is reported to be found in Burgenland Romani (Halwachs 1998a; Elšík 2000a). In this variety, the animate xenoclitic *o*-masculines are gradually taking over the oikoclitic inflection (as *grofo* in Table 33), while the inanimate oikoclitic *o*-masculines are optionally inflected as the xenoclitic nouns (as *khoro* in Table 33). According to Elšík (2000a: 24), a possible outcome of this situation could be that the animate *o*-masculines would inflect as oikoclitic nouns, while the inanimate *o*-masculines as the xenoclitic nouns. By contrast, in Hungarian Vend Romani the innovative change affected only the oikoclitic nouns, irrespective of their animacy. The xenoclitic nouns have been only exceptionally attested with the inherited oblique suffix *-es-* (e.g the xenoclitic *somséd-es-*

⁵⁵ Personal communication with a native speaker of the dialect, December 2014.

‘neighbour-OBL’), mostly from speakers with a lower competence in Romani. I have decided therefore not to indicate it in Table 32.

	originally oikoclitic		originally xenoclitic	
	<i>raklo</i> ‘boy’	<i>khoro</i> ‘jug’	<i>grofo</i> ‘earl’	<i>boklo</i> ‘window’
NOM.SG	-o	-o	-o	-o
OBL.SG	-es-	-es- ~ -os-	-os- ~ -es-	-os- (-es-)

Table 33 Inflection of the Burgenland Romani *o*-masculines; adapted from Elšik (2000a: 23)

Figure 11 shows the proportion of animated and inanimate oikoclitic nouns in relation to the oikoclitic and xenoclitic oblique suffixes (-es-/-os-) in Somogy and Zala Romani.

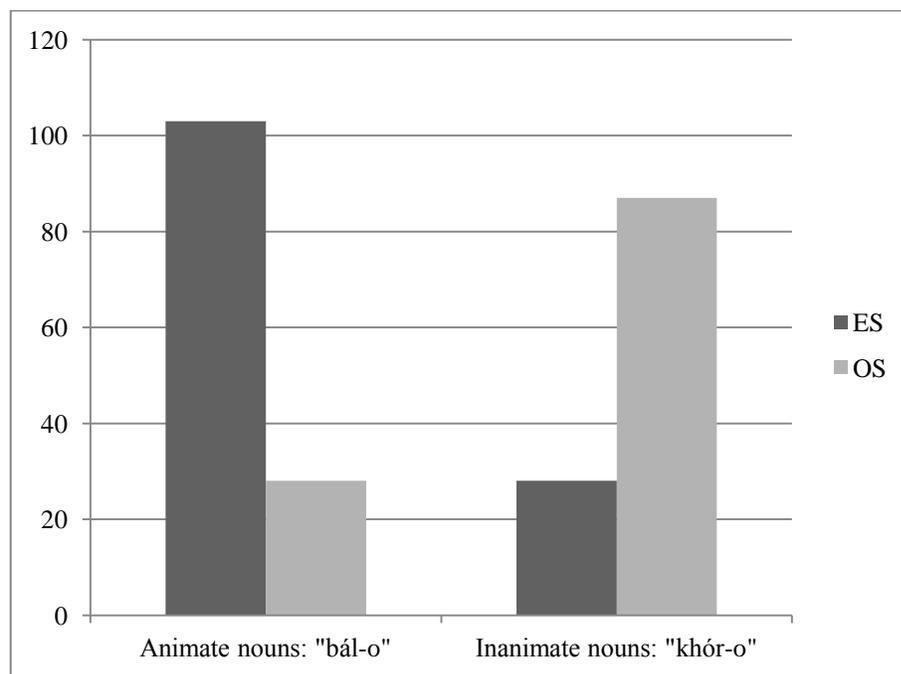


Figure 11 Proportion of in/animated nouns in relation to the oblique suffixes -es-/-os- in Somogy and Zala Romani

As it may be observed, the inanimate *o*-masculines such as *khóro* (Table 32) tend to inflect for the oblique case as xenoclitic nouns (i.e. *khór-os-*), while the animated *o*-masculines such as *bálo* (Table 32) are more resistant to this innovative change, i.e., the inherited oblique form *bál-es-* outnumbers the innovative form *bál-os-*. The disproportion of the innovative change

shows some similarity with the Burgenland Romani example. But again, the Hungarian Vend Romani data suggest that the xenoclitic markers entered, although unevenly, both animate and inanimate nouns of the oikoclitic paradigm, while the reverse change has not been attested in my data.

The irregular accusative form *gra* ‘horse.ACC’ (cf. KR *graste*), which is homonymous with the nominative form, is found in some varieties of Somogy, Zala and Sopron, e.g. *astárd’um mre gra* ‘catch.PRT.1SG my.OBL horse.ACC’. In some varieties of Somogy, the irregular oblique form *té-j-is-* (< *té*) ‘tea’ may exceptionally occur alongside the regular *té-s-*. In contrast to KR, the plural form of *čhib* ‘language, tongue’ is regular in Zala and Vas Romani, as well as in some varieties of Somogy: PL *čhib-d’a*, cf. SG/PL *čhip* in KR. In Zala Romani, the stem of the xenoclitic *i*-masculines is extended with *v* before the plural suffix *-d’a* and the oblique plural *-en-*, e.g. PL *báči-v-d’a* < *báči* (< H *bácsi*) ‘uncle’, cf. PL *bači-k* in KR.

The German plural *-en* has been preserved in various German nouns in the Hungarian Vend Romani varieties. Apart from the most common plural form in *-en*, *cigrétl-en* < *cigrétl-i* ‘cigarette’, we find for instance *minut-en* (< G *Minute-n*) ‘minutes’ in Nemesapáti (Zala), *motor-en* (< G *Motor-en*) ‘cars’ in Nagykanizsa (Zala), *film-en* ‘movies’ (< H *film* ‘movie’) in Táska (Somogy), *cájt-en* (< G *Zeit-en*) ‘times’ in Homokszentgyörgy (Somogy), *šír-en* (< G *Geschirr-e*) ‘dishes’ in Tarany (Somogy), *áringl-en* (< G *Ohrring-e*) ‘earrings’ in Baté and Vásárosdombó (Somogy), and more. The number of German plural nouns in *-en* is the highest in Vas Romani, most probably due to the proximity to Austria (and therefore a more intense German contact). In Vas Romani, the German-origin plural has also been encountered in the inherited *mirikl-en* ‘pearls’ (< *mirikl-i* ‘pearl’). What is more interesting, however, is that the plural formed by *-en* shows some productivity in the adjacent varieties of Tarany and Görgeteg (both Somogy). That is to say that in these Somogy varieties several Hungarian loanwords have been attested with the *-en* plural, such as *doboz-en* ‘boxes’ (< H *doboz* ‘box’), *paplaň-en* ‘duvets’ (< H dial. *paplaň* ‘duvet’), *hajčatt-en* ‘hairgrips’ (< H *hajcsatt* ‘hairgrip’), *hangser-en* ‘musical instruments’ (< H *hangszer* ‘musical instrument’), *čavar-en* ‘screws’ (< H *csavar* ‘screw’), *ostáj-en* ‘classes’ (< H *osztály* ‘class’), or *tank-en* ‘tanks’ (< H *tank* ‘tank’). Nevertheless, the Hungarian-borrowed plural suffix *-(V)k* predominates over the older, German-borrowed, suffix also in these varieties.

The German plural suffix was also borrowed in the form of *-ini*, such as in the masculine nouns *bicigl-ini* (< *bicigl-i*) and *cigrétl-ini* (< *cigrétl-i*) in Veszprém Romani.

Burgenland Romani employs the same suffix in combination with a limited number of nouns with final *l* in their singular forms (Halwachs 1998a), e.g. *šnicl-ini* ‘schnitzel-PL’, cf. singular *šnicl* < G *Schnitzel* ‘schnitzel’. The other German nouns in final consonants which are morphologically unadapted in Burgenland Romani employ the original suffix of xenoclitic nouns *-ča*, as in *turnir-ča* ‘tournament-PL’, cf. the singular *turnir*. The same pattern is found in Prekmurje Romani, where the Slovenian nouns in final consonants take the original suffix of xenoclitic nouns *-(j)a* (Antauer 2010), as in *vinograd’a* ‘vineyard-PL’ (cf. the singular *vinograd*). Thus, in contrast to KR, the recently borrowed nouns in Burgenland and Prekmurje Romani take the original plural suffixes (Table 34).

	Burgenland <i>-l</i> ‘Schnitzel’	Burgenland <i>-C</i> ‘tournament’	Prekmurje ‘shepherd’	Hungarian Vend ‘mayor’
NOM.SG	<i>šnicl</i>	<i>turnir</i>	<i>pastir</i>	<i>polgármešter</i>
NOM.PL	<i>šnicl-ini</i>	<i>turnir-ča</i>	<i>pastir-d’a</i>	<i>polgármešter-ek</i>
ACC.SG	-	-	<i>pastir</i>	<i>polgármešter</i>
ACC.PL	-	-	<i>pastird’en</i>	<i>polgármešter-ek</i>
OBL.SG	<i>šnicl-es/is-</i>	<i>turnir-is-</i>	<i>pastir-is-</i>	<i>polgármešter-is-</i>
OBL.PL	<i>šnicl-en-</i>	<i>turnir-en-</i>	<i>pastird’en-</i>	<i>polgármešter-en-</i>

Table 34 Inflection of borrowed *C*-final nouns integrated as masculines into Burgenland, Prekmurje and Hungarian Vend Romani⁵⁶

In Prekmurje Romani, the accusative form of the recently borrowed animate nouns is zero marked (*pastir* ‘shepherd.(ACC)’), whereas the original xenoclitic suffix is added in plural (*pastird’en* ‘shepherd.ACC.PL’). No animate loanword in a final consonant has been attested in the data for Burgenland Romani (Halwachs 1998a).

The oblique suffixes of the *C*-final Prekmurje and Burgenland Romani loan-nouns are the original xenoclitic ones (SG *-is-*, *-os-*, PL *-en-*). The German-borrowed nouns in final *l* may take both the xenoclitic (*-is-*, *-os-*) and inherited oblique suffixes (*-es-*) in the singular,

⁵⁶ The data for Burgenland and Prekmurje Romani are from Halwachs (1998a, 2002) and Antauer (2010) respectively. I consulted Samanta Baranja, a native speaker of Prekmurje Romani, with regard to the inflectional paradigm of the loanword *pastir*.

e.g. *šnicl-is-ke* alongside *šnicl-es-ke* ‘schnitzel-OBL-DAT’. This variation is caused by the progressive loss of the xenoclitic–oikoclitic dichotomy in Burgenland Romani (Elšík 2000a).

The same development is found in xenoclitic C-final feminine nouns in Prekmurje and Burgenland Romani (Table 35).

	Burgenland	Prekmurje	Hungarian Vend
	‘travel’	‘secret’	‘crab’
NOM.SG	<i>roas</i>	<i>skrivnost</i>	<i>rák</i>
NOM.PL	<i>roas-ča</i>	<i>skrivnost'-a</i>	* <i>rák-ok</i>
ACC.SG	-	-	<i>rák</i>
ACC.PL	-	-	<i>rák-ok</i>
OBL.SG	<i>roas-a-</i>	<i>skrivnost'-a-</i>	<i>rák-a-</i>
OBL.PL	<i>roas-en-</i>	* <i>skrivnost'-en-</i>	<i>rák-en-</i>

Table 35 Inflection of borrowed C-final nouns integrated as feminines into Burgenland, Prekmurje and Hungarian Vend Romani⁵⁷

It may be observed in Table 35 that Vend Romani differs from Burgenland and Prekmurje Romani in the realization of the nominative plural. Whereas the latter two take the original suffixes, Vend Romani borrows the noun together with the Hungarian plural suffix.

4.1.4 Names of localities

The names of municipalities, countries and continents are generally taken over from Hungarian without being morphologically adapted. The exceptions found are the adapted *Taraň-a* from H *Tarany*, the semicalqued *Tikni Baráti* from H *Kisbaráti*, or the South Slavic-borrowed *Ninčko* ‘Germany’. The names of continents and countries in final *a* become feminines in KR, such as *Afrika* (< H *Afrika*) ‘Africa’, *Amerika* (< H *Amerika*) ‘America’, *Anglija* (< H *Anglia*) ‘England’, *Austrija* (< H *Ausztria*) ‘Austria’, *Indija* (< H *India*) ‘India’,

⁵⁷ The data for Burgenland and Prekmurje Romani are from Halwachs (1998a, 2002) and Antauer (2010) respectively. No animate loanword in final consonant is found for Burgenland Romani (Halwachs 1998a); and Antauer (2010), on the other hand, does not give an example of it in Prekmurje Romani.

Románija (< H *Románia*) ‘Romania’, *Serbija* (< H *Szerbia*) ‘Serbia’, or *Slovákija* (< H *Szlovákia*) ‘Slovakia’. The compounded country names with the component *-ország* in Hungarian are integrated into the class of *C*-final masculines, such as *Čeország* (< H *Csehország*) ‘Czech Republic’, *Mađarország* (< H *Magyarország*) ‘Hungary’, *Orosország* (< H *Oroszország*) ‘Russia’, and *Švédország* (< H *Svédország*) ‘Sweden’. In contrast, the *C*-final names of municipalities often became feminines in KR, such as *Kišbajum* from H *Kisbajom*, *Debrecen* from H *Debrecen*, or *Požoň* from H *Pozsony* ‘Bratislava’.

Names of localities in other varieties of Vend Romani

Hungarian towns and villages are known under their local Hungarian names among the Vend Roma. These names are often shorter than the official names, as for instance *Kapoš* (cf. Hung. *Kaposvár*) or *Pešt* (cf. *Budapest*). Like in KR, the names of localities are borrowed without being morphologically or phonologically adapted into Romani. Exception is *Kaniž-i* (cf. H *Nagykanizsa*) recorded in Nagykanizsa (Zala), though this form may also be a dialectal form. The locative form *Peštate* (< *Pešt*, cf. H *Budapest*) is used as nominative in several Vend Romani varieties. In addition, the German-borrowed name *Finkhin* (< G *Fünfkirchen*) as a name of the Hungarian town Pécs was passively known by some speakers in Somogy.

4.2 Adjectives

4.2.1 Integration of adjectives

The adjectival adaptation markers are *-n-* (exceptionally *-utn-*), *-m-n-* (< **-v-n-*) and *-ast-* (Table 36). The distribution of these markers is conditioned by the phonological quality of the loan-adjective’s final sound. Hungarian adjectives with stem-final consonants are generally adapted by *-n-*. If the stem ends in a geminate, it becomes degeminated before the adaptation marker, e.g. *éret-n-o* < H *érett* ‘mature’, *friš-n-o* < H *friss* ‘fresh’. The geminate *nn* may arise, on the other hand, by adding the marker to a stem with final *n*, such as in *idegen-n-o* < H *idegen* ‘strange, foreign’ (see 3.1.4). In some rare instances, the geminate may undergo dissimilation; cf. the above example with *idegem-no*. If the stem ends in a consonant cluster, this cluster may become simplified before the application of the adaptation marker, e.g. *sen-n-o* < H *szent* ‘holy, saint’. The Hungarian adjectives with stem-final consonant may

exceptionally be adapted by *-utn-* instead of *-n-*, e.g. *gazdag-utn-o* ~ *gazdag-n-o* < H *gazdag* ‘rich’, *özved’-utn-o* ~ *özved’-n-o* < H *özvegy* ‘widow’. The marker *-utn-* originally has the function of a derivational marker (see 4.2.2).

STEM-FINAL	MARKER	e.g.
consonant	<i>-n-</i>	<i>vidám-n-o</i> < H <i>vidám</i> ‘merry’
	* <i>-utn-</i>	<i>boldog-utn-o</i> < H <i>boldog</i> ‘happy’
a > 0	<i>-ast-</i>	<i>barn-ast-o</i> < H <i>barna</i> ‘brown’
e > 0		<i>bisk-ast-o</i> < H <i>büszke</i> ‘proud’
as > 0	<i>-(as)t-</i>	<i>kop-as-t-o</i> < H <i>kopasz</i> ‘bald-headed, hairless’
ó	<i>-mn-</i>	<i>utolsó-mn-o</i> < H <i>utolsó</i> ‘last’
ő		<i>középő-mn-o</i> < H <i>középső</i> ‘middle’
ű		<i>šűrű-mn-o</i> < H <i>sűrű</i> ‘thick’

Table 36 Adaptation of Hungarian adjectives

Hungarian adjectives in *a* and *e* take the South Slavic-origin marker *-ast-*, while the stem-final vowel is elided. The *C*-final adjective *kopasz* is seemingly adapted irregularly with *-t-* instead of the regular marker *-n-* of adjectives in stem-final consonant, e.g. *kopas-t-o* vs. **kopas-n-o* ‘bald-headed, hairless’. It seems that the formal analogy of the stem in final *-as* with the suffix *-ast-* triggers the re-analysis of the stem boundary as *kopa-*. Although the marker *-ast-* was originally extracted from Slavic loanwords, it has been retained only in the Slavic-origin *šenkl-ast-o* (< probably the dial. form of Slovenian *škilast*) ‘cross-eyed’ and *grundl-ast-o* (< probably the dial. form of Slovenian *kodrast*) ‘curly’.

The Hungarian adjectives with final long vowel take the adaptation marker *-mn-*.

Some Hungarian adjectives are borrowed into KR without any morphological or phonological changes, e.g. *amerikaji* < H *amerikai* ‘American’, *angol* < H *angol* ‘English’, *büske* (alongside *bisk-ast-*) < H *büszke* ‘proud’. These adjectives do not inflect for gender, number or case (see 4.2.3). Similarly, several German nouns are unadapted in KR, such as the *V*-final *fitti* < G dial. *fiati* ‘ready’ and *C*-final *fajst* < G dial. *fejst* ‘solid, tough, proper’, *lajt* < G *leicht* ‘light’, and *níder* < G *nieder* ‘low’. These loan-adjectives are also uninflected (see 4.2.3). It is interesting that in many instances the inflected form of the German adjectives (in final *-i*) were borrowed instead of the base stem. Examples are *práni* (< G dial. inflected form

praun-i, cf. base form *praun*) ‘brown’, or *rajni* (< G dial. inflected form *rein-i*, cf. base form *rein*) ‘clean’. It is possible that the inherited *hór-i* (< **hór*) ‘deep’ took a final *-i* analogically to the monosyllabic German-origin adjectives in final *-i*.

Integration of adjectives in other varieties of Vend Romani

The *C*-final adjectives are adapted with *-n-* in all varieties of Vend Romani. The adaptation suffix *-ast-* is absent in Vas and Veszprém Romani, where the adjectives in short vowels are mainly unadapted, e.g. *néma* < H *néma* ‘mute’, *sürke* < H *szürke* ‘grey’. Unadapted are also the adjectives in final long vowel in Zala, Vas and Veszprém Romani, such as the Hungarian-origin *šürü* < H *sűrű* ‘thick’ in Vas and Zala Romani, or the German-origin *fró* < G *froh* ‘happy’ in Sopron and Veszprém Romani.

4.2.2 Derivation of adjectives

The only example of derivation of adjectives by prefixation found in the data is *bi-londo* ‘saltless’, consisting of the prefix with privative meaning *bi-* plus the adjective *londo* ‘salty’. In the same function, we find more commonly the imported Hungarian suffixes *-talan* ~ *-telen*, *-tlan* ~ *-tlen*, *-atlan* ~ *-etlen*, e.g. *boldog-talan-no* < H *boldog-talan* ‘unhappy’. The derivational marker *-(V)š* (< H *-(V)s*) expressing quality was also imported from Hungarian, e.g. *baráčág-oš-n-o* < H *barátság-os* ‘friendly’. On the other hand, the imported Hungarian marker of comparison *-só* ~ *-ső* as well as the marker *-i* denoting place and time occurs less frequently in the data, e.g. *utol-šó-mn-o* < H *utol-só* ‘last’ or *čalád-i* < H *čalád-i* ‘family’s’. The marker *-čk-* seems to have entered KR through the South Slavic borrowing *nin-čk-o* (< S; e.g. Serbian *nema-čk-*) ‘German’.

KR has altogether eleven derivational markers and their allomorphs which are used to derive adjectives, as shown in Table 37. Some of these markers are interchangeable in combination with some words, such as *-ún-* and *-itik-* (e.g. *krušk-ún-o* ~ *krušk-itik-o* < *kruška* ‘pear’), or *-ošn-* and *-itik-* (e.g. *akhor-ošn-o* ~ *akhor-itik-o* < *ákhor* ‘nut’). However, individual derivational markers tend to occur in certain semantic areas. For instance, the marker *-(j)án-* tends to derive adjectives from the names of animals (*sáp-án-o* < *sap* ‘snake’, *réj-án-o* < *ré(j)* ‘deer’), while the marker *-(j)ikán-* from nouns denoting humans, such as *raj-ikán-o* < *raj*

‘noble man’, *lumň-ikán-o* < *lumn-i* ‘woman’. The marker *-án-* is also employed in the ethnic noun *rom* ‘Rom’ > *rom-án-o*. The adjectives from local adverbs and the nouns denoting materials and plants are generally derived with *-ún-*, e.g. *upr-ún-o* < *upr-e* ‘up’, *srast-ún-o* < *srast-a* ‘iron’, or *slív-ún-o* < *slív-a* ‘plum’. Furthermore, the marker *-utn-* is generally added to the stem of temporal adverbs (*táh-utn-o* < *táh-a* ‘tomorrow’, *lan-utn-o* < *lan-i* ‘last year’). An exception is the form *koraj-utn-o* ‘early’ derived from the Hungarian adjective *korai* ‘early’, where the stem-final vowel cluster *ai* changed to *aj* before the application of the derivational suffix. The stem of the form *adís-utn-o* ‘today’s’ (< *adí* ‘today’) has preserved the final *s* of the older **av-dives*^{B94} (> *a-dí*) ‘today’. Adjectives are derived from vocalic verbs by adding the suffix *-gutn-* to the perfective stem, such as *daran-gutn-o* ‘timid’ (< PFV stem *daran-* ‘affraid’) and *lažan-gutn-o* ‘shy’ (< PFV stem *lažan-* ‘ashamed’). The marker *-(j)ál-* generally creates adjectives with reference to certain physical or mental state, e.g. *hev-d’ál-o* < *hev* ‘hole’, *lindr-ál-o* < *lindr-a* ‘sleep’. On the other hand, the older derivational marker *-vál-* ~ *-fál-* has been preserved only in *čhor-vál-o* < *čhor* ‘moustache’, *khan-vál-o* < *khan* ‘smell’ and *rat-fál-o* < *rat* ‘blood’.

MARKER	PRODUCTIVITY	DERIVED FROM
<i>-(j)án-</i>	✓	animal noun
<i>-(j)ikán-</i>	✗	human noun
<i>-ún-</i>	✓	local adverb, material and plant noun
<i>-utn-</i>	✓	temporal adverb
<i>-gutn-</i>		vocalic verb
<i>-(j)ál-</i>	✓	physical and mental state noun
<i>-vál-</i> ~ <i>-fál-</i>		
<i>-itik-</i>	✓	material, plant and ethnic noun
<i>-jítik-</i>		
<i>-ik-</i>	✗	ethnic noun
<i>-ck-</i>	✗	ethnic noun
<i>-ošn-</i>	✓	various
<i>-sorošn-</i> ~ <i>-serešn-</i>	✓	numeral

Table 37 Derivational markers of adjectives

A widely used marker is the Greek-borrowed *-itik-* (Matras 2002:197), which derives adjectives from materials, plants and ethnic nouns, e.g. *réz-itik-o* < *réz* ‘copper’, *akhor-itik-o* < *ákhor* ‘nut’, *serb-itik-o* < *serb-(o)* ‘Serbian’. The last vowel of the base form, to which the marker is added, is generally dropped, e.g. *kečk-itik-o* < *kečk-e* ~ *kečk-a* ‘goat’. Exceptions are the Hungarian-borrowed adjectives *mod’oró-jitik-o* < *mod’oró* ‘hazelnut’ and *amerika-jitik-o* < *Amerika* ‘America’, where the stem-final vowel is retained and followed by the marker *-jitik-*. The Greek-origin marker *-ik-* is found only in the ethnic noun *ungr-ik-o* < **ungr-o* ‘Hungarian’. Some other ethnic nouns are derived with *-ck-*, such as *kopaná-ck-o* ‘Boyash’ < *kopan-a* ‘trough’, *židóf-ck-o* ‘Jewish’ < *židó*, cf. **židó-v-* ‘Jew’, *špañol-ck-o* ‘Spanish’ < *špañol* ‘Spanish’. It is possible that the South Slavic derivational suffix *-sk-* is found in these derived forms.

The imported Hungarian suffix *-(V)s* mentioned above has various allomorphs, e.g. *hab-oš-n-o* < H *hab-os* ‘foamy’, *sin-eš-n-o* ‘colored’, *láz-aš-n-o* < H *láz-as* ‘fevered’, *dijó-š-n-o* < H *dió-s* ‘nutty’. Out of the Hungarian borrowings of this type, only the form *-oš* was extracted together with the adaptation suffix *-n-*. This formant is used for deriving adjectives from a number of inherited verbs in KR, such as *masek-ošn-o* ‘month’s’ < *masek* ‘month’, *akhor-ošn-o* ‘nutty’ < *ákhor* ‘nut’, *kiral-ošn-o* ‘of cottage cheese’ < *kíral* ‘cottage cheese’, *kój-oš-n-o* ‘thingummy’ < *kój* ‘thingummy’. Even a few Hungarian-borrowed adjectives employ the marker *-oš-n-*, where a different allomorph of the Hungarian derivational suffix would be expected, e.g. *jeg-oš-n-o* (cf. H *jeg-es*) ‘icy’ < *jeg-o* ‘ice’, or *šeb-oš-n-o* (cf. H *seb-es*) ‘wounded’ < *šeb-o* ‘wound’. Since there is no evidence that these forms would have been directly borrowed from the local Hungarian dialect, it seems that the extracted suffix *-ošn-*, originally attached to inherited nouns, is getting extended to the Hungarian-borrowed bases as well.

The formant *-sor-ošn-* ~ *-ser-ešn-*, which comprises the borrowed multiplicative suffix *-sor-* ~ *-ser-* (< H *-szor-* ~ *-ször-*) plus the derivational marker *-ošn-*, derives adjectives from numerals. It has been attested only with the numeral two: *duj-sorošn-o* ~ *duj-serešn-o* (cf. H *két-szeres*) ‘double’ < *duj* ‘two’. The former form shows, according to the Hungarian rule of vowel harmony, that the inherited base with back vowel *duj* requires the suffix with back vowels *-sorošn-*, too. In contrast, the variant *duj-serešn-* seems to semi-calque the Hungarian form *két-szeres*, irrespective of the rule of vowel harmony.

Derivation of adjectives in other varieties of Vend Romani

The derivational morphology of other varieties of Vend Romani does not differ considerably from the one described for KR. The privative prefix precedes in addition the adjective *bastál-o* ‘lucky’ in the Somogy Romani varieties of Kaposmérő and Görgeteg, and also the genitive form *vójengr-o* ‘happy’ in Vas Romani.

The marker *-utn-* is also quite common in adjectives derived from local adverbs in several varieties of Somogy Romani. This suffix is exceptionally attached to borrowed adjectives in final *g* (i.e. as an adaptation suffix), which is most probably an analogy to the derived forms by *-gutno* from vocalic verbs. Examples are *gazdag-utn-o* < H *gazdag* ‘rich’ in KR, *boldog-utn-o* < H *boldog* ‘happy’ in Tarany (Somogy), *rég-utn-o* < H *rég-i* ‘old’ in Táska (Somogy), or *dug-utn-o* < S *dug* ‘long’ in Zala Romani.

The interrogative *ki’i* ‘how many/much’ takes the marker *-itik-* in the meaning ‘which/what day of the month’ in Zala and Vas Romani, i.e. *ki’-itik-o*. This form resulted from the merging of the Hungarian suffix *-Vdik-* in the same function with the Romani derivational suffix *-itik-* due to their similarity in form, cf. *ki’-itik-o* with the H *hány-adik-a* ‘what day of the month’.

The extracted form of the Hungarian suffix *-(V)s* is *-ášn-* in Zala Romani, e.g. *bríg-ášn-o* < *bríg-a* ‘sorrow’, *ásv-ášn-o* < **ásv-in* ‘tear’, *kóv-ášn-o* < *kóv-a* ‘thingummy’.

4.2.3 Inflection of adjectives

There are three inflectional classes of KR adjectives, one that is inflected and two others uninflected (Table 38).

Adjectives in final		
	vowel	consonant
Inflected	<i>Class 1</i>	
Uninflected	<i>Class 2</i>	<i>Class 3</i>

Table 38 Inflectional classes of adjectives

Class 1 comprises adjectives (e.g. the inherited *bár-o* ‘big-M.SG; big’, the borrowed *id’ešn-o* ‘skillful-M.SG; skillful’) and participles (e.g. the inherited *kerd-o* ‘make.PTC-M.SG; made’, the borrowed *rágim-o* ‘gnaw.PTC-M.SG; gnawed’) in final vowel that agree with their head nouns in gender, number and Layer I case inflection, e.g. *bar-o kher* ‘big-M.SG house.M.SG; big house’, *bar-e khér-a* ‘big-PL house-PL; big houses’. The head noun in other than nominative case requires the dependent adjective to be in oblique case, e.g. *bar-a kopajaha* ‘big-F.SG.OBL stick.INS; with a big stick’. Table 39 summarizes the inflectional paradigm of adjectives pertaining to Class 1.

	M.SG ⁵⁸	F.SG	PL
NOM	<i>bár-o</i>	<i>bár-i</i>	<i>bár-e</i>
OBL	<i>bár-e</i>	<i>bár-a</i>	<i>bár-e</i>

Table 39 Inflection of adjectives of the Class 1

Class 2 comprises several recently borrowed German and Hungarian adjectives in final vowel which do not inflect for gender, number or case, such as *fajni* < G *fein* ‘fine’ or *büske* < H *büszke* ‘proud’. Consider the following example, in which the German-borrowed adjective *ajšti* modifies the head noun in nominative (34) and instrumental case (35).

(34)^{LQCR} *mri ajšti lumni*
 my.F.SG first wife
 my first wife

(35)^{LQCR} *mra ajšti lumňaha*
 my.OBL.F.SG first wife.INS
 with my first wife

The adjective *pherde* ‘full’ seems to be the only inherited V-final adjective which retains the same form irrespective of gender, number or case, e.g. *i pherde kúči* ‘DEF.F full cup; a full cup’, cf. **i pherd-i kúči* ‘DEF.F full-F cup’.

⁵⁸ The only adjective which does not have a corresponding masculine form in my data is *khámn-i* [pregnant-F.SG] ‘pregnant’.

Class 3 subsumes adjectives in final consonant. That is, the comparative and superlative forms derived by *-éder*, the inherited adjectives *aver* ~ *ár* ‘different, another’, *god’ár* ‘smart’, *kuč* ‘expensive’ and *šukár* ‘beautiful’, and several German and Hungarian borrowings, such as *níder* < G *nieder* ‘low’ or *angol* < H *angol* ‘English’.

The adjective *as-ó* (< **asav-o*) ‘such’ has irregular inflection, as it is shown in Table 40. The irregularity lies in the various contractions which have developed in KR (see 3.1.8).

	M.SG	F.SG	PL
NOM	<i>as-ó</i> ~ <i>os-ó</i>	<i>as-aj</i>	<i>as-aj</i>
	< <i>*as-avo</i>	< <i>*as-avi</i>	< <i>*as-ave</i>
OBL	<i>as-aj</i>	<i>as-aj</i>	<i>as-aj</i>
	< <i>*as-ave</i>	< <i>*as-ava</i>	< <i>*as-ave</i>

Table 40 Inflection of the irregular adjective *as-* ‘such’

Inflection of adjectives in other varieties of Vend Romani

The three inflectional classes of adjectives introduced above are also found in other varieties of Vend Romani. Unlike in KR, the adjective *pherd-o* ‘full’ is inflected in Vas and Veszprém Romani. The adjective *hojám-n-o* ‘angry’, which inflects for gender and number, is derived from the participial form *hojám-* by the adaptation suffix *-n-* in Vas Romani and Homokszentgyörgy (Somogy). The xenoclitic participles are inflected in most varieties of Vend Romani (e.g. *feštim-* ‘coloured’), but not in Zala Romani (e.g. *feš-tíme* ‘coloured’).

The nominative masculine form of the irregular adjective *as-* ‘such’ is the original *as-avo* in Veszprém and Sopron Romani, and the apocoped *as-av* in Vas Romani. More interestingly, this adjective is uninflected in the Somogy Romani varieties of Tarany, Kaposmérő and Lengyeltóti. While the former variety employs the contracted form *as-é*, the latter two varieties use the form *as-ej*.

4.2.4 Comparison

The comparative form of adjectives is formed by the suffix *-éder*, and the superlative form by the borrowed prefix *leg-* (< H *leg-*) which is attached to the comparative form (Table 41).

	COMPARATIVE	SUPERLATIVE
<i>bár-o</i> ‘big’	<i>bar-éder</i>	<i>leg-bar-éder</i>
<i>id’ešn-o</i> ‘skillful’	<i>id’ešn-éder</i>	<i>leg-id’ešn-éder</i>
<i>rajn-i</i> ‘clean’	<i>rajn-éder</i>	<i>leg-rajn-éder</i>
<i>kuč</i> ‘expensive’	<i>kuč-éder</i>	<i>leg-kuč-éder</i>
<i>fajs ~ fajst</i> ‘proper’	<i>fajst-éder</i>	<i>leg-fajst-éder</i>

Table 41 Comparative and superlative forms of adjectives

The comparative and superlative degree of the adjective *lácho* ‘good’ is the suppletive *féder* ‘better’ and *leg-féder* ‘the best’, respectively. The comparative form of the indefinite *but* ‘a lot, many’ is also irregular, being marked by the suffix *-er* (i.e. *bút-er* ‘more’). Only a few Hungarian adjectives are borrowed together with the Hungarian comparative suffix *-Vbb* (> *-Vb*), and adapted into KR by *-n-*, e.g. *óčo-b-n-o* ‘cheap-more-AM-M.SG’ (cf. H dial. *ócsó-bb* ‘cheap-more’) ‘cheaper’. In addition, the inherited suffix *-éder* is exceptionally attached to the Hungarian comparative form, e.g. *óčo-b-n-éder* ‘cheap-more-AM-more; cheaper’. In such formations, the Hungarian comparative suffix becomes redundant.

The comparative and superlative forms of adjectives correspond to the comparative and superlative forms of adverbs, respectively, e.g. *furčast-éder* ‘more strange, more strangely’, *leg-dilin-éder* ‘more stupid, more stupidly’. The comparative adverbs *ere-féder* ‘further to this direction’ and *ora-féder* ‘further to that direction’ are composed of a Hungarian-borrowed demonstrative, *ora* (< H *arra*) ‘that way’ or *ere* (< H *erre*) ‘this way’, and the adverb *féder* ‘more’. The literal translation of these comparatives could be ‘more to the this/that way’. The corresponding Hungarian forms are *err-ébb* ‘this_way-more’ and *arr-ább* ‘that_way-more’, which are composed of the respective demonstratives and the Hungarian comparative suffix *-Vbb*.

Comparison in other varieties of Vend Romani

The stem-final *h* is preserved in the comparative forms of the borrowed adjectives *gá* (< G dial. *gach*) ‘fast’ and *švó* (< G dial. *schwoch*) ‘weak’, i.e. *gáh-éder* in Kaposmérő (Somogy) and *švoh-éder* in Nikla (Somogy).

Instead of *feder* ‘better’, the regular comparative form *lačh-éder* (< *lačh-o* ‘good’) was systematically used by a speaker of Tarany Romani (Somogy), perhaps due to the speaker’s lower proficiency in Romani.

4.3 Adverbs

The present chapter provides an overview of the most common local, temporal, deverbal, manner, anaphoric and causal adverbs found in KR, with special focus on their forms and development.

4.3.1 Local adverbs

Table 42 lists the basic local adverbs of KR by their localization and orientation, based on Elšík and Matras (2006: 242).

	DIRECTIVE	STATIVE	SEPARATIVE
	to :: toward		
Inessive	<i>ánd-e :: ánd-e cuj</i>	<i>edej-ánd-e</i>	<i>fen ándr-al</i>
Extraessive	<i>ár-i :: ár cuj</i>	<i>ávr-al</i>	<i>fen ávr-al</i>
Superior	<i>upr-e :: upr-e cuj</i>	<i>upr-e</i>	<i>fen upr-al</i>
Inferior	<i>tél-e :: tél cuj</i>	<i>tél-e</i>	<i>fe(n) tél-al</i>
Anterior	<i>ángl-e :: ángl-e cuj</i>	<i>ángj-al</i>	<i>fen ángj-al</i>
Posterior	<i>pál-e :: pál cuj</i>	<i>pál-al</i>	<i>fe(n) pál-al</i>
Proximate	<i>páš-e</i>	<i>páš-e</i>	<i>fe(n) páš-al</i>
Medial	<i>(ando) maškár-al</i>	<i>(ando) maškár-al</i>	<i>fenal maškár-al</i>

Table 42 Localization and orientation of KR local adverbs

The basic localizations of the KR local adverbs match the basic localizations reconstructed for Early Romani (ibid). Three orientations may be distinguished in most of the localizations: directive, stative and separative. The directive adverbs, in addition, display two types of orientation, which are relative to the nature of the movement. The first type specifies

a movement *to* a location, while the second type refers to a movement *towards* a location. Local adverbs have identical stems in all orientations, as for instance the superior stem *upr-*.

The directive adverbs generally take the suffix *-e*, the separative adverbs the historical ablative suffix *-al* (see e.g. Matras 2002: 42), while the stative adverbs employ either the *-e* or the *-al* suffix (Table 43).

	DIRECTIVE	STATIVE	SEPARATIVE
	to :: toward		
Common	<i>-e</i> :: ADV <i>cuj</i>	<i>-e, -al</i>	<i>fen</i> ADV- <i>al</i>
Less common	<i>-i, -al</i>		

Table 43 Marking of local adverbs

Irregular are the extraessive and medial forms with directive localization. The former has preserved the inherited *-i* suffix which follows the contracted stem *ár-* (*ár-i* ‘out’ < **avr-i*), while the latter employs the ablative form *maškár-al* ‘into the middle’ in all orientations.

Thus, the orientation is generally marked by means of suffixation. The suffixed form is combined with the postposition *cuj* (< G dial. *zui, zua*) ‘toward’ in directive, and with the preposition *fe ~ fen* (< G *von*) ‘from’ in separative orientation. The same adpositional phrases are reserved for directive and separative orientations in the German dialects spoken in eastern Austria (e.g. Styrian *obi zua* ‘lit. down towards; towards down’; standard G *von unter* ‘lit. from below’). The ablative form *fen-al* ‘from’, which is derived from the preposition *fen*, has been encountered only alongside the adverb *maškáral* ‘in the middle’. This ablative form of the preposition is generally combined with names of localities, as the ablative meaning in those cases is encoded only on the preposition, such as in *fenal péč* ‘from Pécs’ (see chapter 5.1.4).

As an innovation, the stative adverbs of extraessive, anterior, posterior and medial localizations in KR take the ablative suffix typical to separative adverbs. On the other hand, the *-e* suffix has been preserved in the inessive, superior, inferior and proximate localizations of stative adverbs. The inessive adverb *edej-ánde* ‘lit. here-inside’ is a compound which inaccurately calques the Hungarian *ide-benn* ‘lit. hither-inside; inside’. The corresponding translation would be **órde-ánde* ‘lit. hither-inside’. The deictics *edej* ‘here’ and *odoj* ‘there’ are optionally preposed to other spatial adverbs as well, and thus providing those adverbs with

additional information on proximity in relation to the speaker, e.g. *odoj ánde* ‘(there) inside’, *odoj ávral* ‘(there) outside’.

A local adverb of Slavic-origin is *prik* (< S *preko*), which competes with the Hungarian-borrowed *sembe* (< H *szembe*) ‘on the other side’. Further local adverbs are the borrowed *körbe* (< H *körbe*) and *körü* (< H *körül*) ‘around’, and the inherited *dúr* ‘far away’ and *dúral* ‘from far away’ indicating distance (Table 44).

	DIRECTIVE	STATIVE	SEPARATIVE
	to :: towards		
Oppositive	–	<i>prik</i> ~ <i>sembe</i>	–
Circumlative	<i>körbe</i>	<i>körü</i>	–
Distant	<i>dúr</i>	<i>dúr</i>	(<i>fen</i>) <i>dúral</i>
Other	<i>khér</i> :: <i>khér cuj</i>	<i>khér</i> ~ <i>odoj-khér</i>	<i>fe(n) khéral</i>
	<i>thán</i>	<i>thán</i>	<i>thánal</i>
	–	–	<i>fe(n) čáčal</i>
	–	<i>vid'ik</i>	<i>vid'ikal</i>

Table 44 Further local adverbs

Local adverbs derived from nouns are the directive/stative *khér* ‘to/at home’ and *thán* ‘to/at a place’, and their separative pairs *fen khéral* ‘from home’ and *thánal* ‘from a place’, respectively. The stative form *khér* is optionally replaceable by the compound *odoj-khér* ‘lit. there-at_home’, which is also inaccurately transferred from the Hungarian *oda-haza* ‘lit. thither-at_home’, cf. **ód'a-khér* ‘lit. thither-inside’.

The separative adverb *fe(n) čáč-al* ‘from the right side’ denoting the right direction is derived from the adjective *čáč-o* ‘right, real’, while the directive and stative forms are expressed by the Hungarian loanword *jobra* (< H *jobbra*) ‘to/on the right side’. The opposite direction is denoted only by loanwords, i.e. *barra* (< H *balra*) ‘to/on the left side’, *barru* (< H *balrul*) ‘from the left side’. The borrowed adverb *vid'ik* (< dial. form of H *végig*) is used in the meaning ‘everywhere’, and the derived form *vid'ik-al* in the meaning ‘from everywhere’. Further borrowed adverbs are *fút* (< G *fort*, G dial. *fuat*) ‘away’, *krót* (< G *gerade*, G dial. *grod*) ‘straight’ and *mind'ár* (< H *mindjárt*) ‘right’ (36).

- (36)^{LQCR} *edej áčhel mind'ár uzar mande.*
 here live.3SG right next_to 1SG.LOC
 S/he lives right next to me.

Local adverbs with *od-* (e.g. *odoj* ‘there’) and *ok-* (e.g. *ók* ‘there!’) stems will be discussed in section 4.5.3.

