Univerzita Karlova v Praze

Filozofická fakulta

Ústav obecné lingvistiky

Obecná lingvistika

Zuzana Bodnárová

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

Teze

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

2015

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.

Chapter 1 – Introduction – introduces the ethnonyms used for/by the Vend Roma, the geographical distribution of the Vend Romani speakers, the dialect classification of Vend Romani, and the source of data and methodology.

According to the most established models of dialect classification (e.g. Boretzky 1999; Matras 2002), 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, and the varieties spoken in the region of Prekmurje in northern Slovenia. 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). The Vend Romani speakers live in Western Transdanubia, and in the western part of Central and Southern Transdanubia. Based on my field research, I estimate the total number of localities with Vend Romani speakers at around 75.

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 was introduced into the Romani literature by the Hungarian linguist Vekerdi (1984). Instead, the speakers call themselves by the professionym *köszörüs* 'Grinder' or less commonly *drótos* 'Tinker' in Hungarian, and *rom* 'Rom' in Romani. Only some speakers were familiar with the term 'Vend' or 'Vendel', considering it to be either an exonym or a subdivision within the group. It is questionable whether the term 'Vend' indicates that the Vend Roma used to live in the Hugarian-Slovene border region called *Vendvidék* 'lit. the land of Vends (= ethnic Slovenes)', or it has been transmitted to the Roma from the surrounding population of ethnic Slovenes only in Somogy. In the Vas county, 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. The Roma speaking Vend Romani in Veszprém have reported to use the designation *Sinti*; while they are called by some Sinti (i.e. Northwestern) Romani speakers *hinsznári*. 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.

The Vend Roma are divided into several smaller kin groups which are called *fajta* 'kin', *banda* 'group', *faj* 'race' or *nemzet* 'nation' in Hungarian, and are referred to by the Hungarian-borrowed terms *banda* 'group' and *fajta* 'kin', less commonly *nemzečég* 'kin', in Romani. Such kin groups or *fajta* are the *boboši* (< *bobo* 'bean'), *prahoši* (< *praho* 'dust'), *žukláši* (< *žúkel* 'dog'), *pataváši* (< *patavo* 'foot-rag'), *feňó* 'pine tree', etc. Furthermore, the Vend Roma perceive the existence of five other Romani groups: Boyash, Vlax Roma, Sinti, Hungarian Roma and Beggars.

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

The SECOND stage dealt with the data collection in the area of research. 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. The field research has been 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–). 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. 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. 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 THIRD stage aimed to deepen the research in one selected locality, on the variety of which the grammatical description of 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 the Somogy county. Kisbajom has in total approximately 450 residents, of which 79, according to the Hungarian Population Census of 2001,¹ claimed Roma ethnic identity, and 46 considered Romani as their mother tongue.

The FOURTH stage comprised the transcription of audio-recordings obtained during the field research, as well as the transcription of existing published sources.

Chapter 2 – Sociolinguistic vitality – deals with the factors influencing the sociolinguistic vitality of Vend Romani, using the UNESCO's (2003) Language Vitality and Endangerment evaluation guideline elaborated by an expert group. There are nine evaluative factors that determine the viability of a language:

¹ See <u>http://www.nepszamlalas2001.hu/</u>. The results published in the Hungarian Population Census from 2001 do not differentiate between varying Romani groups.

1. *Intergenerational language transmission*: Most local varieties of Vend Romani are severely endangered because the youngest speakers are of grandparental and older generations. Vend Romani is critically endangered especially in Vas, Veszprém and Győr-Moson-Sopron. The youngest speakers of these regions are generally of great-grandparental generation. Mixed marriages have become prevalent just recently, which is another factor influencing language transmission.

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).

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. The language shift towards Hungarian is a relatively recent development. 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. It may be thus concluded that only a minority of the population speaks the language.

4. *Trends in existing language domains*: 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.

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.

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.

7. *Governmental and institutional language attitudes and policies*: Romani as a minority language, irrespective of its dialects, is explicitly protected by the Hungarian government. However, teaching of Romani, as well the related standardization and codification efforts are based on a single variety of Vlax Romani. The Vlax Romani varieties are mainly unintelligible to the South Central Romani-speaking Vend Roma.

8. *Community members' attitudes toward their own language*: Almost all speakers are indifferent whether Romani is getting lost because they prefer to use Hungarian in their everyday life. Only the secret function of Romani has been reported by a few consultants as an advantage of Romani speakers.

9. *Amount and quality of documentation*: Vend Romani is only fragmentarily documented. Its documentation includes a brief grammatical sketch, a single word list, and some short fragmentary texts, mainly with no available audio recordings.

The following three chapters are devoted to the grammatical description. Hereinafter, I will present only some selected findings of the thesis.

Chapter 3 – Phonology – discusses the vowel and consonant inventory, the stress pattern and some morpho-phonological processes typical for Vend Romani.

