

# The nominal morphology of Lovari from an analogical perspective

**Márton András Baló**

Research Institute for Linguistics,  
Hungarian Academy of Sciences  
baloam@gmail.com

**Abstract:** The paper attempts to provide evidence that analogy-based approaches make language change, as well as unstable and variegated forms and word classes easier to grasp than they would be in a traditional synchronic framework or through a rule-based diachronic analysis. The example is the declension of the Lovari dialect of Romani, a dialectally most diverse Indo-European language that is often exposed to contact-related influences. A unique feature of Romani, the strict split between the morphology of inherited and borrowed vocabulary is seen in a new light if we examine the possible analogical processes behind the apparent erosion of this system, and the seemingly high number of inflectional nominal paradigms can be reduced to just two.

**Keywords:** analogy; nouns; paradigms; inflection; Romani

## 1. Introduction

The Romani language, due to the circumstances in which it is used, frequently borrows lexical items. Borrowing mostly happens from the local language, that is, the one spoken by the majority society surrounding the given Roma community and has been particularly intensive since the appearance of the Roma in Byzantium and their dispersal in Europe. Romani is originally divided into dialects on a geographical basis (cf. e.g., Matras 2005; Bakker & Matras 1997; Miklosich 1872–1880).

The dialects established in that manner are split into further varieties due to further migration; thus, for instance, whereas Lovari was originally spoken in western Romania, it is possible to talk about Hungarian and Austrian Lovari, which coexist and interfere with the Romungro and the Burgenland Romani varieties, respectively (both belong to the Central dialect group as opposed to Lovari, which is a member of the Vlax dialects). There can be so many differences between two, fairly distinct dialects that the speakers might switch to a language that they both speak alongside Romani (Boretzky 1995).



**Figure 1:** The estimated areas where Lovari and the Central dialects are spoken (cf. the project “The Linguistic Atlas of Central Romani” and Matras 2005)

In addition to the area marked in Figure 1, there are Romani communities in North and South America and even Australia (cf. Fraser 1992; Salo & Salo 1985), including groups speaking Lovari, where they arrived via Western-Europe as a result of a second wave of migration from Eastern-Europe (Romania, Serbia, Hungary, Slovakia) in the second half of the 19th century, following the abolition of slavery in Romania. This continued in the 1990s, after the fall of communism. The linguistic situation in these communities is beyond the limits of the present study.

In the Romani lexicon and declension, there is a very clear-cut distinction between inherited and borrowed vocabulary in terms of morphological patterns and paradigms. No morphophonological reason exists which would justify the difference between the oblique stem of *āro* ‘flour’, which is *āres-* and the oblique stem of *fōro* ‘town’, which is *fōros-*.<sup>1</sup> There is, however, an

<sup>1</sup> This unique feature is called “thematicity” in the terminology of Romani linguistics. The term itself is borrowed from Indo-European linguistics but with an unrelated meaning: in Romani linguistics, it does not refer to the presence or absence of a thematic vowel or consonant; it simply refers to the difference between the ways inherited

ostensible blurring or mixing of the two different inflectional patterns, as noted most prominently by Elšík (2000), and the anomalies cannot possibly be explained with a traditional rule-based approach. There are several references to analogy in Matras (2002) and in Elšík (2000), but only as a secondary phenomenon. However, if we examine the blurring from a primarily analogical perspective, we can see how language changes in real time and how categories based on earlier historical developments become obscure or fuzzy.

The erosion of the distinction between the inherited and the borrowed parts of the lexicon is even more interesting if we take into consideration that this has been one of the few but very solid features that all dialects have shared. We will have a closer look at these processes through the example of the Lovari dialect as spoken in Hungary, while giving examples from other dialects as well.

## 2. Analogy

The term *analogy* will be used in the broad, Saussurean sense throughout the paper: “an analogical form is a form made on the model of one or more other forms” (Saussure [1916] 1966, 161). His classic example is the eventual spread of the rhotacised oblique onto the nominative in Latin (*idem.*):

- (1)  $\bar{o}r\bar{a}t\bar{o}rem : \bar{o}r\bar{a}tor = hon\bar{o}rem : x$   
 $x = honor$

Here, the former nominative singular *honōs* is replaced by a more regular form, *honor*. This pattern, also called four-part or proportional analogy (cf. e.g., Kraska-Szlenk 2007) is the very pattern we encounter in Romani declension, where forms such as *fōros-* are replaced by the more regular and, strangely enough, also more conservative *fōres-*.

Rung (2011) gives a very detailed overview of analogy-based approaches. He notes that structural linguists (Sapir 1921 and Bloomfield 1933 in particular) maintained that analogy had a great significance: new utterances are created based on an analogy with previously uttered or heard words and sentences. In other words, patterns and exemplars, already existing in our minds, serve as bases for new forms or old ones undergoing some sort of change. The governing forces of language may

---

and borrowed lexical items are treated morphologically. The former is referred to as “thematic”, while the latter is called “athematic”.

thus be seen as surface patterns, or, in other words, “closely related surface forms” (Bybee 1985, 49–50): constructions or abstract schemas, which are complex instances of form and meaning (Goldberg 1995; Booij 2010), similarly to the notion of sign taken in the original, Saussurean sense: “the linguistic sign unites [...] a concept and a sound-image” (Saussure [1916] 1966, 66). While constructional approaches maintain the creative aspect of language, they “generally recognise that grammars don’t generate sentences, speakers do” (Goldberg 2006, 22). The same is true for word formation, where “patterns can be seen as abstract schemas that generalize over sets of existing complex words” (Booij 2007, 34). But patterns do not only exist in syntax and word formation; they are also present in inflectional paradigms, as can be seen in Romani.