Local adverbs in other varieties of Vend Romani

The system of local adverbs of other Vend Romani varieties roughly agrees with the one described for KR. However, a difference is found in some western varieties (Zala, Vas, Burgenland, Prekmurje), where the ablative forms *tél-al* ‘below’ and *upr-al* ‘above’ occur also in stative orientation, cf. KR *tél-e* and *upr-e*. What is more, the ablative forms *ángjal* ~ *ánglal* ‘in front’ and *pálal* ‘behind’ are also preferred in directive orientation in almost the same geographical area (Zala, Sopron, Veszprem, Burgenland, Prekmurje). The presence of the deictic element in *edej-ánde* ‘lit. here-inside; inside’ is typical of the Vend Romani varieties of Hungary, thus being absent in Burgenland and Prekmurje Romani (cf. *ánde* ‘inside’).

Zala and Prekmurje Romani lack the German-origin postposition *cuj* ‘towards’ in directive orientation. The separative adverb does not require preposition in Zala, Veszprém and Burgenland Romani, and the prepositional phrase is only optional in some varieties of Somogy, Sopron and Zala Romani. The KR preposition of separative adverbs *fe(n)* has the form *fa* (< G dial. *fa*) in Sopron and Vas Romani, while the preposition *zar* ‘from’ has been attested in Prekmurje Romani.

The adverb for the meaning ‘away’ is the German-borrowed *fút* in Somogy, Veszprém and Sopron Romani, and the compounded *k-rik* (< **jekh-rig* ‘lit. one-side’) in Zala, Vas and Prekmurje Romani, as well as in some peripheral varieties of Somogy Romani. As for Burgenland Romani, it borrowed the form *bejg* (< G *weg*) from its present contact language.

4.3.2 Temporal adverbs

Adverbs denoting time are mostly calqued (a) or borrowed (b), and less commonly inherited (c).

- a. *avri-kor* (cf. H *más-kor*) ‘lit. other-at; another time’
berša sám (cf. H *évek-szám*) ‘lit. years number; for ages’
dúr (cajt) (cf. G *lange Zeit*) ‘lit. long_time; for a long time’
na čilla (cf. H *nem-rég*) ‘lit. not-long_ago; recently’
na dúr (cf. H *nem-sokára* ‘lit. not-for_long’) ‘awhile’
- b. *ariňňa* ‘so long’ < H *annyira* ‘to such extent’
azúta < H *azóta* ‘since then’
korán < H *korán* ‘early’
közbe < H dial. *közbe* ‘in the meantime’
maj < H dial. *maj* ‘then, later’
mind’ár < H dial. *mindjár* ‘immediately’
mindig < H *mindig* ‘always’
néha < H *néha* ‘sometimes’
núrunt < unknown ‘constantly’
órák hossat < H *órák hosszat* ‘for hours’
šoha < H *soha* ‘never’
špót < G dial. *spot* ‘late’
šürüjen < H dial. *sürüjen* ‘often’
vid’ik < dial. form of H *végig* ‘all the time’
- c. *águn* ‘at first’
akán ‘now’
búter ‘more, anymore’
butfar ‘often’
čilla ‘long ago’
jefkar ‘once’
sig ‘soon’

By contrast, the temporal adverbs indicating deictically days or a part of the day are mainly inherited, only exceptionally borrowed. The inherited adverbs comprise *ič* ‘yesterday’ and *táha* ‘tomorrow’, the compounds *a-dí* ‘lit. this-day; today’, *epaš-i-rat* ‘lit. half-DEF-night; at midnight’, *prik-o-ič* ‘lit. over-DEF-yesterday; day before yesterday’ and *prik-o-táha* ‘lit. over-DEF-tomorrow; day after tomorrow’, and the derived *divés-e* ‘during the day’ (cf. **dives*

‘day’) and *kija-rát-i* ‘in the evening’ (cf. **kija-* ‘towards’, **ráti* ‘at night’). Interestingly, the feminine article *i* is involved in the adverbial forms of the noun *rat* ‘night’, i.e. in *i rat* ‘lit. DEF night; at night’ and *epaš-i-rat* (see above). An example of a borrowed adverb is *plán* (< dial. form of S *podne*) ‘at noon’.

Adverbs indicating the days of the week are mainly borrowed from Hungarian (1, 2, 4–6 in Table 45).

INHERITED	BORROWED
1	<i>hetfi-n</i> < H <i>hétfő-n</i> ‘on Monday’
2	<i>kedd-en</i> < H <i>kedd-en</i> ‘on Tuesday’
3	<i>srid-ón</i> < S <i>sred-a</i> ‘Wednesday’ <i>serdá-n</i> < H <i>szerdá-n</i> ‘on Wednesday’
4	<i>čütörtök-ön</i> < H <i>csütörtök-ön</i> ‘on Thursday’
5	<i>parašt’-ún</i> ‘on Friday’ <i>péntek-en</i> < H <i>péntek-en</i> ‘on Friday’
6	<i>subot-ón</i> < S <i>subot-a</i> ‘Saturday’
7	<i>kurk-e</i> ‘on Sunday’

Table 45 Days of the week

The adverbial suffix *-ón* (< **-on-e*) is reserved for the Slavic-origin names of the week (3 and 6), while the suffix *-ún* (< **-un-e*) is employed in the Greek-borrowed *parašt’-ún* (< *parašt’u-va* ‘friday’). The seventh day of the week is formed by the inherited suffix *-e*, i.e. *kurk-e* ‘on sunday’ < *kurk-o* ‘sunday’. The meaning ‘at the weekend’ is expressed by the borrowed *hédvégé-n* (< H *hétvégé-n*, cf. *hétvége* ‘weekend’).

Adverbs referring to months are often imported from Hungarian together with the inessive case markers *-ba* ~ *-be* (e.g. *január-ba* < H dial. *január-ba* ‘in January’), or alternatively a prepositional phrase is used (see 5.1.5).

The main seasons are the inherited *linaj* ‘in summer’ and *d’evénde* ‘in winter’. The former is especially interesting since it seems not to have resulted from the apocopated form of the derived **lináj-e* ‘in summer’, i.e. from **lináj-0* ‘summer’. The latter form is derived by the locative suffix *-e* from the non-contracted form **d’evend* ‘winter’. The transitional seasons are the borrowed *tavas-sal* ‘in spring’ (< H *tavas-szal*, cf. *tavas* ‘spring’) and *ős-sel* ‘in autumn’ (< H *ős-szel*, cf. *ősz* ‘autumn’), the forms of which brought the Hungarian

instrumental marker *-val* ~ *-vel*⁵⁹ into Romani. The meaning ‘last year’ is expressed by the loanwords *lani* (< S *lani*) or *tavaj* (< H *tavaly*).

Temporal adverbs in other varieties of Vend Romani

There are only minor differences between the temporal adverbs of other Vend Romani varieties and the ones introduced above for KR. For instance, in Zala Romani the adverbial suffix *-ón* is added also to the name of the fourth and, optionally, to the fifth day of the week, i.e. *čütörtök-ón* (cf. H *csütörtök-ön*) ‘on Thursday’, *parašt’-ón* ~ *parašt’-ún* (< *parašt’-uva*) ‘on Friday’.

4.3.3 De-verbal adverbs

The marker *-undar* ~ *-lundar*, which is used to form adverbs from verbs, is restricted lexically in KR. It has been attested only in the following adverbs:

- a. *beš-undar* ‘sitting’ < *béš-* ‘to sit’
khel-undar ‘singing’ < *khél-* ‘to dance’
rov-lundar ‘crying’ < *rov-* ‘to cry’
- b. *terd-undar* ‘standing’ < *térd-o* ‘standing’
- c. *ása-undar* ‘laughably’ < *ása-* ‘to laugh’

The marker is added to the present stem of *C*-verbs (a), or to the adjectival stem (b). The stem-final *a* of *V*-verbs is elided before the marker (c). The form *rov-lundar* ‘crying’ is irregular, as an additional *l* is inserted between the verb stem *rov-* ‘to cry’ and the regular marker *-undar*.

De-verbal adverbs in other varieties of Vend Romani

The deverbal adverbial marker *-undar* is found in the central part of Somogy and in Vas and Veszprém Romani. The form *-undar* competes with *-indar* in Zala Romani, e.g. *rov-indar* ‘crying’ > *rov-* ‘to cry’. The formant *-und-ón* is typical for some northern peripheral varieties of Somogy, and the formant *-un-dór* for some southern peripheral varieties of Somogy. The

⁵⁹ As a rule, the first consonant of the instrumental marker assimilates to the preceding consonant.

former contains the adverbial suffix *-ón-*, while the latter most probably resulted from the contamination of the older *-undar* with the recent *-undón-*.

4.3.4 Manner adverbs

Manner adverbs provide information about how an event is performed. They can be either derived or non-derived. The most common non-derived adverbs comprise some inherited (a) and a number of borrowed adverbs originating from German (b) and Hungarian (c).

- a. *afka* ‘in this way’
épaš ‘by the middle’
féder ‘better, more’
- b. *fajs* < G *fein* ‘fine’
fitti < G *fertig* ‘ready’
gá < G dial. *gach* ‘quickly’
rajn < G *rein* ‘totally; cleanly’
gonc < G *ganz* ‘totally; quite’
- c. *alig* < H *alig* ‘hardly, barely’
bistoš < H *biztos(an)* ‘surely’
čembe < H *csendben* ‘quietly’
dalog < H *gyalog* ‘on foot’
inkáb < H *inkább* ‘rather’
ösevisa < H *összevissza* ‘criss-cross’
rembe < H *rendben* ‘all right’
tista < H dial. *tiszta* ‘totally’
külön < H *külön* ‘distinctly’

Non-derived are also the Slavic-borrowed adverbs *édva* (< S *jedva*) ‘hardly’ and *silom* ‘intentionally’. The latter has preserved the Slavic instrumental case marker *-om*, cf. S *sil-a* ‘force’ > S *sil-om* ‘by force’.

The adjectives in final consonants, including the comparatives and superlatives, have identical forms as adverbs, e.g. *god’ár* ‘smart, wise; wisely’, *šukár* ‘nice; nicely’, *féder*

'better'. The Hungarian derivational suffixes *-Vn* and *-ul ~ -ül* (> KR *-u ~ -ü*) were introduced into KR through some borrowed adverbs (d–e):

- | | |
|-----------------------------------------------------|---------------------------------------------------------|
| d. <i>ind'-en</i> < H <i> ingy-en</i> 'for free' | e. <i>serb-ü</i> < H <i>serb-ül</i> 'Serbian' |
| <i>furčá-n</i> < H <i>furcsá-n</i> 'strangely' | <i>slovák-u</i> < H <i>szlovák-ul</i> 'Slovak' |
| <i>nad'-on</i> < H <i>nagy-on</i> 'very, very much' | <i>váratlan-u</i> < H <i>váratlan-ul</i> 'unexpectedly' |
| <i>halk-an</i> < H <i>halk-an</i> 'quietly', etc. | |

On the other hand, the KR markers that serve to derive adverbs from adjectives are *-a*, *-e*, *-ón* (< **-on-e*), and *-án* (< **-án-e*, unclear origin). The marker *-a* is used to form adverbs denoting ethnic groups (e.g. *kopanáck-a* < *kopanáck-o* 'Boyash', *lahitik-a* < *lahitik-o* 'Vlax', *ninčk-a* < *ninčk-o* 'German', *ungrik-a* < *ungrik-o* 'Hungarian'), while the marker *-e* is employed (f) in several inherited and (g) a few Slavic-origin adjectives.

- | | |
|------------------------------------------------|----------------------------------------------------|
| f. <i>lók-e</i> 'slowly' < <i>lók-o</i> 'slow' | g. <i>erd'av-e</i> 'badly' < <i>erd'av-o</i> 'bad' |
| <i>phár-e</i> 'hardly' < <i>phár-o</i> 'hard' | <i>lasn-e</i> 'cheaply' < <i>lasn-o</i> 'cheap' |

The adverbs that are derived from Hungarian-borrowed adjectives take the suffix *-ón* (h). It seems, however, that the function of *-e* is gradually overtaken by the marker *-ón*. As show the examples below, the extracted *-ón* has also been found (i) in a number of inherited adjectives, as well as (j) in the Slavic-borrowed adverb *mirn-ón*. Several inherited adverbs have fixed forms either in *-e* or *-ón*, while others may take both derivational suffixes, such as the inherited adjective *phár-o* 'hard': *phár-e ~ phár-ón* 'hardly'.

- | |
|-------------------------------------------------------------------|
| h. <i>árvast-ón</i> 'orphan-like' < <i>árvast-o</i> 'orphan' |
| <i>furčast-ón</i> 'strangely' < <i>furčast-o</i> 'strange' |
| <i>serelmešn-ón</i> 'amorously' < <i>*serelmešn-o</i> 'enamoured' |
| <i>ňomorutn-ón</i> 'miserably' < <i>ňomorutn-o</i> 'miserable' |
| i. <i>náng-ón</i> 'nakedly' < <i>náng-o</i> 'naked' |
| <i>térn-ón</i> 'in one's youth' < <i>térn-o</i> 'young' |
| <i>živd-ón</i> 'lively' < <i>živd-o</i> 'alive' |
| j. <i>mirn-ón</i> 'peacefully' < <i>*mirn-o</i> 'peaceful' |

The marker *-án* has been attested only in *dilin-án* ‘stupidly’ < *dilin-o* ‘stupid’. It is not clear whether it originates from the *-án* allomorph of the Hungarian suffix *-Vn*, or it is formed analogically to the zero-marked adverbs with denominal adjective stems in *-(ik)án*, such as:

čorikán-O ‘poorly’ < **čorikán-e*, cf. adjective **čorikán-o* ‘poor’
román-O ‘Romani’ < **román-e*, cf. adjective *román-o* ‘Romani’

Manner adverbs in other varieties of Vend Romani

In some Somogy Romani varieties, there is an expansion of the suffix *-ón* at the expense of the adverbial marker of ethnic groups, e.g. *lahitik-ón* (cf. KR *lahitik-a*) ‘Vlax Romani’, *nimčk-ón* (cf. KR *nimčk-a*) ‘German’. Interestingly, the same marker is attached to the reduced stem of the inherited adjective **korkór-o* ‘lonely’ in the Somogy Romani varieties of Kálmánca and Homokszentgyörgy, i.e., *kork-ón* ‘alone’.

4.3.5 Anaphoric adverbs

Anaphoric adverbs are used to refer to an antecedent. For instance the anaphoric adverbs with local meaning *ánde* and *ándral* refer back to the noun *staklo* ‘bottle’ in (37); and the adverb *upral* is related to *hábori* ‘war’ in (38), as it was mentioned earlier.

- (37)^{RM} *Adá staklo, t' adá čhi upro ostolo! Loli mol ánde* [= *ando staklo*] *ovla.*
Taj at'i, ki'i ár pínah' ándral [= *ándral o staklo*], *mindig at'i pál pherd'ola.*
 This bottle, put it on the table! Red vine will be **in it** [= in the bottle].
 The amount, that they will drink **from it** [= from the bottle], will fill up again.

- (38)^{NAR} *Čilla afka na vakernahi asó történeti, sar sína ando hábori.*
Adalenge adá tabu síne. Ón na vakernah' upral [= *upral o hábori*].
 In the past they did not tell stories about the war.
 For them it was a taboo. They did not use to speak about it [= about the war].

Anaphoric adverbs are substitutable by prepositional phrases, and are mostly homonymous with the respective local and temporal adverbs, as it is illustrated by the following examples:

ánde ‘in it’ < *ánde* ‘inside’, cf. preposition *ande* ‘in’

upre ‘on it’ < *upre* ‘up, above’, cf. preposition *upre* ‘on’

The anaphoric adverb *váše* ‘for it’ is an exception, as the corresponding causal adverb is built using the prepositional phrase *vaš odá* ‘that’s why, therefore’, cf. preposition *vaš* ‘for’. Furthermore, the anaphoric adverb *cuj* ‘to it’ is identical with the postposition *cuj* ‘toward, -ward’ used in local adverbs with directive localization (4.3.1), as well as with the verbal particle *cuj* ‘to’ 4.7.6. This adverb is replaceable by a phrase involving the preposition *uze* ‘to’ (cf. **úze*):

(39)^{NAR} *akor páň tuke t’ ánav? ánav tuke **cuj** [= uzo páň]! mer odá čulo hi.*

Should I bring you (more) water? I bring you **to it** [= to the water]! Because you have only little.

Anaphoric adverbs in other varieties of Vend Romani

Other varieties of Vend Romani possess an almost identical series of anaphoric adverbs as KR. The only difference seems to be that in place of the anaphoric adverb *cuj* ‘to it’ we find the Slavic-origin *úze* in Zala and Prekmurje Romani.

4.4 Numerals and quantifiers

4.4.1 Basic numerals

The basic numerals of KR are the inherited *d’ékh* ~ *jékh*⁶⁰ ‘one’, *duj* ‘two’, *trin* ‘three’, *štár* ‘four’, *pándž* ‘five’, *šóv* ‘six’, *děš* ‘ten’, *šel* ‘hundred’ and *biš* ‘twenty’, the Greek-borrowed *éfta* ‘seven’, *ófto* ‘eight’, *eňa* ~ *eňňa* ‘nine’ and *tranda* ‘thirty’, and the Hungarian-borrowed *nulla* (< H *nulla*) ‘zero’, *ezeri* (< H *ezer*) ‘thousand’ and *milijó* (< H *millió*) ‘million’. The numeral *děš* has a short vowel when it functions as a determiner, e.g. *deš berš* ‘ten years’. The fractions are the inherited *epaš* ‘half’ and the Hungarian-origin *ned’ed* (< H *negyed*) ‘quarter’.

⁶⁰ The numeral ‘one’ has the form *ékh* when used as a determiner.

The speakers did not have any difficulty to translate from Hungarian to Romani numerals below hundred. Nonetheless, they tended to switch to Hungarian in the flow of the speech for the numerals above twenty (the code-switch is underlined):

- (40)^{NAR} *Sin osó, hod' man sin dešuduj gra, pándž biki, bákre,
harminchat bále, kaňha, háromsáshatvan álat man sin.
There was a time when I had twelve horses, five bulls, sheep,
thirty-six pigs; I had three hundred and sixty animals.*

Code-switching to Hungarian is also common when referring (41) to time and (42) date, or (43) decimal numbers.

- (41)^{NAR} *Inkáb pal o kilenc óra, pal o kilenc.
It is better after nine o'clock, after nine.*

- (42)^{NAR} *Mri daj ezer kilencáz neđvenéd'be úli.
My mother was born in nineteen forty-four.*

- (43)^{NAR} *Na afka sin sar akának: Ék máro duj taj trin šel forint hi.
Hát kettő hus taj kettő harminc taj három harminc sin ék máro.
It was different from what is now: A loaf of bread costs two and three hundred forints.
It used to be two twenty and two thirty and three thirty.*

The code-switching for numerals may bring Hungarian case suffixes into KR, as for instance the inessive suffix *-ba* 'in' (< H *-ban*) in the following example:

- (44)^{NAR} *And adá cilo gav asó d'iv nána sar akor hatfanhárom-ba.
There has not been as much snow in this village as back in the year sixty-three.*

4.4.2 Formation of numerals

The basic numerals are mainly compounded in order to form higher numerals, as it is summarized and exemplified in Table 46.

	MARKER	e.g.
11–16	-u-	<i>deš-u-duj</i> 12
17–19	–	<i>deš éfta</i> 17
21–26	-u-	<i>biš-u-dúj</i> 22
27–29	–	<i>biš éfta</i> 17
31–36	-u- ~ *-o-	<i>trand-u-duj</i> , <i>*trand-o-duj</i> 32
37–39	–	<i>trand- éfta</i> 37
40, 50, ..., 100	-var- ~ -val-	<i>štar-val-děš</i> 40, <i>deš-var-děš</i> 100
200, ..., 900		<i>šóv šel</i> 600
101–199	–	<i>šel jék</i> 101, <i>šel dešujék</i> 111
	<i>taj</i>	<i>šel taj jék</i> 101, <i>šel taj dešujék</i> 111

Table 46 Compound numerals

The numerals 11–16, 21–26 and 31–36 are compounds of tens and ones, which are connected by the conjunction *-u-*. Instead of *-u-*, the conjunction *-o-* has been attested in numerals 31–36 in the speech of a younger KR speaker. It is however questionable whether this form is only an idiolect feature or an optional variant to the form with *-u-*. The conjunction is generally dropped before numbers in the initial vowel, namely before *éfta* ‘seven’, *ófto* ‘eight’ and *eňa* ‘nine’. This therefore affects the compounds 17–19, 27–29 and 37–39. The other tens are composed of the multiplicative form of the numbers and the number ‘ten’, e.g. *štar-val-děš* ‘lit. four-times-ten; forty’. It is interesting that alongside the lexical *šel* ‘hundred’ we find also the compound *deš-var-děš* ‘lit. ten-times-ten; a hundred’. The hundreds are composed of the basic numeral and the singular noun *šel* ‘hundred’. The conjunction *taj* ‘and’ is only optionally required in compounds of hundreds and basic numerals, as well as those of hundreds and tens.

4.4.3 Inflection of numerals

Numerals are generally uninflected for gender, number and oblique case when they function as determiners:

- (45)^{NAR} *kindum te dešuštár baličhen.*
 buy.PRT.ISG also fourteen pig.ACC.PL

I also bought fourteen pigs.

By contrast, the numerals are inflected when acting as nouns:

- (46)^{NAR} *óv korkóro žalahi, eňň-en děš-en laklahi.*
3SG.M alone go.IMP.F.3SG nine-ACC ten-ACC find.IMP.F.3SG
He used to go alone, and he would find nine or ten (hedgehogs).

The inflectional stems of the numerals are generally *C*-final (e.g. *duj-* ‘two’, *eňň-* < *eňňa* ‘nine’), and they are inflected as oikoclitic nouns. An exception is the numeral *éfta* ‘seven’, which is inflected according to the xenoclitic noun *vajda* ‘leader’ (see 4.1.3.3):

- (47)^{NAR} *so le éft-an phosingerd'a.*
all DEF.OBL.PL seven-ACC.PL stab.PRT.3SG
S/he stabbed all seven (people).

4.4.4 Ordinals

Ordinals are created by means of the suffix *-t-* which seems to be productive: *duj-t-o* ‘second’ < *duj* ‘two’, *štár-t-o* ‘fourth’ < *štár* ‘four’, *šóv-t-o* ‘sixth’ < *šóv* ‘six’, *éfta-t-o* ‘seventh’ < *éfta* ‘seven’. The ordinal ‘first’ is the non-derived *ajšti*, borrowed from the German *erste*. The form *trit-t-o* ‘third’ is irregular, having resulted from the spontaneous gemination of the intervocalic *t* (see 3.1.4). The Hungarian suffix of ordinals *-Vdik* is imported into KR through the borrowed interrogative *háň-adik-a* ‘which day of the month’, and the borrowed adjectives denoting school grades, e.g. *het-edik-ešn-o* ‘in the seventh grade’. The ordinals in the position of determiners are inflected in agreement with the inflectional paradigm of adjectives ending in a vowel (M.SG *-o*, F.SG *-i*, OBL.F.SG *-a*, PL and OBL.M *-e*), e.g.:

- (48)^{NAR} *má i dujt-i vaj tritt-i ovla bótoškiňa.*
already DEF.F.SG second-F.SG or third-F.SG COP.FUT.3SG salesgirl
Already the second or third (girl) will become a salesgirl.

Ordinals are also inflected when functioning as nouns, as for instance the feminine noun ‘the third one’ in the following example:

- (49)^{LQCR} *ola* *tritt-a-kér-o* *ungriko* *hi*.
 DEF.OBL.F.SG third-OBL.F.SG-GEN-M.SG Hungarian COP.3
 The third one's (husband) is Hungarian.

Ordinals that act as masculine nouns are treated as xenoclitic *o*-masculines (see 4.1.3.3):

- (50)^{LQCR} *le* *tritt-os-kér-i* *gáži* *hi*.
 DEF.OBL.M.SG third-OBL.M.SG-GEN-F.SG non-Romani_woman COP.3
 The third one's (wife) is a non-Romani woman.

The ordinal marker is also attached to the indefinite determiner *ár* (< *aver*) 'other', i.e. *ár-to* (< **aver-to*). This form has been attested only alongside the noun *dí* 'day', i.e. *árto dí* 'next day'. The ordinal marker following *ki'i* 'how much/many' gives rise to the interrogative *ki'i-to* 'which'.

4.4.5 Multiplicatives

Multiplicatives are formed by the suffix *-var*, such as *éfta-var* 'seven times' < *éfta* 'seven'. The final sound is omitted before the application of the suffix in *du-var* < *duj* 'two' and *šó-var* < *šóv* 'six'. Other irregular forms include *tri-jal* 'three times' (< **tri-val* < **trin-var*) and *jefkar* 'once' (< **jek-var*). The former resulted from the sound change *iva* > *ija*, while the latter underwent the metathesis *kv* > *vk* with an additional devoicing. The voiceless counterpart of the multiplicative suffix *-far* is attested only in *but-far* 'many times' (< *but* 'a lot'). The multiplicative marker is encountered in the indefinite numeral *at'i-var* 'so many times' (< *at'i* 'so many/much') and interrogative *ki'i-var* 'how many times' (< *ki'i* 'how many/much'), which alternate with the forms *at'i-rval* and *ki'i-rval*, respectively. The latter forms most probably resulted from an analogy to the form *štár-val* 'four times' (< *štár* 'four'), where the morpheme boundary was re-analysed as *štá-rval* (cf. *štár-val*). As it has been mentioned in section 4.4.2, the tens are based on the multiplicative forms of basic numerals, e.g. *štár-val-děš* 'forty'.

4.4.6 Quantifiers

The quantifiers *čul-o* ‘(a) few, (a) little’ and *but* ‘many, a lot’ are used to modify nouns. The latter is uninflected, while the former agrees in gender, number and case with the noun it quantifies, e.g. *čul-e bal* ‘little-PL hair.PL; little hair’. The uninflected quantifiers *nad'on* (< H *nagyon*) and *igen* (< H *igen*) ‘very, a lot’ and the inherited (*ek*) *klig* and (*ek*) *khajt* ‘(a) few, (a) little’ are used to modify verbs, adjectives and adverbs, e.g. *nad'on šukár hi* ‘lit. very beautiful is; s/he is very beautiful’. The latter two quantifiers have also been attested as noun modifiers, e.g. *ek klig hovéli* ‘a few live coals’, *ek khajt páň* ‘a little water’. Further quantifiers are *aťi* ‘so many/much’, its multiplicative form *aťivar ~ aťival* ‘so many times’, and the compounded *d'ék-duj* ‘lit. one-two; a couple of, some’. The latter form competes with the borrowed *apár* (< archaic H *apár*). For further indefinite numerals see section 4.5.6.

Numerals in other varieties of Vend Romani

Many speakers of Hungarian Vend Romani found it difficult to translate numerals above thirty. By contrast, counting in Romani even for hundreds and thousands seemed to be more natural for the speakers of Prekmurje and Burgenland Romani. Unlike in KR, the conjunction *-o-* is employed in the numbers 11–16 and 21–26 in the neighbouring varieties of Németsfalu (Zala), Szakonyfalu (Vas) and Oberwart (Burgenland), e.g. *deš-o-trin* 13, *biš-o-trin* 23.

The temporal adverb with ordinal-multiplicative meaning *lejsti-moj ~ lajc-muj* ‘last time’, which contains the German noun *Mal* ‘time’, is used only outside Somogy. In Somogy, the Hungarian loanword *utojára* (< H *utoljára*) occurs. The quantifier *nad'on* ‘very’ is more frequent in Somogy, while the quantifier *igen* ‘very’ predominates in the western varieties of Vend Romani. The paucal quantifier ‘a few, a little’ is *khlig* in Zala and Prekmurje Romani, *ekhnaj* in Burgenland Romani, while elsewhere the forms *khlig ~ klig* and *khajt* alternate.

4.5 Pronouns

This chapter introduces the origin, form and position of the personal, reflexive, reciprocal, expletive, demonstrative, interrogative and indefinite pronouns of KR. Given that the formation of demonstrative adverbs is similar to that of the demonstrative pronouns, I will discuss them together in the chapter called Demonstratives. For the same reason, the indefinite

and interrogative adverbs are discussed together with the indefinite and interrogative pronouns, respectively.

4.5.1 Personal and reflexive pronouns

The set of personal pronouns and the respective case paradigms are summarized in Table 47.

	1SG	2SG	3SG.M	3SG.F	1PL	2PL	3PL
NOM	<i>me</i>	<i>tu</i>	<i>óv</i>	<i>ój</i>	<i>amen</i>	<i>tumen</i>	<i>ón</i>
ACC	<i>man</i>	<i>tut</i>	<i>l-e</i>	<i>l-a</i>	<i>amen</i>	<i>tumen</i>	<i>l-en</i>
DAT	<i>man-ge</i>	<i>tu-ke</i>	<i>l-es-ke</i>	<i>l-a-ke</i>	<i>amen-ge</i>	<i>tumen-ge</i>	<i>l-en-ge</i>
LOC	<i>man-de</i>	<i>tu-te</i>	<i>l-es-te</i>	<i>l-a-te</i>	<i>amen-de</i>	<i>tumen-de</i>	<i>l-en-de</i>
ABL	<i>man-dar</i>	<i>tu-tar</i>	<i>l-es-tar</i>	<i>l-a-tar</i>	<i>amen-dar</i>	<i>tumen-dar</i>	<i>l-en-dar</i>
INS	<i>man-ca</i>	<i>tu-ha</i>	<i>l-e(-)ha</i>	<i>l-a-ha</i>	<i>amen-ca</i>	<i>tumen-ca</i>	<i>l-en-ca</i>

Table 47 Personal pronouns

As it can be observed, the gender is distinguished only in the third-person singular. Similarly to the nouns, the personal pronouns have nominative and oblique forms. The first and second-person plural pronouns have homonymous forms for both nominative and oblique. The third-person pronouns have the suppletive oblique stems *les-* (cf. NOM *óv*), *la-* (cf. NOM *ój*) and *len-* (cf. NOM *ón*), which resembles the oblique form of the demonstratives and definite article (Matras 2002: 100). The oblique stem of the pronoun is formally homonymous to the accusative form, except of the second-person singular (i.e. *tu-*, cf. ACC *tut*) and masculine third-person singular (i.e. *les-*, cf. ACC *le*). The Layer II case markers are added to the oblique stem of the pronouns. In non-verbal predications, the non-emphatic clitic pronouns *-lo* ‘he’, *-li* ‘she’ and *-le* ‘they’ are employed (see also 4.7.4).

KR possesses reflexive pronouns only for the third person. These are the accusative singular *pe* (OBL *pes-*) and the accusative plural *pumen* (OBL *pumen-*). The reflexive pronouns are primarily used to express that an action affects the subject itself, e.g. *umláda pe* ‘hang.PRT.3SG REFL.3SG; s/he hanged up himself’. They are also frequently attested in fixed idioms or phrases which are often calqued from Hungarian, as for instance *fít peske žal* ‘away REFL.3SG.DAT go.3SG; goes (for him/herself)’, *lel pe* ‘take REFL.3SG; shakes the dust off his/her feet’, or *mér- peske* ‘die REFL.3SG.DAT; to die (for him/herself)’ (51).

- (51)^{NAR} *o čoro Feri, odá má múlo peske, čoro.*
 DEF poor Feri that.M already die.PRT.3SG REFL.3SG.DAT poor
 Poor Feri, he already died (for himself), the poor guy.

The genitive forms of the personal pronouns are used as possessive pronouns. These pronouns agree in gender, number and case with the possessed object. Only the third person forms are formed regularly, where the genitive marker (SG *-kér-* ~ *-kr-*, PL *-gér-* ~ *-gr-*) is attached to the oblique form of the suppletive stem *l-*. On the other hand, the stems of the personal pronouns are reduced in the first and second persons (1SG *m-*, 2SG *t-*, 1PL *am-* and 2PL *tum-*). The reduced stem, or ‘base stem’ as it is called by Elšík (2000b), is then followed by the irregular genitive marker *-r-* in singular and *-ár-* in plural (Table 48).

	1SG	2SG	3SG.M	3SG.F	1PL	2PL	3PL
GEN	<i>m-r-o</i>	<i>t-r-o</i>	<i>l-es-kér-o</i>	<i>l-a-kér-o</i>	<i>am-ár-o</i>	<i>tum-ár-o</i>	<i>l-en-gér-o</i>
			<i>l-es-kr-o</i>	<i>l-a-kr-o</i>			<i>l-en-gr-o</i>

Table 48 Genitive form of personal pronouns

The genitive form of the reflexive pronoun is *pr-o* ‘his/her own’ in singular, and *pumár-o* ‘their own’ in plural. Both forms are derived from the respective reflexive pronouns analogically with the second-person possessive forms *tr-o* ‘your.SG-M.SG’ and *tumár-o* ‘your.PL-M.SG’. The genitive form of the reciprocal pronoun is formed regularly by the genitive marker *-kér-* ~ *-kr-*, i.e. *ékhávres-kér-o* ~ *ékhávres-kr-o* ‘each_other.OBL-GEN-M.SG; each other’s’ (see 4.5.2).

Personal and reflexive pronouns in other varieties of Vend Romani

Unlike in KR, the subject markers *-lo*, *-li* and *-le* are allowed to be attached only to the present copula form *hi* ‘is’ in the vast majority of varieties of Zala and Prekmurje Romani.

The singular reflexive pronoun *pe* has the alternative variant *pe-t*, which involves the Hungarian accusative suffix *-t* in Kapuvár (Sopron).⁶¹ Moreover, only this innovative variant has been attested in the neighbouring variety of Fertőrákos (Sopron).

In some (especially of southern) peripheral varieties of Somogy, the first and second-person possessive pronouns *mr-o* and *tr-o* may optionally be reduced to *m-o* and *t-o*, respectively, e.g. *mo kher* ‘my house’, *to kher* ‘your house’.

4.5.2 Reciprocal pronouns

The reciprocal pronoun is composed of the numeral *ékh* ‘one’ and the pronoun *áver* ‘other’. It does not have a nominative form, and only the masculine form has been attested in oblique, i.e. *ékhávr-es-*, cf. OBL.F **ékh-ávra-*. The accusative form is the *s*-less *ékhávre* ‘each other (52), one another (53)’.

(52)^{NAR} *na muknahi, hod' len ékhávr-e.*
 NEG allow.IMPF.3PL COMP marry.3PL each_other-ACC.SG
 They did not allow them to marry each other.

(53)^{NAR} *afka úle adala róma, hod' ékhávres-ke na pá'tan.*
 so become.PRT.3PL these Roma COMP one_another-DAT NEG trust
 So it came that these Roma do not trust one another.

The singular form of the reciprocal is used also in plural reference although the plural form *ékhávren-* ‘one another’ has also been attested in the data once.

Reciprocal pronouns in other varieties of Vend Romani

The form of the reciprocal pronoun varies from variety to variety. The most frequent forms are *d'ékh-ávre*, *jékh-ávre*, *ékh-ávre*, *kh-ávre* and the metathesized *kh-ráve*, which are composed of the numeral ‘one’ and the pronoun ‘other’. In contrast, the forms *d'ék-ekh-ávre*, *jék-ekh-ávre*

⁶¹ Alternatively, the form *pe-t* may have emerged analogically with the second-person accusative form *tu-t*, cf. *tu* ‘you’ (Viktor Elšík, personal communication, 28 September 2014).

and the metathesized *jék-ekh-ráve* comprise twice the numeral ‘one’ alongside the pronoun ‘other’.

4.5.3 Demonstratives

In this chapter I will adapt the terminology regarding the demonstratives used in Matras (2002: 103–6) and Elšík and Matras (2006: 75). The basic KR demonstratives distinguish between the two distance scales of proximate and remote demonstratives. The distinction between the two sets of demonstratives lies in the initial vowel component, which is *a-* for proximate and *o-* for remote demonstratives (Table 49).

		SG.M	SG.F	PL
NOM	proximate plain	<i>a-d-á</i>	<i>a-ja</i>	<i>a-d-al-a</i>
	remote plain	<i>o-d-á</i>	<i>o-ja</i>	<i>o-d-ol-a</i>
	proximate specific/contrastive	<i>á-k a-d-á</i>	<i>á-k a-ja</i>	<i>á-k a-dal-a</i>
	remote specific/contrastive	<i>o-k-á</i>	<i>o-k-ija</i>	<i>o-k-ol-a</i>
OBL	Proximate plain	<i>a-d-al-e</i>	<i>a-d-al-a</i>	<i>a-d-al-e</i>
	Remote plain	<i>o-d-ol-e</i>	<i>o-d-ol-a</i>	<i>o-d-ol-e</i>
	proximate specific/contrastive	<i>á-k ad-al-e</i>	<i>á-k ad-al-a</i>	<i>á-k ad-al-e</i>
	remote specific/contrastive	<i>o-k-ol-e</i>	<i>o-k-ol-a</i>	<i>o-k-ol-e</i>

Table 49 Demonstrative pronouns

The vowel morpheme indicating distance is prefixed to the root. The form of the root is *-d-* for general deictic reference, and *-k-* for specific or contrastive reference. The specific demonstrative has the meaning ‘precisely/exactly this/that’ (54), while the contrastive demonstrative may be translated as ‘this/that other’ (55).

- (54)^{NAR} *ár len phost'áren taj o gojó ár čhiden, oká han.*
 out 3PL.ACC gouge.3PL and DEF eyeball out throw.3PL exactly_that.M eat.3PL
 They gouge out (the swine’s eyes) and throw out the eyeball, and that is (exactly) what they eat.

- (55)^{LQCR} *adá kašt mod'orójitiko hi, oká meg akhoritiko.*
 this.M tree hazel-nut COP.3 that_other.M and walnut
 This is a hazel-nut tree, and that other is a walnut tree.

The nominative plural and oblique stems *adal-* and *odol-* may optionally be syncopated to *adl-* and *odl-*, respectively. In addition, the consonant cluster *dl* may be assimilated to *ll*, i.e. *all-* and *oll-*.

KR possesses a set of genuine specific/contrastive demonstratives only for remote distance, i.e. demonstratives having the stem *ok-*. As for the proximate distance, the phrase comprising the static deictic *ák* ‘here’ and the respective demonstrative determiner⁶² is applied, such as *ák adá* (cf. **aká*) in (56).

- (56)^{NAR} *akébor lo sin sar ak_adá.*
 such_size 3SG.M COP.3 like exactly_this
 S/he was such size like (exactly) this one (here).

Table 50 represents the set of KR deictics according to their stems:

	AD-	AK-	OD-	OK
determiner	<i>ad-á</i>	<i>ák ad-á</i>	<i>od-á</i>	<i>ok-á</i>
spatial stative deictic	<i>ed-ej</i> ‘here’	<i>ák</i> ‘here!’	<i>od-oj</i> ‘there’	<i>ók</i> ‘there!’
spatial directive deictic	(see below, <i>órde</i>)		<i>ód'-a</i> ‘thither’	
spatial separative deictic	<i>á-thar</i> ‘from here’		<i>ó-thar</i> ‘from there’	
size	<i>ad-ebor</i> this/that size	<i>ak-ebor</i> this/that size		–
side	?	<i>ak-arig</i> on this side	<i>od-orig</i> on that side	<i>ok-orig</i> on that side

Table 50 Deictics

⁶² The demonstrative determiners are demonstratives that modify nouns. They are used attributively, and as nouns as well.

The distance meaning is realized on the stem in the demonstrative determiners, stative and separative deictics and in the demonstratives for the lexicalized meaning ‘(on) this/that side’. By contrast, the demonstratives indicating size refer to both proximal and remote distance, while having the initial vowel *a* typical only for proximal distance. The separative deictics employ the historical *a*-stem in proximal and *ó*-stem in remote distance. Presumably, the temporal adverb *ok-ondak* ‘last time’ comprises also the inherited *ok*- stem, which precedes the borrowed adverb *onda-k* (< S dial. *ondak*, cf. standard *onda*) ‘then’.

Furthermore, only the *a*- stem appears in the demonstrative expressing quality (*a-só* ‘such’), manner (*a-fka* ‘so’), quantity (*a-ti* ‘so much/many’), as well as in the multiplicative demonstrative *a-tivar* ~ *a-tirval* ‘so often’. The directive deictic of proximal distance is *ó-rde* ‘hither’, which is borrowed from Ossetian (Matras 2002: 24). The prolative demonstratives are the recently borrowed *ere* (< Hung. *erre*) ‘this way’ and *ora* (< Hung. dial. *óra*) ‘that way’, and their separative forms *er-al* ‘from this way’ and *or-al* ‘from that way’, being formed by the old ablative suffix *-al*.

Like in Hungarian (Kenesei 1998: 276), the demonstrative pronoun compounded with the particle *ud'an* denotes identity in KR, such as in *ud'an-odá* ‘the same’, *ud'an-odoj* ‘on the same place’, *ud'an-óthar* ‘from the same place’, *ud'an-akebor* ‘the same size’, *ud'an-asó* ‘the same (quality)’, etc.

Demonstratives in other varieties of Vend Romani

The set of demonstrative pronouns in other varieties of Vend Romani differs mainly in that the plural and oblique form of demonstrative pronouns with *-d-* stem may be syncopated only in KR and in a few other varieties of Somogy Romani (i.e. *adl-* and *odl-*), while elsewhere only the full stem is allowed, i.e. *adal-* and *odol-*. The reduced stem in the separative deictics *á-thar* ‘from here’ *ó-thar* ‘from there’ is in particular characteristic to the varieties in the central part of Somogy, cf. with *adá-thar* and *odó-thar* found in other Vend Romani varieties.

The form of the stative deictic is the inherited *adaj* ‘here’ in a few peripheral varieties of Somogy Romani and in some western varieties of Vend Romani (Vas, Veszprem, Burgenland Romani), while in other varieties the forms *adej* or *edej* are preferred. The sharp distinction between the directive and stative orientations in stative deictics (i.e. *ód'a* ‘to there’ vs. *odoj* ‘there’; *órde* ‘to here’ vs. *adaj* ‘here’) is blurred in the peripheral variety of Vásárosdombó (Baranya), where the distal deictic *odoj* indicates both directive and stative

orientation (i.e. *odoj* ‘(to) there’), while the proximal deictic pair can be used interchangeably (i.e. *órde* ~ *adej* ‘(to) here’). It seems that Vásárosdombó has been influenced by its adjacent variety of Versend (Baranya), where both directive and stative orientations are expressed by the deictic *adaj* for proximal and *odoj* for distal localization (Bodnárová 2009: 60). A similar on-going development is attested also in Burgenland and Prekmurje Romani.

4.5.4 Expletives

The use of expletives is quite widespread in KR. They occur in hesitation pauses when the speaker searches for an appropriate word, or they substitute certain parts of speech, but not contributing to the meaning of the sentence. Table 51 presents the expletive pronouns, nouns and adjectives found in KR.