Vend Romani has preserved all Early Romani consonant phonemes including the aspirates (cf. Matras 2002: 56), but not the voiceless velar fricative /x/. This sound has been replaced by the glottal fricative /h/ due to South Slavic or/and Hungarian influence (Elšík et al. 1999: 295–297). Hungarian seems to have brought only the affricate /dz/ into Vend Romani. The aspirated sounds are phonologically distinctive compared to their non-aspirated counterparts. It is interesting that 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 Bavarian origin, e.g. *khafé* (< G *Kaffee* [k^hafe:]) 'coffee'.

Consonant clusters may occur in every position of the word. The most widespread are the clusters consisting of two consonants, while those with three consonants occur mostly in medial, and rarely in initial position. Most consonants have their long counterparts. The existence of consonant length in Vend Romani is most probably triggered by the prolonged contact with Hungarian, that is, with a language with distinctive consonant length. Long consonants generally occur in word-medial position, while they are not allowed in initial position. Geminates are rather rare in final position, as they become de-geminated when adapted into Romani.

A distinctive feature of Vend Romani 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 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 has probably first resulted in hiatus and later in various diphthongs depending on the adjacent vowel, e.g. *tavel > *tael 'cooks'. In Vend Romani the diphthongs were replaced by a long vowel after the loss of German contact.

The vowel system of Vend Romani comprises fourteen vowel phonemes, seven of which are long. The front rounded short and long vowel pairs are reserved for the recent Hungarian loanwords. The phoneme /a/ may become realized as [a] and [p], 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* [bpra:t] 'friend'. The slightly rounded variant [p], 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 Romani, such as *balaton* [bplpton] < H *Balaton* [bplpton] 'Lake Balaton'. 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 [10:ro] 'flour', hijába [hija:bb] < H hijába [hi⁽ⁱ⁾a:bp] 'in vain'. 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 /p 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. Thus, Vend Romani 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 introduction of vowel length into Vend Romani has most probably been triggered by prolonged contact with Hungarian, a language that has length opposition. 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. The thesis describes how the vowel length distinction, after its reintroduction into Vend Romani, may have spread due to the process of analogical extension. In Vend Romani, this process accounts for the development of long vowels in the monosyllabic possessive pronouns in combination with the words for 'mother' and 'father', as well as the vowel length distinction between the imperfect and irrealis markers. The latter case is especially interesting in that the vowel of the imperfective suffix has become long throughout the irrealis paradigm (i.e. $-\dot{a}hi$), while the suffix with short vowel (i.e. -ahi) has been reserved exclusively for the imperfect. This means that vowel length has become functionally relevant, and has therefore enriched the grammatical structure of Vend Romani.

Chapter 4 – Morphology – is devoted to the strategy of loanword incorporation as well as to the derivational and inflectional morphology.

Vend Romani 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. Vend Romani differs from other closely related Romani dialects in that it has a significantly higher number of feminine loan-nouns. The motivation for adapting several Hungarian nouns into the class of Vend Romani feminine nouns may be explained by the strong German influence in the past. That is, a number of German consonant-final nouns have been borrowed without any adaptation markers into the class of either feminine or masculine Vend Romani nouns, depending on their gender value in German. As a result, Vend Romani possessed both feminine and masculine consonant-final German loan-nouns. 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 gender, the speakers have randomly assigned to the Hungarian consonant-final nouns either masculine or feminine gender when borrowed into Romani, according to the innovative adaptation pattern which have had developed through the German-contact. This strategy has subsequently led to a large increase of the number of feminine loan-nouns in Vend Romani.

Nouns are inflected for gender, number and case. Following Matras (2002: 78– 80), two distinct layers in the nominal case system 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 oblique stem agrees with the accusative, while the final *s* is preserved in the masculine singular forms, e.g. *murš-e* 'man.ACC', cf. *murš-es-* 'man.OBL-'. A striking development in Vend Romani is that the accusative forms of the recently borrowed *C*-final animate nouns are homonymous with their nominative forms, both in the singular and the plural. 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.

Two verbal stems are distinguished in Vend Romani (see Matras 2002: 135f.): 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, or with the verbal root extended by adaptation or valency markers. The perfective verbal stem is formed by means of the perfective marker either from the verbal root, or from the derived or adapted form of the verbal root. Individual inflectional classes are distinguished for both present and perfective verbal stems. The present stem is used to form the present, future, imperfect and imperative, while the preterite and conditional irrealis are based on the perfective stem. An innovation in Vend Romani is that the future may be optionally expressed by an analytical construction which involves the borrowed auxilary *fogin*- (< H *fog*) 'will' and the infinitive form of the verb. The auxilary fogin- is obligatorily inflected for future tense. Thus, the future is double marked: by the auxiliary fogin- on the one hand, and by the inherited future marker -a on the other. The construction with fogin- seems to be more frequent in the elicited data, while the inherited future marking predominates in the spontaneous narratives.

Another innovative strategy in Vend Romani, which undoubtedly requires further research, is that the borrowed verbs adapted by the marker -in- sometimes cease to be inflected when the relevant inflection is evident from the context. The uninflected form, which may substitute any inflectional form of the verb, corresponds to the inflectional stem. 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.