Paradigms, that is, a set of forms belonging to the same lexeme (cf. Wurzel 1989, 52), form an important part of analogical theories as the similarity of combinations of form and function is a significant characteristic of paradigms, and analogy maintains paradigmatic uniformity (Albright 2009; Eddington 2006). Similarity in grammatical function involves similarity in form, or, in other words, “it is natural for related concepts to be designated by related sounds” (Humboldt 1999, 71). Similarity can be measured in terms of the surface forms (the actual identity of phoneme sequences) or based on the extent to which the defined features of words are alike (Rung 2011).

Analogy is closely related to the concept of patterns and (ir)regularity in that it “supposes a model and its regular imitation” (Saussure [1916] 1966, 161) and these models, or patterns, can also be seen as constructions. Analogy rests on statistical evidence; analogical force depends on the frequency of the pattern in question. A pattern with higher type or token frequency is more powerful, and competing patterns result in instability. On the other hand, less frequent forms are more prone to undergo analogical change. We can also say that analogical change depends on the productivity of the given pattern. Productivity, in turn, is determined by frequency. Bybee (1995) argues that high type frequency is more important based on research showing that innovative forms are based on word classes which are comprised of a large number of members. Forms of high token frequency, on the other hand, remain autonomous and unanalysed.

The variation caused by competing patterns is further enhanced by the diversity of dialects, the lack of a written standard and frequent borrowing in Romani.

If we dismiss the dichotomy of underlying and surface representations and do not try to force one and single origin on the different surface forms,

considering the surface forms themselves instead, the variation within a paradigmatic schema will be much less problematic and exceptions can easily be integrated in the system. In an analogy-based framework, we can disregard another dichotomy, namely the opposition between the diachronic and the synchronic approach to language, as language change, which is usually part of diachronic descriptions, is palpable in synchronic terms through the unstable and variegated forms. We can also incorporate the notion of rules, although not taken in the generative sense, but following van Marle (1990), who bears upon analogy as a synchronic force and argues that “the speakers of a language have the capacity to construct rules on the basis of the existing words” (*ibid.*, 267), called “rule-creating creativity”.

A practical model for measuring analogical effects is Analogical Modeling of Language (cf. Skousen 1989; 1992; Skousen et al. 2002), where patterns are represented by a dataset of exemplars, that is, a set of words whose similarity to the form in question is strictly predefined. The exemplars are arranged into supracontexts based on the distribution of previously chosen variables. Homogeneous supracontexts form the analogical set which is used to predict the outcome.

### 3. A brief overview of inflection in Romani

Romani is considered an agglutinative language with a moderate but fairly strict case system, originally inherited from Indo-Aryan, later lost and regained. In order to provide a brief outline of the declension, let us take two words, one masculine and one feminine (Table 1).

From a descriptive aspect, the nominative, the accusative and the vocative forms are unique (the other forms are ostensibly attached to the accusative form), see Table 2.

This is justified by the historical evolution of Romani declension, too. The nominative, accusative and vocative cases are preserved, whereas the other five cases (dative, locative, ablative, instrumental, genitive) were lost. They were replaced by postpositions, which were eventually re-analysed as proper case suffixes (Friedman 1991a;b). These case suffixes are uniform all over the system, irrespective of gender. The accusative and vocative forms are regular and come by default, the former depending on the gender. The only variables are the nominative endings, but the plural ending can easily be derived from the singular ending. Otherwise, Romani is a prepositional language, as illustrated in Table 3 below.

**Table 1:** A brief outline of Romani declension

|     | kermó m. 'worm' |           | čirikli f. 'bird' |             |
|-----|-----------------|-----------|-------------------|-------------|
|     | Singular        | Plural    | Singular          | Plural      |
| N   | kermó           | kermé     | čiriklí           | čiriklá     |
| A   | kermés          | kermén    | čiriklá           | čiriklán    |
| D   | kerméske        | kerméngé  | čirikláke         | čiriklángé  |
| L   | kerméste        | kerménde  | čirikláte         | čiriklánde  |
| Abl | kerméstar       | kerméndar | čiriklátar        | čiriklándar |
| I   | kermésa         | kerménca  | čiriklása         | čirikláncia |
| G   | kermésk-        | kerméng-  | čiriklák-         | čirikláng-  |
| V   | kermá           | kermále   | čiriklá           | čiriklále   |

**Table 2:** The nominative, accusative and vocative forms

|            | Masculine singular | Masculine plural | Feminine singular | Feminine plural |
|------------|--------------------|------------------|-------------------|-----------------|
| Nominative | kerm- + -o         | kerm- + -e       | čirikl- + -i      | čirikl- + -a    |
| Accusative | kerm- + -es        | kerm- + -en      | čirikl- + -a      | čirikl- + -an   |
| Vocative   | kerm- + -a         | kerm- + -ale     | čirikl- + -a      | čirikl + -ale   |

**Table 3:** Prepositions in Romani

|                             |              |                                    |
|-----------------------------|--------------|------------------------------------|
| angla 'before, in front of' |              | anglo kher 'in front of the house' |
| andar 'from'                | kher 'house' | andar o kher 'from the house'      |
| tela 'under'                | o 'the'      | telo kašt 'under the tree'         |
| pe 'on'                     | kašt 'tree'  | po kašt 'in the tree'              |

#### 4. The analysis of Romani inflection

Matras (2002) presents an analytical model of Romani nouns, in which the surface form of an inflected noun is assumed to consist of layers, similarly to Indo-Aryan as described by Masica (1991) or the blocks of realisation rules in Paradigm Function Morphology (Stump 2001). We will see, however, that these layers are simply unnecessary for Romani.

Three layers are posited in this framework, namely Layer I, composed of nominal and oblique endings,<sup>2</sup> Layer II, which comprises the actual case endings, and a set of adpositions named Layer III. Layer I endings are attached directly to the nominal base. The declension class of a noun can traditionally be seen from the Layer I oblique ending, which is in turn determined by several factors to be discussed below, but as we will see, many questions arise related to these factors as well as the different declension classes. Layer II endings are case suffixes attached to the Layer I form of a noun. All in all, Romani distinguishes eight different cases: nominative, accusative, dative, genitive, ablative, locative, instrumental and vocative. Most of the case suffixes are fixed in form (although they are subject to some variation among the dialects), showing only voice assimilation, and are added to the oblique stem.