EXPLETIVE	SG.M	SG.F	PL
pronoun	<i>oko</i>	<i>oko</i> ~ <i>oki</i>	<i>oko</i> ~ <i>okój</i>
noun	<i>kova</i>	<i>kova</i> ~ <i>kój</i>	<i>kova</i> ~ <i>kój</i>
adjective	<i>kóvašno</i>	<i>kóvašni</i>	<i>kóvašne</i>
	<i>kójošno</i>	<i>kójošni</i>	<i>kójošne</i>

Table 51 Expletives

The expletive pronoun *oko*, which is overrepresented in my data, often appears as a redundant element of the clause. That is, it does not perform a syntactic role, and it is not used in hesitation pauses either. I suppose that the use of this pronoun is connected to certain style of speech. Consider the example (57)a, where *oko* takes only the place of the definite article, cf. (57)b.

(57)^{LQCR}

- a. *lačho than hi tut and' oko fóro.*
 good place COP.3 2SG.ACC in EXPLtown
- b. *lačho than hi tut and-o fóro.*
 good place COP.3 2SG.ACC in-DEF.M.SG town
 You have a good job in the town.

The expletive pronoun *oko* resembles the demonstrative determiner with remote specific/contrastive meaning, i.e. *oká* (see 4.5.3). Indeed, both types of demonstratives share the same stem *ok-*, but their developments slightly differ. The masculine singular form of the expletive pronoun resulted from apocope (*oko* < **oko-va*), while the respective demonstrative pronoun from contraction (*ok-á* < **ok-ova*). The feminine singular expletive form *oki* is probably the reduced form of **o-k-i-ja*, and the plural form *ok-ó-j* emerged from the form **ok-o-vi*.

It is noteworthy that the expletive pronouns (58), as well as the expletive nouns (59), are ceasing to be inflected in KR.

- (58)^{LQCR} *o phábi upral oko phabalind'a kédas.*
 DEF apple.PL from EXPLapple-tree.PL pick.1PL
 The apples we used to pick from the apple-tree.

- (59)^{LQCR} *fer čhite i phuv valasaj kova, vaj buzaha, vaj árpaha.*
 away put.PRT.3PL DEF field some.PL EXPL or wheat.INS or barley.INS
 They seeded the fields with some thingummy, or with wheat, or with barley.

The expletive adverb has the form *kój-itik-ón*, being formed by both the derivational suffix of adjectives *-itik-* and adverbs *-ón* from the stem *kóv-*, i.e. **kóv-itik-ón* > *kój-itik-ón*. On the other hand, there has not been attested an expletive verb with *-k-* stem in my data. Instead, the verb *kér-* ‘to do, make’ is used as an expletive. Consider the following example where the expletive *kér-* substitutes the verb *žangav-* ‘to wake’ in (60), and the verb *astár-* ‘to hold’ in (61).

- (60)^{LQCR} *sako ra'aha éfta órengé kэрav upre le fat'ú.*
 every morning seven o'clock.DAT do.1SG up DEF.OBL child.ACC
 Every morning at seven I wake my child up.

- (61)^{NAR} *tél len erősakolinen, tél kéren, tél pumen lenca piján.*
 down 3PL.ACC rape.3PL, down do.3PL, down REFL.3PL 3PL.INS do_fellatio.CAUS.3PL
 They rape them, (hold) them down, and force them to have oral sex.

Expletives in other varieties of Vend Romani

The expletive pronoun *oko* is found only in Somogy Romani. On the other hand, the expletive noun *kova* (occasionally *kój* in F and PL) is also present in Zala, Vas and Veszprém Romani. The expletive *ókaj* found in the adjacent varieties of Kálmánca and Homokszentgyörgy (both Somogy) seems to be only a hesitation word, as it can be followed by any kind of part of speech. The verb *kér-* is used as an expletive in other varieties of Vend Romani as well.

4.5.5 Interrogatives

The personal interrogative pronoun *ko* ‘who’ and the impersonal *so* ‘what; which’ have only singular forms. The former has the irregular oblique form *kas-* (cf. **kos-*, see also ACC *kas* ‘whom’), and the irregular instrumental form *kas-a-ha* (cf. **ka(-)ha*) ‘with whom’. The further inflected forms of this pronoun are regular, such as the genitive *kas-kér-o* ‘whose’. The impersonal *so* has the regular oblique form *sos-*. The inflectional forms of this pronoun include, for instance, the instrumental *so(-)ha* ‘with what’, the dative *sos-ke* ‘why’ which refers to cause and reason, and the locative prepositional form *vaš sos-te* ‘for what’. The interrogative denoting quality is *savo ~ só* (F/PL *saj*) ‘what, what kind of’ with the non-contracted oblique forms *sav-es-* (OBL.M.SG), *sav-a-* (OBL.F.SG), *sav-en-* (OBL.PL). As data suggest, the interrogatives referring to size (*kébor* ‘what size’), quantity (*kiṭi* ‘how much/many’), location (*káj* ‘where, to where’ and *khatar* ‘from where’), and manner (*sar* ‘how’) are uninflected. Further uninflected interrogative is *kada* ‘when’ which most probably arose through the contamination of the original **kana* by the Slavic *kad* ‘when’. KR also borrowed the interrogative *sajt* (< G *seit*) ‘since, since when’ from German, and some others from Hungarian, such as the temporal *meddig* (< H *meddig*) ‘how long, till when’, *háñadika* (< H *hányadika*) ‘what day of the month’, *mijuta* (< H *mióta*), and the local *mere* (< H *merre*) ‘which way’. The ordinal *kiṭi-t-o* ‘which’ and the multiplicative *kiṭi-var ~ kiṭi-rval* ‘how many times, how often’ are derived from the interrogative *kiṭi* ‘how much/many’. The borrowed *mer-e* ‘which way’ take the ablative form *mer-al* for the meaning ‘from where, from which side’. The compounded interrogative *sar-hod’* ‘how come’ (semi-)calques the Hungarian *hogy-hogy* ‘lit. how-that’, e.g.:

(62)^{NAR} *t' akor sar-hod' órde ájal?*
 and then how_come here come.PRT.2SG
 And then how come you came here?

The origin of the temporal interrogative *kiriňňa* ‘how long, till when’, which alternates with *meddig*, is unclear. It supposedly developed from the contamination of the inherited *ki-t'i* ‘how much/many’ and the Hungarian *a-nyira* ‘so much, to such extent’ with an additional metathesis, i.e. **ki-ňňira > ki-riňňa*.

The vast majority of interrogatives are also used as relativizers in KR (see 5.3.3).

Interrogatives in other varieties of Vend Romani

In Lengyeltóti (Somogy), the interrogative form *kada* competes with the form *kad* ‘when’. The interrogative *kiriňňa* (also *kiriňa* and *kirija*) ‘how long, till when’ has been attested in most varieties of Vend Romani, except of in Zala and Prekmurje Romani. Zala Romani makes use of the form (*ži*) *kiňi-dig*⁶³ in the meaning ‘till when’, and the form *só dugo* ‘lit. what long’ in the meaning ‘how long’. In Prekmurje Romani, the former meaning is expressed by the prepositional phrase *ži kada* ‘lit. until when’, and the latter by the phrase *sar dugo* ‘lit. how long’. The Zala Romani interrogative *ki'i-tik-o* ‘which; what day of the month’ is formed by means of the Hungarian adjectival suffix *-itik-* in contrast to the ordinal form found in KR, i.e. *ki'i-t-o*.

4.5.6 Indefinites

The specific, free-choice and negative indefinites are formed by the borrowed morphemes *vala-* (< H *vala-*), *akár-* (< H *akár-*) and *ni-* (< S *ni-*), respectively, by prefixing to the respective pronouns and adverbs (Table 52).

⁶³ This form has probably emerged from the contamination of *kiňi-ra* (= *ki-t'i* ‘how many/much’ + *a-ňira* ‘to such extent’) and the Hungarian interrogative *med-dig* ‘till when’.

	SPECIFIC	FREE-CHOICE	NEGATIVE
who	<i>vala-ko</i>	<i>akár-ko</i>	<i>ni-ko</i>
what	<i>vala-so</i>	<i>akár-so</i>	<i>ništ(a)</i>
which	<i>vala-só</i>	<i>akár-só</i>	<i>ni-só</i>
how	<i>vala-sar</i>	<i>akár-sar</i>	<i>ni-sar</i>
(to) where	<i>vala-kháj</i>	<i>akár-kháj</i>	<i>ni-kháj</i>
from where	<i>vala-khatar</i>	<i>akár-khatar</i>	<i>ni-khatar</i>
which way	<i>valamere</i>	<i>akármere</i>	?
when	<i>vala-kada ~ valamikor</i>	<i>akár-kada ~ akármikor</i>	<i>šoha</i>
how many/much	<i>vala-ki'i</i>	<i>akár-ki'i</i>	<i>ni-ki'i</i>

Table 52 Specific, free-choice and negative indefinite pronouns and adverbs

On the other hand, the specific indefinites *valamere* (< H *valamerre*) ‘to some way’ and *valamikor* (< H *valamikor*) ‘sometime’, the free-choice indefinites *akármere* (< H *akármerre*) ‘to whichever way’ and *akármikor* (< H *akármikor*) ‘whenever’, and the negative indefinites *ništ ~ ništa* (< S *ništa*) ‘nothing’ and *šoha* (< H *šoha*) ‘never’ are borrowed.

The indefinites referring to something ‘other’ are formed either by the contracted *ár* (< *aver*) or the syncopated *avr-i* (< OBL **aver-e*; possibly with the final sound change **avr-e* > *avr-i*), as it is shown in Table 53.

	UNIVERSAL	OTHER
how	?	<i>avri-jal, síš</i> < G dial. <i>sist</i>
(to) where	<i>víd'ik</i>	<i>ár-thán</i>
from where	<i>víd'ikal</i>	<i>ár-thánal</i>
when	<i>mindig</i>	<i>avri-kor</i>

Table 53 Universal and ‘other’ indefinite adverbs

The extracted suffix *-kor* (< H *-kor*) occurs in the indefinite *avri-kor* ‘another time’, while the origin of the suffix *-jal* in *avri-jal* ‘another way’ is unclear.

The universals⁶⁴ are the borrowed *vid'ik* (< H *végig*) ‘everywhere’ and *mindig* (< H *mindig*) ‘always’, and the internally derived *vid'ik-al* ‘from everywhere’ (< *vid'ik* ‘everywhere’). Further universal indefinites are *sako* ‘everyone’, *sako d'ékh* ‘each one’ and *cile* ‘everyone’ in personal reference, and the forms *sa* ‘everything’ and *sako d'ékh* ‘everything’ in impersonal reference (Table 54).

	UNIVERSAL	OTHER
Personal	<i>sako</i> <i>sako d'ékh</i> <i>cile</i>	<i>áver ~ ár</i>
Impersonal	<i>sa</i> <i>sako d'ékh</i>	<i>áver ~ ár</i>

Table 54 Universal and ‘other’ indefinite pronouns

On the other hand, the universal quantifiers used in both personal and impersonal reference are the inflected *sak-o* ‘every, each’, *sak-o* (OBL *sak-on-*) *d'ékh* ‘each, every single’ and *cil-o* ‘whole, all’, while the universal determiner of numerals is the uninflected *so* ‘all’, e.g. *so trin phrála* ‘all three brothers’ (Table 55).

	UNIVERSAL	OTHER
(im)personal	<i>sak-o</i> <i>sak-o d'ékh</i> <i>cil-o</i>	<i>aver ~ ár</i>
before numeral	<i>so</i>	–

Table 55 Universal and ‘other’ determiners

The indefinite determiner and nominal for the meaning ‘(an)other’ is *áver ~ ár*, as it is shown in Table 54–55.

⁶⁴ Universal pronouns refer to an entire group or any member of that group. They correspond to the series of pronouns *every-* found in English.

Indefinites in other varieties of Vend Romani

In Hungarian Vend Romani, the marker of free-choice indefinites is the Hungarian-borrowed *akár-* (*kár-* in Prekmurje and Burgenland Romani), which occasionally alternates with the also Hungarian-borrowed *bár-* (< *bár-*). Furthermore, the South Slavic free-choice marker *-gudi* (< S *-god*) is used in some southern varieties of Prekmurje.

In place of the KR indefinite *avrijal* ‘otherwise’, we find *amut* (< H *amúgy*) in Szakonyfalu (Vas) and in Zala and Prekmurje Romani, and *árčán* (< **aver-čand-es*) in Veszprém Romani. The origin of the form *čand* is unknown. In several varieties of Vend Romani, the universal indefinite has the form *sake-thán* ‘lit. every-in_place; everywhere’ besides *vid'ik*.

4.6 Articles

The definite article is declinable, and the masculine singular form is homonymous with the plural form (see Table 56).

	M.SG	F.SG	PL
NOM	<i>o</i>	<i>i</i>	<i>o</i>
OBL	<i>(o)-l-e</i>	<i>(o)-l-a</i>	<i>(o)-l-e</i>

Table 56 Definite article

On the other hand, feminine nouns require the definite article *i* in nominative singular, and the article *la ~ ola* in oblique singular. The oblique stem *l-* predominates over the stem *ol-*. However, further research is needed to determine whether there is some functional difference between the two stems. The definite article obligatorily accompanies proper nouns (63), which is only an optional possibility in Hungarian, the dominant contact language of KR.

- (63)^{NAR} *i* *Melinda sin,* *o* *Gusztí, taj odá* *Marcel.*
DEF.F.SG Melinda COP.PRT.3, DEF.M.SG Gusztí and that.M Marcel
There was Melinda, Gusztí and that one, Marcel.

Furthermore, the prepositional phrases generally comprise definite article, such as *andral o ninčko* ‘from DEF.M.SG Germany; from Germany’. The article causes that the final sound is dropped in prepositions in final *-e* (cf. Elšík et. al. 1999: 375). In addition, the definite article has been sporadically found in prepositional phrases where the noun is determined by a numeral (64), which is a rather unusual construction in other South Central varieties (cf. *ibid.*):

(64)^{NAR} *tel o éfta berš, tel o éfta dí.*
 during DEF.PL seven year during DEF.PL seven day
 During seven years and seven days.

The indefinite article is represented by the indeclinable *ék*⁶⁵ (65), having arisen from the numeral *jékh ~ d'ékh* ‘one’ following the loss of the initial sound *j ~ d'*.

(65)^{NAR} *duj még hi ék ár muršestar.*
 two more COP.PRS.3SG a other man.ABL
 There are two more (children) of another man.

While the numeral ‘one’ may optionally have the aphaeresised form *ékh* as a determiner, the indefinite article does not seem to take the form *jékh* or *d'ékh* with initial prothesis.

Articles in other varieties of Vend Romani

The *ol-* variant of the definite article’s oblique stem is absent in most varieties of Vend Romani, but it occurs sporadically in a few Somogy Romani varieties.

⁶⁵ The indefinite article *ék* is not inflected and the aspiration is generally lost in the word-final position (see 3.1.2), therefore I transcribe it without aspiration. In contrast, I transcribe the numeral *jékh ~ d'ékh ~ ékh* with final aspirate, as the aspiration is preserved in its inflected forms, e.g. SG.M.OBL *jékhes-*.

4.7 Verbs

4.7.1 Adaptation of verbs

The marker *-ín-* is a means to adapt verbs of South Slavic, German and Hungarian origin, e.g. *pis-ín-* < S *pis-ati* ‘to write’, *fíš-ín-* < G *fisch-en* ‘to fish’, *pihen-ín-* < H *pihen* ‘to rest’, *érez-ín-* < H *érez* (cf. inflectional stem *érez-*) ‘to feel’. This adaptation marker is attached to the inflectional stem of the source form. The adaptation marker has the form *-án-* in *tres-án-* (< S *tres-ti* ‘to tremble’), which could have also resulted from the contamination of the common marker *-ín-* (i.e. **tres-ín-* ‘to tremble’) by the stem-final *a* of the vocalic verb *rezda-* (cf. PRT *rezdá-n-*) with the same meaning. The marker *-isaj-* is generally used to adapt middle verbs (see 4.7.2.6) originating in borrowed verbs, e.g. *kezd-isaj-ov-* < *kezd-ín-* (< H *kezd*) ‘to start’.

Adaptation of verbs in other varieties of Vend Romani

The marker to adapt loan-verbs is *-ín-* also in other Vend Romani varieties. In addition to *tres-án-* (see above), the adaptation marker *-(j)án-* is found in the verb *vrišt-án-* ~ *vrišč-án-* (< S *vrisk-ati*) ‘to scream’ in most Vend Romani varieties, and in the verb *tin-án-* (of unknown origin) ‘to tremble’ in the Somogy Romani varieties of Kaposmérő and Görgeteg.

4.7.2 Verb formation

The formation of new verbs is primarily based on suffixation in KR. The inherited suffixes are employed in causative (see 4.7.2.2), factitive (see 4.7.2.3), iterative (see 4.7.2.5) and middle verb forms (see 4.7.2.6), while the borrowed suffixes are preferred in denominal derivations (see 4.7.2.4). Compounding is not a productive means to form verbs in KR, as most of the compounds are either inherited or calqued from Hungarian (see 4.7.2.8).

4.7.2.1 Imported derivational markers

A number of Hungarian derivational markers were imported (within lexical borrowings) into KR through borrowings, such as the inchoative *-od-* ~ *-ed-* and *-ul-*, the reflexive *-koz-* and *-kod-* ~ *-ked-* ~ *-köd-*, and the markers of denominal derivations *-(V)l* and *-(V)z*:

-od- < H -od-	<i>józan-od-ín-</i>	< H <i>józan-od-ik</i> ‘to sober’
	‡ <i>*józan-no</i>	‡ H <i>józan</i> ‘sober’
-ed- < H -ed-	<i>telep-ed-ín-</i>	< H <i>telep-ed-ik</i> ‘to settle’
	‡ <i>telep</i>	‡ H <i>telep</i> ‘settlement’
-ul- < H -ul-	<i>ném-ul-ín-</i>	< H <i>ném-ul</i> ‘to become mute’
	‡ <i>ném-asto</i>	‡ H <i>ném-a</i> ‘mute’
-koz- < H -koz-	<i>čodá-koz-ín-</i>	< H <i>csodál-koz-ik</i> ‘to wonder’
	‡ <i>*čodál-ín-</i>	‡ H <i>csodál</i> ‘to admire’
-kod- < H -kod-	<i>panas-kod-ín-</i>	< H <i>panasz-kod-ik</i> ‘to complain’
	‡ <i>*panasol-ín-</i>	‡ H <i>(el)panasz-ol</i> ‘to complain of’
-ked- < H -ked-	<i>kétel-ked-ín-</i>	< H <i>kétel-ked-ik</i> ‘to be in doubt’
	‡ <i>*kétel-ín-</i>	‡ H <i>kétel-i</i> ‘to doubt’
-köd- < H -köd-	<i>d’ülöl-köd-ín-</i>	< H <i>gyűlölköd-ik</i> ‘to feel hate’
	‡ <i>*d’ülöl-ín-</i>	‡ H <i>gyűlöl</i> ‘to hate’
-(V)l- < H -(V)l-	<i>horg-ol-ín-</i>	< H <i>horg-ol</i> ‘to crochet’
	‡ <i>*horg-ol</i>	‡ H <i>horog</i> ‘hook’
-(V)z- < H -(V)z-	<i>ut-az-ín-</i>	< H <i>ut-az-ik</i> ‘to travel’
	‡ <i>*út-o</i>	‡ H <i>út</i> ‘road, way’

Imported are also the Hungarian factitive (-*itt-*), causative (-*tat-* ~ -*tet-*), and iterative suffixes (-*gat-* ~ -*get-*):

-itt- < H -ít-	<i>sor-itt-ín-</i>	< H <i>szor-ít</i> ‘to press’
	‡ <i>*sor-ul-ín-</i>	‡ H <i>szor-ul</i> ‘to press’
-tat- < H -tat-	<i>šajná-tat-ín-</i>	< H <i>sajnál-tat</i> ‘to make so. feel sorry’
	‡ <i>šajnál-ín-</i>	‡ H <i>sajnál</i> ‘to feel sorry’
-tet- < H -tet-	<i>legel-tet-ín-</i>	< H <i>legel-tet</i> ‘to (let) graze’
	‡ <i>*legel-ín-</i>	‡ H <i>legel</i> ‘to graze’
-gat- < H -gat-	<i>hal-gat-ín-</i>	< H <i>hall-gat</i> ‘to listen’
	‡ <i>*hall-ín-</i>	‡ H <i>hall</i> ‘to hear’

-get- < H -get- *besé-get-ín-* < H *beszél-get* ‘to talk’
 † **besél-ín-* † H *beszél* ‘to speak’

4.7.2.2 Causatives

Causatives are transitive verbs which are derived from verbal roots (Hübschmannová & Bubeník 1997: 135), expressing that the agent makes another participant perform an action. More precisely, the causee is the subject of causatives derived from transitive verbs, and it is the subject-agent of causatives derived from intransitive verbs (Matras 2002: 121). The causatives in KR are formed by means of the inherited marker *-áv-* from consonantal verbs (a, b), and by the marker *-v-* from vocalic verbs (c, d). These markers are added to the root of transitive (b) or intransitive verbs (a, c, d).

- a. *khel-áv-* ‘to make so. dance’ < *khél-* ‘to dance’
- b. *an-áv-* ‘to send so. for sth.’ < *án-* ‘to bring’
- c. *asá-v-* ‘to make so. laugh’ < *ása-* ‘to laugh’
- d. *rezdá-v-* ‘to make so. tremble’ < *rezda-* ‘to tremble’

The causative marker is added to the non-contracted stem in case of contracted verbs, e.g. *phrav-áv-* ‘to make so. open’ < *phráv-* ‘to open’ or *pij-áv-* ‘to make so. drink’ < *pij-* ‘to drink’. The verb *ha-* ‘to eat’ has the irregular causative form *ha-h-áv-* ‘to feed’, where the root *ha-* became reduplicated. Irregular is further the causative form *l-ev-áv-* ‘to make so. take (a picture)’ derived from the verb *l-* ‘to take, carry’, because the present stem is extended by *-ev-* before the application of the causative suffix.

The inherited derivational marker *-áv-* is also applied to several Hungarian-borrowed causative forms consisting of the Hungarian causative marker *-tat-* ~ *-tet-*:

šajná-tat-in-áv- < H *sajnál-tat* ‘to make so. feel sorry’
 † *šajnál-ín-* † H *sajnál* ‘to feel sorry’
dobú-tat-in-áv- < H *dobol-tat* ‘to make so. drum’
 † **dobul-ín-* † H *dobol* ‘to drum’

These forms express single causation, although they involve both the Hungarian and the inherited causative markers. Derivations where only the Hungarian causative marker *-tat-* ~ *-tet-* appears, e.g. *šajná-tat-ín-* ‘to make so. feel sorry’, are less common. These forms were obviously borrowed with causative meaning. The second or double causatives, which are causatives derived from causatives, are unattested in my KR data, though they are common in the northern varieties of South Central Romani (Hübschmannová & Bubeník 1997: 142).

Some causatives are formed by the marker *-ár-* ~ *-jár-* ~ *-tár-* ~ *-dár-* which is otherwise a common derivational marker of factitives (see 4.7.2.3). This marker triggers palatalisation on the morpheme boundary (see 3.1.7), except of in *beš-ár-* (f, cf. e).

- e. *rov-d'ár-* ‘to make so. cry’ < *rov-* ‘to cry’
sov-d'ár- ‘to make so. sleep’ < *sov-* ‘to sleep’
- f. *beš-ár-* ‘to make so. sit’ < *běš-* ‘to sit’

4.7.2.3 Factitives

Factitives are transitive verbs which are derived from nouns and adjectives (Hübschmannová & Bubeník 1997: 135). They generally indicate that the agent causes a change of state in another participant. The most common marker to derive factitives in KR is *-ár-* ~ *-jár-* ~ *-tár-* ~ *-dár-* (see 3.1.7). Examples of deadjectival factitives are:

- tat'ár-* ‘to warm’ < *tát-o* ‘warm’
loj-ár- ‘to redden’ < *lól-o* ‘red’
šuž-ár- ‘to clean’ < *šúž-o* ‘clean’
khamň-ár- ‘to make pregnant’ < *khámn-i* ‘pregnant’

De-nominal derivations are less frequent, represented e.g by *hev-d'ár-* ‘to hole’ (< *hév* ‘hole’), *lon-d'ár-* ‘to salt’ (< *lon* ‘salt’), or *khaň-ár-* (< *khan* ‘smell’) which has the lexicalized meaning ‘to fart’. The meaning of the factitive *nand'ár-* ‘to bath’ also differs from the corresponding adjective *náng-o* ‘naked’. The factitive marker has been preserved in the adjective *kerd'ár-d-o* ‘hot’, with no corresponding factitive verb form: **kerd'ár-*, cf. the middle verb *kerd'ov-* ‘to boil’.

The use of the factitive marker *-isár-* is limited. It has only been attested in *rajn-isár-* ‘to clean’ which draws on the German-borrowed adjective *rajn-i* ‘clean’, and in *harn-isár-* ‘to shorten’, having been derived from the inherited adjective *hárn-o* ‘short’. The regular form *harň-ár-* is unattested in the data.

Some factitives have causative meaning, such as *terd'-ár-* ‘to make so. stop’ (< *térd-o* ‘standing’) and *pašt'-ár-* ‘to lay so. down’ (< *pášt-o* ‘lying’).

4.7.2.4 *De-nominal derivations*

The denominal marker *-ál-* (< H *-VI-*) was extracted from Hungarian borrowings, together with the adaptation marker *-ín-*. The formant *-ál-ín-* is productive, as it also derives verbs from pre-Hungarian stems: *máčh-ál-ín-* ‘to fish’ < *máčh-o* ‘fish’, *pišot-ál-ín-* ‘to play accordion’ < *pišot* ‘accordion’, *šój-ál-ín-* ‘to whistle’ < *šój* ‘whistle’, *grábl-ál-ín-* ‘to rake’ < S *grablj-e* ‘rake’.

The adaptation markers may also be involved in denominal derivations. The marker *-ín-* is used to derive verbs from the nouns *čohán-i* ‘witch’ (> *čohaň-ín-* ‘to bewitch, charm’), *kárd-i* ‘shot’ (> *kárd'-ín-* alongside *kárd'-ál-ín-* ‘to shoot’) and *ásv-in* ‘teardrops’ (> *ásv-ín-* ‘to shed tears’), while the marker *-án- ~ -ján- ~ -d'án-* is involved in the denominal factitive derivations *hoj-án-* ‘to annoy’ (< *hól-i* ‘anger’), *loš-án-* ‘to be happy, glad’ (< *loš* ‘happiness’), and *thuv-d'án-* ‘to smoke’ (< *thuv* ‘smoke’).

4.7.2.5 *Iteratives*

Iteratives are verbs derived from verbs, which indicate that an action occurs frequently, or it is intensive or long-lasting. The iterative markers in KR are the inherited *-kér-* and *-gér-*. These markers are extended with the morpheme *-(i)n-*, originally an adaptation marker (see 4.7.1), when attached to inherited verb stems. More precisely, the vocalic verbs take the iterative formant *-n-gér-* (e.g. *urd'a-ngér-* ‘to flit, fly about’ < *urd'a-* ‘to fly’), while the consonantal verbs take one of the formants *-in-gér-*, *-in-kér-* and *-gér-* (Table 57). There are no examples of iterative derivations from middle verbs in my data.

	MONOSYLLABIC	POLYSYLLABIC
Inherited	<i>-in-gér-</i>	<i>-in-kér-</i>
Borrowed (<i>-ín-</i>)	– ⁶⁶	<i>-gér-</i>

Table 57 Iterative markers of consonantal verbs

Two factors play a role in the choice of the iterative marker: a) the number of syllables in the verb stem and b) whether the word is inherited or borrowed. Inherited verbs with monosyllabic stem take the formant *-in-gér-* (a), and those with more than one syllable in their stem's the formant *-in-kér-* (b).

- a. *khos-ingér-* ‘to wipe repeatedly’ < *khós-* ‘to wipe’
phuč-ingér- ‘to ask repeatedly’ < *phúč-* ‘to ask’
náš-ingér- ‘to run about’ < *náš-* ‘to run’
- b. *čhun.gar-inkér-* ‘to spit repeatedly’ < *čhun.gár-* ‘to spit’
ter.d'ar-inkér- ‘to stop repeatedly’ < *ter.d'ár-* ‘to stop’

The derivational marker is added to the non-contracted stem of contracted verbs, such as in *sikav-inkér-* ‘to teach from time to time’ < *sikav-* ‘to teach’, *čalav-inkér-* ‘to hit repeatedly’ < *čalav-* ‘to hit’, but not in *phos-ingér-* ‘to stab repeatedly’ < *phosav-* ‘to stab’, where the marker typical for inherited monosyllabic verbs is applied. The iterative form *čhin-gér-* is irregular, having been derived from the monosyllabic inherited verb *čhín-* ‘to cut’. The expected form would be **čhin-in-gér-*. The irregular iterative suffix *-ér-* is used to derive the iterative form *phag-ér-* ‘to break repeatedly’ from the verb *phág-* ‘to break’.

Loanwords adapted by *-ín-* take the iterative marker *-gér-*, e.g.:

- fárin-gér-* ‘to jump repeatedly’ < *fárín-* ‘to jump’
pruťin-gér- ‘to kick repeatedly’ < *pruťín-* ‘to kick’
truskin-gér- ‘to sneeze repeatedly’ < *truskín-* ‘to sneeze’

⁶⁶ The loan-verbs are always polysyllabic, since the adaptation marker *-ín-* constitutes an additional syllable.

The non-extended form of the marker *-kér-* has also been retained in the *d*-verbs⁶⁷ *čumi-kér- com* ‘to kiss repeatedly’ (< *čumíd-* ‘to kiss’) and *čhi-kér-* ‘to throw repeatedly’ (< *čhíd-* ‘to throw’), where the stem-final *d* is dropped. In contrast, the inherited monosyllabic *d*-verb *cíd-* ‘to pull’ takes the suffix *-inkér-*, which marker is typical of polysyllabic verbs, i.e. *cid-inkér-* ‘to pull repeatedly’. The motivation for using this irregular form is unclear.

The German-origin verbal particle *um* is also used to provide the verb with iterative meaning (see 4.7.6).

4.7.2.6 Middle verbs

Middle verbs are formed by the markers *-ov-* ~ *-jov-* and *-isaj-ov-* in KR. They function mostly as anticausatives and passives when derived from verbs, and inchoatives when derived from adjectives (Elšík & Matras 2006: 211). Middle verbs are intransitive verbs, except of *haj-ov-* ‘to understand’ (ibid: 432) and a few other middle forms accompanied by a verbal particle, which calque a corresponding Hungarian expression, e.g. *neki térd’-ov-* ‘lit. against stand; to set about’ in (66).

- (66)^{RM} *i čhaj neki_terđini te phagérel o akhora.*
 DEF girl set_about.PRT.3SG.F COMP crack.INF DEF walnut.PL
 The girl set about to crack the walnuts.

The marker *-ov-* ~ *-jov-* is generally used to derive middle verbs from inherited participial (a) and adjectival stems (b), while its extended form *-isaj-ov-* tends to be used with borrowed transitive verb stems (c) or adjectival stems (d).

- | | | |
|----|------------------------------------------------|-----------------------------------------------------------|
| a. | <i>šund’-ov-</i> ‘to sound’ | < PTC <i>šund-</i> ‘listened’ |
| | <i>phrád’-ov-</i> ‘to open’ | < PTC <i>phrád-</i> ‘open’ |
| b. | <i>lój-ov-</i> ‘to redden’ | < <i>lól-o</i> ‘red’ |
| | <i>khiň-ov-</i> ‘to get tired’ | < <i>khiň-o</i> ‘tired’ |
| c. | <i>kezd-isaj-ov-</i> ‘to start (intransitive)’ | < <i>kezd-ín-</i> < H <i>kezd</i> ‘to start (transitive)’ |

⁶⁷ The *d*-verbs are historical compounds including the verb *d-* ‘to give’ (see e.g. Matras 2002: 119).

- d. *erd'av-isaj-ov-* ‘to worsen’ < *erd'av-o* < S *rdjav* ‘bad’
ném-isaj-ov- ‘to become mute’ < *ném-ast-o* < H *néma* ‘mute’

An irregular form is the deadjectival derivation *šit-isaj-ov-* ‘to grow dark’, where the adjectival root *šitit-n-o* ‘dark’ became reduced to *šit-*. The corresponding transitive form of the intransitive *naťh-ov-* ‘to pass, pass away’ is unattested in the sample.

The marker *-isaj-ov-* has been occasionally attested also with inherited stems, e.g. *azd-isaj-ov-* ‘to lift’ < *azd-* ‘to lift’, *khand-isaj-ov-* (alongside *khand-*) ‘to stink’, *barval-isaj-ov-* (alongside *barvaj-ov-*) ‘to become rich’ < *barvál-o* ‘rich’, *čor-isaj-ov-* ‘to become poor’ < *čor-o* ‘poor’, *kor-isaj-ov-* ‘to go blind’ (< *kor-o* ‘blind’), or *kuč-isaj-ov-* ‘to become expensive’ < *kuč* ‘expensive’. What is noteworthy is the form *parvard-isaj-ov-* ‘to grow up’ which is based on the participial form of the factitive *parvár-* ‘to raise’ (cf. PTC *parvar-d-*), and *darand-isaj-ov-* ‘to get frightened’ which is derived from the participial form *dara-n-* (< *dára-* ‘to fear’) extended by *d-*: *dara-n-d-*. The middle form *žut'ov-* ‘to become yellow’ is irregular, because it is derived from the Slavic-origin adjective *žut-o* ‘yellow’ by the derivational marker reserved for inherited stems.

Some middle verbs have a lexicalized meaning in KR, such as *díťh-ov-* ‘to appear, seem’ (cf. *díkh-* ‘to see, watch’), *čhord'ov- com* ‘to ruin’ and *čhord'ov- ánde* ‘to darken, cloud’ (cf. *čhór-* ‘to pour’), *kerd'ov-* ‘to boil’ (cf. *kér-* ‘to do, make’), or *nand'ov-* ‘to bath’ (cf. *náng-o* ‘naked’). The middle form *rúšť'ov-* (< PTC *rúšť-*) alternates with the corresponding intransitive verb *rúš-* ‘to be angry’.

4.7.2.7 Analytic constructions

Intransitivity may also be expressed by analytic construction. The analytic reflexives involve the verb and the first and second person accusative pronouns (*man* 1SG, *tut* 2SG, *amen* 1PL, *tumen* 2PL) or the third-person reflexive pronouns *pe* (67) and *pumen* (68).

- (67)^{LQCR} *fer pumen síte.*
 VP REFL.3PL sleep.PRT.3PL
 They fell asleep.

- (68)^{RM} *lel pe o čoro rom (...)*
 take.3SG REFL.3SG DEF poor Rom

The poor Rom shakes the dust off his feet...

Only a few borrowed verbs have been attested in the data as part of the analytic reflexive construction, e.g.:

Transitive *idegeš-itt-ín-* < H *ideges-ít* ‘to irritate’
Intransitive *idegeš-itt-ín-* REFL cf. H *ideges-ked-ik* ‘to get nervous’

The example above shows that only the transitive form was borrowed from Hungarian, while the corresponding intransitive form is expressed analytically.

Another means to express intransitivity in KR is to use the finite form of the copula together with the participial form of transitive (69)–(70), or intransitive verbs (71).

(69)^{LQCR} *šéste=lo* *sin* *phosinger-d-o*.
 VP=M.3SG COP.PRT.3 stab-PTC-M.3SG
 He was stabbed to death.

(70)^{NAR} *t'* *odá* *pál* *ó-l* *hahá-d-o* *le* *žuklenca*.
 also that.M then become-3SG feed-PTC-3SG DEF.OBL dog.INS
 Then that is also fed to the dogs.

(71)^{LQCR} *uzar* *leste* *sum-ahi* *béš-t-i*
 next_to 3SG.M.LOC COP.1SG-PRT sit-PTC-F.3SG
 I was sitting next to him.

4.7.2.8 *Compounds and collocations*

The *d*-verbs are historical compounds of the verb *d-* ‘to give’ (see e.g. Matras 2002: 119), examples of which are *az-d-* ‘to lift’, *cí-d-* ‘to pull’, *čumí-d-* ‘to kiss’, *čhán-d-* ‘to vomit’, *čhí-d-* ‘to throw’, *ké-d-* ‘to collect’, *khán-d-* ‘to stink’, *lá-d-* ‘to drive’, *phan-d-* ‘to bind’, *phú-d-* ‘to blow’, *ró-d-* ‘to look for’, *trá-d-* ‘to drive; bend’, *ušťí-d-* ‘to get’. KR has also several lexicalized verb-noun collocations formed by the verb *d-* ‘to give’ (*d- at'háli* ‘to bewitch’, *d- kárd'i* ‘lit. give shot; to shoot’, *d- kölčön* ‘lit. give loan; to lend’, *d- óbok* ‘lit. give attention; to pay attention’, *d- ril* ‘lit. give fart; to fart’, *d- šój* ‘lit. give whistle; to whistle’, *d- o bríšind* ‘lit.

give rain; to rain’, *d- o d’iv* ‘lit. give snow; to snow’), *kér-* ‘to do’ (*kér- búti* ‘lit. do work; to work’, *kér- armáňa* ‘lit. do curse; to curse’, *kér- jag* ‘lit. do fire; to make a fire’, *kér- pherd’as* ‘lit. do joke; to joke’), *l-* ‘to take’ (*l- luft* ‘lit. take breathe; to breathe’, *l- lumňa ~ l- murše* ‘lit. take woman ~ man; to marry’), *pij-* ‘to drink’ (*pij- cigrétli* ‘lit. drink cigarette; to smoke’), *čhiv-* ‘to put’ (*čhiv- víra* ‘lit. put vow; to vow’), *dikh-* ‘to see’ (*dikh- súno* ‘lit. see dream; to dream’), *máng-* ‘to ask for’ (*máng- bočánat* ‘lit. ask_for pardon; to apologize’), and more.

Verb-adjective collocations are *múk- mirno* ‘lit. leave peaceful; to leave alone’ and *pér- khámni* ‘lit. fall pregnant; to become pregnant’, while an example of a verb-verb collocation is *ža- te áčhel* ‘lit. go to live; to move’. The phrases *lák- khér* ‘lit. find to_home; to find the way home’ and *pér- khér* ‘lit. get to_home; to get home’, which consist of a verb in combination with the adverb *khér* ‘to home’, calque on the Hungarian compounds *haza-talál* ‘lit. to_home-find’ and *haza-jut* ‘lit. to_home-get’, respectively.

Verb formation in other varieties of Vend Romani

The devices that serve to form verbs in other varieties of Vend Romani roughly correspond with those described above for KR. A difference is found, for instance, in the formation of the causative form of the verb *béš-* ‘to sit’, which is *beš-ajár-* in Zala Romani and Tarany (Somogy), *beš-av-* in Veszprém Romani and in a few peripheral varieties of Somogy, and *beš-ár-* elsewhere (including KR). In several Somogy Romani varieties, the factitive marker *-isár-*, which derives verbs from adjectives in KR, has been found attached to the participial stem *darand-* (i.e. *darand-isár-* ‘to frighten’). In place of the denominal marker *-álín-* we find the also Hungarian-extracted marker *-ázín-* especially beyond Somogy. There has been attested the form *kárj-ázín-* ‘to shoot’ in Zala Romani, but *kárj-álín-* in Somogy and Veszprém Romani, or *šój-ázín-* ‘to whistle’ in Sopron, Vas, Veszprém and Zala Romani, but *šój-álín-* in Somogy Romani.

The marker *-ingér-*, which is homonymous with the iterative marker (see 4.7.2.5), is attached to the adjective *rajn-i* ‘clean’ in order to form the factitive *rajn-ingér-* ‘to tidy up’ in Vásárosdombó (Baranya) and Csokonyavisonta (Somogy). The iterative forms of polysyllabic inherited verbs have been attested only in some Somogy varieties. These varieties employ the marker *-ingér-* for both monosyllabic and polysyllabic verbs (e.g. *čhin-gér-* ‘to tear’, *va.ker-ingér-* < *va.kér-* ‘to talk’) in contrast to KR where, on the other hand, the voice opposition -

ingér- (monosyllabic) vs. *-inkér-* (polysyllabic) has been developed, e.g., *ter.d'ar-ingér-* < *ter.d'ár-* ‘to stop’.

An interesting example of a middle verb is *dis-ňov-* ~ *dis-jov-* ‘to dawn’ attested in some varieties of Somogy and Zala Romani, because it is composed of the contracted noun *dis* (< **dives* ‘day’) and the marker *-ňov-* ~ *-jov-*. The expansion of the derivational marker *-isaj-ov-* at the expense of *-ov-* is typical also to other Vend Romani varieties in Hungary. The marker *-isaj-ov-* occurs mostly in *kor-isaj-ov-* ‘to go blind’ (< *kor-o* ‘blind’) and *kašuk-isaj-ov-* ‘to turn deaf’ (< *kašuk-o* ‘deaf’). On the other hand, the marker *-ov-* ~ *-jov-* has been only rarely attested with borrowed stems, e.g. *ésň-ov-* ‘to grow grey’ (< *ésn-o* ‘grey’) in Zala and Vas Romani.

The analytic passive constructions are also typical to Burgenland Romani, as they seem to have been triggered by German contact (Halwachs 2002: 40, Matras 2002: 128).

4.7.3 Verb inflection

Two verbal stems are distinguished in KR (Matras 2002: 135–136): The present and the perfective verbal stem. For each type of stem there is an individual set of person (1, 2, 3) and number (singular and plural) concord markers. The **present verbal stem** is either identical with the verbal root (e.g. *kér-* ‘to do’), or with the verbal root extended by adaptation (e.g. *kezd-ín-* ‘to start’) or valency markers (e.g. the middle verb *kerd'-ov-* ‘to boil’). The **perfective verbal stem** is formed by means of the perfective marker either from the verbal root (e.g. *kerd-* ‘do-PFV-’ < *kér-* ‘to do’), or from the derived (e.g. *khel-á-d-* ‘dance-CAUS-PFV’ < *khel-áv-* ‘to make so. dance’) or adapted form of the verbal root (e.g. *kezd-in-d-* ‘start-AM-PFV-’ < *kezd-ín-* ‘to start’). Individual inflectional classes are distinguished for both present and perfective verbal stems.

The present stem is used to form the present (see 4.7.3.2), future (see 4.7.3.3), imperfect (see 4.7.3.4) and imperative (see 4.7.3.5), while the preterite (see 4.7.3.7) and conditional irrealis (see 4.7.3.8) are based on the perfective stem (Table 58).