The development of verbal particles in Vend Romani is triggered by language contact with Hungarian and/or German. 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 complex verbs, which consist of a verbal particle plus verb, are generally (semi-)calques of Hungarian complex verbs, regardless of the particle's origin. A few complex verbs that have no equivalent counterparts in Hungarian are modelled on German complex verbs. An interesting case is the verbal particle um which – apart from expressing resultativity – is used to mark the repetition or the frequency of an action, i.e. iterativity. The origin of um as an iterative marker is probably from the German dialect verbal particle uma (dum) ['ome, ome'dom] [iterative]. It seems that due to the similarity of the German (dialect) forms um (resultative) and uma(dum) (iterative), these particles merged in Vend Romani. The development of a verbal particle with an iterative meaning is an interesting contact phenomenon, since Vend Romani also has the productive inherited iterative suffixes -(*in*)kér- and -(*in*)gér-. Nonetheless, the suffixed verbs occur more frequently than the complex 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.

The thesis compares the Romani verbal particles in Somogy and Zala Romani as well as those in the related varieties of Burgenland and Prekmurje Romani. It is shown that Somogy Romani has a set of verbal particles that are very similar to Burgenland Romani; and vice versa, the verbal particles in Zala Romani are highly similar to those that occur in Prekmurje Romani. 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, 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.

Chapter 5 – Syntax – introduces the word order and the basic and complex syntactic structures.

Vend Romani has several adpositions for expressing various spatial relations. Vend Romani differs from other South Central Romani varieties in that the inessive preposition *ande* 'in, to' is often omitted before proper names of localities. In other words, these localisations are zero marked in Vend Romani. For instance, the noun *Kišbajom* 'Kisbajom (proper name)' may have the following spatial meanings: the inessive stative 'in Kisbajom' and the inessive directive 'to Kisbajom'. The prepositional phrase with *ande* 'in, to' was more favoured in the elicited data, while the unmarked expression was preferred in the spontaneous language data.

Vend Romani borrowed the adversative coordinators de (< H de) 'but' and ham (< H hanem) 'but' from Hungarian. They are inserted between the coordinands (as called in Haspelmath 2007). The adversative de 'but' is generally used after the first coordinated expression with affirmative meaning, while the substitutive hanem is solely reserved for coordinations in which the first coordinated expression is negated. These coordinators are homonymous with the frequently used connective markers de and ham. In this function, de is interchangeable with ham. The use of ham as a connective marker is an innovation in Vend Romani, since the corresponding Hungarian marker hanem does not have such a function.

Chapter 6 – Lexicon – analyses the basic lexicon of Vend Romani with a focus on the post-Greek borrowings.

The core vocabulary of Vend Romani is of Indo-Aryan origin. There are also a number of lexical roots of Greek, and some others of Iranian and Armenian origin. These roots are to a varying extent shared by all present-day Romani dialects. Vend Romani has been further shaped by its recent contact languages: Slavic, German and Hungarian.

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.

The most interesting development in Vend Romani is that the kinship terms **rom* 'husband' and **romni* 'wife' were replaced by *murš* (original meaning 'man') and *žuvli* (original meaning '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'. The terminology used for human beings include *murš* and *mánuš* meaning 'man', and *žuvli, lumni* and *manušni* meaning 'woman'. The terms denoting non-

Romani ethnicity are $g\dot{a}zo / g\dot{o}ri$ 'non-Romani man' and $g\dot{a}zi$ '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. 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'. On the other hand, the ethnic identity is still encoded on the original terms *ráklo* 'non-Romani boy' and *rákli* 'non-Romani girl'. To summarize, the Vend Romani terms originally referring only to Roma have been replaced by an ethnic-indifferent term (*čhá* 'Romani boy' > *fat'ú* 'boy'; **romni* 'Romani woman' > *lumni* ~ *žuvli* 'woman'), or through the change of their semantic value they have become ethnic-indifferent (*čhaj* *'Romani girl' > 'girl'). By contrast, the terms encoding non-Romani ethnic belonging have been preserved together with their semantic value.

References

- 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.
- 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 internazionalen Kontext. Klagenfurt: Drava. 277–390.
- Elšík, Viktor. 2008–. 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.
- Imre, Samu. 1971. A felsőőri nyelvjárás. Budapest: Akadémia kiadó.
- Matras, Yaron. 2002. *Romani: A Linguistic Introduction*. Cambridge: Cambridge University Press.
- Morgan, David L. 2008. Snowball Sampling. In: Given, Lisa M. (ed.) The SAGE Encyclopedia of Qualitative Research Methods. Thousand oaks: Sage publications, 815–816
- UNESCO. 2003. Language vitality and endangerment. Ad Hoc Expert Group on Endangered Languages. Paris: UNESCO
- Vekerdi, József. 1984. The Vend Gypsy dialect in Hungary. Acta Linguistica Academiae Scientiarum Hungaricae, Volume 34, 65–86.