**Table 4:** The general layout of a Romani noun

|              |      |                |
|--------------|------|----------------|
| Lemma        | ānró | ‘egg’          |
| Nominal base | ānr- | stem           |
| Layer I      | -és- | oblique marker |
| Layer II     | -te  | locative       |
| Layer III    | ande | locative?      |

As we can see from the example, the terms used are ambiguous. The use of the term “nominal” is redundant if we say that there is no other stem. That is what we apparently see, as both the nominal and the oblique (and the vocative, for that matter) endings attach to this. Strictly speaking, the oblique “stem” is not a stem, but it is derived from the nominal base.

- (2) *bakr-* + *-o* → nom. *bakró* ‘sheep’  
*bakr-* + *-es* → obl. *bakrés-* ‘sheep’  
*bakr-* + *-a* → voc. *bakrá* ‘sheep’

It would therefore be sufficient to posit one single stem which serves as the basis for all other forms of the given noun. Elšík (2000), on the other hand, proposes to differentiate between BSA (base-stem affixation) and OSA (oblique-stem affixation) languages. The former refers to languages where the cases are marked with individual suffixes; the latter means that

<sup>2</sup> For the sake of simplicity, I will adopt the terms “nominal” and “oblique” used in Romani linguistics in the paper, although they might as well just be dubbed “short” and “long” stems.

the case suffixes are attached to an oblique stem. Romani belongs to the OSA languages, as opposed to Hungarian, for example, where case suffixes directly follow the stem, i.e., the nominative form, without mediation (nom. *bárány* ‘sheep’ → loc. *bárányban* ‘sheep’). If we treat the nominative and the vocative independently, this could indeed be a possible analysis. Blake (2000; 2001) also make reference to an oblique stem “which serves to set the nominative off from the other cases” (Blake 2001, 42). A similar example is Lezgian (Blake 2000 based on Mel’čuk 1986), where the bare oblique stem functions as the ergative case. Elšík (2000) also mentions Daghestanian languages, where the ergative case is unmarked, similarly to Romani, where the unmarked case is the accusative, which is derived from the oblique stem by an identity process. As for the vocative, Matras (2002) notes that the vocative forms can be found “alongside” the three layers and “connect directly to the nominal base” (*ibid.*, 80). But even then, the nominative-oblique dichotomy remains.

From a synchronic perspective, it is more economical to say that there are two bases: the nominal base, marked by a zero morpheme: *bakr-* +  $\emptyset$ , and an oblique base, marked by a suffix: *bakr-* + *es* and *sokr* + *os*. These could then in turn serve as bases for the nominative and the vocative on the one hand, and the rest of the cases on the other. Thus, we could treat all the cases in the same way, even if historically the vocative is of different origin, and we would not have to deal with all the diverse nominative endings as Layer I elements.

**Table 5:** Singular case markers in Lovari

| Base    | Case       | Form                       | Base    | Case         | Form |
|---------|------------|----------------------------|---------|--------------|------|
| nominal | nominative | - $\emptyset$ , -o, -i, -a | oblique | locative     | -te  |
| nominal | vocative   | -a                         | oblique | ablative     | -tar |
| oblique | accusative | - $\emptyset$              | oblique | instrumental | -sa  |
| oblique | dative     | -ke                        | oblique | genitive     | -k-  |

The exact status of the Romani genitive (it behaves like an adjective in many aspects) is not relevant here but it is discussed in detail, among others, by Elšík (1997); Grumet (1985a) and Koptjevskaja-Tamm (2000).

The term Layer III is also misleading, and its use may only be justified by historical reasons. Layer III in Romani derives from Indo-Aryan Layer III indeed, but while the latter is “**potentially** mediated” (Masica 1991, 234) by a Layer II element, the former has a “preposed position”



(Matras 2002, 80), and so it has no contact with the other layers. In actual fact, it consists of prepositions, which are used completely independently of Layers I and II and by now they are always followed by the nominative: *ando kher* ‘in the house’.<sup>3</sup> For the purposes of our analysis, they do not need to be considered as an inflectional layer.

**Table 6:** The current state of Layer III in Lovari

| Preposition | Definite article | Noun nom.    |
|-------------|------------------|--------------|
| ande ‘in’   | o                | kher ‘house’ |

We have now arrived at the conclusion that – if we disregard the concept of layers derived from the Indo-Aryan heritage for a moment – it looks economical to say that there are two base forms and a set of suffixes, some of which attach to the nominal base, while others attach to the oblique base. This is in line with the notion of stem space as described by Bonami & Boyé (2006). There is one single lexeme equipped with a stem space with two slots.

We will now turn to the factors determining which declension class a given noun belongs to. These factors are mostly hypothetical and have little to do with what information native speakers might store in their minds (cf. Blevins & Blevins 2009b). If one is aware of all the background information concerning a word, one can inflect the noun according to the “rules”, but if only the surface similarities and differences are considered (cf. e.g., Kálmán et al. 2012), the variation and the erosion of the inherited–borrowed dichotomy can be explained more easily.

Sixteen declension classes are listed for Hungarian Lovari in Hutterer & Mészáros (1967) based on gender, animacy, the nominative ending and the inherited or borrowed nature of the word. Some grammars, e.g., Matras (2002), consider palatalisation as an additional, separate criterion. The high number of classes serves as a motivation for the revision of these criteria, after which we will see that the information regarding animacy

<sup>3</sup> With regard to Austrian Lovari, we find the form *ande bute beršende* ‘in many years’ in Cech & Heinschink (1999a), which testifies the diachronic development of the layers and the existence of variation among different varieties of the same dialect. They also note the existence of the form *ande bute beršen*; here, the disappearance of the locative case marker points towards the erosion of redundancy finally achieved in Hungarian Lovari, where the use of the nominative case after Layer III adpositions eliminates the “double” locative.

and the nominative ending is redundant, and the declension is determined by grammatical gender, and, to a lesser extent, the above-mentioned inherited–borrowed distinction. Tables 7 and 8 below show the reduced system.