TAM	MARKING	Example in 2SG
PRS	zero	<i>phénes</i> ‘you say’
FUT	PRS + <i>-a</i>	<i>pheneh-a</i> ‘you will say’
IMPF	PRS + <i>-ahi</i>	<i>pheneh-ahi</i> ‘you were saying’
IMP	stem-0	<i>phen</i> ‘say!’
PRT	zero	<i>phend'al</i> ‘you said’
IRR	PRT + <i>-áhi</i>	<i>phend'al-áhi</i> ‘you would have said’

Table 58 Tense-aspect-mood marking of lexical verbs

The tense-aspect-mood values are organized in KR as follows (based on Elšík & Matras 2006: 188–189 and Matras 2002: 151–159, see Table 59): The perfective **aspect** is expressed by the form which I refer to as preterite here. This form implies the completion of an action, mostly with past reference. However, the preterite may also denote hypothetical completion with future reference (see 4.7.3.7). The present is unmarked for **tense**, while the future (*-a*) and imperfect (*-ahi*) tenses are marked by suffixation on the present form. The aspectual distinction is realized only in the past, where the perfectivity is encoded on the preterite form, while the imperfect may be characterised by the absence of perfectivity (Matras 2002: 152). The conditional irrealis conflates both the perfective aspect and the (unreal) past tense, being formed from the preterite form by means of the irrealis suffix *-áhi* (originally an imperfective suffix, see 3.2.5.4). The **mood** is expressed by the imperative, the singular form of which mostly corresponds to the inflectional stem, and the plural form to the corresponding present indicative form. The functions of the individual tense-aspect-mood categories are dealt with in the respective sections. The tense-aspect-mood paradigm of the copula, which slightly differs from that of the lexical verbs, are discussed in section 4.7.4.

	TENSE	ASPECT
PRS	non-remote	non-perfective
FUT	non-remote	non-perfective
IMPF	remote	non-perfective
PRT	non-remote	perfective
IRR	remote	perfective

Table 59 The functions of indicative tense-aspect-mood values; based on Elšík and Matras (2006: 188)

4.7.3.1 Present stem

Following Matras (2002: 135–136), I will distinguish inflectional classes with present verbal stems according to the vowel component by which the concord markers are bound to the stem (see 4.7.3.2). In this way we can distinguish four verbal inflectional classes in KR: The class of 1) consonantal (= *C*-verbs), 2) vocalic (= *V*-verbs), 3) middle (= *MID*-verbs) and 4) contracted verbs. Most verbs belong to the consonantal class, meaning that the verb stem, as the name indicates, ends in a consonant, e.g. *án-* ‘to bring’, *rés-* ‘to reach’, *múk-* ‘to let’, *pisín-* ‘to write’. The vocalic class comprises verbs the stems of which end in the vowel *a*. These are *ása-* ‘to laugh, smile’, *dára-* ‘to be afraid, fear’, *díha-* ‘to take care’, *dúkha-* ‘to hurt’, *láža-pe* ‘to be ashamed’, *páťa-* ‘to believe’, *prasta-* ‘to hurry’, *prása-ári* ‘to ridicule’, *rezda-* ‘to tremble’, *tróma-* ‘to dare’, *urd’a-* ‘to fly’, *ha-* ‘to eat’, and *ža-* ‘to go’. The class of *MID*-verbs includes the verbs formed by the markers *-ov-* ~ *-jov-* and *-isaj-ov-* (see 4.7.2.6), e.g. *térd’-ov-* ‘to stand’.

The fourth class, the class of contracted verbs, may be further divided into four sub-classes: The sub-class of contracted verbs with stem-final 1) *-áv-*, *-óv* and *-úv*, 2) *-iv*, 3) *-ij* and 4) *-am*. Most contracted verbs belong to the first sub-class. This sub-class consists of a number of inherited words in stem-final *-áv* (e.g. *áv-* ‘to come’, *láv-* ‘to comb’), and several internally derived causatives (e.g. *dará-v-* ‘to frighten’ < *dára-* ‘to be afraid’; see 4.7.2.2). The contracted verbs with stem-final *-óv-* are *róv-* ‘to cry’, *sóv-* ‘to sleep’, *thóv-* ‘to wash’, and the verb *óv-* ‘to become; to be born’, which is homonymous with the copula. The *-úv* stem is found in *garív-* ‘to hide’, *harív-* ‘to scratch’, *khív-* ‘to weave’ and *parív-ári* ‘to exchange’. The second sub-class is represented by the verbs *siv-* ‘to sew’, *živ-* ‘to live’ and *čhiv-* ‘to put’. The latter verb is further irregular in that it is reduced to *čh-* in the first person. The third sub-class comprises the verbs *hij-* ‘to defecate’ and *pij-* ‘to drink; suck’, while the fourth sub-class only contains the auxiliary *pekám-* ‘need’.

The difference between the individual sub-classes of contracted verbs is constituted by the allomorphy of the stem (see Table 60). The alternation of the stem is phonologically conditioned (see 3.1.8).

		1	2SG	2PL3
sub-class 1	<i>táv-</i> ‘to cook’	<i>táv-</i>	<i>táj-</i> ~ <i>táv-</i>	<i>tá-</i> ~ <i>táv-</i>
sub-class 2	<i>živ-</i> ‘to live’	<i>žij-</i>	<i>ží-</i> ~ <i>živ-</i>	<i>ží-</i> ~ <i>živ-</i>
sub-class 3	<i>pij-</i> ‘to drink’	<i>pij-</i>	<i>pí-</i>	<i>pí-</i>
sub-class 4	<i>pekám-</i> ‘need’	<i>pekám-</i>	<i>pekaj-</i>	<i>peká-</i> ~ <i>pekam-</i>

Table 60 Stem allomorphy of contracted verbs

The allomorphs are distributed according to the following rule: In the first person, the non-contracted stem is employed, and the *-ij* allomorph of the stem is used in case of verbs belonging to the second sub-class. In the second and third persons, the verbs of the third sub-class have a vocalic stem in *-í*, while the verbal stems of the verbs of the other sub-classes are either vocalic or consonantal.

4.7.3.2 Present tense

The present tense is formed by the present stem which is accompanied by the concord markers indicating person and number. The concord markers are 1SG *-v*, 2SG *-s* (~ *-o*), 3SG *-l*, 1PL *-s*, and 2/3PL *-n* (Table 61).

	V-verbs	C-verbs	MID-verbs
	<i>dára-</i> ‘to fear’	<i>már-</i> ‘to beat’	<i>haj-ov-</i> ‘to understand’
1SG	<i>dára-v</i>	<i>már-av</i>	<i>haj-ov-av</i>
2SG	<i>dára-s</i>	<i>már-e(s)</i>	<i>haj-o-s</i> (~ <i>haj-oj-s</i>)
3SG	<i>dára-l</i>	<i>már-el</i>	<i>haj-o-l</i>
1PL	<i>dára-s</i>	<i>már-as</i>	<i>haj-ov-as</i>
2/3PL	<i>dára-n</i>	<i>már-en</i>	<i>haj-o-n</i>

Table 61 Present tense inflection of V-verbs, C-verbs and MID-verbs

The first-person concord marker is preceded by *-a-* in all classes of verbs, except of the V-verbs. In C-verbs, the second and third-person markers are attached to the stem with *-e-*. The concord markers are extended by *-ov-* in the first person of MID-verbs, and by *-o-* in other persons. The MID-verb *haj-ov-* ‘to understand’ is irregular, since the second-person singular

concord marker may be optionally preceded by *j*, i.e. *haj-oj-s* alongside *haj-o-s* ‘you.SG understand’). This type of stem contraction is typical for the fourth inflectional class, i.e. to the contracted verbs (Table 62).

	<i>-áv ~ -óv ~ -úv</i>	<i>-iv, -ij</i>	<i>-ám</i>
	<i>táv-</i> ‘to cook’	<i>živ-</i> ‘to live’	<i>pekám-</i> ‘need’
1SG	<i>táv-av</i>	<i>živ-av</i>	<i>pekám-av</i>
2SG	<i>táj-s</i>	<i>ží-s</i>	<i>pekáj-(s)</i>
3SG	<i>tá-l</i>	<i>ží-l</i>	<i>peká-l</i>
1PL	<i>táv-as</i>	<i>živ-as</i>	<i>pekám-as</i>
2/3PL	<i>tá-n</i>	<i>ží-n</i>	<i>peká-n</i>

Table 62 Present tense inflection of contracted verbs

The second and third sub-classes of contracted verbs, i.e. the verbs with stem-final *-iv* and *-ij*, have identical inflectional forms in the present tense.

An interesting on-going development in KR is that the final *s* of the second-person singular marker can be optionally dropped. The *s*-less form has been attested in case of some *C*-verbs and in the contracted verb *pekám-* ‘need’, but not in *V*-verbs and *MID*-verbs. This innovative change has occurred more frequently in the narratives (72)–(73) than in the elicited data.

- (72)^{NAR} *só gá márel! má so vaker-e?*
 what fast beat.3SG so what speak-2SG
 What does beat fast (about the heart)? What are you speaking about?

- (73)^{NAR} *na, na pekáj-0 te phénel!*
 NEG NEG need-(2SG) COMP say.INF
 No, you don’t have to say it!

The primary function of the present tense is to refer to an action in the present time. That is to indicate repeated or habitual actions (74), (universal) statements (75) and continuous actions (76).

(74)^{NAR} *má na but hohávav; taj na sokínav te hohál nikaske na.*
 already NEG much lie.1SG and NEG used_to.1SG COMP lie.INF nobody.DAT NEG
 I do not lie much anymore; I do not lie to anybody.

(75)^{NAR} *te čárel o (sic!) mindž, adá žánel, ár na!*
 COMP lick.3SG DEF pussy this.M know.3SG other NEG
 To lick the pussy, that is the only thing he knows.

(76)^{NAR} *žuto húmer, sar akának so távav.*
 yellow pasta how now what cook.1SG
 ‘Yellow pasta’, this is what I am cooking now.

The present form is often used in a future reference in combination with time related words, or when the future meaning is already evident from the context (77).

(77)^{NAR} *táha káj žas, žas khér?*
 tomorrow where go.2SG go.2SG home
 Where are you going tomorrow? Are you going home?

4.7.3.3 Future

The future tense, which is used to refer to a future action, is formed by adding the marker *-a* to the present form of all classes of verbs (Table 63–64).

	C-verbs	V-verbs	MID-verbs
1SG	<i>már-á</i>	<i>dar(-)á</i>	<i>haj-ov-á</i>
2SG	<i>már-eh-a</i>	<i>dara-h-a</i>	<i>haj-o-h-a</i>
3SG	<i>már-l-a</i>	<i>dara-l-a</i>	<i>haj-o-l-a</i>
1PL	<i>már-ah-a</i>	<i>dara-h-a</i>	<i>haj-ov-ah-a</i>
2/3PL	<i>már-n-a</i>	<i>dara-n-a</i>	<i>haj-o-n-a</i>

Table 63 Future tense of C-verbs, V-verbs and MID-verbs

Before the application of the future marker, the personal concord markers underwent the following changes: In the first person singular, the personal marker *-av* and the future marker *-a* merged in *-á*, e.g. *már-á* < **már-av-a* ‘I will beat’. The intervocalic *s* became

debuccalised in the second-person singular and the first-person plural, e.g. *már-eh-a* < **már-es-a* ‘you will beat’. The original sibilant has been only retained in the V-verb *ha-s-a* ‘you/we will eat’ (cf. Elšík *et al.* 1999: 301). Finally, the personal marker in the third person and in the second person plural became syncopated, e.g. *már-l-a* < **már-el-a* ‘s/he will beat’.

	-áv ~ -óv ~ úv	-iv	-ij	-am
1SG	<i>táv-á</i>	<i>živ-á</i>	<i>pij-á</i>	<i>pekam-á</i>
2SG	<i>táj-h-a</i>	<i>ží-h-a</i>	<i>pí-h-a</i>	<i>pekaj-h-a</i>
3SG	<i>táv-l-a</i>	<i>živ-l-a</i>	<i>pí-l-a</i>	<i>pekam-l-a</i>
1PL	<i>táv-ah-a</i>	<i>živ-ah-a</i>	<i>pij-ah-a</i>	<i>pekam-ah-a</i>
2/3PL	<i>táv-n-a</i>	<i>živ-n-a</i>	<i>pí-n-a</i>	<i>pekam-n-a</i>

Table 64 Future tense of contracted verbs

The future may be optionally expressed by means of the borrowed auxiliary *fogín-* (< H *fog*) ‘will’ together with the infinitive form of the verb (see 4.7.5.2). The following examples illustrate the future construction with *fogín-* (78)a and the synthetic future marking (78)b in KR. In Hungarian, future is expressed only analytically (79).

(78)^{LQCR}

- a. *angli khangéri tut fogín-á t' užárel*
 in_front_of.F church you.ACC will-FUT.1SG COMP wait.INF
- b. *angli khangéri tut užar-á*
 in_front_of.F church you.ACC wait-FUT.1SG
 I will wait for you in front of the church.

(79) Hungarian

- a templom előtt foglak várni.*
 DEF church in_front_of will.PRS.1SG>2SG wait.INF
 I will wait for you in front of the church.

It is interesting that the auxiliary *fogín-* is obligatorily inflected for future tense (78)a. Thus, the future is double marked: by the auxiliary *fogín-* on the one hand, and by the inherited

future marker *-a* on the other. The construction with *fogín-* seems to be more frequent in the elicited data, while the inherited future marking predominates in the spontaneous narratives.

Another means to indicate future action is the use of the borrowed adverb *maj(d)* (< H *majd*) with an approximate meaning of ‘at a time in the future’. This time related word also requires a verb in the future form (80).

- (80)^{NAR} *maj me tuke sikavinker-á o cile kipi.*
 ‘will’ 1SG 2SG.DAT show.ITER-FUT.1SG DEF.PL all.PL photo.PL
 I will show you all the photos.

The volition modal *kám-* has a future form also in present reference (81).

- (81)^{NAR} *akán na kam-á te hohál*
 now NEG want-(FUT.1SG) COMP lie.INF
 I do not want to lie now.

4.7.3.4 Imperfect

The imperfect is formed by the marker *-ahi* from the present form of the verb, as it is demonstrated in Table 65 and Table 66.

	C-verbs	V-verbs	MID-verbs
1SG	<i>már-áhi</i>	<i>dar(-)áhi</i>	<i>haj-ov-áhi</i>
2SG	<i>már-eh-ahi</i>	<i>dara-h-ahi</i>	<i>haj-o-h-ahi</i>
3SG	<i>már-l-ahi</i>	<i>dara-l-ahi</i>	<i>haj-o-l-ahi</i>
1PL	<i>már-ah-ahi</i>	<i>dara-h-ahi</i>	<i>haj-ov-ah-ahi</i>
2/3PL	<i>már-n-ahi</i>	<i>dara-n-ahi</i>	<i>haj-o-n-ahi</i>

Table 65 Inflectional markers in imperfect

The person markers followed by the imperfective marker have undergone the same development as in the future form: contraction in the first-person singular (e.g. *már-áhi* < **már-av-ahi* ‘I was beating’), syncope of *e* in the third person and in the second-person plural (e.g. *már-n-ahi* < **már-en-ahi* ‘they were beating’), and debuccalisation of *s* in the second-

person singular and first-person plural (e.g. *már-ah-ahi* < **már-as-ahi* ‘we were beating’). The intervocalic *s* has been preserved only in *ha-s-ahi* ‘you/we were eating’ (cf. Elšík *et al.* 1999: 301).

	<i>-av ~ -ov ~ -uv</i>	<i>-iv</i>	<i>-ij</i>	<i>-am</i>
1SG	<i>garuv-áhi</i>	<i>žij-áhi</i>	<i>pij-áhi</i>	<i>pekam-áhi</i>
2SG	<i>garuj-h-ahi</i>	<i>ží-h-ahi</i>	<i>pí-h-ahi</i>	<i>pekaj-h-ahi</i>
3SG	<i>garuv-l-ahi</i>	<i>živ-l-ahi</i>	<i>pí-l-ahi</i>	<i>pekam-l-ahi</i>
1PL	<i>garuv-ah-ahi</i>	<i>žij-ah-ahi</i>	<i>pij-ah-ahi</i>	<i>pekam-ah-ahi</i>
2/3PL	<i>garuv-n-ahi</i>	<i>živ-n-ahi</i>	<i>pí-n-ahi</i>	<i>pekam-n-ahi</i>

Table 66 Inflectional markers of contracted verbs in imperfect

The imperfect is used to refer to repeated (82), habitual (83) and continuing actions (84), and to states of being (85) with past reference.

(82)^{NAR} *sako ra'aha upre fárinlahi, nášlahi fút te pálinkázinel.*
 every morning VP jump.IMPF.3SG run.IMPF.3SG away COMP booze
 S/he would jump out (from the bed) every morning and run to booze.

(83)^{NAR} *mri baba, oja halahi t' o rókano mas.*
 my grandmother that.F eat.IMPF.3SG also DEF fox meat
 My grandmother, she used to eat even the fox meat.

(84)^{LQCR} *hát so khírinlahi, sako upr' uš'íno.*
 well as scream.IMPF.3SG everyone VP wake.PRT.3SG
 Well, as she was screaming, everyone woke up.

(85)^{NAR} *de oja kuňuva, phendle, búter molahi sar odá kher.*
 but that.F hovel say.PRT.3PL more worth.IMPF.3SG than that.M house
 But, as they said, that hovel was worth more than that house.

The imperfect is further used to encode potential conditional, indicating willingness (86), suggestion (87), condition (88) and possibility with regard to innate ability (89).

(86)^{NAR} *sótár kamlahi te kérel.*
 dictionary want.IMPF.3SG COMP make.INF
 S/he would like to make a dictionary.

(87)^{NAR} *táha afka site ájhahi, hod' akor na sójhahi odoj, hanem edej.*
 tomorrow so should come.IMPF.2SG, COMP then NEG sleep.IMPF.2SG there, but here
 Tomorrow you should come in such a way that you would not sleep there, but here.

(88)^{NAR} *oja korkóri ando temetéši ár na žalahi,*
 that.F alone into.M graveyard out NEG go.IMPF.3SG,
te lakeri men činnahi, ni akor.
 even_if her neck cut.IMPF.3PL, neither then
 She would not go to the graveyard alone, not even if they had cut her neck.

(89)^{LQCR} *hajo-h-ahi serbül vad' ninčka?*
 understand-2SG-IMPF Serbian or German
 Could you speak Serbian or German?

4.7.3.5 Imperative

KR distinguishes imperative forms for the second person singular and plural (Table 67). In the singular, the imperative forms of *C*-verbs and *V*-verbs are homonymous with their respective present stems. These forms are thus zero marked. The *d*-verbs are irregular, as their singular imperative forms are formed by the suffix *-e*. The inflectional stems of the MID-verbs are extended by the derivational marker *-(j)ov-* in both singular and plural.

	MARKER	<i>C</i> -verbs	<i>d</i> -verbs	<i>V</i> -verbs	MID-verbs
2SG	<i>-0</i>	<i>már-0</i>		<i>dara-0</i>	
	<i>-e</i>		<i>ked-e</i>		
	<i>-ov</i>				<i>haj-ov</i>
2PL	<i>-en</i>	<i>már-en</i>	<i>ked-en</i>		
	<i>-n-en</i>			<i>dara-n-en</i>	
	<i>-ov-en</i>				<i>haj-ov-en</i>

Table 67 Second-person imperative forms

The plural imperative marker is *-en* for all types of verbs. This marker is preceded by *-n* in case of *V*-verbs. The verbs *d-* ‘to give’ and *l-* ‘to take’ are irregular, as they take the marker *-e* in the singular (i.e. *d-e* and *l-e*), and the extended marker *-en-en* in the plural (i.e. *d-en-en* and *l-en-en*). Verbs with final postalveolar or palatal sounds also take the imperative marker *-e* in the singular, e.g. *lež-e* < *léž-* ‘to carry’, *urd-e* < *urd-* ‘to dress’, and *ušt-e* < *ušt-* ‘to wake’. The plural imperative forms of these verbs are however the regular *lež-en*, *urd-en*, and *ušt-en*, respectively.

The singular imperative forms of contracted verbs are identical with their present stems. Most contracted verbs have non-contracted stems in *-Vv*. The verbs that belong to the type *-iv* may have both *-iv* and *-ij* stem in the singular. However, the final semi-vowel is generally not realized, such as in *ži* < **žij* ‘live!’ The same development is found in the third sub-class of contracted verbs, i.e. *pi* < **pij* ‘drink!’

	<i>-av ~ -ov ~ -uv</i>	<i>-iv</i>	<i>-ij</i>	<i>-am</i>
2SG	<i>garuv</i>	<i>ži ~ živ</i>	<i>pi</i>	?
2PL	<i>garuv-en</i>	<i>žij-en</i>	<i>pij-en</i>	?

Table 68 Second-person imperative forms of contracted verbs

The plural marker *-en* is attached to the *-Vv* stem of the first sub-class, and to the *-ij* stem of the second and third sub-classes of contracted verbs (Table 68). The imperative form of the fourth sub-class of contracted verbs (i.e. of *pekám-* ‘need’) is unattested in the sample.

The meaning ‘let’s go’ is expressed by the idiom *av žas* ‘come.IMP.2SG go.IMP.1PL’ when only one person is addressed, and by *aven žas* ‘come.IMP.2PL go.IMP.1PL’ when more persons are addressed. The imperative form *le* ‘here you are/go’ (< *l-* ‘to take’) is used when giving something to someone.

4.7.3.6 *Perfective stem*

In line with Matras (2002:_135), I will distinguish inflectional classes with perfective verbal stems by the form of the perfective marker attached to the stem. More precisely, by the form of the unpalatalized perfective marker as it is found in the third person plural. According to

this, KR has five inflectional classes: the *-n-* class, *-d-* (~ *-dl-*) class, *-l-* class, *-t-* ~ *-l-* class, and the *-ín-* class (Table 69).

CLASS	STEM	e.g.	MARKER	
			1/2	3PL
			3SG (transitive)	3SG (intransitive)
<i>-n-</i>	vowel	<i>ása-</i> ‘to laugh’	<i>-ň-</i>	<i>-n-</i>
<i>-d-</i> ~ -	l	<i>khél-</i> ‘to dance’	<i>-d’-</i>	<i>-d-</i>
<i>dl-</i>	r	<i>már-</i> ‘to beat’		
	v > 0	<i>már-áv-</i> ‘to make so. beat’		
	n	<i>án-</i> ‘to bring’	<i>-d’-</i>	<i>-d-</i> ~ <i>-dl-</i>
<i>-l-</i>	kh > k	<i>díkh-</i> ‘to watch’	<i>-j-</i>	<i>-l-</i>
	(n)d	<i>čhánd-</i> ‘to vomit’		
	g	<i>phág-</i> ‘to break’	<i>-j-</i> ~ <i>-d’-</i>	<i>-l-</i>
	m	<i>kám-</i> ‘to love, want’		
	č	<i>phúč-</i> ‘to ask’	<i>-j-</i> ~ <i>-t’-</i>	<i>-l-</i>
	k	<i>pék-</i> ‘to bake’		
<i>-t-</i> ~ <i>-l-</i>	s	<i>rés-</i> ‘to reach’	<i>-t’-</i> ~ <i>-j-</i>	<i>-t-</i> ~ <i>-l-</i>
	š	<i>béš-</i> ‘to sit’		
<i>-ín-</i>	t’	<i>ušt’-</i> ‘to wake’	<i>-ij-</i>	<i>-ín-</i>
	d’	<i>urd’-</i> ‘to dress’		
	čh	<i>áčh-</i> ‘to stay’		
	d (<i>d-</i> verbs)	<i>kéd-</i> ‘to collect’		
	(MID-verbs)	<i>ker-d’-ov-</i> ‘to boil’		

Table 69 Perfective markers

The perfective marker in other forms than the third-person plural is palatalized (see 3.1.7), except for the third-person singular marker of several derived intransitive and unaccusative verbs (Matras 2002: 145, see 4.7.3.7). All *V-*verbs are assigned to the *-n-* class, by employing the unpalatalized allomorph *-n-* in the third person and the palatalized *-ň-* in

other persons. The *C*-verbs in final *l r* and *n* take the perfective marker *-d-* in the third-person plural, and *-d'* in other persons. The same markers are employed in verbs in stem-final *v* (i.e. contracted verbs in *-Vv*). However, before the application of this marker the stem-final *v* became deleted. The *n*-final verbs may optionally also take the marker *-dl-* in the third-person plural, e.g. *phen-d-* ~ *phen-dl-* ‘tell-PFV-’. The *-l-* class includes verbs in final *kh g m ě k* and *d*-verbs in final *nd*, i.e. *čhánd-* ‘to vomit’, *phand-* ‘to bind’, and possibly also *khánd-* ‘to stink’. The third-person plural marker *-l-* has the allomorph *-j-* in other persons, or optionally the allomorph *-d'* in *g*-final and *m*-final verbs, and the allomorph *-t'* in *ě*-final and *k*-final verbs. The *-t-* ~ *-l-* class comprises verbs ending in the sibilants *s* and *š*. The allomorph *-t-* is more frequently used than *-l-*. The corresponding palatalized forms are *-t'* and *-j-* (see 3.1.7), respectively. Finally, the *-ín-* class incorporates the MID-verbs, the *d*-verbs (except of the ones in final *nd*, see above) and the verbs *ušt'* ‘to wake’, *urd'* ‘to dress’ and *áčh-* ‘to stay’. The verbs *d-* (> *d-ij-* ~ *d-in-*) ‘to give’ and *l-* (> *l-ij-* ~ *l-in-*) ‘to take’ also pertain to this class. In MID-verbs, the marker *ij-* ~ *-ín-* is added either to the verbal (e.g. *kezdisaj-ij/in-* < *kezd-isaj-ov-* ‘to start’) or non-verbal stem (e.g. *khiň-ij/in-* < *khiň-ov-* ‘to get tired’). After the application of these markers, the final dental of the stem generally became palatalized (see 3.1.7). The verbal stem of MID-verbs in final *s* and *š* may exceptionally remain unpalatalized, such as in *pašl-ín-e* alongside *pašj-ín-e* ‘they laid’, cf. *pášj-ov-* ‘to lie’.

The perfective stem of the verb *ža-* ‘to go’ is formed from the suppletive root *gé-* by the perfective marker *-j-* ~ *-l-*. Several other verbs exhibit stem alternations, such as the verbs *av-* (> *á-j-* ~ *á-l-*) ‘to come’, *hij-* (> *hin-d'* ~ *hin-d-* ~ *hin-dl-*) ‘to defecate’, *lež-* (> *legé-d'* ~ *legé-d-*) ‘to take, carry’, *mér-* (> *mú-j-* ~ *mú-l-*) ‘to die’, *pér-* (> *pé-j-* ~ *pé-l-*) ‘to fall’, *rov-* (> *rú-ň-* ~ *rú-n-*) ‘to cry’, *sov-* (> *sú-t'* ~ *sú-t-*) ‘to sleep’, and *ov-* (> *ú-j-* ~ *ú-l-*) ‘to become’. The perfective stems of the verbs *ha-* ‘to eat’, *čhiv-* ‘to put’ and *živ-* ‘to live’ (possibly also *siv-* ‘to live’) are also irregular. The *V*-verb *ha-* takes the perfective markers *-j-* ~ *-l-*, even though both of them are expected to be assigned to the *-n-* class. The *v*-final verb *čhiv-* takes the marker *-t'* ~ *-t-* (i.e. *čhi-t'* ~ *čhi-t-*) instead of the regular *-d'* ~ *-d-*. On the other hand, the *v*-final *živ-* employs the (in part regular) markers *-d'* ~ *-j-* ~ *-d-*, but without deleting the stem-final *v*, i.e. *živ-d'* ~ *živ-j-* ~ *živ-d-*. The perfective stem of the verbs *bikn-* ‘to sell’, *mutr-* ‘to urinate’ and *hušn-* ‘to knead’ is the non-syncopated *biken-d-*, *muter-d-*, and *hušen-d-*, respectively.

Perfective markers of oikoclitic *C*-verbs are identical to the respective participial markers. Participial markers of xenoclitic verbs are discussed in chapter 4.7.5.1.

4.7.3.7 Preterite

The preterite is formed from the perfective stem by the following subject concord markers: 1SG *-um*, 2SG *-al*, 3SG *-a*, 1PL *-am*, 2PL *-en*, 3PL *-e*. These markers are directly added to the perfective stem (Table 70). The first-person singular form *-um* (< **-om*) developed due to the sound change *o* > *u*, while the older marker is preserved only in the reduced forms *ph-om* (< **phend'-om*) ‘I told’ and *ž-om* (< **žand'-om*) ‘I knew’ (see 3.1.8). The second-person plural marker **-an* was replaced by *-en* in KR. It resulted from the original marker (**-an*) having been influenced by the third-person plural marker (*-e*), and thus taking over its vowel quality (Elšík & Matras 2006: 122).

	transitive	intransitive
	<i>már-</i> ‘to beat’	<i>phurd'-ov-</i> ‘to grow old’
1SG	<i>márd'-um</i>	<i>phurd'ij-um</i>
2SG	<i>márd'-al</i>	<i>phurd'ij-al</i>
3SG	<i>márd'-a</i>	–
M	–	<i>phurd'in-(o)</i>
F	–	<i>phurd'in-(i)</i>
1PL	<i>márd'-am</i>	<i>phurd'ij-am</i>
2PL	<i>márd'-en</i>	<i>phurd'ij-en</i>
3PL	<i>márd'-e</i>	<i>phurd'in-(e)</i>

Table 70 Preterite forms

The third-person singular forms of intransitive derivations, psych verbs with vocalic stem and a few other intransitive verbs have adjectival inflection (cf. Matras 2002: 122): the marker *-o* stands for the masculine, and the marker *-i* for the feminine gender. More precisely, the gender-specific markers in the third person singular have been attested with the vast majority of *V*-verbs and *MID*-verbs, and with the verbs *av-* ‘to come’ (< *ál-o/i*), *áčh-* ‘to stay’ (< *áčhin-o/i*), *běš-* ‘to sit’ (< *běšt-o/i*), *mér-* ‘to die’ (< *múl-o/i*), *náš-* ‘to run’ (< *náš-t-o/i*), *pekám-* ‘need’ (< *pekáml-o/i*), *pér-* ‘to fall, arrive’ (< *pěl-o/i*), *rov-* ‘to cry’ (< *rún-o/i*), *sov-* ‘to sleep’ (< *sút-o/i*), *ov-* ‘to become’ (< *úl-o/i*), *ušť-* ‘to wake’ (< *ušťin-o/i*), and *urd'-* ‘to dress’ (< *urd'in-o/i*). The intransitive verb *ža-* (> *gél-o/i*) ‘to go’ has been exceptionally attested with the

gender-indifferent third-person marker, i.e. *gěj-a* ‘s/he went’. The participial, unpalatalized stem is also employed in the third-person plural of all verbs.

The gender-specific markers as well as the third-person plural marker of intransitive verbs may optionally be dropped. As a result, the same form may occur in both third-person singular and plural, e.g. *áčhín-0* ‘stay.PRT-(3SG/PL); she/he/they stayed’. The same development is found in the third-person plural of *d*-verbs (e.g. *kedín-0* ‘collect.PRT-(3SG/PL); she/he/they collected’) and the verbs *d*- ‘to give’ (e.g. *dín-0* ‘give.PRT-(3SG/PL); she/he/they gave’) and *l*- ‘to take’ (e.g. *lín-0* ‘take.PRT-(3SG/PL); she/he/they took’). The perfective marker *-ij-* may optionally become reduced to *-j-* when attached to MID-verbs in stem-final *j* in other than the third person, such as in *rumisaj-j-um* < **rumisaj-ij-um* ‘worsen-PFV-1SG; I worsened’.

The preterite form is used to denote actions that were completed in the past, e.g. *astárd’um* and *gejam* in (90), or *dikjal* in (91).

(90)^{NAR} *mre dujen ánde astárd’um, afka gejam kašt t’ ánel.*
 my.PL two.ACC.PL into harness.PRT.1SG so go.PRT.1PL wood COMP bring.INF
 I harnessed my two (horses), and so we went to bring wood.

(91)^{NAR} A: *ko gélo odoj fer? na dikjal?* B: *káj?*
 who go.PRT.3SG there VP NEG see.PRT.2SG where
 A: Who has passed by over there? Have you not seen [them]? B: Where?

The past reference is, however, not inherent in the preterite form, as it also used in the predictive type of conditional sentences with future reference (92).

(92)^{LQCR} *te dikjal valaso, phen mange!*
 if see.PRT.2SG something tell.IMP.2SG 1SG.DAT
 If you see something, tell me!

4.7.3.8 *Conditional irrealis*

The conditional irrealis is formed from the preterite form by the marker *-áhi* in the first and second persons (Table 71).

	transitive	intransitive
	<i>már-</i> ‘to beat’	<i>phurd’-ov-</i> ‘to grow old’
1SG	<i>márd’um-áhi</i>	<i>phurd’ijum-áhi</i>
2SG	<i>márd’al-áhi</i>	<i>phurd’ijal-áhi</i>
3SG	<i>márd’á(-)hi</i>	<i>phurd’ijá(-)hi</i>
1PL	<i>márd’am-áhi</i>	<i>phurd’ijam-áhi</i>
2/3PL	<i>márd’en-áhi</i>	<i>phurd’ijen-áhi</i>

Table 71 Conditional irrealis marker

In the third person singular, the irrealis marker is added to the gender-indifferent personal concord of both transitive and intransitive verbs. The irrealis form of the third-person plural is identical with that of the second-person plural. It has most probably developed as an analogy to the homonymous second and third-person plural forms found in present, future and imperfect (cf. Elšík *et al.* 1999: 358; see the examples below). Thus, the second and third person plural forms are distinguished only in the preterite.

2PL = 3PL

- PRS *máren* ‘you/they beat’
 FUT *márna* ‘you/they will beat’
 IMPF *márnahi* ‘you/they were beating’
 > IRR *márd’enáhi* ‘you/they would have beaten’

The irrealis expresses unreal condition (93) or hypothetical action (94) with past reference.

- (93)^{NAR} *te afka kerd’áhi, sar me phom, hod’ (...)*
 if so do.IRR.3SG how 1SG say.PRT.1SG COMP
 If s/he had done what I said, that...

- (94)^{NAR} *afka le rembe_čhite, hod’ i Madonna ánde šaj áchijáhi.*
 so 3SG.ACC set_in_order.PRT.3PL COMP DEF Madonna in can live.IRR.3SG
 They set (the house) so that even Madonna could have lived in there.

4.7.3.9 Uninflected forms

A unique innovation in KR, which undoubtedly requires further research, is that the borrowed verbs adapted by the marker *-ín-* sometimes cease to be inflected when the relevant inflection is evident from the context. The uninflected form thus corresponds to the inflectional stem. The example sentences (95)–(98) come from a native speaker of KR. The uninflected forms are marked by bold font, the expected inflected forms are indicated in the subsequent brackets, and the verbs which give the context (i.e. the appropriate person and number and tense-aspect-mood category) are underlined:

- (95)^{NAR} *Čak muken la ári te histizín-el čak.*
 just let.3PL 3SG.F.ACC VP COMP throw_tantrum.3SG just
*Kas **érdekelin-0** (< érdekelín-el)? Hát so **histizín-0** (< histizín-el)*
 who.ACC be_interested well what throw_tantrum
 Just let her throw a tantrum. Who cares? Why is she throwing a tantrum?

- (96)^{NAR} *Taj béš-en taj lešín-en taj **fid'elin-0** (< fid'elín-en),*
 and sit.3PL and peer.3PL and pay_attention
hod' te upral lende valaso vakerde.
 COMP if about 3PL.LOC something say.PRT.3PL
 And they are sitting and peering and paying attention to whether someone says something about them.

- (97)^{NAR} *Seretin-lahi, odá nad'on **seretin-0** (< seretin-lahi),*
 like.IMPF.3SG that.M very_much like
hod' ój pe öltözinel, te nad'ulinel, ast nad'on!
 COMP 3SG.F REFL dress.3SG, COMP boast.INF, that very_much
 She liked, she liked very much to dress, to boast, that is what she liked very much!

- (98)^{NAR} *Me lake **rendezin-0** (< rendezin-d'um) t' o verda,*
 1SG 3SG.F.DAT arrange also DEF car
t' o utánfutó me lake rendezin-d'um.
 also DEF truck 1SG 3SG.F.DAT arrange.PRT.1SG
 I arranged the car for her as well, and I also arranged the truck trailer for her.

The uninflected form may substitute any inflectional form of the verb, as for instance the present third-person singular in a. (*érdekelin/-el* ‘be_interested/-PRS.3SG’, *histizin/-el* ‘throw_tantrum/-PRS.3SG’), present third-person plural in b. (*fid’elin/-en* ‘pay_attention/-PRS.3PL’), imperfect third-person singular in c. (*seretin/-lahi* ‘like/-IMPF.3SG’) and preterite first-person singular in d. (*rendezin/-d’um* ‘arrange/-PRT.1SG’).

Verb inflection in other varieties of Vend Romani

Some vocalic verbs have become consonantal especially in Zala and Vas Romani (e.g. *ás-* < *ása-* ‘to laugh’), while some others have consonantal stems only in the present tense, e.g. *láž-el pe* ‘be_ashamed-3SG REFL.3SG; s/he is ashamed’, cf. *laža-l-ahi pe* ‘be_ashamed-3SG-IMPF REFL.3SG; s/he was ashamed’.

The second-person plural marker *-en* is typical to the Somogy Romani varieties, but not to other Vend Romani varieties, where the original form *-an* has been preserved. The loss of final *s* in the second-person singular marker *-es* has also been attested only in some varieties of Somogy Romani, e.g. in Vásárosdombó (Baranya) *baba, so kére?* ‘grandma what do.2SG; Grandma, what are you doing?’. The verb *hajov-* is treated as a contracted verb rather than a MID-verb in other Vend Romani varieties as well, which is evident from the *-j* contraction found in the second-person singular present stem, i.e. *haj-oj-s* ‘you.SG understand’, cf. MID-verb *térd’-o-s* ‘you.SG stand’. In Zala Romani, the second-person singular present stem of contracted verbs is vocalic, e.g. *garú-s* < **garuv-es*, cf. KR *garúj-s* ‘you hide’.

The verbs *žij-* (< *živ-*) ‘to live’, *sij-* (< *siv-*) ‘to sew’ and *čhij-* (< *čhiv-*) ‘to put’ are not contracted in Vas Romani, e.g. *čhijel* ‘puts’, cf. KR *čhil*. Non-contracted form is also used with the verb *pij-* ‘to drink’ in Vas and Zala Romani, and the verb *hij-* ‘to defecate’ in Zala Romani.

The volition modal *kám-* has only the future form in Vend Romani, except for the varieties of Zala and Prekmurje Romani, where both the present and the future forms are distinguished. The future auxiliary *fogin-* is absent in Zala, Vas, Veszprém and Sopron Romani, but it does occur in several varieties of Somogy Romani. Like in KR, the auxiliary is generally marked as a future form. The exception is the Somogy Romani variety of Baté, where the auxiliary is optionally inflected for the present tense, e.g. *fogin-es te žal* ‘will-PRS.2SG to_go’ alongside *fogin-eh-a te žal* ‘will-2SG-FUT to_go; you will go’. Vas Romani frequently uses the present form of the verb in future reference, especially alongside the time

related word *maj* ‘at a time in the future’, e.g. *maj žav* ‘(will) go.PRS.1SG’ alongside *maj žá* ‘(will) go.FUT.1SG; I will go’.

The second-person plural imperative forms of the verbs *d-* ‘to give’ and *l-* ‘to take’ are formed either by *-en* or by its extended form *-en-en* in Zala Romani and in a few peripheral varieties of Somogy Romani: *d-en ~ d-enen* ‘give-IMP.2PL’, *l-en ~ l-enen* ‘take-IMP.2PL’.

In Zala, Veszprém and Vas Romani, the MID-verbs are assigned to the *-íl-* class (e.g. *phurd-íl-e* ‘get_old-PRT-3PL; they grew old’), while the *d-* verbs and the verbs *d-* ‘to give’ and *l-* ‘to take’ belong to the *-ín-* class. Except of Veszprém Romani, the palatalized allomorph of the perfective marker *-ín-* is *-iň-* alongside *-ij-* e.g. *uštíd-iň-a ~ uštíd-ij-a* ‘catch-PFV-3SG; s/he caught’.

The split of the imperfective and irrealis paradigms (IMPF *-ahi*, IRR *-áhi*, see 3.2.5.4) is typical for Somogy, Veszprém and Sopron Romani, while only the short-vowel imperfective suffix is applied in Zala and Vas Romani, as well as in some varieties of Somogy Romani.

The innovative strategy of not inflecting verbs has been only sporadically found in other varieties of Vend Romani. The small number of occurrences in contrast to KR may be explained by the fact that my data of other Vend Romani varieties are mainly elicited. The KR data suggest that the innovative pattern is more likely to occur in spontaneous speech, where the context is given in the narrative, in contrast to the elicited data, where the context is mostly not provided. Examples of such uninflected verbs have been attested, for instance, in the varieties of Lengyel-tóti (99) and Nagykanizsa (100). In the former, the context is given by the plural noun phrase *o phure romňa* ‘old women’, while in the latter by the future first-person form *žá* ‘I will go’, which requires the infinitive form *te šétálinel* ‘to walk’ to be used.

(99)^{LQCR} *o phure romňa seretín-0 (seretín-en) te béšel anglo khéra*
 DEF old woman.PL like COMP sit.INF in_front_of.M house.PL
 Old women like to sit in front of the houses.

(100)^{LQCR} *žá d'ék te šétálin-0 (šétálin-el)*.⁶⁸
 go.FUT.1SG one COMP walk

⁶⁸ The KR expression *žá- d'ék* ‘lit. go one’ is based on the Hungarian phrase *megy egyet* ‘go one.ACC’ which is used to express that a continuous action (e.g. walking, swimming, dancing, etc.) is performed ‘once’, i.e. for a certain period of time. Note that the expression is inaccurately translated to KR, since the element ‘one’ has the nominative form in KR, but the accusative form in Hungarian.

4.7.4 Copula inflection

Three stems are involved in copula inflection: 1) *s- ~ h*, 2) *ov-* and 3) *ú-l-*. Both the present and the preterite copula draw on the *s- ~ h* stem. The present subjunctive, future and conditional are based on the *ov-* stem, while the preterite subjunctive and irrealis are based on the perfective *ú-l-* stem (Table 72).