**Table 7:** Masculine declension in Lovari

| Masculine | bakró ‘sheep’<br>(inherited) |           | sóкро ‘father-in-law’<br>(borrowed) |           |
|-----------|------------------------------|-----------|-------------------------------------|-----------|
|           | Singular                     | Plural    | Singular                            | Plural    |
| N         | bakró                        | bakré     | sóкро                               | sóкруrá   |
| A         | bakrés                       | bakrén    | sóкρός                              | sóкρόн    |
| D         | bakréske                     | bakrénge  | sóкρόςke                            | sóкρόnge  |
| L         | bakréste                     | bakrénde  | sóкρόste                            | sóкρόnde  |
| Abl       | bakréstar                    | bakréndar | sóкρόstar                           | sóкρόndar |
| I         | bakrésa                      | bakrénca  | sóкρόςa                             | sóкρόnca  |
| G         | bakrésk-                     | bakréng-  | sóкρόςk-                            | sóкróng-  |
| V         | bakrá                        | bakrále   | sóкρά                               | sóкrále   |

**Table 8:** Feminine declension in Lovari

| Feminine | kirí ‘ant’<br>(inherited) |          | rāca ‘duck’<br>(borrowed) |          |
|----------|---------------------------|----------|---------------------------|----------|
|          | Singular                  | Plural   | Singular                  | Plural   |
| N        | kirí                      | kirá     | rāca                      | rācí     |
| A        | kirá                      | kirán    | rācá                      | rācán    |
| D        | kiráke                    | kiránge  | rācáke                    | rācánge  |
| L        | kiráte                    | kiránde  | rācáte                    | rācánde  |
| Abl      | kirátar                   | kirándar | rācátar                   | rācándar |
| I        | kirása                    | kiránca  | rācása                    | rācánca  |
| G        | kirák-                    | kiráng-  | rācák-                    | rācáng-  |
| V        | kirá                      | királe   | rācá                      | rācále   |

With regard to the paradigms, that is, the forms in the cells, the nominative ending is not important, as it is always dropped, be it a vowel, as can be seen in the tables, or a zero morpheme, and replaced by the oblique ending; thus nom. sing. *manúš* ‘man’ → obl. *manušés-/manušén-*, following

the pattern of *bakró* ‘sheep’, and nom. sing. *phen* ‘sister’ → obl. *pheňá-/pheňán-*, following the pattern of *kirí* ‘ant’.

The forms in the dark grey cells depend on the animacy status of nouns. The accusative form depends on the animacy of the noun: that of inanimate nouns is syncretic with the nominative (nom. and acc. *kher* ‘house’). The accusative form of animate nouns is the oblique stem (nom. *bāló* ‘pig’ and acc. *bālés*). The ending of the oblique stem does not depend on the animacy status of the given noun. The animacy status of a noun follows from a possible animacy hierarchy (Matras 2002 based on Holzinger 1993 and Hancock 1995) among the nouns, with humans at one end and body parts at the other. Lovari maintains the broadest animacy split, with body parts inflected as animates as well (nom. *vast* ‘hand’ → acc. *vastés*, as opposed to Romungro, a Central dialect spoken in Hungary: nom. *va* → acc. *va*).<sup>4</sup>

Elšík (2000) gives a thorough account of the historical development of Romani nominal paradigms, where he notes that “two criteria are general for all nouns: the gender, and the shape of the base-form suffix” (*op. cit.*, 14). Hutterer and Mészáros (1967) also claim that the classification of a noun relies heavily on three factors: the meaning of the word (natural gender), the nominative ending (grammatical gender) or the meaning and the ending together. The first factor may be easily dismissed and replaced overall by grammatical gender, which is indeed a determining factor, as we can see from the tables; this is only complicated by the inherited–borrowed dichotomy in the masculine, as can be seen in Table 4.

- (3) a. *ānró* m. ‘egg’ → obl. sing. stem *ānrés-*  
 b. *coló* m. ‘blanket’ → obl. sing. stem *colós-*<sup>5</sup>

<sup>4</sup> There have been attempts to formalise this hierarchy. Matras (2002) explains the role of the oblique stem as the marker of the accusative and other functions (the “Independent Oblique”) in terms of topicality. Elšík (2000) claims that the direct object in case of inanimate nouns is expressed by the nominative form and suggests to treat the subject/direct object split as “hyper-paradigmatic”. Thus, there would be no need to postulate separate paradigms based on the animate-inanimate distinction. This idea leaves the choice to pragmatic aspects.

<sup>5</sup> The data presented in the paper mostly come from reliable written sources, which are free from the desire to codify or unify the language (Cech & Heinschink 1999a and b; Hutterer & Mészáros 1967; Vekardi 2000), but all the data were also confirmed and attested by the informant I worked with while writing the paper: Mária Nagy, an elderly, less educated lady, originally from Nagykálló in the north-eastern part of Hungary, currently residing in Budapest.

Gender neutralisation appears in many cases in the nominative; therefore, the nominative endings cannot be considered a determining factor of the declension of a given noun. The consonantal stems or, in other words, the stems ending in a zero morpheme, and words ending in *-i* may either be masculine or feminine. Again, the latter has diachronic reasons: it either goes back to the Proto-Romani inheritance (Elšík 2000) or later loanword adaptation in Lovari using an *-i* marker (Hutterer & Mészáros 1967). However, the declension of these loanwords has always been ambiguous. The words derived by the borrowed agentive suffix *-āri*, also inflect according to the inherited pattern.

- (4) a. *phrāl* m. ‘brother’ → obl. sing. stem *phrālés-*  
 b. *phen* f. ‘sister’ → obl. sing. stem *pheňá-*  
 c. *juhási* m. ‘shepherd’ → obl. sing. stem *juhāsós-*  
 d. *vitézi* m. ‘brave warrior’ → obl. sing. stem *vitézés-*  
 e. *romňí* f. ‘woman’ → obl. sing. stem *romňá-*  
 f. *butjári* m. ‘worker’ → obl. sing. stem *butjāres-*

In Early Romani, there was a distinction between two different kinds of consonant-final masculine classes, whose nominative plural forms were different: *kher* ‘house’ → nom. pl. *kherá*, as opposed to *vast* ‘hand’ → nom. pl. *vast*. This identical plural form has disappeared by now from Lovari, and we find that the plural of *vast* is *vastá*.

Nouns ending in *-o* are exclusively masculine, nouns ending in *-a* are exclusively feminine.<sup>6</sup> The former group is split again as for the oblique stem, because there are inherited and borrowed items among them. The masculine nouns ending in a consonant contain a special subgroup of “abstract nouns, which are characterized by a specific derivational suffix” (Elšík 2000), *-ipen*. Matras (2002) notes that its formants treasure an old oblique form *-ip(e)nas* in many Romani dialects instead of the expected *\*-ip(e)nes*. He adds that the form nevertheless appears in Burgenland Romani “by analogy to the general masculine oblique formation” (Matras 2002, 84; and see also Halwachs 1998) in the form *-ipes*. However, *-ipes* is probably not an alternative formant of *\*-ip(e)nes*. Rather, the loss of the final nasal in Burgenland Roman as well as Lovari and more generally south of the Great Divide (a bundle of isoglosses in central Europe, for more detail cf. Matras 2005) resulted in the surface form *-ipe* (the abstract

<sup>6</sup> Not even this is completely unambiguous as Elšík (2000) mentions Romungro loanwords borrowed from Hungarian such as *komuništa* ‘communist’, which is masculine.

nouns thus becoming the only group with an *-e* as the nominative singular ending), to which the old oblique could not be applied, and so the general, inherited oblique spread onto the abstract nouns containing this particular suffix.

**Table 9:** The pattern of nominative endings in Lovari after the historical developments described by Elšík (2000)

| Ending           | ∅    |     | -o   |           | -i   |           | -e  | -a   |     |
|------------------|------|-----|------|-----------|------|-----------|-----|------|-----|
|                  | m    | f   | m    |           | m    | f         | m   | f    |     |
| Inherited status | +    | +   | +    | -         | +    | -         | +   | +    | -   |
| Oblique singular | -es- | -a- | -es- | -os-/-es- | -es- | -os-/-es- | -a- | -es- | -a- |

There are two important conclusions we can draw from the table: 1. the inherited or borrowed status of feminine nouns is irrelevant with regard to their declension; 2. there is variation in the oblique form of borrowed masculine nouns.

The small number of loans ending in *-u* resulted in the change of the final vowel to *-o* (e.g., original *pápu* ‘grandfather’ → *pápo*) in some varieties. A similar phenomenon can be seen in the Lovari verbal system where verb stems ending in an *-u-*, which are exclusively made up loan verbs, are prone to losing their category and are recategorised as *-i-* stem verbs, which constitute the bigger class of loan verbs (Baló 2011; 2012). The analogical effect based on frequency is conspicuous here and it heavily affects the low number of stems containing an *-u-* all over the Lovari morphology.

A very interesting dichotomy has existed within inherited feminine nouns ending in a consonant, namely that some of them are palatalised<sup>7</sup> in the plural and in the oblique cases, while others are not. The examples below are taken from Elšík (2000).

- (5) a. *žuv* ‘louse’ → obl. sing. *žuvá-*  
 b. *suv* ‘needle’ → obl. sing. *suvjá-*  
 c. *pīri* ‘pot’ → obl. sing. *pīrjá-*

According to Elšík (2000) and Matras (2002), this is of Proto-Romani origin and a result of the infiltration of palatalised forms from other feminine paradigms. We can see double neutralisation here, between the nominative forms like *žuv* and *suv* on the one hand, and between the oblique forms

<sup>7</sup> This phenomenon is called “jotation” in Romani linguistics.

*svjǎ-* and *pǎrǎjǎ-* on the other. Elšík (2000) suggests that feminine nouns jotated in the oblique constitute a mixed class, where the nominative form resembles that of the consonant-final feminines, while the other forms are taken from feminines with a stem-final *-i*, where jotation is obligatory. Matras (2002) adds that “with pre-European feminines ending in a consonant, jotation is analogous, and hence often irregular” (*op.cit.*, 83). Regularity is in the process of being reinstated, however; Elšík (2000) also adds that there is a tendency of de-jotation so as to avoid mixed paradigms, and that is why the originally jotated oblique singular form *svjǎ-* ‘needle’ becomes *suvǎ-*, at least in certain dialects. This is justified by newly collected data, which proves that the process has not stopped, and the words where jotation could easily be triggered by the stem-final *-i* lose the palatalisation: oblique singular *pǎrǎ-* instead of *pǎrǎjǎ-*. This regularisation process consists of a simple analogy again, where the unpalatalised forms spread all over the feminine, except for words where the stem-final *-i* is preceded by a palatalised consonant, like, for example, *romnǎjǎ* ‘woman’, *raklǎjǎ* ‘non-Roma girl’, *anǎrustǎjǎ* ‘ring’, *brǎdǎjǎ* ‘bucket’.<sup>8</sup>

## 5. The inherited–borrowed dichotomy

This peculiar feature, which was first described in detail by Boretzky (1989), is present in several areas of Romani grammar. It is characterised by an ostensible, morphologically marked difference between the inherited and the borrowed vocabulary. In addition to that, even linguistic categories and syntactic structures have been borrowed, for example the prepositional nature, mentioned above, or comparison. For the sake of the present paper, we will focus on the phenomenon as it appears in the nominal inflection.