TAM	STEM	Example in 2SG
PRS	zero	<i>sal</i> ‘you are’
PRT	PRS + <i>-ahi/-áhi</i>	<i>sal-ahi ~ sal-áhi</i> ‘you were’
PRS.SUBJ	zero	<i>ov-es</i> ‘(that) you are’
FUT	PRS.SUBJ + <i>-a</i>	<i>ov-eh-a</i> ‘you will be’
COND	PRS.SUBJ + <i>-ahi</i>	<i>ov-es-ahi</i> ‘you would be’
IMP	stem-0	<i>ov</i> ‘be!’
PRT.SUBJ	zero	<i>új-al</i> ‘(that) you were’
IRR	PRT.SUBJ + <i>-áhi</i>	<i>újal-áhi</i> ‘you would have been’

Table 72 Tense-aspect-mood marking of the copula

The function of the present, future, imperative and conditional irrealis corresponds to the function of the tense-aspect-mood values in lexical verbs. I refer to the past indicative form of the copula with the term preterite. Unlike the lexical verbs, the preterite form of the copula does not encode aspectual meaning. The copula together with the imperfective suffix *-ahi* expresses only the potential conditional value. Thus, in contrast to the lexical verbs, the ‘imperfective form’ of the copula is not used in past reference.

While the present subjunctive forms of lexical verbs are identical with the corresponding present indicative forms, the subjunctive form of the copula is based on the root *ov-* instead of the present suppletive root *s- ~ h-*. The preterite subjunctive form is used only in predictive conditional clauses with future reference (cf. preterite form of lexical verbs), e.g. *újen* in (101). The subjunctive is generally introduced either by the non-factual particles *te* and *hod*, or by the modals *šaj* ‘can’, *náštig* ‘cannot’, *site* ‘have to’ and *eremešt* ‘like’.

- (101)^{LQCR} *te akársar újen urd'im,*
 if anyhow COP.PRT.SUBJ.3PL dressed
ni akor tumen na solgálinen ár ando mojakéro.
 neither then 2PL.ACC NEG serve.3PL out in.M pub
 No matter how you were dressed, they would not serve you in the pub.

The copula has two sets of subject concord markers, similarly to the lexical verbs. In the first and second persons, the *s-* ~ *h-* and *úl-* stems employ the concord markers that attach to the perfective stem of lexical verbs (1SG *-um*, 2SG *-al*, 1PL *-am*, 2PL *-en*). On the other hand, the *ov-* stem requires the same concord markers as the present stem of lexical verbs (1SG *-v*, 2SG/1PL *-s*, 3SG *-l*, 2/3PL *-n*).

The suppletive stem *s-* occurs in both the present and the preterite, and the stems *s-* and *h-* in the third-person present (Table 73).

	PRS	PRT
1SG	<i>s-um</i>	<i>s-um-ahi ~ s-um-áhi</i>
2SG	<i>s-al</i>	<i>s-al-ahi ~ s-al-áhi</i>
3SG	<i>h-i ~ s-i</i> <i>h-i=lo, h-i=li</i>	<i>s-in ~ s-íne</i>
1PL	<i>s-am</i>	<i>s-am-ahi ~ s-am-áhi</i>
2PL	<i>s-en</i>	<i>s-en-ahi ~ s-en-áhi</i>
3PL	<i>h-i ~ s-i</i> <i>h-i=le</i>	<i>s-in ~ s-íne</i>

Table 73 Present and preterite Copula forms

The personal markers of lexical verbs with perfective stems are added to the copula stem in the first and second persons in both the present and the preterite. The present third-person form of the copula is *hi ~ si*. The form *hi* generally follows the subject (102), while the form *si* precedes the subject (103).

- (102)^{LQCR} *tut hi duj fat'uj.*
 2SG.ACC COP.3 two children

You have two children.

- (103)^{LQCR} *si tut jag?*
 COP.3 2SG.ACC light
 Do you have a light?

The form *hi* is often accompanied by the enclitic pronouns *-lo*, *-li* (104) and *-le* in order to indicate masculine, feminine or plural form, respectively.

- (104)^{NAR} *Atád hi=li andi špita.*
 Nagyatád COP.3=3SG.F in.F hospital
 She is in the hospital in Nagyatád.

The preterite first and second-person forms of the copula consist of the present copula forms and the imperfective marker *-ahi*, or its allomorphic variant *-áhi*. In the third person we find the apocopated form *sin* alongside the form *síne*. The latter was originally used to express only the third person plural, i.e. **sín-e* ‘COP.PRT-3PL’, cf. **síň-a* ‘COP.PRT-3SG’. The preterite third-person singular form is thus homonymous with the third-person plural form. This may be explained by an analogy to the homonymous third-person singular and plural forms found in the present, i.e. 3SG/PL *hi* ~ *si* ‘is’.

The copula is negated by means of the particle *na* in the first and second persons. In the third person, the irregular forms *nán* ~ *náne* (PRS) and *nána* (PRT) are encountered. The third-person forms are also frequently accompanied by the enclitic pronouns M *-lo*, F *-li*, PL *-le* (see Table 74).

	PRS	PRT
1/2	<i>na</i> COP	<i>na</i> COP
3SG	<i>nán</i> ~ <i>náne</i> <i>nán=lo</i> , <i>nán=li</i>	<i>nána</i> <i>nána=lo</i> , <i>nána=li</i>
3PL	<i>nán</i> ~ <i>náne</i> <i>nán=le</i>	<i>nána</i> <i>nána=le</i>

Table 74 Copula negation in the present and the preterite

The present subjunctive forms of the copula correspond to the present indicative forms of the homonymous verb *ov-* ‘to become, be born’, which belongs to the first sub-class of contracted verbs. Thus, it is formed by the person and number concord markers from the stem *ov-* (Table 75).

	PRS.SUBJ	FUT	COND
1SG	<i>ov-av</i>	<i>ov-á</i>	<i>ov-á(-)hi</i>
2SG	<i>oj-s</i>	<i>oj-h-a</i>	<i>oj-h-ahi</i>
3SG	<i>ó-l</i>	<i>ov-l-a</i>	<i>ov-l-ahi</i>
1PL	<i>ov-as</i>	<i>ov-ah-a</i>	<i>ov-ah-ahi</i>
2/3PL	<i>ó-n</i>	<i>ov-n-a</i>	<i>ov-n-ahi</i>

Table 75 Present subjunctive, future and (conditional) imperfect forms of the copula

The future tense is formed by the suffix *-a* from the present subjunctive form, while the suffix *-ahi* give rise to the conditional form. The imperative is formed from the *ov-* stem: The inflectional stem *ov* occurs in the second-person singular, and the form *ov-en* in the second-person plural.

The preterite subjunctive forms of the copula are identical to the preterite forms of the lexical verb *ov-* ‘to become, be born’. The conditional irrealis is based on the perfective *ú-l-* stem which is followed by the concord markers and the irrealis suffix *-áhi* (see Table 76).

	PRT.SUBJ	IRR
1SG	<i>új-um</i>	<i>új-um-áhi</i>
2SG	<i>új-al</i>	<i>új-al-áhi</i>
3SG	M <i>úl-o</i>	<i>új-á(-)hi</i>
	F <i>úl-i</i>	
1PL	<i>új-am</i>	<i>új-am-áhi</i>
2PL	<i>új-en</i>	<i>új-en-áhi</i>
3PL	<i>úl-e</i>	<i>új-en-áhi</i>

Table 76 Inflection of the copula

Copula inflection in other varieties of Vend Romani

There are some interesting developments regarding the present and past copula forms in Vend Romani. As the Table 77 shows, the palatalized form of the perfective suffix *-in-* constitutes an integral part of the copula stem in the first and second persons only in Szakonyfalu Romani (marked with light grey), while elsewhere the personal markers are attached to the *s-* stem.

	1SG	2SG	3SG	1PL	2PL	3PL
Somogy	<i>s-om, s-um</i>	<i>s-al</i>	<i>h-i, s-i</i>	<i>s-am</i>	<i>s-en</i>	<i>h-i, s-i</i>
Veszprém/Sopron	<i>s-om</i>	<i>s-al</i>	<i>h-i, s-i</i>	<i>s-am</i>	<i>s-an</i>	<i>h-i, s-i</i>
Burgenland	<i>s-om</i>	<i>s-al</i>	<i>h-i, s-i</i>	<i>s-am</i>	<i>s-an</i>	<i>h-i, s-i</i>
Szakonyfalu (Vas)	<i>siň-om</i>	<i>siň-al</i>	<i>h-i, s-i</i>	<i>siň-am</i>	<i>siň-an</i>	<i>h-i, s-i</i>
Zala	<i>s-om</i>	<i>s-al</i>	<i>h-i, -j</i>	<i>s-am</i>	<i>s-an</i>	<i>h-i, -j</i>
Prekmurje	<i>s-om</i>	<i>s-al</i>	<i>h-i</i>	<i>s-am</i>	<i>s-an</i>	<i>h-i</i>

Table 77 Present copula form

It may be observed that Burgenland Romani has the same set of copula forms as Somogy, Veszprém and Sopron Romani. On the other hand, Zala Romani agrees with Prekmurje Romani in the fact that the *s-* stem has been lost in the third person (marked with dark grey). Moreover, the third-person form *hi* is optionally reduced to *-j* when preceded by a vowel in Zala Romani, e.g. *kašuko -j lo* ‘deaf COP.3 3SG; he is deaf’. The sound change *o > u* in the first-person singular personal marker, as well as the change *a > e* in the second-person plural is typical only to (some) Somogy Romani varieties, including KR.

The past copula form in the first and second-persons is composed of the respective present copula form and the suffix *-ahi/-áhi*, which is optionally *-a* in Prekmurje Romani (e.g. *sam-a ~ sam-ahi* ‘we were’). The stem is generally extended by the perfective marker *-in-* in the third person. More precisely, the marker has the form *-in-* in Somogy, Veszprém, Sopron and Burgenland Romani, *-iň-* in Szakonyfalu (Vas) and Zala Romani, while the two forms alternate in Prekmurje Romani (Table 78).

	3SG	3PL	
Somogy	<i>s-ín-(e)</i>	← <i>s-ín-(e)</i>	
Veszprém/Sopron	<i>s-ín-(a)</i>	→ <i>s-ín-(a)</i>	
Burgenland	<i>s-ín-a</i>	→ <i>s-ín-a</i>	
Szakonyfalu (Vas)	<i>s-iň-a</i>	<i>s-iň-an-ahi</i>	← 2PL
Zala	<i>s-iň-a</i>	<i>s-an-ahi</i>	← 2PL
Prekmurje	<i>s-ij-a</i>	<i>s-an-a(hi)</i>	← 2PL
	<i>s-iň-a(hi)</i>		

Table 78 Past copula form

As it is evident from the table above, there is a tendency toward an analogical change in the third person, especially in the third-person plural. That is, the plural form was generalized to the singular in Somogy Romani (SG/PL *sín-e*, cf. SG **sín-a*), while the singular form was taken over to replace the original plural form in Veszprém, Sopron and Burgenland Romani (SG/PL *sín-a*, cf. SG **sín-e*). Moreover, in place of the third-person plural copula form of Szakonyfalu (Vas), Zala and Prekmurje Romani we find the copula form typical for the second-person plural, i.e. 2/3PL *sanahi*. In the third-person singular, the entire form of the imperfective suffix *-ahi* occurs only optionally in Prekmurje Romani.

The negative third-person present copula form is *nán-(e)* in Somogy and Zala Romani, while only the apocopated form *nán* has been encountered in other Vend Romani varieties. Furthermore, the third-person past copula form is generally *nán-a* for both singular and plural, while in some peripheral varieties of Somogy either the form *nán-ahi* occurs, or the two forms alternate.

4.7.5 Non-finite forms

The following section deals with forms that are not marked for tense and person concord. The KR non-finite forms include the participles and the infinitive. Analytic constructions, in which the participles take part, are discussed in section 4.7.2.7.

4.7.5.1 Participles

KR distinguishes active and passive participles. Active participles are involved in the formation of the third-person preterite forms of certain intransitive verbs (see 4.7.3.7). These participles are formed by the perfective markers and inflected as adjectives, e.g. M.SG *áčh-in-o* ‘he stayed’, F.SG *áčh-in-i* ‘she stayed’, PL *áčh-in-e* ‘they stayed’. Passive participles are generally formed by the Greek-origin participial marker *-im-* from xenoclitic transitive verbs (verbs adapted by *-ín-* and *-án-*), and by the unpalatalized perfective markers (*-d-* ~ *-dl-*, *-l-*, *-t-* ~ *-l-*) from oikoclitic transitive verbs. The exceptions are the inherited *d*-verbs and the verbs *d-* ‘to give’ and *l-* ‘to take’, since they employ the participial marker *-im-* typical for xenoclitic verbs (see Table 79). The motivation for this irregularity may be found in the fact that the perfective marker of these verbs (i.e. *-in-*) is formally similar to the xenoclitic participial marker *-im-*.

CLASS	STEM	MARKER	e.g.
oikoclitic	l, r	<i>-d-</i>	<i>ker-d-o</i> ‘done’ < <i>kér-</i> ‘to do’
	v > 0		<i>thó-d-o</i> ‘washed’ < <i>thóv-</i> ‘to wash’
	n	<i>-d-</i> ~ <i>-dl-</i>	<i>čhin-d(l)-o</i> ‘cut’ < <i>čhín-</i> ‘to cut’
	č, g, kh > k, k, m	<i>-l-</i>	<i>pék-l-o</i> ‘baked’ < <i>pék-</i> ‘to bake’
	s, š	<i>-t-</i> ~ <i>-l-</i>	<i>rés-t-o</i> ~ <i>rés-l-o</i> ‘reached’ < <i>rés-</i> ‘to reach’
	d	<i>-im-</i>	<i>čhid-im-o</i> ‘thrown’ < <i>čhíd-</i> ‘to throw’
xenoclitic		<i>-im-</i>	<i>fešt-im-o</i> ‘painted’ < <i>fešt-ín-</i> ‘to paint’

Table 79 Participial markers

Inflection of participles is discussed in chapter 4.2.3.

4.7.5.2 Infinitive

Infinitive is used in modal constructions with the same subject. The complement clause of modal constructions involves the infinitive, while the modal encodes the person and number and tense-aspect-mood categories. The infinitive is based on the third-person singular subjunctive present form, and preceded by the non-factual complementizer *te*, e.g. *te žal* and *te ródel* in (105).

(105)^{LQCR} *külföld kamahahi te žal búti te ródel.*
 abroad want.IMPF.1PL COMP go.INF job COMP look_for.INF
 We wanted to go abroad to look for a job.

The infinitive construction may precede (106) or follow (107) the finite verb, depending on the focus.

(106)^{LQCR} *akán ni te vakérel na žánel.*
 now neither COMP speak.INFNEG know.3SG
 Now s/he cannot even speak.

(107)^{LQCR} *na žánav te vakérel ninčka.*
 NEG know.1SG COMP speak.INF German
 I do not speak German.

Non-finite forms in other varieties of Vend Romani

In all varieties of Vend Romani, the infinitive is composed of the complementizer *te* and the present third-person singular form. The form of the participial and gerund markers, on the other hand, varies only insignificantly across the varieties. The borrowed verbs generally take the xenoclitic marker *-im-*, but not in Tarany Romani (Somogy), where the oikoclitic markers seem to be productive, e.g. *fešt-in-dl-o* ‘painted’ < *fešt-in-* ‘to paint’, cf. KR *fešt-im-o*. It is questionable whether we are dealing with an internal innovation here, or whether this generalization was caused by the low Romani language proficiency of the Tarany Romani speaker.

4.7.6 Verbal particles⁶⁹

The development of verbal particles (in short particles; also referred to as verbal coparticles, preverbs, coverbs, verbal prefixes or prefixal preverbs) in some dialects of Romani is triggered by language contact with Hungarian and/or German (Elšík *et al.* 1999, Iglá 1992, Matras 2002,

⁶⁹ This chapter is a shorter version of the author’s manuscript (Bodnárová & Wiedner 2015a; submitted for publication).

Schrammel 2005). These verbal particles can be separated from a verb and result in a change in the aktionsart, aspect and/or the meaning of a verb. The particle verbs, which consist of a verbal particle plus verb, are generally (semi-)calques of Hungarian particle verbs, regardless of the particle's origin:

com čumíd- 'lit. together kiss' < H *össze-csókol* 'lit. together-kiss; to kiss (fully)'
cuj díkh- 'lit. to(ward) see' < H *hozzá-lát* 'lit. to(ward)-see; to start to do'
donde ža- 'lit. apart go' < H *szét-megy* 'lit. apart-go; to break up'
upre kér- 'lit. up make' < H *fel-csinál* 'lit. up-make; to make pregnant'

A few particle verbs that have no equivalent counterparts in Hungarian calque on German particle verbs, such as *ánk pij-*:

ánk pij- 'lit. on smoke/drink' < G *an-rauchen* 'lit. on-smoke; to light a cigarette'
**upre thár-* 'lit. onto burn' cf. H *rá-gyújt* 'lit. onto-burn; to light a cigarette'

Only a limited number of KR particle verbs seem to copy neither Hungarian expressions, nor German expressions:

com pér- 'lit. together fall; to meet' cf. H *találkozik* (no verbal particle)
 cf. G *sich treffen* (no verbal particle)
cuj čhiv- 'lit. closed put; to close' cf. H *be-zár* 'lit. into-close'
 cf. G *zu-machen* 'lit. closed-make'
com kér- 'lit. together make; to clean up' cf. H *ki-takarít* 'lit. out-clean'
 cf. G *zusammen-räumen* 'lit. together-clear', etc.

These particle verbs were most probably motivated by older German dialect expressions that are no longer in use. Thus, the verbal particles are not productive in KR because they are not used to create new particle verbs that are independent of the source languages (cf. Schrammel 2005:108).

The verbal particles encountered in KR are borrowed and/or calqued from German and Hungarian. There are ten verbal particles of German origin (Table 80). Some of these particles

are primarily used to indicate direction, such as *com*, *cuj*, *fenont*, *fút* and *bajk*, while the particles *link*, *fer*, *ánk*, *mit* and *um* change the aktionsart, aspect, or the verb meaning. In addition, the particles *com*, *cuj*, *fenont*, *fút*, *fer* and *um* (iterative) are used to calque the Hungarian verbal particles, while *link*, *bajk*, *ánk*, *mit* and *um* (resultative) occur only in a limited number of particle verbs that are all German calques.

	H	MEANING
<i>com</i>	<i>össze, egybe</i>	together, into one
<i>cuj</i>	<i>hózzá</i>	towards, to, closed, shut
<i>fenont</i>	<i>szét, széjjel</i>	apart
<i>link</i>	<i>el</i>	[resultative]
<i>bajk</i>	<i>el</i>	away
<i>fút</i>	<i>el</i>	away
<i>fer</i>	<i>el</i>	[resultative]
<i>um</i>	<i>el</i>	[resultative]
	–	[iterative]
<i>ánk</i>	–	at, on, to
<i>mit</i>	–	with, by

Table 80 German-borrowed verbal particles

The particle *com* originates from the dialect form [tsɔ̃m] of the German verbal particle *zusammen*, while the particle *cuj* is borrowed from the *Hianzisch* dialect (also known as the *Heanzisch* or *Hoanzisch dialect*) of German spoken in the Austro-Hungarian border region. This is evident from one of the most characteristic features of this dialect, diphthongisation *u* > *uj* (Hannabauer 2007: 30–31; Wiesinger 1967: 126–127); cf. G *zu* > Hianzisch [tsɔ̃] > KR *cuj* ‘towards, to’. The German calques *cuj špírín-* (cf. G *zu-sperren*) ‘to close, to lock’ and *cuj učár-* (cf. G *zu-decken*) ‘to cover up’ do not have an equivalent in Hungarian.

Some German-origin particles have no corresponding counterparts in German. For instance, the particle *link* is probably extracted from the German verbal expression [li:ŋ-g-lɔsŋ] ‘left’ (cf. G *liegen-ge-(lassen)* ‘lie-PRF-leave’), consisting of the dialect verb for ‘to lie’ plus the prefix of the past participle *g-* (cf. Standard G *ge-*). This form was extracted and

grammaticalised as a verbal particle in KR. An alternative explanation could be that the dialect verb [li:ŋ] (< G *liegen* ‘to lie’) was extracted from the present form of the infinitive [li:ŋ-lɔsŋ], while the final velar nasal of the verb [li:ŋ] was interpreted by speakers as /nk/ or /ng/. It also follows that the particle *link* appears primarily in the German calque *link múk-* (cf. G *liegen-lassen*) ‘to leave’. In addition, this particle has been encountered with the verb *áčh-* ‘to stay, to remain’ (*link áčh-* ‘has been canceled, not to take place’), which is probably calqued on the Hungarian *el-marad* with the same meaning. The particle *fenont* (also attested as *feront* / *fenon* / *fenom*) appears to originate in the German dialect form of the reciprocal pronoun: *fa-r-anaunt*⁷⁰ [fɛrɛ'nɔʊnt] ‘from each other’.

The Hungarian verbal particle *el*, specifying the direction (‘away’) or the result state of an action (Dékány 2008: 3), is calqued by means of the German-origin particles *fút*, *fer* and *bajk*. Similarly to German, the KR particles *fút* (< G dial. [fʊɣt], cf. G *fort* ‘away’) and *bajk* (< G dial. [vɛɪŋ], cf. G *weg* ‘away’) have directional meaning, while the particle *fer* (< G *ver-*[resultative]) indicates resultativity. However, owing to influence from Hungarian, the particle *fút* occasionally has the resultative meaning (for example, *fút čór-* ‘to steal’, *fút kišérin-* ‘to accompany’); and vice versa, the particle *fer* rarely may have directional meaning (for example, *fer trád-* ‘to chase away’). It seems therefore that the strong Hungarian language influence on KR is progressively blurring the original functional difference between these particles. Finally, the particle *bajk* can be used only with three inherited verbs, calquing German expressions: *bajk žal* (cf. G *weg-gehen*) ‘to go away’, *bajk l-* (cf. G *weg-nehmen*) ‘to take away’ and *bajk čhiv-* (cf. G *weg-legen*) ‘to put away’.

The verbal particle *um* denotes the outcome of an action, that is resultativity. The only particle verb where the particle has resultative function is *um pér-* (cf. G *um-fallen*) ‘to fall’, which is a direct calque from German. The same particle is used to mark the repetition or the frequency of an action, i.e. iterativity; e.g. (108)–(109).

- (108)^{LQCR} *cilo paloplán um truskinlahi.*
 whole.M afternoon ITER sneeze.IMPF.3SG
 S/he was sneezing a lot the whole afternoon.

⁷⁰ The Standard German equivalent is ‘voneinander’, a hiatus avoiding /t/ occurs only in the dialect (cf. Zehetner 1985: 88).

(109)^{LQCR} *site um phíres, mert fad'ínes.*
 should ITER walk.2SG because freeze.2SG
 You should walk a lot, otherwise you are going to freeze.

The origin of *um* as an iterative marker is probably from the German dialect verbal particle *uma (dum)* ['ʊmə, ʊmə'dʊm] [iterative]. It seems that due to the similarity of the German (dialect) forms *um* (resultative) and *uma(dum)* (iterative), these particles merged in KR. The development of a verbal particle with an iterative meaning is an interesting contact phenomenon, since KR also has the productive inherited iterative suffixes *-(in)kér-* and *-(in)gér-*. Thus we may find *náš-ingér-* 'run-ITER-3SG' alongside *um náš-* 'ITER run; to run around'. Nonetheless, the suffixed verbs occur more frequently than the particle verbs with the particle *um*. On the other hand, the iterative meaning of verbs is also marked in Hungarian by suffixation (*-gat-*, *-get-*). This may therefore imply that the current language contact with Hungarian – and this particular typological similarity – reinforces the use of the inherited suffixes.

Finally, the German-borrowed particle *ánk* (< G dial. [ãŋ, ɒŋ], cf. G *an* 'at, on, to') appears with the inherited verb *píj-* 'to drink' (*ánk píj-*, cf. G *an-rauchen*) 'to light a cigarette'), while the particle *mit* (< G *mit* 'with, by') is bound to the verb *cíd-* 'to pull': *mit cíd-* (cf. G *mit-bringen*) 'to bring'.

Six Hungarian-borrowed verbal particles are attested in KR (Table 81).

	H	MEANING
<i>körü(l) / kiri(l)</i>	<i>körül</i>	around
<i>esbe</i>	<i>észre</i>	to the mind
<i>abba</i>	<i>abba</i>	into it
<i>neki</i>	<i>neki</i>	against
<i>pujsto</i>	<i>tönkre</i>	into ruin
<i>meg</i>	<i>meg</i>	[perfective]

Table 81 Hungarian-borrowed verbal particles

The most common borrowed verbal particles are *körü(l) / kiri(l)* that has the spatial meaning 'around'. The particle *esbe* occurs only in the particle verb *esbe l-* (cf. H dial. *észbe-*

vesz) ‘to notice’, and the particle *abba* only in the expression *abba múk-* (cf. H *abba-hagy*) ‘to stop doing’. In addition, the particle *neki* occurs in the data in combination with two verbs: *neki žal* ‘to go against’ and *neki náš-* ‘to run against’. This particle is probably used with other verbs of movement as well, although, due to its specificity it is not attested in the data. The particle *pujsto*, which replaces the Hungarian verbal particle *tönkre* ‘into ruin’, in the Hungarian-calqued expressions, seems to be borrowed from the Hungarian noun *puszta* ‘waste’ (compare H *puszt-ul* ‘to decay, to perish, to waste away’). Furthermore, the perfectivity of the verb can be expressed by the borrowed verbal particle *meg* (< H *meg*). However, this verbal particle is only optionally ‘translated’ into KR from Hungarian, e.g. (110)a vs. (110)b, cf. the corresponding Hungarian expression (111).

(110)^{LQCR}

- a. *meg le mange parikerd'a.*
 VP 3SG.M 1SG.DAT thank.PRT.3SG
- b. *parikerd'a mange.*
 thank.PRT.3SG 1SG.DAT
 S/he thanked me (for it).

(111) Hungarian

meg-köszönte nekem.
 VP-thank.PRT.3SG 1SG.DAT
 S/he thanked me for it.

The only verbal particle of Slavic origin that occurs in the data is *prik* (compare the South Slavic local adverb, *preko*), occurring in verbs such as *prik astár-* ‘to embrace’, *prik fárín-* ‘to jump over’, or *prik másín-* ‘to climb (over)’.

KR has also a number of particles of inherited matter which were mostly derived from local adverbs (Table 82). Furthermore, the verbal particle *šé-s-te* is the contracted form of the locative *šér-es-te* ‘head-OBL.SG-LOC’ (< *šéro* ‘head’), calquing the respective Hungarian particle *agy-on* ‘lit. brain/head-on; on the head’] ‘to death, to excess’ (cf. Ladányi 2000: 116), e.g. *šéste čhín-* ‘to strike to death’, *šéste phosingér-* ‘to stab to death’. The verbal particle *náši* occurs only in the particle verbs *léž- náši* ‘to elope’ and *ža- náši* ‘to escape’ with lexicalized meaning.

	H	MEANING
<i>órde</i>	<i>ide</i>	(to) here
<i>óđ'a</i>	<i>oda</i>	(to) there
<i>ángle</i>	<i>elő, előre</i>	forward
<i>pál(e)</i>	<i>vissza, hátra</i>	(to the) back
<i>tél(e)</i>	<i>le, alá</i>	down, (to) under
<i>upre</i>	<i>fel, rá</i>	up, onto
<i>ár(i)</i>	<i>ki</i>	out
<i>ánde</i>	<i>be, bele</i>	in, into
<i>šéste</i>	<i>agyon</i>	to death, to excess
<i>náši</i>	–	–

Table 82 Calqued verbal particles

Verbal particles in other varieties of Vend Romani

The particle *dond(e)* ‘apart’, which in some localities alternates with the German-borrowed *fenont*, occurs in almost all varieties of Vend Romani, but not in KR. It has its origin in the locative form of the numeral *duj*, ‘two’:⁷¹ *duj-en-de* ‘two-OBL.PL-LOC’ (> contracted *dond(e)*), resulting, for example, in *donde žal* ‘to break up, separate’.

There is a significant difference between the verbal particles of Somogy and Zala Romani (Table 83). Unlike Somogy Romani, Zala Romani has no verbal particles that have been borrowed from German. The functions of the Somogy Romani particles that have been borrowed from German, such as *fenont* ‘apart’, and *com* ‘together, into one’, are adapted by the calqued *donde* (see above), and *khetán* (< **jekhe-than-e* ‘(to) one-place’), respectively. In the directional meaning ‘away’, we find the inherited *kh-rik* (< **jekh-rig*), however, the usage of this particle is very limited. The Slavic-origin *úze* (< South Slavic *uz*) occurs in the meaning of ‘towards’. Furthermore, Zala Romani does not mark the verb for resultativity. On the other hand, the Hungarian-origin particles are similar to those in Somogy Romani: *esbe* ‘to the mind’, *keril* ‘around’, *meg* [perfective], and the recent loanword *tönkre* (< H *tönkre*), alongside *pujsto* ‘into ruin’.

⁷¹ Viktor Elšík, personal communication, March 2013. On the other hand, Vekerdi (1984: 74) proposes the etymology *donde* ‘up to the point’, to be of Serbo-Croatian origin, without indicating the source form.

The German-origin particle *fer* with resultative meaning is being progressively replaced by the Hungarian *el*, especially in some varieties of central Somogy. There are also varieties that only have the newly-borrowed Hungarian particle *el*, while the ‘earlier’ German-origin particle, *fer*, has remained more vital in the periphery. It is therefore clear that the borrowing of the Hungarian particle *el* was triggered by the existence of the earlier German-origin particle *fer*. The particle *fer* is employed in Prekmurje Romani, while in Burgenland Romani, the German dialect form *fa-* (< G dial. [fa] < G *ver-*) occurs. This particle has become extinct without traces in Zala Romani.

By comparing the Romani verbal particles in Somogy (represented by the set of particles found in KR) and Zala Romani as well as those in the related varieties of Burgenland and Prekmurje Romani (Table 83), we discover that Somogy Romani has a set of verbal particles that are very similar to Burgenland Romani (indicated in light grey); and vice versa, the verbal particles in Zala Romani are highly similar to those that occur in Prekmurje Romani (indicated in dark grey).

#	MEANING	ORIGIN	VERBAL PARTICLES	SR	BR	ZR	PR
1	(to) there	calque	<i>ód'a, óča, ódža</i>	✓	✓	✓	✓
2	(to the) back	calque	<i>pál(e)</i>	✓	✓	✓	✓
3	in(to)	calque	<i>ánde</i>	✓	✓	✓	✓
4	out	calque	<i>ár(i)</i>	✓	✓	✓	✓
5	up, onto	calque	<i>upre</i>	✓	✓	✓	✓
6	down	calque	<i>tél(e)</i>	✓	✓	✓	✓
7	forward	calque	<i>ángle</i>	✓	✓	✓	✓
8	(to) here	calque	<i>órde</i>	✓	✓	✓	✓
9		G	<i>dahea</i>		✓		
10	through, over	S	<i>prik</i>	✓	✓	✓	✓
11	into ruin	H	<i>pujsto</i>	✓	✓	✓	✓
12		H	<i>tönkre</i>			✓	
13	to the mind	H	<i>esbe</i>	✓	✓	✓	✓
14	apart	calque	<i>dond(e)</i>	✓	✓	✓	✓

#	MEANING	ORIGIN	VERBAL PARTICLES	SR	BR	ZR	PR
15		G	<i>fenont, fenaunt</i>	✓	✓ ⁷²		
16	around	H	<i>körü(l), kiril, kirin</i>	✓		✓	✓
17		G	<i>um</i>		✓		
18	towards, to	G	<i>cuj</i>	✓	✓		
19		S	<i>úze</i>		✓	✓	✓
20	away	calque	<i>khrik, krik</i>			✓	✓
21		G	<i>fuat, fút</i>	✓	✓		
22		G	<i>bejg, bajk</i>	✓	✓		
23	[resultative]	G	<i>fer, fa-</i> [prefix]	✓	✓		✓
24		G	<i>um</i>	✓	✓		
25		G	<i>link</i>	✓			
26		H	<i>el</i>	✓			
27	together, into one	calque	<i>(e)khetán</i>		✓	✓	✓
28		G	<i>com, cam</i>	✓	✓		
29	at, on	G	<i>aun, ánk</i>	✓	✓		
30	to death	calque	<i>šéste</i>	✓	✓		
31	[iterative]	G	<i>um</i>	✓			
32	with, by	G	<i>mit</i>	✓			
33	into it	calque	<i>and' odá</i>			✓	
34		H	<i>abba</i>	✓			
35	[perfective]	H	<i>meg</i>	✓		✓	
36	against	H	<i>neki</i>	✓			

Table 83 Verbal particles in SR (= Somogy Romani), BR (= Burgenland Romani), ZR (= Zala Romani) and PR (= Prekmurje Romani); ✓ present

The Romani varieties spoken in Somogy and Burgenland share German-borrowed particles (# 15, 18, 21-24, 28-29) as well as a calqued particle (# 30). In contrast, the Romani varieties that are spoken in the two neighbouring regions of Hungary and Slovenia, Zala and Prekmurje Romani, share calqued particles (# 20, 27) and a South Slavic particle (# 19),

⁷² The German-origin particle *fenaunt* has been attested only once in Knobloch (1953: 32): *tšūri las taj fēnáunt me pharā'was* 'we take the knife and cut it apart'.

whereas both varieties lack most German particles. However, an absence of German-borrowed particles, except for the particle *fer* in Prekmurje Romani, does not necessarily imply a lack of German influence. For instance, the South Central varieties spoken in Slovakia also tend to calque the verbal particles rather than borrow them from Hungarian, which is the recent contact language of that dialect group. In addition, evidence suggests that Burgenland Romani is in a transitional position, as alongside German particles (# 18–28) it also possesses Slavic and calqued verbal particles (# 19–27) typical for the Prekmurje–Zala Romani subgroup. With regard to similarities, all the compared varieties share calqued particles (# 1–8, 14), Hungarian-origin particles (# 11, 13) as well as a particle originating from South Slavic (# 10).

The fact that most of the shared particles of Somogy and Burgenland Romani have been borrowed from German implies that both varieties have been in an intensive language contact with German. Considering the current geographical location of the speakers (1.5–3), it can be suggested that the speakers of Somogy Romani migrated from the cross-border regions of Hungary and Austria to their current location. The very similar development of verbal particles in Zala and Prekmurje Romani indicates that the speakers of these two varieties have also been in mutual contact throughout the past.

5 Syntax

5.1 Case marking

In this chapter I will follow the terminology of Elšík and Matras (2006: 218–238). They distinguish between ‘core’ and ‘adverbial’ case roles, which are defined ‘as grammatical relations and/or thematic roles encoded by inflectional cases and adpositions, with the exception of local and temporal case relations’ (ibid: 219). According to them, the core case roles include the Subject, the Experiencer, the Predicative, the Object, the Recipient, the Possessee, Possessor, and the Adnominal and External Possessor. I will deal with the first five case roles in section 5.1.1, and with the last four case roles in section 5.1.2. Section 5.1.6 is devoted to the adverbial case roles which include the Benefactive, Goal, Comitative, Instrument, Reference, Source, Material, Origin, Partitive, Reason, Privative, Substitutive, and Exceptive roles. The Comparative (standard of comparison) and Equative roles (standard of equation), classified also as adverbial case roles in Elšík and Matras (2006), are dealt with in section 0. The local and the temporal case relations are examined in section 5.1.4 and 5.1.5, respectively.

Since I analysed the data using the *Linguistic Database for the Documentation of Central European Romani* (2008–a), where the elicited and transcribed sentences of the *Linguistic Questionnaire for the Documentation of Central European Romani* (2008–b; see 1.7.2.1) are tagged for semantic and syntactic functions, I have decided to proceed from **function to form** in presenting my data. The summary of functions of synthetic case markers (i.e. form to function) is found in Table 91 at the end of this section.

5.1.1 Basic syntactic structures

The case roles listed in Table 84 correspond to those in Elšík and Matras (2006: 218–238). According to their use, the term Subject refers to the canonical transitive or intransitive subject, the Experiencer to the non-canonical subject which experiences or undergoes the effect of an action, the Predicative to the nominal predicate of the clause, and the Object to the direct object. The Recipient is understood here as an animate being that receives ‘something concrete transferred to its sphere of control’ (Kittilä 2005: 274). In KR, these case roles are generally marked by synthetic cases (Table 84).

CASE ROLE	MARKING	e.g.
Subject	NOM	(112)
Experiencer	DAT	(118)
	ACC	(119)
Predicative	NOM	(114)
Predicative: complement	DAT	(see examples in the text below)
	<i>ande, upre</i>	(120)–(121)
Object: inanimate	NOM	(113)
Object: animate	ACC	(115)
Recipient	DAT	(117)
Recipient: ‘to give’	ACC (~ DAT)	(116)

Table 84 Basic syntactic structures

The Subject (112), the inanimate Object (113) and the nominal predicate of the clause (114) are expressed by the nominative case.

(112)^{LQCR} *me na sum khér.*
 1SG.NOM NEG COP.1SG at_home
 I am not at home.

(113)^{LQCR} *dikjum o kher.*
 see.PRT.1SG DEF house.NOM
 I saw the house.

(114)^{LQCR} *lakro moštómno_dad pádári hi.*
 her step-father doctor.NOM COP.3
 Her stepfather is a doctor.

The accusative case is used to mark the animate Object (115).

(115)^{LQCR} *dikjum ole manuše.*
 see.PRT.1SG DEF.OBL man.ACC
 I saw the man.

The Recipient of the verb *d-* ‘to give’ is often expressed by the accusative (116) instead of the dative case, which is otherwise the most commonly used case to express the Recipient role (117) in KR.

- (116)^{LQCR} *dijum mra babuka čati.*
 give.PRT.1SG my.OBL lover.ACC hairgrip.PL
 I gave my lover hairgrips.

- (117)^{LQCR} *na kámes tu mange.*
 NEG owe.2SG 2SG 1SG.DAT
 You don't owe me.

In clauses expressing change of state, the predicate complement is marked by the dative, which replicates the corresponding Hungarian structure. Examples are *pál pérel muršeske* ‘turn.3SG man.DAT; he turns into man’, *sikjol átóserelőske* ‘study.3SG car_mechanic.DAT; he is studying to be a car mechanic’, or *géli bótoškiňake* ‘go.PRT.3SG saleswoman.DAT; she started to work as a saleswoman’. The dative case is also reserved for the Experiencer arguments (118), except of some body-state expressions (119), where the accusative case is used. The dative is further required by the Experiencer with the verbs *teccín-* ‘to like’ and *izlín-* ‘to taste’.

- (118)^{LQCR} *erd'avo tuke ovlahi.*
 bad 2SG.DAT COP.COND.3SG
 You would feel bad.

- (119)^{LQCR} *šil hi man.*
 cold COP.3 1SG.ACC
 I am cold.

The prepositions *ande* (120) or *upre* (121) are used to introduce the complements of the change of quality verbs.

(120)^{NAR} *ék máro ando štár čhindle.*
 a bread into four cut.PRT.3PL
 They cut up bread into four pieces.

(121)^{LQCR} *tél čhinda la ka'aha lakere bal upro hárno.*
 down cut.PRT.3SG DEF.OBL scissor.INS her.PL hair on short
 S/he cut her hair short with the scissors.

5.1.2 Possessive constructions

Possessive constructions in KR include the predicative (also called clausal), adnominal (also called attributive) and external constructions. The term predicative possession refers to a ‘syntactic construction whose function is to predicate the existence of a possessive relation’ (Baldi & Nuti: 246). On the other hand, adnominal possession comprises ‘nominal phrases which are linked together according to certain parameters, such as word order or the presence/absence of possessive marker(s)’ (ibid). As defined by Heine (1997: 143), the adnominal/attributive possession is ‘presupposed’, while the predicative possession is ‘asserted’. The term external possession, as defined by Payne and Barshi (1999: 3), refers to a construction ‘in which a semantic possessor-possessum relation is expressed by coding the possessor as a core grammatical relation of the verb and in a constituent separate from that which contains the possessum’. A cross-dialectal comparison of the Romani external constructions is found in Crevels and Bakker (2000).

CASE ROLE	MARKING	e.g.
Possessee	NOM	(122)
Possessee: physical/mental state	INS	(125)–(126)
Possessor	ACC (~ DAT)	(122), ~ (123)
Possessor: physical contact	LOC	(124)
Possessor: Adnominal	GEN	(127)–(131)
Possessor: External	DAT	(132)a, (133)
Possessor: Adnominal-external	DAT+GEN	(134)–(135)

Table 85 Case marking in possessive constructions

According to Elšík and Matras (2006: 219), the relevant case roles related to possession for Romani include the Possessee and Possessor in predicative constructions, as well as the Adnominal and External Possessor. Table 85 shows the case marking reserved for the individual case roles in KR.

5.1.2.1 *Predicative possession*

In KR, predicative possessive constructions involve the possessor in accusative, the Possessee (also called possessum or possessed object) in nominative, and the copula, which agrees with the possessee in person and number, e.g. (122). The copula may have different tense-aspect-mood values.

- (122)^{LQCR} *si man duj phěňa*
 COP.3 1SG.ACC two sister.PL.NOM
 I have two sisters.

Due to the influence of Hungarian, the possessor in a predicative possession may exceptionally be marked for the dative, such as *lake* in (123).

- (123)^{NAR} *úče šorkengeri bočkora lake sin.*
 high-heeled boot 3SG.F.DAT COP.PRT.3
 She had boots with high heels.

The possessor takes the locative case when it is involved in a physical contact with the possessed object, such as the possessor *leste* in (124).

- (124)^{NAR} *akebor kópal leste síne!*
 this_size stick 3SG.LOC COP.PRT.3
 He had such a stick!

In predicative constructions, the Possessee is often in instrumental case when it refers to physical (125) or mental attributes (126). This type of possession requires the Possessor to be in the nominative case, e.g. the clitic pronoun *lo* in (125), and *li* in (126).

(125)^{NAR} *zelene út'henca=lo sin*
 green.OBL eye.PL.INS=3SG.M COP.PRT.3
 He had green eyes.

(126)^{LQCR} *lačhe vód'iha=li síne.*
 good.OBL heart.INS=3SG.F COP.PRT.3
 She was kind-hearted.

5.1.2.2 Adnominal possession

Adnominal possessive constructions in KR require the Possessor to be marked for genitive, irrespective of the Possessor's animacy, e.g. *mri* in (127), and *verdaskero* in (128).

(127)^{NAR} *aja mri phen hi*
 this.F 1SG.GEN sister COP.3
 She is my sister.

(128)^{LQCR} *phad'ino le verdaskero kereko.*
 break.PRT.3SG.M DEF.OBL cart.GEN wheel
 The wheel of the cart broke.

Adnominally expressed are also possessive constructions referring to age (129) and body-parts being in pain (130).

(129)^{LQCR} *ék masekero murš_fat'ú*
 a month.GEN boy
 a one-month old boy

(130)^{LQCR} *dúkhal lakro pér.*
 hurt.3SG 3SG.F.GEN belly
 Her belly hurts.

Note that the Possessor in the genitive case serves as the head noun of the preceding adjective(s). For instance, the genitive Possessor (or possessive determiner) *manuš-es-ker-o* is the head noun of the adjective *phur-e* in (131).