The difference between the inherited and borrowed inflectional patterns is most conspicuous in the masculine oblique. Diachronically, even the borrowing pattern was borrowed from Greek (Bakker 1997), and many of the first borrowings also came from Greek.<sup>9</sup> Therefore, the declension of borrowed words in Early Romani showed a variety of oblique affixes,

<sup>8</sup> Palatalisation in this position seems to be very common for the alveolar stops, nasal and approximant.

<sup>9</sup> Borrowing on the structural level is not rare in Romani: similarly to the borrowing pattern which was borrowed from Greek, there is a plural marker borrowed from Romanian, namely *-uri*, which often appears in the form *-ura* and attaches to borrowed nouns, but never to inherited ones: *juhǎsi* ‘shepherd’ → pl. *juhǎsurǎ*, *sókro* ‘father-in-law’ → pl. *sokrurǎ*.

as far as the vowel is concerned: *-os*, *-is*, *-us* (Elšík 2000).<sup>10</sup> The most obvious effect of analogy is the almost complete disappearance of the different vowels in Lovari (e.g., *sápuj* ‘soap’ → obl. sing. *sapujós-* from Early Romani \**sápu*(*s*) → obl. sing. \**sapunís-*), preserving only the oblique in *-os*, which has thus become the only oblique affix of borrowed words.<sup>11</sup> The literature regarding the *-is* ending is slightly contradictory. Elšík (2000) considers the oblique singular in *-is* to be “a later development” at one point (*op.cit.*, 18), replacing the originally inherited form \**sapunés-*. On the other hand, he also says (*ibid.*, 23) that, at a later stage in the history of Romani declension, the original *-is* ending was replaced by *-os*, e.g., *doktóri* ‘doctor’ → obl. sing. *doktorós-*, instead of the original \**doktorís-*. This only justifies that “forms are [...] preserved because they are constantly renewed by analogy” (Saussure [1916] 1966, 172).

Loanword markers, such as *-i* and *-o* can be used simultaneously, thus *tudóšo* ‘scientist’ may coexist with *tudóši*, from Hungarian *tudós*. The oblique endings may vary, too, but not necessarily related to the nominal ending: *tudóšo/tudóši* → obl. sing. *tudóšós-* as well as *tudóšo/tudóši* → obl. sing. *tudóšés-*.

As a well-known example, let us take a look at the case of the Greek-derived word *fóro* ‘town’. We learn that the forms of both the singular and the plural oblique stems are ambiguous: they may be *fórés-/fóros-* and *fóren-/fóron-*, respectively. As Elšík (2000) states, diachronically, *fóros-* replaced *fórés-*, so that the oblique form could resemble the nominative singular. This process, however, goes against the basic layout of the inherited inflection, where the oblique singular stem ends in *-es-*, no matter what the nominative ending is (for example nominative singular *báló* ‘pig’ and oblique singular stem *báles-*). The case is more likely to be that the loss of the word-final consonant resulted in a form similar to many inherited nouns, and the oblique form is slowly taking on the inherited pattern, too – or at least re-acquiring it. This is what the Saussurean analogical proportion applied to the Lovari oblique stem looks like:

<sup>10</sup> Elšík (2000) also adds that the marker has thus lost its monomorphemic nature. However, as Baló (2012) notes with regard to the Lovari verbal system, the bimorphemic or monomorphemic nature of these markers becomes insignificant if we look at the analogy-based processes which have taken place and are taking place in the language. For instance, in Estonian where the partitive singular *lukku* of *lukk* ‘lock’ implies the short illative singular *lukku*, alongside *lukusse*, “even though neither *lukk* nor *-u* can be associated with the grammatical meaning “partitive” or “illative” (Ackerman et al. 2009, 56). Their inference is that the deduction of new forms is facilitated by the knowledge of other forms.

<sup>11</sup> Although see the remarks regarding words with a stem-final *-u*.

- (6) *báló* : *bálás-* = *főro* : *x*  
*x* = *főrés-*

As for the oblique plural stem, Matras (2002) claims that *főrén-* became *főrón-*, possibly based on an analogy to the nominative singular. On the other hand, Hutterer and Mészáros (1967) mention that the original form of the oblique plural stem of *sókro* ‘father-in-law’, from Romanian *socru*, is *sokrón-*, but it appears more and more frequently in the form *sokrén-*, and even the oblique singular stem can be *sokrés-* instead of *sokrós-*, as attested by informants. A similar process to the singular might have taken place in the plural as well.

**Table 10:** The pattern of nominative endings in Lovari after the historical developments described by Elšik (2000)

| Marker                              | -es-                  | -os-                  | -en-                  | -on-                  | -a-                   | -an-                  |
|-------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Function                            | oblique singular      | oblique singular      | oblique plural        | oblique plural        | oblique singular      | oblique plural        |
| Possible gender                     | masculine             | masculine             | masculine             | masculine             | feminine              | feminine              |
| Inherited or borrowed?              | inherited<br>borrowed | inherited<br>borrowed | inherited<br>borrowed | inherited<br>borrowed | inherited<br>borrowed | inherited<br>borrowed |
| Possible nominative singular ending | ∅, -o, -i             | ∅, -o, -i             | ∅, -o, -i             | ∅, -o, -i             | ∅, -a, -i             | ∅, -a, -i             |
| Possible nominative plural ending   | -e, -a, ura           | -ura, -a              | -e, -a, ura           | -ura, -a              | -a, -i                | -a, -i                |

The feminine oblique plural was historically *-en-*, which is renewed in the Vlach dialects, possibly by analogy to the nominative plural, which ends in an *-a* (Matras 2002 based on Boretzky 1994). This is not generally true if we consider the data from Austrian Lovari; Cech and Heinschink (1999a) note that *-en-* is possible, too, for inherited words, and only *-en-* is possible for borrowed words. Considering this fact from a synchronic point of view, we might say that the *-en-* appears because it is “typical” of the oblique plural (cf. the masculine).