- (131)^{NAR} *phur-e* *manuš-es-ker-o* *ánav*
 old-OBL.M.SG man-OBL.M.SG-GEN-M.SG name
 the name of the old man

5.1.2.3 External Possession

External possessive constructions occur only sporadically in KR, in the cases where the Possessor is encoded with the dative case (132)a. However, in similar possessive constructions the speakers favoured the adnominal (132)b instead of the external construction (132)a.

- (132)^{LQCR}
- a. *upro va lake síne i angrusti.*
 on.M hand 3SG.F.DAT COP.PRT.3 DEF ring
- b. *upre lakro va síne i angrusti.*
 on.M 3SG.F.GEN hand COP.PRT.3 DEF ring
 She had a ring on her hand!

The external construction is frequent when it denotes kinship relation, e.g. (133).

- (133)^{NAR} *moštohatešvér hi mange.*
 step-brother COP.3 1SG.DAT
 He is my step-brother.

The external construction may also be accompanied by an adnominal construction, which has been attested only in the spontaneous language data (134)–(135).

- (134)^{NAR} *mange účo hi mro vérňomáš.*
 1SG.DAT high COP.3 1SG.GEN blood_pressure
 I have a high blood pressure.

- (135)^{NAR} *tuke trí daj hi?*
 2SG.DAT 2SG.GEN mother COP.3
 Is she your mother?

5.1.3 Comparative constructions

Comparative constructions in KR are generally formed by the comparative particle *sar* ‘than’ (136), by means of which the subject of comparison (such as *leskero kher*) and the standard of comparison (such as *mro*) is compared. The standard of comparison is marked for the same case as the subject of comparison.

- (136)^{LQCR} *leskero kher baréder hi sar mro.*
3SG.M.GEN house bigger COP.3 than 1SG.GEN
His house is bigger than mine.

If the subject of comparison functions as the grammatical subject of the clause, the standard of comparison may also be marked by the ablative (137). In these constructions, the subject of comparison is in the nominative case, while the comparative particle is absent. Note that the quantitative difference (such as *pándž beršenca*) of the objects that are compared is marked for the instrumental.

- (137)^{LQCR} *lakero phral pándž beršenca hi phuréder lestar.*
3SG.F.GEN brother five year.PL.INS COP.3 older 3SG.M.ABL
Her brother is five years older than him.

The particle *sar* is obligatorily used in constructions in which two equal noun phrases are being compared (138). Like in unequal comparison (137), the standard of comparison receives the same case as the subject of comparison, e.g. the accusative case in (138).

- (138)^{NAR} *grundlaste žute bal hi la sar tut.*
curly blond hair COP.3 3SG.F.ACC as 2SG.ACC
She has as blonde hair as you.

The equality of the two compared entities may be emphasised by demonstratives, as for instance by *asó* ‘such’ in (139).

- (139)^{NAR} *asó báro sin sar tu.*
such.M big COP.PRT.3 as 2SG
He was as big as you.

5.1.4 Local adverbials

In this section, I will deal only with noun phrases (e.g. *Pešt-ate* ‘Budapest-LOC; in/to Budapest’) and adpositional phrases expressing spatial relations (e.g. *ando vodro* ‘in_the bed; in the bed’). The local (non-phrasal) adverbs are discussed in section 4.3.1. I will adapt the categorisation of local relations proposed by Elšík and Matras (2006: 239–242) for Romani, where each localisation value denotes distinct spatial position between the ‘figure object’ and the ‘ground object’. The localisations relevant for KR include the inessive (‘inside of’), contact-superior (‘on the surface of’), adessive (‘at’), proximate (‘by’), superior (‘over, above’), translative-perlative (‘across, over; through’), inferior (‘under, below’), anterior (‘in the front of’), posterior (‘behind’), medial (‘between, among’), and circumlative (‘around’) (see Table 86). In addition, I have added the localisation termed as ‘direction’, which encodes the meaning ‘in the direction of’.

Three orientations may be distinguished in most of the localisations, i.e. stative, directive and separative (Elšík & Matras 2006: 271–273). As it may be observed in Table 86, individual adpositional forms are distinguished for the stative/directive and the separative orientations. More precisely, the prepositions have mostly ablative forms with *-al/-ar* in the separative role, while the simple, non-derived, forms in the stative and the directive role. Only the ablative form has been attested in the stative/directive of proximate, anterior, medial, superior, translative-perlative and circumlative localisations (i.e. *uz-ar*, *angj-al*, *mašk-ar*, *ped-ar*), while the ablative form alternates with the base (non-ablative) form in inferior and posterior stative localisations, i.e. *tel* ~ *telal*, *pal* ~ *palal*. The original preposition *kija* (< **ke* ‘at, to’, cf. Elšík & Matras 2006: 265) has been fossilized in the temporal adverb *kija-ráti* ‘to-night; in the evening’. The separative orientation is unattested in KR for proximate, superior, translative-perlative, posterior, medial and circumlative localisations. The stative orientation in the direction localisation (‘in the direction of’) is absent.

The adpositional phrases denoting location involve several inherited (*ande*, *andral*, *upre*, *upral*, *pedar*, *tel*, *telal*, *angjal*, *pal*, *palal*, *maškar*) and borrowed adpositions (*uz-e/uz-ar* < S *uz*, *prik* < S *prek*, *körü* < H *körül*, *mer/mer-al* – unclear origin, cf. Elšík *et al.* 1999: 375). The adpositions *uz-ar* and *mer-al* are internally derived by means of the historical ablative marker *-al* and its variant *-ar* from the borrowed roots *uz* and *mer*, respectively. Diachronically, the vast majority of these adpositions are derived from local adverbs (see 4.3.1).

LOCALISATION & ORIENTATION	MARKING	e.g.
Inessive stative/directive	<i>ande</i>	<i>ando vodro</i> ‘in/to the bed’
Inessive separative	<i>andral</i>	<i>andral o vodro</i> ‘out of the bed’
Contact-superior stative/directive	<i>upre</i>	<i>upro vodro</i> ‘on/onto the bed’
Contact-superior separative	<i>upral</i>	<i>upral o vodro</i> ‘from the bed’
Adessive stative/directive	<i>uze</i>	<i>uzo vodro</i> ‘at/by the bed’
Adessive separative	<i>uzar</i>	<i>uzar o vodro</i> ‘from the bed’
Proximate stative/directive	<i>uzar</i>	<i>uzar o vodro</i> ‘next to the bed’
Superior stative/directive	<i>pedar ~ prik</i>	<i>pedar ~ prik o vodro</i> ‘above/over the bed’
Translative-perlative stative/directive	<i>pedar ~ prik</i>	<i>pedar ~ prik o vodro</i> ‘over/through the bed’
Inferior directive	<i>tel</i>	<i>tel o vodro</i> ‘under the bed’
Inferior stative	<i>tel ~ telal</i>	<i>tel ~ telal o vodro</i> ‘under the bed’
Inferior separative	<i>telal</i>	<i>telal o vodro</i> ‘from under the bed’
Anterior stative/directive/separative	<i>angjal</i>	<i>angjal o vodro</i> ‘(from) in front of the bed’
Posterior directive	<i>pal</i>	<i>pal o vodro</i> ‘behind the bed’
Posterior stative	<i>pal ~ palal</i>	<i>pal ~ palal o vodro</i> ‘behind the bed’
Medial stative/directive	<i>maškar</i>	<i>maškar o vodri</i> ‘between the beds’
Direction directive	<i>mer</i>	<i>mer o vodro</i> ‘towards the bed’
Direction separative	<i>meral</i>	<i>meral o vodro</i> ‘from the direction of the bed’
Circumlative stative/directive	<i>körü</i> <i>pedar</i>	<i>~ körü ~ pedar o vodro</i> ‘around the bed’

Table 86 Marking of local adverbials

The adpositions are placed before the noun phrase. The prepositions *andral*, *angjal*, *pedar* and *maškar* may become shortened before the definite article to *andr-*, *angl-*, *pedr-* and

maškr-, respectively, e.g. *angjal o kher* ~ *anglo kher* ‘in front of the house’, cf. *angjal mro kher* ‘in front of my house’, not **angle mro kher*.

Marking of names of localities in the inessive localisation may differ from the marking of other local adverbials (see Table 87).

LOCALISATION & ORIENTATION	MARKING	e.g.
Inessive stative/directive	<i>ande</i>	<i>andi Kišbajom</i> ‘in/to Kisbajom’
	zero-marked	<i>Kišbajom</i> ‘in/to Kisbajom’
Inessive separative	<i>andral</i>	<i>andral i Kišbajom</i> ‘from Kisbajom’
	<i>fenal</i> ~ <i>fen</i> ~ <i>fe</i>	<i>fenal Kišbajom</i> ‘from Kisbajom’

Table 87 Marking of names of localities

In KR, the stative/directive inessive preposition *ande* ‘in, to’ is often omitted before proper names of localities. In other words, these localisations are zero marked. Examples are the stative *Vesprim med’e* and *Pirit* in (140), and the directive *Kišbajom* in (141).

(140)^{NAR} *mró dad meg Vesprim med’e, Pirit hi temetim.*
 my father and Veszprém county Nagypirit COP.3 buried
 And my father is buried in the Veszprém County, in Nagypirit.

(141)^{NAR} *fút indulinde órde Kišbajom, ere mer Kutaš ále.*
 VP start_off.PRT.3PL here Kisbajom here in_direction_to Kutas come.PRT.3PL
 They came here to Kisbajom, so they came in the direction of Kutas.

The prepositional phrase with the inessive *ande* ‘in, to’ was more favoured in the elicited data, while the unmarked expression was preferred in the spontaneous language data. The separative inessive form of the names of localities is expressed by the prepositional phrase employing either the inherited preposition *andral* ‘from’ or the internally derived *fen-al* (less commonly *fen* ~ *fe*; < G *von*) ‘from’, e.g. *andral o Ninčko* ‘from Germany’, *fenal Kapoš* ‘from Nagykapos’.⁷³

⁷³ Note that the separative preposition *fenal* does not require the definite article as compared to *andral*.

The stative/directive inessive form of the Hungarian capital *Pešta* ‘Budapest’ is marked by the locative *Pešta-te* ‘Budapest-LOC; in/to Budapest’, while the separative inessive is formed by the locative form together with the separative prepositions *fenal* or *andral*, i.e. *fenal* ~ *andral Pešta-te* ‘from Budapest-LOC; from Budapest’. Exceptionally, also the noun *gav* ‘village’ was marked for the locative case in stative/directive inessive; however, this form requires in addition the preposition *ande*: *ande gaves-te* ‘in/to village-LOC’ ~ *ando gav* ‘in/to village; in/to the village’. Interestingly, in the directive inessive localisation the borrowed noun *iškola* ‘school’ has been attested once as a locative *iškolen-de* ‘school.PL-LOC; to the schools’ (142).

- (142)^{NAR} *mer štár fa’uj sin, taj štár iškolende phirnahi.*
 because four children COP.PRT.3 and four school.LOC attend.IMPF.3PL
 Because they were four children, and so they attended four schools.

Both locative forms, *gaveste* and *iškolende*, have been attested in the sample of spontaneous speech. The locative is also employed in the idiomatic phrase *čhív- šúkende* ‘put dry.LOC; to change diaper’.

The dative case marking is used for the directive meaning ‘to move against something’ (143)–(144).

- (143)^{NAR} *leskero dumo ole grasteske, me meg le kašteske man mükav.*
 his back DEF.OBL.horse.DAT 1SG and DEF.OBL tree.DAT 1SG.ACC lean.1SG
 He leaned against the horse, and I leaned against the tree.

- (144)^{RM} *čalav neki le vudareske!*
 throw.IMP.2SG against DEF.OBL door.DAT
 Throw it against the door!

Local adverbials in other varieties of Vend Romani

The anterior preposition has the non-ablative form *angl-* ‘in front of’ in Zala, Prekmurje and Burgenland Romani, while in several other varieties both *angl-* and *angl-al* has been attested. In a few northern varieties of Somogy Romani, only the derived form *prek-al* ‘through’ occurs, while the original form *prek* is found elsewhere. These forms may optionally be

replaced with *pedar* (~ *pedr-*) in some Somogy Romani varieties and in Burgenland Romani. The stative and/or directive inessive forms of local names are zero marked also in Burgenland and Prekmurje, pointing to the fact that we are dealing with a shared innovation in Vend Romani.

5.1.5 Temporal adverbials

Simultaneous temporal relation is marked by noun phrases, adpositional phrases or synthetic cases in KR. Further (non-phrasal) temporal adverbs are discussed in section 4.3.2.

Noun phrases are quite common in the simultaneous relation. These phrases are marked for the nominative case. The determiners *árto* ‘next’, *sak-o* ‘every’ and *cil-o* ‘whole’, and the adverbs *adi* ‘today’ and *ič* ‘yesterday’ are employed with various parts of the day (e.g. *árto dí ra’aha* ‘next day in the morning’, *ič pal o plán* ‘yesterday afternoon’), while the determiners *adá* ‘this’, *cil-o* ‘whole’, *sak-o* ‘every’, *oká* ‘last/next’ and the adverb *lani* ‘last year’ are used with years (e.g. *cilo berš* ‘whole year’), months (e.g. *lani ando januári* ‘last year in January’) and seasons (e.g. *oká berš línaj* ‘last year in summer’).

Simultaneous relation is encoded by adpositions with months of the year (*ande* ‘in’, e.g. *ando január* ‘in January’, see also 4.3.2) and the parts of the day *angl-o plán* ‘before_the noon; forenoon’ and *pal o plán* ‘lit. after the noon; afternoon’. The use of the dative case in the simultaneous relation can be found with clock time. More precisely, the dative form of the borrowed noun *óra* (< H *óra*) ‘o’clock’ is used together with the numeral referring to the respective time, e.g. *jékh óra-ke* ‘one o’clock-DAT; at one o’clock’, *šóv óren-ge* ‘six o’clock-DAT; at six o’clock’. To emphasise that an event has to take place at an exact time, the prepositional phrase consisting of the preposition *upre* ‘at’ is employed, e.g. (145).

(145)^{NAR} *av* *upro šóv óri!*
 come.IMP.2SG on six o’clock
 Come at six o’clock!

Further Layer II case markers are found in the lexicalised forms *ra’a-ha* ‘night-INS; in the morning’ and *kijarátis-ker-o* ‘evening-GEN-M.SG; towards the evening’. The adverb *pál-al* ‘then, later’ is formed by the historical ablative suffix *-al* from the inherited root *pal* ‘after,

behind’. The meaning ‘at age’ is expressed by marking the noun *berš* with the instrumental case, e.g. *dešušov beršenca* ‘at the age of sixteen’ in (146).

- (146)^{LQCR} *dešušov beršenca naši _géli.*
 sixteen year.PL.INS run_away.PRT.3SG.F
 She ran away at the age of sixteen.

Temporal relations that are anterior (‘before; until; ago’) or posterior (‘after; since; in’) to a specified point in time, or specify the extent (‘within, during; for’) of an action, are generally expressed by prepositional phrases in KR (Table 88). The only exception is the inherited particle *ži* ‘until’, which is employed in anterior-durative relations (‘until’). The terminology of temporal relations used in this section is based on Haspelmath (1997; followed also by Elšík and Matras 2006).

TEMPORAL RELATION	MARKING	e.g.
Anterior-sequence	<i>angle</i>	<i>angl’ odá</i> ‘before that’
Posterior-sequence	<i>pal</i>	<i>pal o but berša</i> ‘after many years’
Anterior-durative	<i>ži</i>	<i>ži kijaráti</i> ‘until the evening’
Posterior-durative	<i>uzar ~ sajt</i>	<i>uzar ~ sajt o hábori</i> ‘since the war’
Anterior-distance	zero-marked	(see below)
	<i>ezelöt</i> ‘ago’ + INS	<i>ezelöt ék maseka-ha</i> ‘a month ago’
Posterior-distance	<i>tel</i>	<i>tel ék masek</i> ‘in a month’
Telic-extent	<i>tel</i>	<i>tel ék óra</i> ‘within/during an hour’
Atelic-extent	zero-marked	(see below)
	<i>upre</i>	<i>upr’ ék kurko</i> ‘for a week’

Table 88 Marking of the non-simultaneous temporal relations

The prepositions used are the inherited *angle* ‘before’, *pal* ‘after’, the derived *uz-ar* (from the S *uz*) ‘since, from’, the calqued *tel* (cf. H *alatt*) ‘in, within, during’ and *upre* (cf. H -*ra/re*) ‘for’, and the borrowed *sajt* (< G *seit*) ‘since’ and *ezelöt* (< H postposition *ezelött*) ‘ago’. The anterior-distance *ezelöt*, which is a postposition in Hungarian, was grammaticalised as a preposition in KR. However, it may still be rarely found postposed to its complement. The

noun in the adpositional phrase with *ezelöt* is marked for the instrumental case. The Slavic-borrowed *uzar* competes with the German-borrowed *sajt* in KR, and it seems to prevail in the speech of the younger generation. Interesting is that one of my consultants would often switch from Romani to Hungarian in order to denote the posterior-durative expression ‘since a year ago, it's been a year’, as in (147)–(148) (the code-switch to Hungarian is underlined).

(147)^{NAR} *odóle hi má hus husonöt éve jagošitfán.*
 that.M.ACC COP.3 already twenty twenty-five year-since driving_licence
 He has had a driving licence for twenty or twenty-five years.

(148)^{NAR} *ed' jó éve, hod' samáh' uzo temető.*
 one good year-since COMP COP.PRT.1PL at.M graveyard
 It's been a year since we were at the graveyard.

The anterior-distance (149) and atelic-extent relations (150) are zero marked; that is, the meaning is given only by the context:

(149)^{LQCR} *so le díjenca but berša má com štrajtind'um.*
 all DEF.OBL two.INS many year.PL already VP quarrell.PRT.1SG
 I quarrelled with both of them already many years ago.

(150)^{NAR} *dešupándž berš samáhi khetán.*
 fifteen year COP.PRT.1PL together
 We were together for fifteen years.

The prepositions *ezelöt* ‘ago’ and *upre* for anterior-distance and atelic-extent relations, respectively, occur only sporadically.

Non-simultaneous temporal relation in other varieties of Vend Romani

In Sopron Romani, the clock time is expressed by the numeral accompanied with the Hungarian-borrowed *óra-kor* ‘o'clock-at; at (...) o'clock’. The group of Zala and Prekmurje Romani and the adjacent variety of Szakonyfalu (Vas) are linked together by the genitive marking of the expression ‘in the morning’ *raťas-kr-o* ‘morning-GEN-M.SG’ (< *raťaha*

‘morning’) against the instrumental marking *rat’a-ha* ‘night-INS’ (< *rat* ‘night’) found beyond this area. Outside Somogy, the adverbial meaning ‘afternoon’ expressed by the prepositional phrase requires the noun to be in the locative case, e.g. *pal plána-te* ‘after noon-LOC’, cf. KR *pal o plán* ‘lit. after the noon; afternoon’.

In Hungarian Vend Romani, the greatest variation of case marking is found in the posterior-durative time relation: The inherited ablative marking *-tar/-dar* appears as a minor variant in Vásárosdombó (Baranya) and Lengyeltóti (Somogy), while it is being systematically used in Veszprém Romani, e.g. *račas-tar* ‘since the morning’, *táhas-tar* ‘from tomorrow’, *júnijušis-tar* ‘since June’. The preposition *fe* (< G *von*) is common in Nagykanizsa (Zala), and its variant *fa* in Szakonyfalu (Vas). In some Somogy Romani varieties, the prepositions *uzar* and *sajt* (see above) are in free variation with the Hungarian-borrowed postposition *óta* ~ *úta* (< H *óta*). This postposition seems to have merged with the head noun *rat’aha* ‘morning’ in *rat’ah-úta* ‘lit. morning-from; from the morning’, the form of which is attested only in Homokszentgyörgy Romani (Somogy). The preposition *sajt* has the form *sajder* in Zala and *sejder* in Prekmurje Romani, which has probably its origin in the German *seither* ‘since that time’.

The posterior-durative relation is encoded on the adverbial expression *tikna-varijal* ~ *tikno-varijal* ‘since childhood’ in the neighbouring Somogy Romani varieties of Homokszentgyörgy, Görgeteg and Tarany. Although the adjectival component *tikn-o* ‘small’ is clearly recognizable, the exact origin of this word is unknown.

5.1.6 Other adverbials

Adverbials other than spatial, temporal and causal are marked by either inflectional cases or adpositions (see Table 89). The inflectional case marking is primarily used to encode the Benefactive (‘for the benefit of’), Comitative/Instrument (‘with’) and Source/Origin (‘from’) case roles, and it is only secondarily used in the Material (‘from’) and Goal (‘for’) roles. The inflectional marking in the latter two roles has been frequently attested in the spontaneous language data, while the adpositional marking was preferred in the elicited data (see below).

CASE ROLE	MARKING	e.g.
Benefactive	DAT	(151)
Comitative/Instrument	INS	(152)–(153)
Source/Origin	ABL	(154)
Material	INS	(155)
Goal	DAT	(156)

Table 89 Inflectional marking of other adverbials

(151)^{LQCR} *kind'a peske még jék šero.*
 buy.PRT.3SG REFL.3SG.DAT yet one beer
 He bought himself another beer.

(152)^{LQCR} *dešutrin čhájenca*
 thirteen girl.INS
 with thirteen girls

(153)^{LQCR} *ole báreha*
 DEF.OBL stone.INS
 with the stone

(154)^{LQCR} *adala virági pre piránostar uštidiya.*
 these flower.PL own.OBL lover.ABL get.PRT.3SG
 She got these flowers from her lover.

(155)^{LQCR} *kañhane maseha kerde zumi.*
 chicken.OBL meat.INS make.PRT.3PL soup
 They prepared a soup of chicken meat.

(156)^{LQCR} *adí lónge čhind'am tél i čár.*
 today money.PL.DAT cut.PRT.1PL down DEF grass
 Today we cut the grass for money.

The beneficiary of an action is obligatorily marked with the dative case (151). In addition, the dative is sporadically used to encode the goal of an action (156). The noun *drom*

is obligatorily in dative when used in the phrase *indulín- dromeske* ‘start_off road.DAT; to take to the road’.

The instrumental case is applied in the Comitative/Instrument case role (152)–(153), and less commonly in the Material role (155). The instrumental case is taken by the argument of the verbs *čájov-* ‘to eat one's fill’, *pér- com* ‘to meet’, *phír-* ‘to date’, *porotálín-* ‘to chat’, *sohajár-* ‘to marry’, *štrajtín-* ‘to quarrel’, *vakér-* ‘to speak’, and *vigzín-* ‘to finish’. The instrumental case is required by the phrases *ov- pherde* ‘to be full of’ (e.g. *pherde hi pišumenca* ‘full is fleas.INS; it is full of fleas’), *már- át'henca* ‘beat eyes.INS; to bewitch (lit. to beat with eyes)’, and *dikh- súno* ‘to have a dream of’ (e.g. *mrá daha dikjum súno* ‘my mother.INS saw.1SG dream; I had a dream about my mother’). The same case is applied when the object refers to certain knowledge (157), ability or possession.

- (157)^{NAR} *pándž čhip'tenca te akárkáj gejal,*
 five language.PL.INS COMP wherever go.PRT.2SG,
bistoš hod' upre tut len.
 surely COMP VP 2SG.ACC employ.3PL
 With (your knowledge of) five languages wherever you go, you will surely get a job.

The ablative case is used to mark the source of the verbs *čór-* ‘to steal’, *dára-* ‘to be afraid of’, *l-* ‘to take’, *máng-* ‘to ask, beg’, *šún-* ‘to hear’, *uštíd-* ‘to get’, and *phúič-* ‘to ask’. The source of the latter verb may be also marked with the accusative, e.g. *phuč'um tut* ‘ask.PRT.1SG you.ACC’ ~ *phuč'um tutar* ‘ask.PRT.1SG you.ABL; I asked you’. The ablative has been further attested in the phrases ‘to have a baby by’ (158), ‘to take someone by the hand’ (159), and ‘to have a bread with something on top’ (160).

- (158)^{LQCR} *fa'uj la sin ole nincostar.*
 children 3SG.F.ACC COP.PRT.3 DEF.OBL soldier.ABL
 She had children by a soldier.

- (159)^{LQCR} *astárd'a la vastestar.*
 grab.PRT.3SG 3SG.F.ACC hand.ABL
 S/he took her by the hand.

(160)^{NAR} *si balvastar máro, has?*
 COP.3 bacon.ABL bread eat.2SG
 There is bread with bacon; do you want to eat it?

Adpositional marking is employed in the Partitive (‘a part of’), Privative (‘without’), Exceptive (‘except of; all but’) and Substitutive (‘instead of’) case roles, and it is also a frequent means to express the Material (‘from’), Reference (‘about’), Goal (‘for’), and Reason (‘because of’) roles (see Table 90).

CASE ROLE	MARKING	e.g.
Partitive	<i>maškar</i>	(161)
Privative	<i>bi</i> (~ <i>mist</i>)	(162)
Exceptive	<i>bi</i>	(163)
Substitutive	<i>mist</i>	(164)
Material	<i>andral</i>	(165)
Reference	<i>upral</i>	(166)
Goal	<i>vaš, mist</i>	(167)
	<i>upre</i>	(168)
	<i>pal</i>	(169)
Reason	<i>mist</i>	(170)
	<i>ande</i>	(171)

Table 90 Adpositional marking of other adverbials

The preposition *maškar* ‘among’ is used in Partitive, the *bi* ‘without, except of, all but’ in Privative and Exceptive, the *mist* ‘instead’ in Substitutive (and sporadically in Privative), the *andral* ‘from’ in Material, and the preposition *upral* ‘about’ in Reference constructions.

(161)^{LQCR} *i jék thúli sin maškar o rákja.*
 DEF one fat COP.PRT.3 of DEF non-Romani_girl.PL
 One of the non-Romani girls was fat.

- (162)^{LQCR} *bi late erd'avo hi.*
 without 3SG.F.LOC bad COP.3
 It is bad without her.
- (163)^{LQCR} *bi mro papu sako fút gélo.*
 all_but my grandfather everybody away go.PRT.3SG.M
 Everyone went away but my grandfather.
- (164)^{LQCR} *mist i viršli gullípe kind'a.*
 instead DEF sausage sweets buy.PRT.3SG
 S/he bought sweets instead of sausage.
- (165)^{LQCR} *andral o srasta hi.*
 from DEF iron COP.3
 It is made of iron.
- (166)^{LQCR} *upral o muršikane bú'a vakéren.*
 about DEF male stuff.PL speak.3PL
 They speak about male stuff.

There are various prepositions found in the Goal and Reason case roles. The preposition *mist* (167)–(170) may occur in both meanings. The prepositions *upre* (168) and *pal* (169) are used in addition in the Goal, and the preposition *ande* (171) in the Reason role.

- (167)^{LQCR} *žav vaš/mist o thud.*
 go.1SG for DEF milk
 I am going to get milk.
- (168)^{LQCR} *upre mro sületéšnap uštidiyum.*
 for my birthday get.PRT.1SG
 I got it for my birthday.
- (169)^{LQCR} *žas pal leste.*
 go.1PL for 3SG.M.LOC
 Let's go and get him!

(170)^{LQCR} *mist late nasvajino.*
 because_of 3SG.F.LOC get_ill.PRT.3SG.M
 She got ill because of him.

(171)^{LQCR} *ande pri holi rovlahi.*
 in.DEF own anger cry.IMP.3SG
 S/he was crying of anger.

Furthermore, the preposition *ande* is required after verbs such as *kám- ánde* ‘to fall in love with’, *kételkedín-* ‘to doubt in’, *mér-* ‘to die of’, *páta-* ‘to believe in’, *pér-* ‘to cost’ (172), and *pomožín-* ‘to help’.

(172)^{LQCR} *ando kiči péli aja láda?*
 in.DEF how_much cost.PRT.3SG.F this.F box
 How much did this box cost you?

The same preposition appears in the phrases *a(v)- ando šéro* ‘to bethink of’, *o(v)- ando jékh* ‘to be together’ (173), *phén- ando át’ha* ‘to tell to the face’, *čumíd- ando muj* ‘to kiss on the lips’, and ‘to have something around the neck’ (174).

(173)^{NAR} *de odoj még ando jék samáhi.*
 but there still in.M one COP.PRT.1PL
 But there we were still together.

(174)^{LQCR} *šál hi leske andi men.*
 scarf COP.3 3SG.M.DAT in.F neck
 He has a scarf around his neck.

The latter two phrases are sometimes formed by the preposition *upre*, i.e. *upri músi* ‘in the hands’ and *upri men* ‘around the neck’. Apart from these, we find the preposition *upre* in the phrases *phén- upre* ‘the word for’ (175) and *upre amaro sámó* (cf. *H szám-unk-ra* ‘part-our-on’) ‘for/to us’ (176).

(175)^{LQCR} *sar phénen upro ‘hed’, de, román?*
 how say.3PL on.M mountain ah in_Romani

What is the word for ‘mountain’, ah, in Romani?

(176)^{LQCR} *upr’ amaro sámó adá baro tistelet hi.*
for our part this.M big honour COP.3
It is a big honour for us.

This preposition may occur in combination with the verb phrases *astár- i phudimni* ‘to point the gun at’, *čhi(v)- víra* ‘to swear on’, *čodákozín-* ‘to wonder’, *emléksín-* ‘to remember’, *gondolín-* ‘to think about’, *hašonlittín-* ‘to look like’, *khél-* ‘to play (an instrument)’, and *rúš-* ‘to be angry with’.

Other adverbials in other varieties of Vend Romani

The privative preposition *bi* alternates with *mist* ‘without’ only in some Somogy Romani varieties. Beyond this area, the Privative role is generally expressed by *bi*. In Zala and Veszprém Romani, the preposition *vaš* has also the Substitutive meaning (cf. KR *mist*) in addition to the Reason and Goal case roles. The borrowed preposition *dú ~ dúh* (< G *durch*) is used also in the Reason role in the eastern periphery of Somogy as well as in Prekmurje. Moreover, the speaker of Csokonyvisonta Romani used the form *dú vaš* in the Reason role, while the Homokszentgyörgy Romani speaker employed the preposition *dú(h)* also in the Privative and Exceptive roles.

MARKING	CASE ROLE
NOM	Subject Object: inanimate Predicative Possessee
ACC	Object: animate Recipient: 'to give' Experiencer Possessor
DAT	Benefactive Goal Recipient Predicative: complement Experiencer (Possessor) Possessor: External Temporal adverbial: clock time
ABL	Source/Origin Standard of comparison (subject of comparison = grammatical subject)
LOC	Possessor: physical contact Prepositional case (pronouns) Local adverbial: Inessive: 'Budapest'
INS	Comitative/Instrument Causee in causative constructions (Material) Possessee: physical/mental state Temporal adverbial: 'at age' Temporal adverbial: 'ago' (<i>ezelöt</i> + INS)
GEN	Possessor: Adnominal

Table 91 Synthetic case marking: Form to function

5.2 Particles and interjections

The vast majority of KR particles are borrowed from the local Hungarian dialect. The most commonly used ones are *azér* (< H *azért*) ‘even so; still’ (177), *his ~ isen* (< H *hiszen*) ‘well’, *körübelü* (< H *körülbelül*) ‘around, about’, *legaláb* (< H *legalább*) ‘at least’ (178), *megin* (< H *megint*) ~ *ujra* (< H *újra*) ‘again’, *mégiš* (< H *mégis*) ‘after all, however’, *talán* (< H *talán*) ‘maybe’, *ud’hod* (< H *úgyhogy*) *dehát* (< H *dehát*) ~ *hát* ‘so, well’ (178) and *ud’iš* (< H *úgy is*) ‘anyway, either way’.

(177)^{RM} *na nađ'on kamlahi, de azér sikáđ'a leske.*
 NEG very want.IMP.3SG but even_so show.PRT.3SG 3SG.M.DAT
 S/he did not really want to, but despite that s/he showed him.

(178)^{NAR} *hát sikav tut ánde legaláb!*
 well introduce.IMP.2SG 2SG.ACC VP at_least
 Well, at least introduce yourself!

The borrowed affirmative and negation particles are the *ám* (< H *ám*) (179) ~ *dehod* (< H *dehogy*) ~ *perse* (< H *persze*) ~ *hát* (< H coll. *hát*) ‘of course, sure’, *igen* (< H *igen*) ‘yes’, *ja* (< H coll. *ja*) ‘yeah’, *jó* (< H *jó*) ~ *jóvan* (< H *jól van*) ‘all right’, *dehod’iš* (< H *dehogy is*) ‘nope’, *ed’átalán* (< H *egyáltalán*) ‘at all’ and *nem* (< H *nem*) ‘no’.

(179)^{RM} *hát na afka sin sar akán ám, hod’ urak h’ o fa’uj.*
 well NEG so COP.PRT.3 how now of_course COMP lord.PL COP.3 DEF children
 Well, it was different before, of course! Not like nowadays that the children are ‘lords’.

Several focus particles and phasal adverbs are borrowed from Hungarian, such as *čak* (< H *csak*) ‘only, just’, *majnem* (< H *majdnem*) ‘almost, nearly’, *má* (< H *már*) ‘already’, *még* (< H *még*) ‘still, yet, so far; even; more; else’ and *még mindig* (< H *még mindig*) ‘still’. The phasal adverbs *még na* (cf. H *még nem*) ‘not yet’ and *má na* (cf. H *már nem*) ‘not any more’ are semi-calqued. Like in Hungarian, the particle *čak* appears also in the meaning ‘nothing but’ (180) and ‘may perhaps’ (181).

(180)^{NAR} *adá čak o čorípe!*
 this.M nothing_but DEF poverty
 This is nothing but poverty!

(181)^{NAR} *čak žanav má aťi!*
 may_perhaps know.1SG (already) so_much
 I may perhaps know that!

It is also used as a politeness marker in imperative sentences (182), following the Hungarian pattern.

(182)^{RM} *phénel i phuróri: dik čak órde!*
 say.3SG DEF old_woman look.IMP.2SG please here
 The old woman said: Please take a look here!

Similarly to Hungarian, the imperative value of the sentence is often reinforced by the particle *má* ‘already’, as in (183).

(183)^{RM} *tu meg na aťi kotrin má!*
 2SG and NEG so_much drag.IMP.2SG (already)
 And you stop dragging around!

Among the earlier borrowings we find *krót* (< G dial. *grod*) ‘just, exactly’, *ni* ‘neither’ (< S *ni*) and *méguli* (< probably from Slavic) ‘even’. The negative focus particle *ném* ‘neither, not even’ seems to have resulted from the merger of *ni* ‘neither’ and *még* ‘even’, i.e. **né-m* < **né-mig* < **ni-még*.⁷⁴ It is used alongside the negated predicate, e.g. (184).

(184)^{NAR} *ném le grastenge le na des ód'a.*
 not_even DEF.OBL horse.PL.DAT 3SG.M.ACC NEG give.2SG VP
 You wouldn't even give it to the horses.

⁷⁴ Compare it with the form *ne-mik* (< **né-mig* < **ni-még*) found in Burgenland Romani (Halwachs & Ambrosch 2002: 64).

The corresponding Hungarian form is *sem* ‘neither, not even’ which is, in contrast to the KR particle *ném*, postposed to its head. Furthermore, KR uses the inherited negator *na* ‘not’ in the meaning ‘not at all’ where Hungarian would, again, employ the particle *sem* ‘not at all’. In this function, the particle *na* follows the Hungarian word order, as it occupies the final position of the sentence, as in (185)–(187).

(185)^{LQCR} *nán la d'ék phral na.*
 COP.NEG.3 3SG.F.ACC one brother at_all
 She does not have a brother at all.

(186)^{NAR} *o Puška na bántínel nikas na.*
 DEF Puska NEG hurt.3SG nobody.ACC at_all
 Puška (name of the horse) would not hurt anybody at all.

(187)^{NAR} *ón na válinde fer šoha na.*
 3PL NEG divorce.PRT.3PL VP never at_all
 They were never divorced at all.

The phrase *ništa na* ‘nothing at all’ is also frequently used in KR, an example of which is:

(188)^{NAR} *na kamna ništa na; afka hi=le sar so tél čhindlo kašt.*
 NEG want.FUT.3PL nothing at_all so COP.3=3PL how what down cut tree
 They do not want anything at all; they are like a cut-down tree.

The discourse particles found in KR are often borrowed from Hungarian, or more precisely from the local Hungarian dialect. The most frequently used ones are *asisem* (< H *azt hiszem*) ‘I think/suppose so’, *hod' még* (< H *hogy még*) ‘so much, as much as’ (189), *minded'* (< H *mindegy*) ‘no matter’, *mond'uk* (< H *mondjuk*) ‘so to say’, *nem baj* (< H *nem baj*) ‘no problem’, *nem čoda* (< H *nem csoda*) ‘no wonder’ and *ud'e* (< H *ugye*) ‘eh, is that not so’ (190).

(189)^{NAR} *hát at'i ásáňam, hod'_még!*
 so so_much laugh.PRT.1PL as_much_as

So we were laughing as much as we could.

- (190)^{NAR} *phénav lake: ud'e pháre hi?*
tell.1SG 3SG.F.DAT is_that_not_so heavy COP.3
I am telling her: It is heavy, isn't it?

A frequently used discourse marker is *de* (< H dial. *de*), the function of which is to introduce a remark (191), call attention (192) or express hesitation in order to search for an appropriate word (193).

- (191)^{NAR} *de, akor vakeras román!*
well then speak.1PL in_Romani
Well, let's speak Romani then!

- (192)^{NAR} *de, ók žalék murš!*
look there go.3SG a man
Look, there is a man walking!

- (193)^{NAR} *sar phénen upr' odá kova, de!*
how say.3PL on that.M EXPL wait
How is that thingummy called, wait!

This particle has also been found in some imperative clauses with emphatic focus, such as in (194).

- (194)^{NAR} *phénel pre dadeske: av de žas!*
tell.3SG own.OBL father.DAT come.IMP.2SG (let's) go.1PL
S/he told to his/her father: Come on, let's go!

The particle *de* 'but' is also often used as a connective marker (195). In this function, it is interchangeable with *ham* (< H *hanem*), which has been attested only in the spontaneous data (196). The use of *ham* as a connective marker is an innovation in KR, since the corresponding Hungarian marker *hanem* does not have such a function (see 5.3.5).

(195)^{NAR} *de dikhes, tipikuš beáš!*
 but see.2SG typical Boyash
 (But) you see, he is a typical Boyash!

(196)^{NAR} *ham órde šun!*
 but here listen.IMP.2SG
 (But) listen to me!

The interjections that express various emotions in KR include the borrowed *ó* (< H *ó*) ‘oh: surprise’ (197), *ehe* (< H dial. *ehe*) ‘hey: attention’, *jaj* ~ *jáj* ‘ouch: pain; ah: realisation’ (198), *hú* (< H *hú*) ‘gee: wonder, surprise, pleasure’ (199), and *na* (< H *na*) ~ *no* (< H *no*) ‘well, come on: pleading, order, etc.’ (200).

(197)^{NAR} *ó, hát ój nad'on d'áral!*
 oh well 3SG.F very be_afraid.3SG
 Oh, well, she is very afraid!

(198)^{NAR} *jáj, o Marcel, odóle prindžáres?*
 ah DEF Marcel that.M.ACC know.2SG
 Ah yeah, Marcel, do you know him?

(199)^{NAR} ZB: *de si ase, ko pá'an ánde?* M: *hú, de kitì!*
 but COP.3 such_(people) who believe.3PL in_it gee and how_many
 But is there anyone who believes in it? M: Gee, and how many!

(200)^{NAR} *phen lenge te áchen, no!*
 tell.IMP.2SG 3PL.DAT COMP stop.3PL come_on
 Tell them to stop it, come on!

The inherited particles include *papal* ‘again’, the affirmative *he* ‘yes’, the negation *na* ‘no(t)’, *na'a* ‘nope’ and *ma* ‘do not!’, the focus *te* ‘also, too, as well’, and the comparative *sar* ‘as, like, than’ and *sar te* ‘as if’. The expression *so hi* ‘lit. what is; what?’ is reduced either to *só* or *hó* when used emphatically, which may be seen as an interjection with the meaning ‘eh, huh?’ (as in Q36–37).

(Q36)

M [to W]: *Užár, so lake phennahi?*

W: *Só?*

M: *Sar akharnahi la Laciskera da?*

M [to W]: Wait, what was her name?

W: Eh?

M: What was the name of Laci's mother?

(Q37)

ZB: *So hi leskro ánav?*

F: *Hó?*

M: *Sar phéne... so h' o Guszti?*

W: *Ó, hát me na žánav.*

ZB: What's his name?

F: Eh?

M: How do you call... what is the Guszti
[i.e. what is Guszti's surname]?

W: Hmm, I don't know this.

Finally, the discourse markers *adale-ha* (cf. H *ez-zel*) 'this-INS; hereby' (201), *odole-ha* (cf. H *az-zal*) 'that-with; thereby', *upr' odá* 'on that' (cf. H *er-re* 'this-on') 'thereupon', *upr' adá* 'on this' (cf. H *er-re* 'this-on') 'hereupon' and *t' akor* (cf. H *és akkor*) 'and then' are calqued from Hungarian.

(201)^{RM} *adaleha lija pe o borzo, fút gélo.*
this.M.INS take.PRT.3SG REFL.3SG DEF hedgehog away go.PRT.3SG.M
The hedgehog hereby shook the dust from his legs and went away.

Particles and interjections in other varieties of Vend Romani

The particle *méguli* 'even' is sporadically attested throughout the Vend Romani area, having also the forms *moguli* and *muguli*. The negative focus particle has the form *nimi* in Zala and Prekmurje Romani, and *menik* in Burgenland Romani.

5.3 Complex syntactic structures

5.3.1 Modal constructions

Two types of modal constructions should be distinguished in KR. The first type involves the finite form of the modal verb which is accompanied by the complementizer *te* and the infinitive form of the verb, e.g. *pekál te phénel* 'need.3SG COMP say.INF; s/he needs to say',

kezdínav te ról ‘start.1SG COMP cry.INF; I start to cry’. The modal verb may either precede or follow the infinitive construction, depending on the focus, e.g. *kezdind’a te ról* ‘start.PRT.3SG COMP cry.INF; s/he started to cry’ vs. *te ról kezdind’a* ‘COMP cry.INF start.PRT.3SG; to cry, that's what s/he started to do’. Other modal verbs are the inherited *kám-* ‘to want, like’, *žán-* ‘to know, be able to, can’ and *tróma-* ‘to be allowed to; must’, the borrowed *tránín-* (< G *sich trauen*) ‘to dare’ and *próbálín-* (< H *próbál*) ‘to try’, and the calqued *abba múk-* (cf. H *abba hagy* ‘VP let’) ‘to stop doing’. The second type of modal constructions is composed of an uninflected modal which, on the other hand, requires a finite verb, e.g. *šaj sojs* ‘can sleep.2SG; you can sleep’, or *eremešt hordínav* ‘like wear.1SG; I like to wear’. Other uninflected modals are the possibility modal *šaj* ‘can’ and the necessity modal *site* ‘have to, should, must’. The latter is composed of the third-person present indicative copula form *si* and the complementizer *te* (see Elšík & Matras 2009: 289).