- (7) a. *romňí* ‘woman’ → obl. pl. *romňén/romňán*  
 b. *vórba* ‘word’ from Romanian *vorbă* → obl. pl. *vorbén*

If we look at the masculine now, we can see that it is also fairly uniform. Some anomaly only occurs among the nominative plural endings. Elšík (2000) mentions a similar anomaly with regard to Burgenland Roman: inherited and borrowed masculine nouns with a stem-final *-o* differ only in their oblique singular and nominative plural forms there and calls it “interaction between thematic and athematic classes” (*op.cit.*, 23). In Lovari, the situation is somewhat different: even if the nominative plural ending is *-ura* (e.g., *fōrurá*), the oblique endings may be either *-es-/-en-* or *-es-/-os-*.

## 6. Conclusion

In summary, we can say that basically there are two declension classes in Lovari: one masculine and one feminine. There is only one factor which slightly alters this: the inherited–borrowed dichotomy, described in detail by Boretzky (1989) and Bakker (1997) among others. However, as Elšík (2000) already notes, it is becoming blurred. Let us conclude by summing up the possible processes behind this phenomenon.

In the case of Romani, it is difficult to determine what the original forms of a certain word were exactly. What we can see here is that the forms *fōrén-*, *fōrés-*, *sokrén-* and *sokrés-* are in use, and they are spreading, weakening the role of the forms in *-os-/-on-*, which suggests that the inherited classes seem to exert an analogical force on the borrowings, at least as far as the masculine is concerned.

This can be related to the fact that many borrowings become obscure; for a bilingual speaker of Hungarian and Lovari, the words *tudóšo* and *juhási* might be transparent borrowings;<sup>12</sup> older borrowings, like *fōro* and *sókro* might become more integrated into the system. Generalisations may be made based on surface patterns that are “stronger” in some aspect; this might be the case for the historically deeper-rooted inherited pattern which is in constant opposition with the borrowed pattern, which is not as old but has become well-established due to the high degree of contact Romani has been exposed to.

<sup>12</sup> As one Romungro informant remarked to the author during his fieldwork, referring to a lexical item borrowed from Hungarian but inflected in Romani: “we can only say this in Hungarian”.

## References

- Ackerman, Farrell, James P. Blevins and Robert Malouf. 2009. Parts and wholes: Implicative patterns in inflectional paradigms. In Blevins & Blevins (2009a, 54–82).
- Albright, Adam. 2009. Modelling analogy as probabilistic grammar. In Blevins & Blevins (2009a, 185–213).
- Bakker, Peter. 1997. Athematic morphology in Romani: The borrowing of a borrowing pattern. In Matras et al. (1997, 1–21).
- Bakker, Peter and Yaron Matras. 1997. Introduction. In Matras et al. (1997, VII–XXX).
- Baló, András Márton. 2011. A current trend or a historic remnant? The case of a Lovari verb-forming suffix. *Poznań Studies in Contemporary Linguistics* 47. 264–282.
- Baló, András Márton. 2012. Arguments from lovari loan-verb adaptation for an analogy-based analysis of verbal systems. In Kiefer et al. (2012, 3–22).
- Blake, Barry J. 2000. Case. In G. Booij, C. Lehmann and J. Mugdan (eds.) *Morphology. An international handbook on inflection and word-formation*. Berlin & New York: Mouton de Gruyter. 1073–1090.
- Blake, Barry J. 2001. *Case*. Cambridge: Cambridge University Press.
- Blevins, James P. and Juliette Blevins (eds.). 2009a. *Analogy in grammar: Form and acquisition*. Oxford: Oxford University Press.
- Blevins, James P. and Juliette Blevins. 2009b. Introduction: Analogy in grammar. In Blevins & Blevins (2009a, 1–12).
- Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart and Winston.
- Bonami, Olivier and Gilles Boyé. 2006. Deriving inflectional irregularity. In S. Müller (ed.) *Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar, Varna*. Stanford: CSLI Publications. 361–380.
- Booij, Geert. 2007. Construction morphology and the lexicon. In F. Montermini, G. Boyé and N. Hathout (eds.) *Selected proceedings of the 5th Décembrettes: Morphology in Toulouse*. Somerville: Cascadilla Proceedings Project. 34–44.
- Booij, Geert. 2010. *Construction morphology*. Oxford: Oxford University Press.
- Boretzky, Norbert. 1989. Zum Interferenzverhalten des Romani (Verbreitete und ungewöhnliche Phänomene). *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42. 357–374.
- Boretzky, Norbert. 1994. *Romani. Grammatik des Kalderaš-Dialekts mit Texten und Glossar*. Wiesbaden: Harrassowitz.
- Boretzky, Norbert. 1995. Interdialectal interference in Romani. In Y. Matras (ed.) *Romani in contact. The history, structure and sociology of a language*. Amsterdam & Philadelphia: John Benjamins. 69–94.
- Bybee, Joan L. 1985. *Morphology. A study of the relation between meaning and form*. Amsterdam & Philadelphia: John Benjamins.
- Cech, Petra and Moses F. Heinschink. 1999a. *Basisgrammatik. Arbeitsbericht 1a des Projekts Kodifizierung der Romanes-Variante der österreichischen Lovara* (hrsgg. von Dieter W. Halwachs). Wien: Verein Romano Centro.
- Cech, Petra and Moses F. Heinschink. 1999b. *Wörterbuch. Arbeitsbericht 3 des Projekts Kodifizierung der Romanes-Variante der österreichischen Lovara* (hrsgg. von Dieter W. Halwachs). Wien: Verein Romano Centro.