To express pretence, the speakers used to calque the Hungarian phrase *úgy tesz mint aki* ‘lit. do like who; to pretend’: *kér- sar ko* ‘lit. do like who; to pretend’. Exceptionally, the adverb *kamukéri* has also been used together with the previous expression, i.e. (*kér-*) *sar ko kamukéri* ‘lit. (do) like who by_pretence’. These expressions are followed by the subjunctive form of the verb, e.g. (202).

(202)^{NAR} *oja meg gud’ulínel, taj sar ko kamukéri mutrel.*
 that.F and sit_on_hunkers.3SG and like who by_pretence piss.3SG
 And she is sitting on her hunkers and pretending to pee.

5.3.2 Complement clauses

Following the terminology of Matras (2002: 179ff.; 2004), we can distinguish between the complementisers used in factual (or real) and non-factual (non-real) predications, based on the truth-value of the dependent clause. The manipulation and purpose clauses are positioned in the middle of the ‘factuality continuum’. The complementisers that are used to introduce complement clauses in KR are listed in Table 92.

TYPE	MARKING	e.g.
Factual	<i>hod'</i> ~ zero marked	(203) ~ (204)
Factual: Polar	<i>hod'</i>	(205)
Factual: Interrogative	(<i>hod'</i> +) <i>interrogative</i>	(206)
Non-factual: Manipulation	<i>hod' te</i> (~ <i>te</i>)	(207) ~ (208)
Non-factual: Modal	<i>te</i>	(see 5.3.1)

Table 92 Complement clauses

The factual subordinations are introduced by the borrowed complementizer *hod'* (< H *hogy*) ‘that’ (203), which may be occasionally dropped, as in (204). The complementizer *hod'* is obligatory in polar subordinate clauses (205), while it is only optionally used alongside the interrogative word in interrogative subordinate clauses (206). The verb in factual subordinations agrees in person and number with the subject, and it is marked for tense, aspect and mood.

(203)^{LQCR} *phom, hod' oja bóta deš órenge phrád'ol.*
 tell.PRT.1SG COMP that.F shop ten o'clock.DAT open.3SG
 I said that the shop opens at ten.

(204)^{LQCR} *phénen, 0 lačho than hi tut ando fóro.*
 say.3PL good place COP.3 2SG.ACC in.M town
 They say that you have a good job in the town.

(205)^{LQCR} *na žánav, hod' buza hi, zabo, vad' rožo hi.*
 NEG know.1SG COMP wheat COP.3 oat or rye COP.3
 I do not know whether it is wheat, oat or rye.

(206)^{LQCR} *žánes, (hod') mere site khér žas?*
 know.2SG (COMP) where_to should home go.2SG
 Do you know where you should go?

Non-factual subordinations with manipulative meaning involve the factual complementizer *hod'* and the non-factual *te*, being subsequently accompanied by the subjunctive form of the verb (207). The complementizer *hod'* is sometimes omitted (208).

Interrogative clauses with an underlying manipulation meaning are used to express also politeness. However, the manipulative verb is systematically dropped and only the non-factual *te* is used to introduce the request, as in (209)–(210). The modal constructions are also formed by the non-factual *te* (see 5.3.1).

(207)^{LQCR} *niko na phend'a, hod' adá te kéres.*
 nobody NEG tell.PRT.3SG COMP this.M COMP do.2SG
 No one told you to do this!

(208)^{LQCR} *phom te kínel dúdum.*
 tell.PRT.1SG COMP buy.3SG pumpkin
 I told (him/her) to buy a pumpkin.

(209)^{LQCR} *te pomožínav renbe te čhíl tumaro kher?*
 COMP help.1SG order COMP put.INF your house
 (May I / Would you like me to) help you to clean the house?

(210)^{NAR} *ki'i tuke te čhórav?*
 how_much 2SG.DAT COMP pour.1SG
 How much (may I / do you want me to) pour for you?

Complement clauses in other varieties of Vend Romani

The factual complementizer is the original *káj* in Szakonyfalu (Vas) and in some Burgenland and Prekmurje varieties. The most commonly used non-factual complementizer is *neka* and *káj neka* in Szakonyfalu (Vas), *kaj te* in Burgenland, and *ne, neka, kaj* and *kaj ne* in Prekmurje Romani.

5.3.3 Relative clauses

Relative clauses in KR are introduced by relativizers which correspond to interrogative pronouns and adverbs (see 4.3, 4.5.5). The relative pronoun *so* (OBL *sos-*) ‘which’ refers primarily to inanimate nouns (211), and less commonly to animate nouns (212). It is often used in prepositional phrases, where it is marked for the locative case (213). In relation to

animate nouns, the relativizer *ko* (OBL *kas-*, INS *kasaha*) ‘who’ is the most commonly used one (214). This relativizer inflects for case (215).

(211)^{NAR} *oja kuňuva, so lake kezdind'um te kérel (...)*
 that.F hovel which 3SG.F.DAT start.PRT.1SG COMP do.INF
 That hovel which I started to build for her (...)

(212)^{NAR} *mró dad, so andral o németország álo (...)*
 my father which from DEF Germany/Austria come.PRT.3SG.M
 My father who came from Germany/Austria (...)

(213)^{LQCR} *o sane_pórd'a, andral soste kéren i gój.*
 DEF small_intestine.PL from what.LOC do.3PL DEF gój
 The small intestine, from which they prepare the gój (= traditional meal).

(214)^{NAR} *si duj lumňa vaj trin, ko phénel (...)*
 COP.3 two woman.PL or three who say.3SG
 There is one or two women who say that (...)

(215)^{LQCR} *na žánav, kaské le te dav pále.*
 NEG know.1SG who.DAT 3SG.M.ACC COMP give.1SG back
 I do not know whom I should return it to.

The relativizer *sav-o* ‘which one/kind of’, which is marked for gender, number and case, may be applied to both animate (216) and inanimate nouns (217).

(216)^{NAR} *phuč, hod' savo h' odá,*
 ask COMP which.M COP.3 that.M
sav-o čak román žánel te vakérel!
 which-M.SG only in_Romani know.3SG COMP speak.INF
 Ask (them) which one is the one who speaks only Romani!

(217)^{LQCR} *odola ponnaha žas, sav-i sigéder ál?*
 that.OBL train.INS go.2SG which-F.SG earlier come.3SG
 We are going by that train which comes earlier?

Resumptive pronouns are not used in KR.

5.3.4 Adverbial clauses

5.3.4.1 Temporal clauses

The terminology of temporal relations used in this section follows the terminology in Elšík and Matras (2006; based on Haspelmath 1997). In KR, temporal adverbial clauses are introduced by subordinators which are mostly borrowed from the contact languages. The subordinator *sar* ‘when’ is based on the inherited interrogative *sar* ‘how’ (see 4.5.5), while the subordinator *ži még* comprise the inherited preposition *ži* ‘until’ (see 5.1.5). The verb has finite form in temporal adverbial subordinations.

TYPE	MARKING	MEANING	e.g.
Simultaneous: punctual	<i>kada, sar</i>	when, just as	(218), (222)
Simultaneous: durative	<i>kada, sar, ži még</i>	while, as long as	(221), (223)
Simultaneous: habitual	<i>kada, akárkada</i>	every time, whenever	(227)
Anterior: sequence	<i>kada, mijelöt, mire</i>	before, by the time	(225)–(226)
Anterior: durative	<i>ži még, amég</i>	until	(219)
Posterior: sequence	<i>kada, sar</i>	after, as	(224)
Posterior: durative	<i>sajt, mijuta</i>	since	(220)

Table 93 Temporal adverbial clauses

The most commonly used subordinator for introducing temporal adverbial clauses is the Slavic-origin *kada* ‘when’ (see Table 93). It is used in the simultaneous (‘when (218); while; every time’) and the anterior (‘before’) and posterior sequence (‘after’) time reference.

(218)^{LQCR} *daráno,* *kada le* *dikja.*
 be_scared.PRT.3SG.M when 3SG.M.ACC see.PRT.3SG
 He got scared when he saw her.

Kada is not applied in temporal clauses with durative meaning (i.e. ‘until; since’). Instead, the semicalque *ži még* (< H dial *amég*) ‘until’ (219) or the corresponding loanword

amég is used with anterior reference, and the loanwords *sajt* (< G *seit*) (220) or *mijuta* (< H *mióta*) ‘since’ with posterior reference.

(219)^{NAR} *ži_még lakeri daj na géli andi bóta, (...) šaj búti_ker'd'a.*
 as_long_as her mother NEG go.PRT.3SG.F in.F shop can work.PRT.3SG
 As long as her mother wouldn't come to the shop, she could work.

(220)^{NAR} *sajt újum, azúta román vakérav.*
 since born.PRT.1SG since_then in_Romani speak.1SG
 Since I was born, I speak Romani.

The subordinator *ži még* can also be used in simultaneous durative (‘while, as long as’) time relation, such as in (221).

(221)^{NAR} *ži_még hi khul grastáno, addig hi te čirikli.*
 as_long_as COP.3 shit horse until_then COP.3 also bird
 As long as there is horse shit, there is also a bird.

The subordinator *sar* may optionally be employed in clauses with simultaneous punctual (‘when (222), just as’), simultaneous durative (‘while (223), as long as’) and posterior sequence relations (‘after, as (224)’). The same structure is found also in Hungarian.

(222)^{NAR} *sar ánde gejam, ed'bü la prindžárd'um.*
 how in go.PRT.1PL immediately 3SG.F.ACC recognize.PRT.1SG
 When we entered, I immediately recognized her.

(223)^{LQCR} *kezdind'a te del o d'iv, sar užarahahi.*
 start.PRT.3SG COMP give.INF DEF snow how wait.IMPF.1PL
 It started snowing while we were waiting.

(224)^{LQCR} *sar ár phrád'a o vúdar, kezdind'a te del o brišind.*
 how VP open.PRT.3SG DEF door start.PRT.3SG COMP give.INF DEF rain
 Just as s/he opened the door, it started raining.

The anterior sequence (‘before’) clauses are generally preceded by the borrowed subordinator *mijelöt* (< H *mielőtt*) (225) or *mire* (< H dial. *mire*) (226).

(225)^{LQCR} *mijelöt órde áli te áčhel, áthar dúr áčháhi.*
 before here come.PRT.3SG.F COMP live.INF from_here far_away live.IMPF.3SG
 Before she came to live here she had lived far away from here.

(226)^{NAR} *mire má aňaköňvezinde le, má báro sin.*
 before already register.PRT.3PL 3SG.M.ACC already big COP.PRT.3
 He had already grown adult before they registered him.

The simultaneous habitual (‘every time, whenever’) relation is expressed by the semicalque *akárkada* (< H *akármikor*) (227), which has been attested only in the elicited data.

(227)^{LQCR} *akárkada ásal, o cile róma šúnen.*
 whenever laugh.3SG DEF all Roma hear.3PL
 Whenever s/he laughs, all Roma can hear him.

5.3.4.2 Conditional clauses

KR differentiates between realis, potential and irrealis (also called counterfactual) conditionality, similarly to other Romani dialects (Elšík and Matras 2006: 204). The difference lies in the tense marking found in the two clauses of the conditional sentence, in the protasis (i.e. dependent clause) and apodosis (i.e. main clause). The realis conditional refers to situations which are very likely to happen. In potential conditional, the condition is possible, but very unlikely to be fulfilled. Finally, the irrealis involves a condition which is impossible to be fulfilled, since it refers to the past.

In all types of conditional sentences, the subordinator *te* ‘if, in case, etc.’ is used to introduce the protasis. Although the speakers systematically used the semicalqued *még te* (cf. H *még-ha*) ‘even if’ 0 in concessive-conditional clauses during the elicited speech session, only the original *te* (229) has been attested in the spontaneous narratives.

(228)^{LQCR} *még_te asaj párne ováhi sar ék gážo,*
 even_if such white COP.COND.1SG like a non-Rom
ni akor n' uštidáhi búi
 neither then NEG get.IMPF.1SG job
 Even if I was as white as a non-Rom, I wouldn't get a job.

(229)^{NAR} *te leske prik i bul araň ár álo,*
 even_if 3SG.M.DAT through DEF anus gold out come.PRT.3SG.M
t' akor adá phénen: (...) cigány.
 also then this.M say.3PL Gypsy
 Even if he defecates gold, they will keep saying: He is a Gypsy.

Realis constructions involve the protasis with the preterite, e.g. *úlo* in (230) and *ájal* in (231), and the apodosis either with the present, e.g. *kérav* in (230), or the future form of the verb, e.g. *dikhá* in (231). The copula in the apodosis always has the future form, such as *ová* in (232)a, cf. (232)b.

(230)^{RM} *akárso te úlo odá, kérav le.*
 whatever if COP.SUBJ.PRT.3SG.M that.M do.1SG 3SG.M.ACC
 Whatever it is, I will do it.

(231)^{LQCR} *te órde ájal, akor dikhá tut.*
 if hither come.PRT.2SG then see.FUT.1SG 2SG.ACC
 If you come, I shall see you.

(232)^{LQCR}
 a. *te but thud pijum, zoráli ová.*
 if a_lot milk drink.PRT.1SG strong COP.FUT.1SG
 b. **te but thud pijum, zoráli sum.*
 if a_lot milk drink.PRT.1SG strong COP.PRS.1SG
 If I drink a lot of milk, I will be strong.

In potential conditional sentences, the imperfect is employed in both protasis and apodosis, e.g. *ovnahi* and *dáhi* in (233). Similarly, the same tense, namely the conditional irrealis, occurs in both parts of the irrealis conditional sentences, e.g. *ujumáhi* and *uštidijumáhi* in (234).

(233)^{LQCR} *te man ovnahi lój, tuke len dáhi.*
 if 1SG.ACC COP.COND.3PL money 2PL.DAT 3PL.ACC give.IMPF.1SG
 If I had some money I would give it to you.

(234)^{LQCR} *te na ujumáhi nasváli, uštídjumáhi búti.*
 if NEG COP.IRR.1SG sick get.IRR.1SG job
 If I had not been sick, I would have got a job.

5.3.4.3 Other adverbial clauses

Causal clauses employ the borrowed subordinator *mer* (< H *mer*) ‘because’ (235). These clauses may sporadically be introduced by the function word *so* ‘as, what’ (236).

(235)^{LQCR} *žav te sól, mer khiňijum.*
 go.1SG COMP sleep.INF because get_tired.PRT.1SG
 I am going to sleep because I'm tired.

(236)^{LQCR} *hát so khírinlahi, sako upr' ušt'ino.*
 well as scream.IMPF.3SG everybody VP wake.PRT.3SG.M
 Well, as she was screaming, everybody woke up.

Like in most Romani dialects (Matras 2002: 183), purpose clauses that involve a predication expressing movement of an agent take the original subordinator *te* (237). In other purpose clauses the composed subordinator *hod' te* (238) (cf. H *hogy*) ‘in order to, so that’ is employed.

(237)^{LQCR} *gejum uz odá gážo lój te mánjel.*
 go.PRT.1SG to that.M non-Rom money COMP ask.3SG
 I went to the non-Rom to ask for money.

(238)^{NAR} *(i močka) chungárlahi upre, hod' te rajnisajol tél o šebo.*
 (DEF quid) spit.IMPF.3SG on COMP become_clean.3SG VP DEF wound
 S/he used to spit the quid on it in order to clean the wound.

Adverbial clauses that denote the circumstance (239) and the extent (240) of an action are introduced by the borrowed subordinator *hod'* 'that'.

(239)^{LQCR} *valasar fút žandle te žal,*
 somehow away manage.PRT.3PL COMP go.INF
hod' na po'inde ári pumaro adóšág.
 COMP NEG pay VP REFL.3PL.GEN debt
 Somehow they managed to leave without paying their debt.

(240)^{LQCR} *ati hája, hod' čhánda pe.*
 so_much eat.PRT.3SG COMP vomit REFL.3SG
 S/he ate so much that she vomited.

Adverbial clauses in other varieties of Vend Romani

In other than Somogy Romani varieties, the Hungarian-borrowed durative subordinator *amég* (also attested as *amíg, mīg, még*) is clearly preferred to the calqued *ži még* 'as long as, until'. The German-borrowed posterior-durative subordinator *sajt* is attested only in KR and Zala Romani. In the former area, it has the form *sajder* (< G *seither* 'since that time').

In Táska (Somogy), the Hungarian-borrowed causal subordinator *mer* 'because' is in free variation with the, also borrowed, subordinator *min* (< H *mint* 'as, like'). It is interesting that, in addition to Táska, the subordinator *min* is found only in Burgenland Romani. In adverbial clauses of purpose we find the borrowed *neka* (< S *neka*) in Tarany (Somogy), Szakonyfalu (Vas) and in the varieties of Zala. Its reduced form *ne* is attested in Zala, where it is often used together with *hod'* 'that', i.e. *hod' ne*. The complementizer *hod'* 'that' is absent in Szakonyfalu (Vas) where the inherited *káj* is used instead. The two subordinators, *hod'* and *káj*, alternate in Sopron Romani.

5.3.5 Coordination

In this section I will adapt the terminology and classification of coordinators found in Haspelmath (2007). The main coordinators (= coordinating conjunctions) of KR include the conjunctive ('and'), the disjunctive ('or'), the adversative ('but'), and the causal ('for') coordinators (Table 94). KR has also a set of so-called contrastive coordinators that involve

two coordinators, such as the conjunctive *te – te* ('both – and'), the negative *ni – ni* ('neither – nor'), and the disjunctive *vaj – vaj* ('either – or'). The coordinators are used to connect both constituents and clauses, i.e. coordinands (as called in Haspelmath 2007).

TYPE	PHRASAL	CLAUSAL
Conjunctive	<i>A taj B</i>	<i>A taj B</i>
Conjunctive: contrastive	<i>te A te B</i>	<i>te A te B</i>
Conjunctive: oppositive	–	<i>A, B1 meg B2</i> <i>A, B1 pedig B2</i>
Negative: contrastive	<i>ni A ni B</i> <i>ni A ni B na</i>	<i>ni (not) A ni (not) B</i>
Disjunctive	<i>A vaj B</i>	<i>A vaj B</i>
Disjunctive: contrastive	<i>vaj A vaj B</i>	<i>vaj A vaj B</i>
Adversative	<i>A de B</i>	<i>A de B</i>
Adversative: substitutive	(not) <i>A hanem B</i>	(not) <i>A hanem B</i>
Causal	–	<i>mist' odá'adá</i>

Table 94 Coordinating conjunctions

The conjunctive coordinators are the inherited *taj* 'and' and *te – te* 'both – and', and the borrowed *meg* (< H dial. *meg*) and *pedig* (< H *pedig*) 'and'. The conjunction *taj* is placed between the constituents and clauses it links together. In case of multiple coordination, it is used to connect only the last two elements (241).

- (241)^{NAR} *ár nána len: ék ziha, 0 ék plasta, taj osó nejlón.*
 else COP.NEG.PRT.3 3PL.ACC a duvet a bed-sheet and such nylon
 They didn't have anything else: a duvet, a bed-sheet, and such a nylon.

The contrastive conjunction *te* is, on the other hand, preposed to both coordinands, as it is illustrated in (242).

- (242)^{NAR} *te román žanlahi, te ninčka, taj te ungrika.*
 also in_Romani know.IMPF.3SG also in_German and also in_Hungarian

He spoke both Romani and German, and also Hungarian.

The oppositive *meg* and *pedig* is used when ‘there is a contrast between the two coordinands, but no conflicting expectations’ (Haspelmath 2007: 26). Like in Hungarian (Kenesei *et al.* 1998: 102), both *meg* and *pedig* are preceded by the first topicalized phrase of the last coordinated clause, e.g. by the subject ‘mother’ in (243).

- (243)^{NAR} *o dad le fa'unca paššolahi,*
 DEF father DEF.OBL son.PL.INS lie.IMPF.3SG
*i daj **meg** le čhájenca.*
 DEF mother and DEF.OBL daughter.PL.INS
 The father used to sleep together with the sons, and the mother with the daughters.

The contrastive negative coordinator is the Slavic-borrowed *ni – ni* ‘neither ... nor’. Similarly to the contrastive *te – te* ‘both – and’, *ni* is preposed to all constituents that are coordinated (244).

- (244)^{NAR} *ni abrak, ni hábe, ni píbe, ništ na delahi.*
 neither forage neither food neither drink nothing NEG give.IMPF.3SG
 Neither forage, nor food, nor drink, he didn't give anything.

If the second coordinand is emphasised, it is also followed by the negation particle *na* ‘not’, such as in (245).

- (245)^{NAR} *odoleskero fa'ú na ovla **ni** gážo, de **ni** rom **na**.*
 that.M.GEN son NEG COP.FUT.3SG neither non-Rom but neither Rom NEG
 His son would not become a non-Rom, but neither a Rom.

The predicates of the coordinated clauses are always negated, as it is underlined in (246).

- (246)^{NAR} *me **ni** na áčháhi árhán, (...)*
 1SG neither NEG live.IMPF.1SG elsewhere

ni mri daj nána šoha árhán.
 neither my mother COP.NEG.PRT.3 never elsewhere
 Neither I lived elsewhere, (...) nor my mother lived elsewhere.

KR retains the original disjunctive conjunction *vaj*, instead of which we may sporadically find the Hungarian-borrowed *vad'* (< H *vagy*) 'or'. It is positioned between the two coordinands (247), and in contrastive use it is preposed to each coordinand (248).

(247)^{NAR} *bi cukro le pijav má děš vaj dešuduj berš.*
 without sugar 3SG.M.ACC drink.1SG already ten or twelve year
 I have drunk it without sugar already for ten or twelve years.

(248)^{NAR} *and' odóla meg akán vaj duj, vaj d'ék murš_fatú sin.*
 in that.PL and now or two or one boy COP.PRT.3
 And among those siblings you had either one or two boys.

The coordinator *vaj* may sporadically be omitted between two constituents, i.e. A 0 B, as in (249).

(249)^{NAR} *taj még duj méter 0 trin fárinel (...)*
 and more two meter three jump.3SG
 And he jumps two or three metres higher (...)

The adversative coordinators *de* (< H *de*) 'but' and *ham ~ hanem* (both < H *hanem*) 'but' are borrowed from Hungarian.⁷⁵ They are inserted between the coordinands. The adversative *de* 'but' is generally used after the first coordinated expression with affirmative meaning, such as in (250).

(250)^{NAR} *phom te hal, de na hája!*
 tell.PRT.1SG COMP eat.3SG but NEG eat.PRT.3SG
 I told him/her to eat, but s/he didn't eat!

⁷⁵ The function of *de* and *ham* 'but' as discourse markers is discussed in section 5.2.

On the other hand, the substitutive *hanem* is solely reserved for coordinations in which the first coordinated expression is negated (251).

- (251)^{NAR} *akor na sójhahi odoj, hanem edej.*
 then NEG sleep.IMPF.2SG there but here
 Then you are not going to sleep there, but (rather) here.

In sentences like (252), the standard adversative *de* may also be allowed. Consider the following example with the conjunction *de*, where the substitutive *hanem* would be expected:

- (252)^{NAR} *na andral i Kišbajom, de andral o Mad'arország.*
 NEG from DEF Kisbajom but from DEF Hungary
 Not from Kisbajom, but from Hungary.

Causal conjunction is used to mark the consequence of an action. In this function, KR uses the prepositional phrase *mist' odá* 'that's why' which calques the Hungarian conjunction *az-ért* 'lit. that-because_of' (253).

- (253)^{NAR} *mist'_odá hi t' akán ando astaribe, ár čórda i átómata.*
 for_that.M COP.3 also now in.M jail VP steal.PRT.3SG DEF cash_machine
 He is now in the jail again for stealing money from the cash machine.

Coordination in other varieties of Vend Romani

In Zala Romani, the conjunctive *te* 'and' is employed in both plain and contrastive coordinations, i.e. *te* (cf. KR *taj*) 'and', *te – te* 'both – and'. Zala, Vas and Veszprém Romani and several Somogy Romani varieties use the conjunction *pa* or *pal* 'and' in oppositive relation. It is unclear whether we are dealing with the original adverbial *pal* 'after', or with the (contamination of the original *pal* and the) South Slavic conjunction *pa* 'and'. Neither of these particles are attested in KR.

The adversative coordinator *de* 'but' competes with the reduced form *ham* (< H *hanem*) 'but' in Sopron, Vas and Veszprém Romani, and in a single variety of Somogy Romani (i.e. in Táska). It is clearly an innovation in Romani, since the corresponding Hungarian form is applied only in substitutive adversative coordinations. Moreover, only the

form *ham* is attested in both plain and substitutive adversative coordinations in Burgenland and Prekmurje Romani.

Beyond Somogy Romani, the causal coordinator is either the Hungarian-borrowed *azért ~ ezért* ‘for; that's why’ or the prepositional phrase *vaš odá* (cf. KR *mist’ odá*). The latter form has also been used by the story-teller of Rézműves (2006) who originally comes from Kisbajom. It may indicate that the idiolect of this speaker had been influenced by the Veszprém varieties, as he was residing in Veszprém at the time of the data collection.

5.4 Word order

5.4.1 Noun phrase

The most typical word order found in noun phrases is that the head noun is preceded by the descriptive adjective, which is, on the other hand, preceded by the article, determiner or quantifier. The genitive possessor is fronted to the head noun, too, e.g. (254).

- (254)^{NAR} *ék balane šéreskero baro močárno fa’u*
 a pig headed.GEN big ugly son
 a very ugly, big, pig-headed son

The head noun may exceptionally precede the determiner, e.g. *grastáno* in (221), or the possessor, e.g. *mro* in (255).

- (255)^{NAR} *adá baro čéderi mro, soha ládáhi, prik gélo (...)*
 this.M big stallion 1SG.GEN what.INS ride.IMP.1SG over go.PRT.3SG.M
 This big stallion, which I used to ride, went over (to someone's yard).

5.4.2 Adpositional phrase

The prepositional phrase is composed of the preposition which precedes the noun phrase, e.g. *uz-* ‘next to’ in (256).

- (256)^{RM} *uze mro rézitiko véš*
 next_to 1SG.GEN copper forest

next to my copper forest

5.4.3 Verb phrase

The preferred word order in verb phrases is VO (257), though the combination OV is also possible (258).

(257)^{NAR} *uštidiĵa hacás forint.*
get.PRT.3SG six_hundred forint
S/he got six hundred forints.

(258)^{NAR} *šov ŝel forint uštidiĵa.*
six hundred forint get.PRT.3SG
S/he got six hundred forints.

In contrast, the copula is typically placed to the final position (259). Less commonly, it can occupy other positions as well (260).

(259)^{NAR} *óv rendeŝno rom hi.*
3SG.M real Rom COP.3
He is a real Rom.

(260)^{NAR} *t' oĵa hi nasváli nad'on.*
also that.M COP.3 sick very
She is also very sick.

The nominal subject precedes the verb in sentences without focus, as *mró dad* in (261), while it is postposed to the focused phrase, as *mró dad* in (262).

(261)^{NAR} *mró dad erd'avéder sin sar o žukéla.*
my father worse COP.PRT.3 than DEF dog.PL
My father was worse than the dogs.

(262)^{NAR} *odá com phánd'a mró dad, upro phiko liĵa.*
that.M VP bind.PRT.3SG my father on.M shoulder take.PRT.3SG
My father bound it, and took it onto his shoulder.

Since the subject is marked on the verb, the nominative pronoun is not obligatory in non-emphatic or non-contrastive contexts. The pronoun with emphatic or contrastive use (i.e. the subject) is generally placed before the verb in the indicative sentences, example of which is (263).

- (263)^{NAR} *ój míli sigéder, mri daj.*
 3SG.F die.PRT.3SG.F earlier my mother
 She died earlier, my mother.

On the other hand, the subject pronoun is allowed to occupy the post-verbal position in interrogative (264) and imperative sentences (265).

- (264)^{RM} *so kéres tu edej adale bákrenca?*
 what do.2SG 2SG here this.PL.OBL sheep.INS
 What are you doing here with these sheep?

- (265)^{RM} *ma ža tu pal late!*
 NEG go.IMP.2SG 2SG after 3SG.F.LOC
 Don't you go to her!

The non-emphatic clitic pronouns *-lo* 'he', *-li* 'she' and *-le* 'they' are employed in non-verbal predications (266), and only rarely in verbal predications (267). They are postposed to the present indicative copula (see 4.7.4), while directly preposed to other copula forms, such as in (266) and (268).

- (266)^{NAR} *prik=lo sin thádo.*
 over=3SG.M COP.PRT.3 burnt
 It was burnt up.

- (267)^{NAR} *sikjolahi=li t' andr' iškola víd'ik.*
 learn.IMP.3SG=3SG.F also from school all_along
 She learnt (languages) also from the school all along.

(268)^{LQCR} *nasválo=lo sin.*
 sick=3SG.M COP.PRT.3
 He was sick.

The reflexive pronoun is generally placed after the finite verb in sentences where the first position is occupied by the subject or the verb itself, e.g. (269).

(269)^{LQCR} *oja lumni urđini pe.*
 that.F woman dress.PRT.3SG.F REFL.3SG
 That woman dressed herself.

On the other hand, the reflexive pronoun is preposed to the finite verb (or to the finite verb directly preceded by the negator *na*) when the predicate is preceded by other constituents of the sentence, such as by *šoha* in (270).

(270)^{NAR} *ój šoha pe na muklahi.*
 3SG.F never REFL.3SG NEG let.IMPF.3SG
 She never let herself.

The reflexive pronoun is inserted between the verbal particle and the verb, e.g. (271).

(271)^{LQCR} *la mojaha tél pe čhord'a.*
 DEF.OBL vine.INS VP REFL.3SG pour.PRT.3SG
 S/he poured the wine to her/himself.

The linear order of verbal particles and verbs generally follows the Hungarian pattern (see e.g. Koopman & Szabolcsi 2000: 10–12). According to that, the verbal particle directly precedes the verb in neutral sentences, i.e. in sentences without focus or negation, e.g. (272)a, cf. H (272)b. On the other hand, the particle follows the verb in sentences with negated or focused phrases, while the preverbal position is occupied by the negation marker or the focused constituent, e.g. by the focused constituent *kašti* in (273)a, cf. H (273)b.

(272)^{LQCR} Somogy Romani

- a. *fer bujind'a pal ék kašt.*
VP hide.PRT.3SG behind a tree
Hungarian
- b. *el-bújt egy fa mögé.*
VP-hide.PRT.3SG a tree behind
S/he hid behind a tree.

(273)^{LQCR} Somogy Romani

- a. *pal ék kašt bujind'a fer.*
behind a tree hide.PRT.3SG VP
Hungarian
- b. *egy fa mögé bújt el.*
a tree behind hide.3SG VP
Behind the tree, s/he hid there.

In contrast to Hungarian, in KR the pronominal direct object can stand between the particle and the verb in neutral sentences. For instance, the pronoun *leske* is inserted between the particle *fer* (resultative aktionsart) and the verb *phénav* ‘into’ in (274)a, an order that is not permissible in Hungarian (274)b.⁷⁶

(274)^{NAR} Somogy Romani

- a. *fer leske phénav i paramisi*
VP 3SG.M.DAT tell.1SG DEF tale
Hungarian
- b. *el-mondom neki a mesét*
VP-tell.1SG 3SG.DAT DEF tale
I'll tell him the tale.

⁷⁶ The same structure was observed by Wogg and Halwachs (1998:18–19) for Burgenland Romani.

5.4.4 Subordinated and interrogative clauses

Subordinated clauses are generally introduced by a subordinator, which can be preceded by a focused constituent, such as by *akársaj búti* in (275). The focused constituent is positioned between two parts of the complex subordinator *hod' te*, e.g. (276) and (277)⁷⁷.

(275)^{RM} *akársaj búti te úli,*
 whatever job if COP.SUBJ.PRT.3SG.F
akársaj meláli vaj erd'avi, kèras le.
 whatever dirty or bad do.1PL 3SG.M.ACC
 Whatever job it is, no matter how dirty or bad it is, we do it.

(276)^{NAR} *még t' odá kannahi, hod' o hábe te ár po'inen.*
 even also that.M want.IMPF.3PL COMP DEF food COMP VP pay.3PL
 They even wanted them to pay for the food.

(277)^{NAR} *o Dél odá te del, hod' tri men te čhínes téle.*
 DEF God that.M COMP give.3SG COMP your.SG neck COMP break.2SG VP
 May the God make you break your neck!

Interrogatives generally occupy the first position of interrogative clauses (278). If the subject is emphasised, it may be preposed to the interrogative (279). Interrogatives may also be preceded by discourse markers, such as by *hát* in 0.

(278)^{LQCR} *káj h' oja lumni?*
 where COP.3 that.F woman
 Where is that woman?

(279)^{NAR} *Taraña mere hi?*
 Tarany which_direction COP.3
 To which direction is Tarany (village in Hungary)?

⁷⁷ Note that the verbal particle may both precede and follow the verb in these subordinated clauses, e.g. *ár* in (276), and *téle* in (277).

(280)^{NAR} *hát sar šaj ól adá?*
so how can COP.SUBJ.3SG this.M
So how is that possible?

The yes-no questions differ from their indicative equivalents only in the intonation, not in word order, e.g. SV in (281), VS in (282).

(281)^{NAR} *tu még osó na dikjal(?)*
2SG yet such NEG see.PRT.2SG
Haven't you seen such a thing? You have not seen such a thing.

(282)^{NAR} *fer pe sítí má i čhaj(?)*
VP REFL.3SG sleep.PRT.3SG.F already DEF girl
Has the girl already fallen asleep? The girl has already fallen asleep [with focus on the verb phrase].

6 Lexicon

The core vocabulary of Romani, including KR, is of Indo-Aryan origin. There are also a number of lexical roots of Greek (Boretzky & Iгла 1994: 332–338), and some others of Iranian and Armenian origin (ibid: 333–338). These roots are to a varying extent shared by all present-day Romani dialects. KR has been further shaped by its recent contact languages: Slavic, German and Hungarian. In the following two sections I will indicate only the source forms of the Slavic, German and Hungarian loanwords, while for the source forms of the Greek loanwords I refer the reader to Boretzky and Iгла (1994: 333–338) and Boretzky (2012). The inherited lexicon including the pre-Greek loanwords will be discussed only for the domain of kinship terms.

6.1 Layers of lexical borrowings

Figure 12 demonstrates the distribution of the inherited lexicon and the recent layers of lexical borrowings, based on the vocabulary of KR (see the Appendix). I have considered all lexemes of inherited matter to be ‘inherited’, even though they were clearly calqued from the contact languages, such as *edej-ánde* (cf. H *ide-benn*) ‘lit. here-inside; inside’. The classification of lexemes according to the source language is based on the origin of the lexical root. For instance, I considered the noun *šógor-kiňa* (cf. H *sógor-nő*) ‘sister-in-law’ to be of Hungarian origin, though it is an internally derived noun with the Slavic-origin suffix *-kiňa*. Similarly, I have classified the verb *máčh-ál-in-* ‘to fish’ as inherited, though it is formed by the Hungarian-origin denominal suffix *-ál-* from the noun *máčh-o* ‘fish’. The term ‘inherited’, as it has already been defined in section 1.2, includes also the Greek loanwords.

The vast majority of Greek loanwords in KR are also found in other South Central varieties, such as *angáli* ‘lap’, *ármi* ‘sour cabbage’, *cipa* ‘skin, leather’, *drom* ‘way, road’, *éfta* ‘seven’, *eňa ~ eňňa* ‘nine’, *fóro* ‘town’, *hij*⁷⁸ ‘to defecate’, *hóli* ‘anger’, *hovéli* ‘live coals’, *irín-* ‘to turn’, *kokalo* ‘bone’, *kopana* ‘trough’, *silavo* ‘pincers’, *kurko* ‘week; Sunday’, *mulhi*

⁷⁸ According to Boretzky and Iгла (1994: 334), the etymology of the verb meaning ‘to defecate’ is doubtful. They propose either Greek or Indo-Aryan-origin.

‘fog’, *ófto* ‘eight’, *óra* (or < H *óra*) ‘hour’, *papal*⁷⁹ ‘again’, *pápin* ‘goose’, *papu* ‘grandfather’, *paramísi* ‘tale’, *paraštuva* ‘Friday’, *pizdi* ‘jelly’, *rumín-* ‘to ruin, spoil’, *sája* ‘saliva’, *sapuňi* ‘soap’, *sirimi* ‘belt’, *svíri* ‘hammer’, *táha* ‘tomorrow’, *tranda* ‘thirty’, *tróma-* ‘to dare’, *zumi* ‘soup’, and *žamba*⁸⁰ ‘frog’. KR has preserved some further Greek loanwords, which are absent or very rare in the northern varieties of South Central Romani. These are the *cukňúdi* ‘nettle’, *harkum* ‘lead’, *karavd’in* ‘crab’, *kukur* ‘hailstone’ and *vrca* ‘hair’.

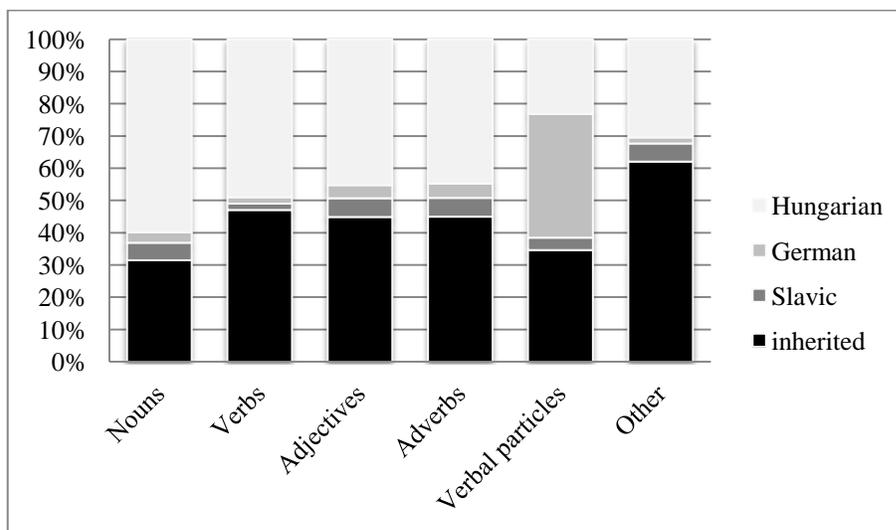


Figure 12 Inherited lexicon and layers of lexical borrowings according to the parts of speech

The table above shows that most KR loanwords originate from Hungarian, while the number of German and Slavic loanwords is much less significant. The only exceptions are the verbal particles, since the German-borrowed verbal particles outnumber both the Hungarian-origin and inherited ones. On the other hand, Hungarian has contributed the most to the nouns, while the inherited lexicon remains relatively strong regarding the verbs, adjectives, adverbs and other parts of speech.

I have counted a total of 81 roots (ca. 127 lexemes) of Slavic origin in my data. According to Vekerdi (1984: 66), Slavic loanwords in Vend Romani are not from Slovenian, but from the *čakavic* dialect (of Croatian), which is spoken in the Austro-Hungarian border region. Nevertheless, Vekerdi mentions only the determiner *cil-o* to be of *čakavic* origin (ibid:

⁷⁹ The origin of the form *papal* is the Greek *pale* (Boretzky and Iglá 1994: 336). The KR *papal* has developed through the reduplication of the stem, i.e. *pale-pale**.

⁸⁰ Vekerdi (2000: 173) proposes the noun *žamba* to be of Serbian origin.

74), while the other Slavic loanwords are labeled as Serbo-Croatian in his article (ibid: 73–75). Nonetheless, South Slavic most probably influenced the ancestor language of KR not only *in situ*, but already in the time of the migration of the Roma through the Balkans. Thus, in case of KR it is convenient to distinguish two layers of Slavic borrowings: 1) an older layer which is shared with the northern varieties of South Central Romani, and 2) a recent layer consisting of loanwords typical only for Vend Romani. The older Slavic layer include loanwords such as the before-mentioned *cil-o* (< S *cio* ~ *cil-o*; Elšík 2009: 271) ‘whole’, *boba* ‘maize’ and *bobo* ‘beans’ (both < S *bob* ‘bean’), *bubreško* (< S *bubreška*, cf. *bubreg* in Vekerdi 2000: 39) ‘kidney’, *duhano* (< S *duhan*) ‘tobacco’, *erd'av-o* ‘bad’ (< S *rđav* ‘rusty, bad, evil’; Elšík 2009: 271), *kruška* (< S *kruška*) ‘pear’, *klédalo* (< S *ogledalo*) ‘mirror’, *klinco* (< S *klinac*) ‘nail’, *lin-o* (< S *len*, *lijen*) ‘lazy’, *molín-* (< S *mol-iti*) ‘to pray’, *nébo* ‘umbrella’ (< S *nebo* ‘sky’), *péťa* (< S *peć*) ‘stove’, *plasta* (< S *plahta*) ‘sheet, bed-sheet’, *praho* (< S *prah*) ‘dust; ash’, *prosto* ‘non-Rom’ (< S *prost* ‘simple, dumb’), *rokla* (< S *roklja*; Vekerdi 2000: 142) ‘skirt’, *staklo* (< S *staklo*) ‘bottle’, *sveci* ‘feast’ (< S *sveto* ‘saint’; Vekerdi 2000: 152), *uzar* (< S *uz*) ‘beside, next to, from’, *víra* (< S *vjera*) ‘vow’, *zelen-o* (< S *zelen*) ‘green’ or *žut-o* (< S *žut*) ‘yellow’. On the other hand, the recent Slavic layer subsume the loanwords *božiťa* (< S *božič*) ‘Christmas’, *češňáko* (< S *česnek*) ‘garlic’, *dokle* (< Slovenian *dokler*) ‘until’, *évda* (< S *jedva*) ‘hardly’, *garvano* ‘sparrow hawk’ (< S *gavran* ‘raven’), *germín-* (< S *grm-eti*) ‘to thunder’, *gráblálín-*⁸¹ ‘to rake’ (< S *grablj-e* ‘rake’), *grobo* (< S *grob*) ‘grave’, *gulubica* (< S *golobica*) ‘*Russula vesca* (kind of mushroom)’, *kada*⁸² (< S *kad*) ‘when’, *kosa* (< S *kosa*) ‘scythe’, *lani* (< S *lani*) ‘last year’, *lasn-e* ‘cheaply’ (< S *lasn-o* ‘easy’), *morkoňi* (< Slovenian dial. *morkova*)⁸³ ‘carrot’, *motika* (< S *motika*) ‘hoe’, *nojo* (< S *gnoj*) ‘dung’, *opruja* (< S *obrva*) ‘eyebrows’, *pádári* ‘doctor’ (< S *padar* ‘quack’), *pékári* (< S *pekar*) ‘baker’, *plán* (< S *plande*; Vekerdi 1984: 74) ‘noon’, *pómod'* (< S *pomoč*) ‘help’, *pupa* (< Slovenian dial. *pup*) ‘bellybutton’, *puručín-* (< S *poroč-ati*) ‘to message’, *silom* (< S *silom*) ‘purposely, intentionally’, *srída* (< S *sreda*) ‘Wednesday’, *šlězinka* (< S *slezina*) ‘spleen’, *trašilo* (< S

⁸¹ The verb *gráblálín-* is formed from the noun **grabl-a* by means of the denominal *-ál-* and the adaptation suffix *-ín-*.