- Eddington, David. 2006. Paradigm uniformity and analogy: The capitalistic versus militaristic debate. *International Journal of English Studies* 6. 1–18.
- Elšík, Viktor. 1997. Towards a morphology-based typology of Romani. In Matras et al. (1997, 23–59).
- Elšík, Viktor. 2000. Romani nominal paradigms: Their structure, diversity, and development. In Elšík & Matras (2000, 9–30).
- Elšík, Viktor and Yaron Matras (eds.). 2000. *Grammatical relations in Romani: The noun phrase*. Amsterdam & Philadelphia: John Benjamins.
- Fraser, Angus. 1992. *The Gypsies*. Oxford & Cambridge, MA: Blackwell.
- Friedman, Victor A. 1991a. Case in Romani: Old grammar and new affixes. *Journal of the Gypsy Lore Society* 5th series 1. 85–102.
- Friedman, Victor A. 1991b. Romani nominal inflection: Cases or postpositions? In M. Grochowski (ed.) *Problemy opisu gramatycznego języków słowiańskich*. Warsaw: Polska Akademia Nauk – Instytut Języka Polskiego. ??–?? PAGES??
- Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: The University of Chicago Press.
- Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Grumet, Joanne. 1985a. On the genitive in Romani. In Grumet (1985b, 84–90).
- Grumet, Joanne (ed.). 1985b. *Papers from the sixth and seventh annual meetings, Gypsy Lore Society, North American Chapter*. New York: Gypsy Lore Society.
- Halwachs, Dieter W. 1998. *Amaro vakeripe Roman hi – Unsere Sprache ist Roman: Texte, Glossar und Grammatik der burgenländischen Romani-Variante*. Klagenfurt: Drava.
- Hancock, Ian. 1995. *A handbook of Vlax Romani*. Columbus, OH: Slavica.
- Holzinger, Daniel. 1993. *Das Romanes: Grammatik und Diskursanalyse der Sprache der Sinte (Innsbrucker Beiträge zur Kulturwissenschaft 85)*. Innsbruck: Verlag des Instituts für Sprachwissenschaft der Universität Innsbruck.
- Humboldt, Wilhelm von. 1999. *On language. On the diversity of human language construction and its influence on the mental development of the human species*. Cambridge: Cambridge, MA.
- Hutterer, Miklós and György Mészáros. 1967. *A lovári cigány dialektus leíró nyelvtana [A descriptive grammar of the Lovari Gypsy dialect]*. Budapest: Magyar Nyelvtudományi Társaság.
- Kálmán, László, Péter Rebrus and Miklós Törkenczy. 2012. Possible and impossible variation in Hungarian. In Kiefer et al. (2012, 23–49).
- Kiefer, Ferenc, Péter Siptár and Mária Ladányi (eds.). 2012. *Current issues in morphological theory: (Ir)Regularity, analogy and frequency*. Amsterdam & Philadelphia: John Benjamins.
- Koptjevskaja-Tamm, Maria. 2000. Romani genitives in cross-linguistic perspective. In Elšík & Matras (2000, 123–149).
- Kraska-Szlenk, Iwona. 2007. *Analogy: The relation between lexicon and grammar*. Munich: Lincom Europa.
- Marle, Jaap van. 1990. Rule-creating creativity: Analogy as a synchronic morphological process. In W. U. Dressler, H. C. Luschützky, O. E. Pfeiffer and J. R. Rennison (eds.) *Contemporary morphology*. Berlin & New York: Mouton de Gruyter. 267–273.

- Masica, Colin P. 1991. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.
- Matras, Y., P. Bakker and H. Kyuchukov (eds.). 1997. *The typology and dialectology of Romani*. Amsterdam & Philadelphia: John Benjamins.
- Matras, Yaron. 2002. *Romani: A linguistic introduction*. Cambridge: Cambridge University Press.
- Matras, Yaron. 2005. The classification of Romani dialects: A geographic–historical perspective. In D. Halwachs and B. Schrammel (eds.) *General and applied Romani linguistics*. Munich: Lincom Europa. 7–26.
- Mel'čuk, Igor A. 1986. Toward a definition of case. In R. D. Brecht and J. S. Levine (eds.) *Case in Slavic*. Columbus, OH: Slavica. 35–85.
- Miklosich, Franz. 1872–1880. *Über die Mundarten und Wanderungen der Zigeuner Europas X–XII*. Wien: Karl Gerold's Sohn.
- Rung, András. 2011. *Magyar főnévi alaktani jelenségek analógiás megközelítésben* [Certain phenomena of Hungarian nominal morphology in an analogy-based framework]. Doctoral dissertation. Eötvös Loránd University, Budapest.
- Salo, Matt T. and Sheila Salo. 1985. Gypsy immigration to the United States. In Grumet (1985b, 85–96).
- Sapir, Edward. 1921. *Language*. New York: Harcourt, Brace and World.
- Saussure, Ferdinand de. [1916] 1966. *Course in general linguistics*. New York: McGraw-Hill.
- Skousen, Royal. 1989. *Analogical modeling of language*. Dordrecht: Kluwer.
- Skousen, Royal. 1992. *Analogy and structure*. Dordrecht: Kluwer.
- Skousen, Royal, Daryle Lonsdale and Dilworth B. Parkinson (eds.). 2002. *Analogical modeling: An exemplar-based approach to language*. Amsterdam & Philadelphia: John Benjamins.
- Stump, Gregory T. 2001. *Inflectional morphology A theory of paradigm structure*. Cambridge: Cambridge University Press.
- Vekerdi, József. 2000. *A comparative dictionary of Gypsy dialects in Hungary*. Budapest: Terebess.
- Wurzel, Wolfgang U. 1989. *Inflectional morphology and naturalness*. Dordrecht: Kluwer.