⁸² The form *kada* has most probably resulted from the contamination of the original **kana* ‘when’ and the Slavic *kad* ‘when’.

⁸³ The intervocalic *v* changes to *j* in KR (see 3.1.8). The KR form *morkoňi* is probably the plural form of **morkoňa*.

strašilo ‘scarecrow’, *tresán-* (< S *tres-ti*) ‘to shake’, *trézvisajov- ári* ‘to sober’ (< S *trezv-o* ‘sober’), *vódín- ári* (< S *vod-iti*) ‘take out, walk out’, *vóra* (< S *dvor*) ‘court, yard’, *vusko* (< Slovenian dial. *vuzak*) ‘narrow’, *zubuňi* (< S *zubun*; Vekerdi 1984: 75) ‘coat’, and possibly also *mégudi* ‘even’. Determining which Slavic dialects were the donor languages of KR would certainly shed light on the migration route of the ancestors of the Kisbajom Roma. This would however require further research.

The number of the German roots is 69 (ca. 99 lexemes) in my sample. These were most probably borrowed from the German dialect spoken in the Austro-Hungarian border area. However, the Hungarian dialect spoken in the same area also borrowed several of these German loanwords. Thus, in many instances it is difficult to determine whether KR directly borrowed the word from German or via Hungarian. Examples of such words are *šlájferi* (cf. G *Schleifer*, H dial. *slájfer*; Imre 1973: 150) ‘grinder’, *štrajtín-* (cf. G *streit-en*; H dial. *strájd-ul*, ibid: 152) ‘to quarrel’, or *klát* (cf. G dial. [k^hladl]; H dial. *klad, kload*; ibid: 100) ‘dress’. These loanwords I have mostly labelled as German. I have classified the German-origin words as Hungarian only if they are also commonly used in the Hungarian language, or being widespread in Hungarian dialects. Examples of these are the KR nouns *špajz* ‘larder’ (cf. H *spájz*, G *Speis*), *firhang* (cf. H dial. *firhang*, G *Vorhang*) ‘curtain’, *jáger* (cf. H dial. *jäger*, G *Jäger*) ‘hunter’, or *štrimfi* (cf. H dial. *strimfi*, G *Strümpfe*) ‘tights’.

Layers of borrowing in other varieties of Vend Romani

Further Greek loanwords in other than KR Romani varieties are the *cimbi* ‘eyebrow’ and *d’erni* ‘file, rubber’ (in Somogy, Zala and Vas), *petala* ‘horseshoe’ (in Somogy and Veszprém), *stádik* (in Vas), and *skámi* ~ *eskámo* (in several Vend Romani varieties). The number of German loanwords is slightly higher in western varieties of Vend Romani, i.e. in Sopron, Veszprém and Vas Romani. In contrast to KR, almost all German-borrowed verbal particles are absent in Zala Romani (see 4.7.6).

6.2 Semantic domains⁸⁴

Having distributed the lexemes by their semantic values, I have found that either the inherited (including Greek) or the Hungarian lexicon is dominant in each semantic domain. The inherited lexicon prevails in the domains denoting human beings, body parts, time and food, drinks and drugs, while Hungarian outnumbers both the inherited lexicon and the German and Slavic loanwords in the rest of the domains (Figure 13). In the following section I will analyze the inherited lexicon only with regard to the kinship terms and human beings. As for the inherited lexicon in other semantic domains, I refer the readers to Elšík (2009) and Matras (2002: 20–30). In addition, I will deal more thoroughly only with those earlier contact languages to which the KR speakers have lost access, i.e. Greek, Slavic and German.

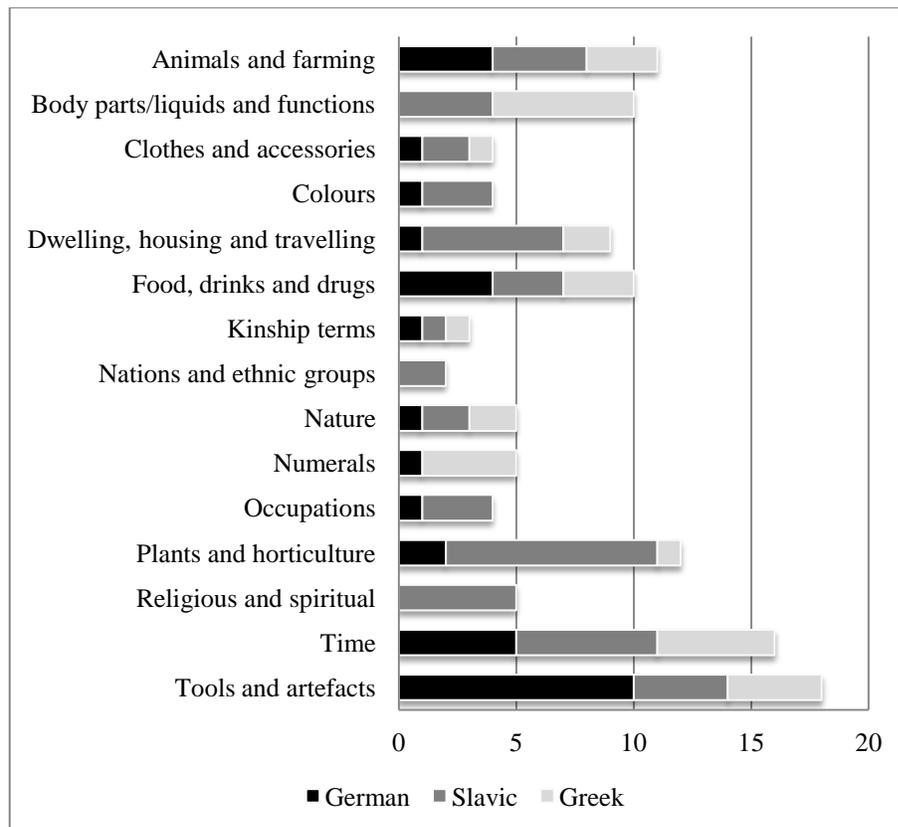


Figure 13 Inherited and borrowed lexicon with regard to some selected semantic domains

The most interesting development in KR is that the **kinship terms** **rom* ‘husband’ and **romni* ‘wife’ were replaced by *murš* (original meaning ‘man’) and *žuvli* (original meaning

⁸⁴ This chapter has drawn on the author’s earlier paper (Bodnárová 2013b).

‘woman’) or *lumni* (original meaning ‘whore’), respectively. The meaning of *rom* was reduced to ‘Romani man’, while the original *romni* ‘Romani woman’ is known only passively by the speakers. The meaning ‘whore’ is now expressed by the borrowed *kurva* (< H *kurva*). Furthermore, the original term **čhavo* ‘child, son’ was forced out by the Hungarian-borrowed *fat’ú* (< H *fattyú*), with the original meaning of ‘bastard’. The female counterpart of *fat’ú* is the original *čhaj* ‘daughter’. Further kinship terms are the original *daj* ‘mother’, *dad* ‘father’, *phral* ‘brother’, *phen* ‘sister’, *žamutro* ‘son-in-law’, *sástro* ‘father-in-law’, *sási* ‘mother-in-law’ and *kirivo*⁸⁵ ‘godfather’, the Greek *papu* ‘grandfather’, the Slavic *baba* (< S *baba*) ‘grandmother’, and the German *móm* (< G *Muhme*) ‘aunt’. All other kinship terms are borrowed from Hungarian, such as *bači* (< H *bácsi*) ‘uncle’, *šógori* (< H *sógor*) ‘brother-in-law’, *unoka* (< H *unoka*) ‘grandchild’ or *unokatešvér* (< H *unokatesvér*) ‘cousin’. The terms denoting ‘family’ and ‘relatives’ are the Hungarian-borrowed *čaláda* (< H *család*) and *nípo* (< H *nép* ‘folk’), respectively.

The terminology used for **human beings** include *muřš* and *mánuš* meaning ‘man’, and *žuvli*, *lumni* and *manušni* meaning ‘woman’. The terms denoting non-Romani ethnicity are *gážo* ‘non-Romani man’ and *gáži* ‘non-Romani woman’. The former has the opposite pair *rom* ‘Romani man’, while the latter has lost its counterpart term **romni* (see above). The Hungarian-borrowed *fat’ú* (see above) covers the meaning ‘boy’, irrespective of the person's ethnic belonging. The term *čhaj* became neutral, meaning that it refers to both Romani and non-Romani girl. I was also addressed in the conversation several times by *čhaj* ‘girl’ besides *manušni* ‘woman’. Thus, the ethnic belonging is mostly specified by means of the adjectives *román-o* ‘Romani’ and *gažikán-o* ‘non-Romani’, e.g. *gažikani čhaj* ‘non-Romani girl’, *gažikano muřš* ‘non-Romani man’, *romani žuvli* ‘Romani woman’. On the other hand, the ethnic identity is still encoded on the original terms *ráklo* ‘non-Romani boy’ and *rákli* ‘non-Romani girl’. The inherited terms *piráno* and its female counterpart *piráni* refer to engaged persons, and less commonly to extramarital lovers. Finally, the collective term ‘people’ is expressed by the inherited *žéne*.

To summarize, the KR terms originally referring only to Roma have been replaced by an ethnic-indifferent term in KR (**čhá* ‘Romani boy’ > *fat’ú* ‘boy’; **romni* ‘Romani woman’ > *lumni* ~ *žuvli* ‘woman’), or through the change of their semantic value they have become

⁸⁵ According to Elšík (2009: 268), based on Mānušs *et al.* (1997: 72), the Romani noun *kirivo* has Kurdish origin.

ethnic-indifferent (*čhaj* *'Romani girl' > 'girl'). By contrast, the terms encoding non-Romani ethnic belonging have been preserved together with their semantic value. This development is illustrated by the following two tables. Table 95 contains the original terms denoting human beings as I have reconstructed them for KR by considering the corresponding terminology found in other closely-related South Central varieties. The mentioned changes in the KR lexicon referring to human beings are portrayed in Table 96 (coloured in grey).

MEANING	ETHNIC-SPECIFIC		ETHNIC-INDIFFERENT
	Romani	non-Romani	
man	<i>rom</i>	<i>gážo</i>	<i>murš</i>
		<i>prosto</i>	<i>mánuš</i>
woman	<i>romni</i>	<i>gáži</i>	<i>žuvli</i>
		<i>prostófkina</i>	<i>manušni</i>
boy	<i>čhá</i>	<i>ráklo</i>	–
girl	<i>čhaj</i>	<i>rákli</i>	–

Table 95 Reconstructed terms denoting human beings

MEANING	ETHNIC-SPECIFIC		ETHNIC-INDIFFERENT
	Romani	non-Romani	
man	<i>rom</i>	<i>gážo</i>	<i>murš</i>
		<i>prosto</i>	<i>mánuš</i>
woman	–	<i>gáži</i>	+ <i>lumni</i> (< *'whore')
	–	<i>prostófkina</i>	<i>žuvli</i>
boy	–	<i>ráklo</i>	+ <i>fat'ú</i> (< H *'bastard')
	–	<i>rákli</i>	<i>čhaj</i>

Table 96 Terms denoting human beings

The Greek layer is the strongest in the domains of body parts and numerals. The Greek-origin **body parts** include the nouns *angáli* 'lap', *cipa* 'skin, leather', *kokalo* 'bone' and

vrca ‘hair’, the body liquid *sája* ‘saliva’, and the body-related function *hij*-⁸⁶ ‘defecate’. The Slavic loans of this domain are the *bubreško*⁸⁷ ‘kidney’, *opruja* ‘eyebrows’, *pupa* ‘bellybutton’ and *šlězinka* ‘spleen’. The Greek **numerals** are *éfta* ‘seven’, *eña ~ eňña* ‘nine’, *ófto* ‘eight’ and *tranda* ‘thirty’. In addition, KR borrowed the ordinal *ajšti* (< G *erste*) ‘first’ from German.

The Slavic loanwords are outstanding in the domains of animals and farming, plants and horticulture, dwelling and housing, religious and spiritual, nations, occupations and colours. Hence, we can draw a conclusion that the ancestors of the KR speakers were exposed to an intensive Slavic contact at the time of settling down, as it is indicated especially by the first three domains. Slavic words related to **animals and farming** are *garvano* ‘sparrow hawk’, *straka* (< S *straka*) ‘magpie’, *nojo* ‘dung’ and *trašilo* ‘scarecrow’. The nouns *éza* (< G *Esel*) ‘donkey’ and *ré* (< G *Reh*) ‘deer, roe’ are borrowed from German; as well as the verbs *fíšín-* (< G *fisch-en*) ‘to fish’ and *jógín-* (< G *jag-en*) ‘to hunt’. KR has preserved the Greek animal names *pápin* ‘goose’, *karavd’in* ‘crab’ and *žamba* ‘frog’. The Slavic-origin lexicon of **plants and horticulture** comprises the plant names *boba* ‘corn, maize’, *bobo* ‘beans’, *češňáko* ‘garlic’, *gulubica* ‘*Russula vesca* (kind of mushroom)’, *kruška* ‘pear’, *morkoňi* ‘carrot’, *slíva* (< S *sljiva*) ‘plum’ and some tools and activities related to horticulture such as *motika* ‘hoe’, *kosa* ‘scythe’ and *gráblálín-* ‘to rake’. The only Greek loanword denoting plant is *cukňúdi* ‘nettle’, while two others, *krumpa* (< G dial. [ǰrompan]) ‘potato’ and *paradajs* (< Austr. G *Paradeiser*) ‘tomato’, are borrowed from German. The domain of **dwelling, housing and travelling** is covered by the Slavic *klédalo* ‘mirror’, *klúčo* (< S *ključ*) ‘key’, *praho* ‘dust; ash’, *péťa* ‘stove’, *plasta* ‘sheet, bed-sheet’ and *vóra* ‘court, yard’, the Greek *drom* ‘way, road’ and *fóro* ‘town’, and the German *rittín- com* (< G *richten* ‘align, arrange’) ‘to make the bed’. Greek and German have not contributed to the **religious and spiritual** terms. By contrast, Slavic brought in KR the nouns *božita* ‘Christmas’, *grobo* ‘grave’, *sveci* ‘feast’ and *víra* ‘vow’, and the verb *molín-* ‘pray’. The terminology related to nations includes the Slavic nouns *ninčko* (< S *nemšk-i*) ‘German’ and *prosto* ‘non-Rom’. The **occupation** terms borrowed from Slavic are the *ninco* ‘soldier’ (< S *nemac* ‘German’), *pádári* ‘doctor’ and *pékári* ‘baker’. The German-origin occupation name *šlajferi* (< G *Schleifer*) ‘grinder’ is also used as a self-

⁸⁶ The etymology of the verb meaning ‘to defecate’ is doubtful (see above).

⁸⁷ Henceforth, I will indicate only those Slavic source forms which have not been mentioned in the previous chapter.

appellation meaning ‘Grinder Rom’ (see 1.3). Finally, the **colours** borrowed from Slavic are *zelen-o* ‘green’, *modr-o* (< S *moder*) ‘blue’ and *žut-o* ‘yellow’, and from German it is *práni* (< G *braun*) ‘brown’.

German loanwords prevail only in the domain of **tools and artefacts**. Some of them are related to the traditional profession of being a grinder (*ampós* < G *Amboß* ‘anvil’, *kráksni* < G dial. [ǵraksn] ‘tool bag’, *nítnen* < G *Nieten* ‘rivets’ and *šujajtli* < G (unidentified) ‘drill’), others are manipulable artefacts (*pon* < G dial. [pəʊn] ‘train’ and *mašin* < G *Maschine* ‘machine’), or tools used in the kitchen (*síderi* < G *Sieb* ‘sieve’, *šír* < G *Geschirr* ‘dish’ and *šlifera* < probably from a dialectal form of the G *Schöpflöffel* ‘ladle’). The noun *khuglina* (< G *Kugel*) ‘bullet, shot’ is also borrowed from German. In this regard, it is interesting to note that KR has an innovative name for the ‘gun’, that is *šóveskéro*⁸⁸ (the genitive form of *šóv* ‘six’). The above mentioned tools and artefacts point to the fact that German was a dominant language when the ancestors of the KR speakers practiced the profession of a grinder, and came in more intense contact with the modern world. The tool and artefact names *klinco* ‘nail’, *nébo* ‘umbrella’, *staklo* ‘bottle’ and *babuka* ‘doll’ are borrowed from Slavic, and the names *kopana* ‘trough’, *silavo* ‘pincers’, *svíri* ‘hammer’ and *sapuňi* ‘soap’ from Greek.

All three contact languages (i.e. Greek, Slavic and German) have largely contributed to the domains of **time and food, drinks and drugs**. The former semantic field comprises the Greek *kurko* ‘week; Sunday’, *óra* (or < H *óra*) ‘hour’, *papal*⁸⁹ ‘again’, *parašt’uva* ‘Friday’ and *táha* ‘tomorrow’, the Slavic *plán* ‘noon’, *srida* ‘Wednesday’, *subota* (< S *subota*) ‘Saturday’, *kada* ‘when’, *lani* ‘last year’ and *dokle* ‘until’, and the German *ajštímuj* (< G *erstemal*) ‘at first, for the first time’, *cajt* (< G *Zeit*) ‘time; weather’, *núrunť* (< G, unclear source form) ‘always’, *sajt* (< G *seit*) ‘since, since when’ and *špót* (< G *spät*) ‘late’. The latter domain contains the Greek *ármi* ‘sour cabbage’, *pizdi* ‘jelly’ and *zumi* ‘soup’, the Slavic *šunka* ‘ham’, *duhano* ‘tobacco’ and *trézvisajov-ári* ‘to sober’, and the German *élo* (< G *Öl*) ‘oil’, *khafě* (< G *Kaffee*) ‘coffee’, *rajs* (< G *Reis*) ‘rice’ and *té* (< G *Tee*) ‘tea’. On the other hand, Greek, Slavic and German have added only few loanwords to the domains of nature and clothes and accessories. The loanwords related to **nature** are the Greek *mulhi* ‘fog’ and *kukur* ‘hailstone’,

⁸⁸ The genitive form of *šóv* in the meaning of ‘gun, revolver’ is also attested in Vlax Romani, e.g. the genitive plural form *šovéngo* (Boretzky and Iglá: 1994: 271).

⁸⁹ The origin of the form *papal* is the Greek *pale* (see above).

the Slavic *germín-* ‘thunder’ and the German *luft* (< G *Luft*) ‘air’. The lexicon for **clothes and accessories** includes the Greek *sirimi* ‘belt’, the Slavic *rokla* ‘skirt’ and *zubuňi* ‘coat’, and the German *klát* (< G *Kleid*) ‘dress’.

Semantic domains in other varieties of Vend Romani

Zala and Prekmurje Romani use the inherited ethnic name *góri* ‘non-Rom’ instead of *gážo*, and the term *čhá* ‘boy, son, child’ instead of the borrowed *fa’ú*. The ethnic name *romni* is used in some Somogy Romani varieties only to address older women, while it has preserved its conservative meaning ‘Romani woman, wife’ in Sopron and Veszprém Romani.

7 Summary

Vend Romani in Hungary has been, so far, a seriously under-documented language. The present dissertation is a first attempt to provide a basic overview of the sociolinguistic situation of this Romani dialect, and to systematically describe its grammar.

The dissertation is pioneer in exploring the wide variety of ethnonyms used by the Vend Roma with reference to their own group and to other Romani groups. In this regard, I have found that the most widespread ethnonym of the Vend Romani group is the occupation name ‘Grinder’. I have also given an account of the subdivision of the Vend Romani group into smaller kin groups, such as *boboši*, *prahoši*, or *žukláši*. In addition, I have drawn attention to the fact that Vend Roma distinguish between two Romani groups the members of which have been monolingual in Hungarian for several generations: the ‘Hungarian Roma’ and the ‘Beggars’. These two groups have been so far treated as one and the same in the Romani literature, mostly under the name ‘Hungarian Roma’ or ‘Romungro’.

During my field research I have explored the entire area where Vend Romani is spoken, which has been only partially known before. As for the number of Vend Roma, I have pointed to the fact that the population census is not an appropriate tool for assessing the issue, knowing that the manifestation of ethnic identity is in particular sensitive in case of the socially marginalized groups of Roma.

The thesis also examines several sociolinguistic variables such as language transmission, domains of language use and language attitudes, which are supported by several quotations of Vend Roma. In the chapter dealing with the extent of documentation of Vend Romani I have analysed especially the work of Szmodics (1827; 1836) and Habsburg (1888; 1890), since these are the earliest sources on Romani in the region where Vend Romani is now spoken. After analysing their data I have come to the conclusion that they had documented the South Central Romani dialect spoken at that time by the ‘Hungarian Roma’ and ‘Beggars’, whose dialect has become extinct in the past few centuries. Thus, the first reliable source on Hungarian Vend Romani is provided by the Hungarian linguist József Vekerdi, and dated only to the second half of the 20th century.

In my dissertation I have put a special focus on the linguistic changes triggered by the Hungarian and/or German contact. I have found, for instance, that the German influence may account for the widespread contraction of sound clusters with intervocalic *v*, diachronically

resulting in a diphthong. The strong influence of Hungarian has subsequently led to the monophthongization of these diphthongs. This feature may be considered as a diagnostic feature of Vend Romani. Furthermore, I have devoted a significant part of the dissertation to the adaptation strategies regarding the Hungarian vowels, consonants, nouns, adjectives and verbs. Noteworthy is the innovative adaptation of the consonant-final Hungarian nouns into Vend Romani. These loan-nouns are neither morphologically nor phonologically adapted, but they still represent an integral part of the Vend Romani lexicon. The accusative forms of these loan-nouns, which do not equal either to the corresponding inherited or to the Hungarian forms, imply that the xenoclitic noun paradigms split into further paradigms, depending on the time of the borrowing. Furthermore, I have found that Vend Romani differs from other closely related Romani dialects in the fact that it has a significantly higher number of feminine loan-nouns. First, Vend Romani borrowed a number of German consonant-final nouns without any adaptation markers into the class of either feminine or masculine nouns, depending on their gender value in German. The speakers of Vend Romani then lost access to German, and their primary contact language became Hungarian. Since Hungarian is a language which does not distinguish between genders, the speakers have randomly assigned either masculine or feminine gender to the Hungarian consonant-final nouns when borrowed into Romani. This strategy has subsequently led to a large increase in the number of feminine loan-nouns in Vend Romani.

Another important finding, which has been only implicitly referred to throughout the thesis, is that Hungarian Vend Romani may be divided into two subgroups: 1) the group of the Zala Romani varieties, and 2) the group of the Somogy, Sopron, Vas and Veszprém Romani varieties. The former group shares several features with the Vend Romani varieties of Prekmurje (Slovenia), while the latter group is closely connected to the Vend Romani varieties spoken in Burgenland (Austria). An illustrative example of this division is the system of verbal particles. While the vast majority of the Zala Romani verbal particles are calqued from Hungarian and/or German, a significant number of the verbal particles found in other Vend Romani varieties of Hungary are directly borrowed from German. In general, Zala and Prekmurje Romani seem to have been exposed to less German influence than other Vend Romani varieties.

The main purpose of the dissertation was to contribute in the form of a sociolinguistic and grammatical description to the documentation of Vend Romani. The importance of this

purpose has gradually increased during my research, especially by realizing that Vend Romani is severely endangered by language shift to Hungarian.

References

- Antauer, Živa. 2010. *Oris slovnične in leksikalne strukture prekmurske romščine (s posebnim ozirom na slovanske izposojenke)* [An outline of the grammatical and lexical structures in Prekmurje Romani (with special focus on the Slavic loanwords)]. MA thesis at the University of Ljubljana, Faculty of Arts.
- Anttila, Raimo. 2003. Analogy: The Warp and Woof of Cognition. In: Brian, Josephs and Richard Janda (eds.) *The Handbook of Historical Linguistics*. Oxford: Oxford University Press. 425–40.
- Baldi, Philip and Andrea Nuti. 2010. Possession. In: Baldi, Philip and Pierluigi Cuzzolin (eds.) *New Perspectives on Historical Latin Syntax. Vol. 3: Constituent Syntax*. Berlin: Mouton de Gruyter. 239–387.
- Bari, Károly. 1999. *Cigány folklór. Magyarország-Románia, I-X* [Gypsy Folklore. Hungary-Romania. I-X] (CD). Budapest: Self-publication.
- Bárkányi, Zsuzsanna and Zoltán Kiss. 2010. A phonetic approach to the phonology of *v*: A case study from Hungarian and Slovak. In: Fuchs, Susanne, Martine Toda and Marzena Żygis (eds.) *Turbulent Sounds. An Interdisciplinary Guide*. Berlin: Mouton de Gruyter. 103–142.
- Blevins, James P. and Juliette Blevins. 2009. Introduction: Analogy in grammar. In: Blevins, James P. and Juliette Blevins (eds.) *Analogy in grammar. Form and acquisition*. Oxford: Oxford University Press. 1–12.
- Bodnárová, Zuzana and Jakob Wiedner. 2015a. A Comparative study of verbal particles in varieties of Vend Romani (submitted for publication in 2014).
- Bodnárová, Zuzana and Jakob Wiedner. 2015b. Analogical extension of vowel length in Vend Romani (submitted for publication in 2014).
- Bodnárová, Zuzana (ed.) 2013c. *Andi paramisi sig džal – Time goes fast in fairytales. Text collection of South Central Romani in Hungary*. (= Grazer Romani Publikationen 3. Grazer Linguistische Monographien). Graz.
- Bodnárová, Zuzana. 2009. *Gramatický náčrt romského dialektu maďarské obce Versend* [A Grammatical Sketch of the Romani Dialect of Versend, Hungary]. MA thesis, Charles University in Prague, Prague.

- Bodnárová, Zuzana. 2013a. Overview of Vend Romani in Hungary. In: Schrammel-Leber, Barbara and Barbara Tiefenbacher (eds.) *Romani V. Papers from the Annual Meeting of the Gypsy Lore Society, Graz 2011*. (= Grazer Romani Publikationen 2. Grazer Linguistische Monographien). Graz. 29–41.
- Bodnárová, Zuzana. 2013b. A magyarországi köszörösök etnikai és rokonsági terminológiája [Ethnic and kinship terminology of the Hungarian Grinders]. *Romológia* 2/2013, 149–162.
- Bodnárová, Zuzana. 2014. Loanword integration: A case study of Kisbajom Romani. *Romani Studies* 24(1), 71–91.
- Borbély, Anna. 2001. A magyarországi beás nyelv helye a román nyelven belül [The place of the Hungarian Boyash language within Romanian]. In: Cserti Csapó, Tibor (ed.) *Cigány nyelvek nemzetközi szemináriuma. Konferenciakötet*. (= Gypsy Studies – Cigány Tanulmányok 6.) Pécs: PTE BTK Romológia tanszék. 75–84.
- Boretzky, Norbert and Birgit Iгла. 1994. *Wörterbuch Romani-Deutsch-Englisch für den südosteuropäischen Raum: Mit einer Grammatik der Dialektvarianten*. Wiesbaden: Harrassowitz.
- Boretzky, Norbert. 1998. Der Romani-Wortschatz in den Romani-Misch-Dialekten (Pararomani). In: Matras, Yaron (ed.) *The Romani Element in Non-Standard Speech*. Wiesbaden: Harrassowitz. 97-132.
- Boretzky, Norbert. 1999. Die Gliederung der Zentralen Dialekte und die Beziehungen zwischen Südlichen Zentralen Dialekten (Romungro) und Südbalkanischen Romani-Dialekten. In: Halwachs, Dieter W. and Florian Menz. *Die Sprache der Roma*. Klagenfurt: Drava Verlag. 193–256.
- Boretzky, Norbert. 2012. *Die lexikalischen Gräzismen des Romani* (= Grazer Linguistische Studien 31). Graz.
- Cech, Petra. 2006. *Dolenjska Romani: The dialect of the Dolenjski Roma in Novo Mesto and Bela Krajina, Slovenia*. Munich: Lincom.
- Crevels, Mily and Peter Bakker. 2000. External possession in Romani. In: Elšík, Viktor and Yaron Matras (eds.) *Grammatical relations in Romani: The noun phrase*. (= Current Issues in Linguistic Theory, 211). Amsterdam & Philadelphia: John Benjamins. 151–185.
- Croft, William. 1990. *Typology and Universals*. Cambridge: Cambridge University Press.

- Dékány, Éva. 2008. *El + verb complex predicates in Hungarian*. In: Svenonius, Peter and Inna Tolskaya (ed.) *Tromsø Working Papers on Language & Linguistics: Nordlyd 35*, Special issue on Complex Predication. Tromsø: CASTL. 1–17.
- Eisenberg, Ann R. 1986. Teasing: verbal play in two Mexicano homes. In: Schieffelin, Bambi B. and Elinor Ochs (eds.) *Language socialization across cultures*. Cambridge: Cambridge University Press. 182–98.
- Elšík, Viktor, Milena Hübschmannová and Hana Šebková. 1999. The Southern Central (ahi-imperfect) Romani dialects of Slovakia and Northern Hungary. In: Halwachs, Dieter W. and Florian Menz (eds.) *Die Sprache der Roma. Perspektiven der Romani-Forschung in Österreich im interdisziplinären und internationalen Kontext*. Klagenfurt: Drava. 277–390.
- Elšík, Viktor and Yaron Matras. 2006. *Markedness and language change: The Romani sample* (= Empirical Approaches to Language Typology, 32). Berlin & New York: Mouton de Gruyter.
- Elšík, Viktor and Yaron Matras. 2009. Modality in Romani. In: Hansen, Björn and Ferdinand de Haan (eds.) *Modals in the Languages of Europe*. Berlin: Mouton De Gruyter. 267–322.
- Elšík, Viktor. 2000a. Romani nominal paradigms: their structure, diversity, and development. In: Elšík, Viktor and Yaron Matras (eds.) *Grammatical relations in Romani: The noun phrase* (= Current Issues in Linguistic Theory, 211). Amsterdam & Philadelphia: John Benjamins. 9–30.
- Elšík, Viktor. 2000b. Dialect variation in Romani personal pronouns. In: Elšík, Viktor and Yaron Matras (eds.) *Grammatical relations in Romani: The noun phrase* (= Current Issues in Linguistic Theory, 211). Amsterdam & Philadelphia: John Benjamins. 65–94.
- Elšík, Viktor. 2007. *Affix Extraction: A Case Study on Hungarian Romani*. PhD thesis at the Charles University in Prague, Faculty of Arts.
- Elšík, Viktor. 2008–a. *Linguistic Database for the Documentation of Central European Romani* [= LDCR]. Charles University in Prague.
- Elšík, Viktor. 2008–b. *Linguistic Questionnaire for the Documentation of Central European Romani* [= LQCR]. Versions 1–4. Language mutations: Czech, Slovak, Polish, Hungarian, Ukrainian, Russian, Slovene, German. Charles University in Prague.

- Elšík, Viktor. 2009. Loanwords in Selice Romani, an Indo-Aryan language of Slovakia. In: Haspelmath, Martin and Tadmor, Uri (eds.) *Loanwords in the World's Languages: A comparative handbook*. Berlin: Mouton De Gruyter. 260–303.
- Elšík, Viktor. 2014. Multal and paucal quantifiers in Central Romani; paper presented at the 11th International Conference on Romani Linguistics, Oslo, 15-17 September.
- Glaeser, Ursula, Dieter W. Halwachs and Mozes F. Heinschink. 1999. Vier Vend-Romani-Texte. *Grazer Linguistische Studien* 51, 31–45.
- Greenberg, Joseph H. 1966. *Language Universals, With Special Reference to Feature Hierarchies*. The Hague: Mouton.
- Grinevald, Collette and Michel Bert. 2011. Speakers and Communities. In: Austin, Peter K. and Julia Sallabank (eds.) *Cambridge Handbook of Endangered Languages*. Cambridge: Cambridge University Press. 45–65.
- Habsburg, József Főherceg. 1888. *Czigány nyelvtan* [Gypsy grammar]. Budapest: Magyar Tudományos Akadémia.
- Habsburg, József Főherceg. 1890. Eredeti cigánylevelek [Original Gypsy letters]. *Egyetemes Philologiai Közlöny*, XIV, 737–768.
- Halwachs, Dieter W. and Michael Wogg. 2002. *Burgenland-Romani (Roman)*. Munich and Newcastle: Lincom.
- Halwachs, Dieter W. 1998a. *Morphologie des Roman. Basisgrammatik der Romani-Variante der Burgenland-Roma, Arbeitsbericht 3a des Projekts Kodifizierung und Didaktisierung des Roman*. Graz: Romani-Projekt.
- Halwachs, Dieter W. 1998b. *Phonologie des Roman, Arbeitsbericht 5 des Projekts Kodifizierung und Didaktisierung des Roman*. Graz: Romani-Projekt.
- Halwachs, Dieter W. 2002. Anmerkungen zur Romani-Variante der Prekmurje. *Grazer Linguistische Studien* 58, 43–54.
- Hannabauer, Franz. 2007. *Erstes Burgenländisches Mundart-Wörterbuch*. Oggau: Burgenländisch-Hianzische Gesellschaft.
- Haspelmath, Martin. 1997. *From Space to Time: Temporal Adverbials in the World's Languages*. (= Lincom Studies in Theoretical Linguistics, 3.) Munich and Newcastle: Lincom Europa.

- Haspelmath, Martin. 2007. Coordination. In: Shopen, Timothy (ed.) *Language typology and syntactic description, vol. II: Complex constructions*. 2nd ed. Cambridge: Cambridge University Press. 1–51.
- Heine, Bernd. 1997. *Possession: cognitive sources, forces, and grammaticalization*. Cambridge: Cambridge University Press.
- Hock, Hans Henrich. 1986. *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter.
- Hungarian Population Census 2001. <http://www.nepszamlalas2001.hu/> (accessed March 7, 2015).
- Hungarian Population Census 2011. <http://www.ksh.hu/nepszamlalas/> (accessed March 7, 2015).
- Hübschmannová, Milena and Vít Bubeník. 1997. Causatives in Slovak and Hungarian Romani. In: Matras, Yaron, Peter Bakker and Hristo Kyuchukov (eds.) *The typology and dialectology of Romani* (= Current Issues in Linguistic Theory, 156). Amsterdam & Philadelphia: John Benjamins. 133–145.
- Hübschmannová, Milena. 1995. Trial and error in written Romani on the pages of Romani periodicals. In: Matras, Yaron (ed.) *Romani in contact: The history, structure and sociology of a language*. Amsterdam: John Benjamin's Publishing Company. 189–205.
- Igla, Birgit. 1992. Entlehnung und Lehnübersetzung deutscher Präfixverben im Sinti. In: Erfurt, Jürgen, Benedikt Jeßing and Matthias Perl (eds.) *Prinzipien des Sprachwandels, I. Vorbereitung. Beiträge zum Leipziger Symposium Prinzipien des Sprachwandels 1991 an der Universität Leipzig*. Bochum: Brockmeyer. 38–56.
- Imre, Samu. 1971. *A felsőőri nyelvjárás* [The dialect of Oberwart]. Budapest: Akadémia kiadó.
- Imre, Samu. 1972. Megfigyelések a magyar nyelvjárások változásáról [Notes on the changes of the Hungarian dialects]. *Általános Nyelvészeti Tanulmányok VIII*, 85–104.
- Imre, Samu. 1973. *Felsőőri tájszótár* [The dialect vocabulary of Oberwart]. Budapest: Akadémia kiadó.
- Jakobson, Roman. 1939. *Signe zéro. Melanges de linguistique offerts a Charles Bally*. Librairie de l'Université, Geneva, 143–152.
- Kanczler, Balázs. 2009. A cigány identitás változása az asszimiláció és a szegregáció tükrében [Change of the Romani identity through assimilation and segregation]. *Publicationes Universitatis Miskolciensis* 14(1), 111–122.

- Kenesei, István, Robert M. Vago and Anna Fenyvesi. 1998. *Hungarian*. London and New York: Routledge.
- Király, Lajos. 2005. *Somogy-zalai nyelvatlasz* [Linguistic Atlas of Somogy and Zala]. Budapest: Magyar Nyelvtudományi Társaság Kiadványai.
- Kittilä, Seppo. 2005. Recipient prominence vs. beneficiary prominence. *Linguistic Typology* 9, 269–297.
- Knobloch, Johann. 1953. *Romāni Texte aus dem Burgenland* (= Burgenländische Forschungen, 24). Eisenstadt.
- Koopman, Hilda and Anna Szabolcsi. 2000. *Verbal Complexes*. Cambridge, MA: MIT Press.
- Kozár M., Mária. 1999. A magyarországi szlovének. Letelepedésük, múltjuk, jelenük [Slovenes in Hungary. Their settling down, history and present]. *Vasi szemle* 53/4, 492–510.
- Ladányi, Mária. 2000. Productivity as a sign of category change: the case of Hungarian verbal prefixes. In: Dressler, Wolfgang U. et al. (eds.) *Morphological Analysis in Comparison*. Amsterdam: John Benjamins. 113–141.
- Mānušs, Leksa, Jānis Neilands and Kārlis Rudevičs. 1997. *Čigānu–latviešu–angļu etimoloģiskā vārdnīca un latviešu–čigānu vārdnīca*. [Gypsy–Latvian–English etymological dictionary and Latvian–Gypsy dictionary.] Rīgā: Zvaigzne ABC.
- Marics, József. 2010. Fénylő csillagok a Magyar Egyház egén [Shining stars on the Hungarian Church's sky]. *Credo* 12/1(114), 4.
- Masica, Colin P. 1991. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.
- Matras, Yaron. 2002. *Romani: A Linguistic Introduction*. Cambridge: Cambridge University Press.
- Matras, Yaron. 2004. Typology, dialectology and the structure of complementation in Romani. In: Kortmann, Bernd (ed.) *Dialectology meets typology*. Berlin: Mouton de Gruyter. 277–304.
- Mészáros, György. 1980. *A Magyarországi szinto cigányok: Történetük és nyelvük*. [The Hungarian Sinti Roma: Their history and language]. Budapest: Magyar nyelvtudományi társaság.
- Miklosich, Franz. 1872–80. *Über die Mundarten und Wanderungen der Zigeuner Europas X–XII*. Vienna: Karl Gerold's Sohn.

- Morgan, David L. 2008. Snowball Sampling. In: Given, Lisa M. (ed.) *The SAGE Encyclopedia of Qualitative Research Methods*. Thousand Oaks, CA: Sage.
- Payne, Doris L. and Immanuel Barshi (eds.) 1999. *External Possession*. Amsterdam: John Benjamins.
- Puchmayer, Anton Jaroslaw. 1821. *Románi Čib, das ist: Grammatik und Wörterbuch der Zigeuner Sprache, nebst einigen Fabeln in derselben. Dazu als Anhang die Hantýrka oder die Čechische Diebessprache*. Prague: Fürst-erzbischöflichen Buchdruckerey.
- Réger, Zita. 1999. Teasing in the linguistic socialization of Gypsy children in Hungary. *Acta Linguistica Hungarica*, Vol. 16 (3–4), 289–315.
- Rézműves, Melinda. 2000. Pal o Šlajferitika roma [About the Grinder Roma]. *Amaro drom* 2, 24–25.
- Rézműves, Melinda. 2006. *O ćorro rrom taj o borzo: Rromane Paramici* [The poor Rom and the hedgehog: Romani tales]. Budapest: Fővárosi Cigány Önkormányzat.
- Schrammel, Barbara. 2005. Borrowed verbal particles and prefixes in Romani: a comparative approach. In: Halwachs, Dieter, Barbara Schrammel and Gerd Ambrosch (eds.) *General and applied Romani linguistics*. Munich: Lincom Europa. 99–113.
- Sowa, Rudolf von. 1887. *Die Mundart der Slovakischen Zigeuner*. Göttingen: Vandenhoeck und Ruprecht's Verlag.
- Štrukelj, Pavla. 1980. *Rom na Slovenskem* [Roma in Slovenia]. Ljubljana: Cankarjeva Založba.
- Szmodics, János. 1827. *Czigány Grammatika* [Gypsy grammar]. Gelse, manuscript.
- Szmodis, János. 1836. A cigány nyelvnek némelly tulajdonságai több más nyelvekéhez képest [Some characteristics of the Gypsy language as compared to many other languages]. Buda: *Tudománytár* 1336. IX, 47–59.
- UNESCO. 2003. *Language vitality and endangerment*. Ad Hoc Expert Group on Endangered Languages. Paris: UNESCO
- Vekerdi, József. 1982. *A magyarországi cigány kutatások története* [A history of research on the Hungarian Roma]. (= *Folklór és Etnográfia* 7). Debrecen: Kossuth Lajos Tudományegyetem.
- Vekerdi, József. 1984. The Vend Gypsy dialect in Hungary. *Acta Linguistica Academiae Scientiarum Hungaricae*, Volume 34, 65–86.

- Vekerdi, József. 1985. *Cigány nyelvjárási népmesék* [Gypsy dialect tales from Hungary I–II]. (= Folklor és etnográfia 19.) Debrecen: Kossuth Lajos Tudományegyetem.
- Vekerdi, József. 2000. *A magyarországi cigány nyelvjárások szótára* [A dictionary of Gypsy dialects in Hungary]. Budapest: Terebess Publications.
- Vollmann, Ralf and Sylvia Moosmüller. 2001. 'Natürliches Driften' im Lautwandel: die Monophthongierung im österreichischen Deutsch. *Zeitschrift für Sprachwissenschaft* 20/1, 42–65.
- Wiesinger, Peter. 1967. *Mundart und Geschichte in der Steiermark. Ein Beitrag zur Dialektgeographie eines österreichischen Bundeslandes*. Marburg: Elwert.
- Winter, Werner. 1969. Analogischer Sprachwandel und semantische Struktur. *Folia Linguistica* 3, 29–45.
- Wogg, Michael and Dieter W. Halwachs. 1998. *Syntax des Roman*. (Arbeitsbericht 6 des Projekts Kodifizierung und Didaktisierung des Roman.) Oberwart: Verein Roma.
- Zehetner, Ludwig. 1985. *Das bairische Dialektbuch*. Munich: Beck.