

**Univerzita Karlova v Praze**

Filozofická fakulta

Ústav lingvistiky a ugrofinistiky

obecná lingvistika – typologie

Viktor Elšík

**Affix Extraction:  
A Case Study on Hungarian Romani**

**Extrakce afixů:  
případová studie k maďarské romštině**

Disertační práce

vedoucí práce – Doc. Zdeněk Starý, CSc.

2006

*Prohlašuji, že jsem disertační práci vykonal samostatně s využitím uvedených pramenů a literatury.*

## ***Abstract in English***

The PhD thesis *Affix Extraction: A Case Study on Hungarian Romani* explores the principles and complexities of the contact-induced mechanism of affix extraction, i.e. of affix borrowing through the mediation of lexical borrowing. The thesis is a case study on affix extraction in Selice Romani, a variety of Romani (Indo-Aryan) that is strongly influenced by Hungarian (Finno-Ugric). After a brief delimitation of the phenomenon of affix extraction and an outline of the contact situation, the thesis describes in some detail several individual instances of extracted Hungarian-origin affixes in Selice Romani. It is claimed that several levels of bilingual morphology and two tiers of potentially constraining factors must be distinguished in order to describe the phenomenon of affix extraction in an adequate manner. The concept of ‘gap filling’ is tested as a potential predictor of affix extraction. It turns out that the contact situation of Selice Romani *vis-à-vis* Hungarian instantiates a stage of borrowing that is characterized by categorially redundant lexical borrowing of morphologically complex forms and affix extraction.

## ***Abstrakt v češtině***

Disertace s titulem *Extrakce afixů: případová studie k maďarské romštině* zkoumá principy kontaktního mechanismu afixální extrakce, tj. přejímání afixů prostřednictvím lexikálních přejímek. Jde o případovou studii afixální extrakce v selické romštině, jedné varietě tohoto indoárijského jazyka silně ovlivněné maďarštinou. Po stručném vymezení fenoménu afixální extrakce a nástinu kontaktní situace je podrobněji popsáno několik jednotlivých případů romských extrahovaných afixů maďarského původu. V práci se rozlišuje několik rovin bilingvní morfologie a dvě vrstvy potenciálně omezujících faktorů. Na materiále selické romštiny je testován pojem tzv. *gap filling* ‚zaplňování mezer‘ jakožto potenciální prediktor afixální extrakce. Kontaktní situace selické romštiny vzhledem k maďarštině je charakterizována kategoriálně redundantním přejímáním morfologicky komplexních slovních tvarů a kategoriálně redundantní extrakcí afixů.

# Affix Extraction: A Case Study on Hungarian Romani

## Table of contents

Table of contents.....	1
1 Introduction.....	3
2 Hungarian Romani of Selice.....	8
2.1 The language and its speakers.....	8
2.2 Typology.....	10
2.3 Lexicon.....	11
2.4 Phonology.....	12
2.5 Syntax.....	14
2.6 Function words.....	18
3 Affix extraction: verbs.....	22
3.1 Inflectional adaptation.....	22
3.2 Morphological categories and lexical borrowing.....	23
3.2.1 Verb inflection.....	23
3.2.2 Possibility verbs.....	25
3.2.3 Frequentative verbs.....	26
3.2.4 Valency-changing verbs.....	27
3.2.5 Factitive and inchoative verbs.....	30
3.2.6 De-nominal verbs.....	32
3.3 Extracted affixes.....	33
3.3.1 The de-nominal suffix <i>-áz-</i> .....	33
3.3.2 The causative suffix <i>-atat-</i> .....	35
3.3.3 The infinitive suffix <i>-ñi</i> .....	38
3.3.4 The loan-verb adaptation suffix <i>-ál-</i> .....	40
4 Affix extraction: nouns.....	42
4.1 Inflectional adaptation.....	42
4.2 Morphological categories and lexical borrowing.....	42
4.2.1 Noun inflection.....	42
4.2.2 Diminutive nouns.....	47
4.3 Extracted affixes.....	49
4.3.1 The separative suffix <i>-tú</i> .....	49
4.3.2 The abstract/collective suffix <i>-(a)šág-</i> .....	51
4.3.3 The action suffix <i>-áš<sup>1</sup> ~ -iš-</i> .....	53
4.3.4 The agentive suffix <i>-áš<sup>2</sup></i> .....	53
4.3.5 The pecunial suffix <i>-eš-</i> .....	56
4.3.6 The tool suffix <i>-ó</i> .....	58
4.3.7 The artificiality prefix <i>mí-</i> .....	62
5 Affix extraction: adjectives.....	64
5.1 Inflectional adaptation.....	64
5.2 Morphological categories and lexical borrowing.....	64
5.2.1 Adjective inflection.....	64

5.2.2	Adjective comparison .....	66
5.2.3	Attenuative and similative adjectives .....	66
5.2.4	Diminutive adjectives .....	68
5.2.5	Deprivative adjectives.....	68
5.2.6	Negative adjectives .....	69
5.2.7	De-nominal adjectives .....	70
5.2.8	De-adverbial adjectives.....	72
5.2.9	De-verbal adjectives.....	73
5.3	Extracted affixes .....	74
5.3.1	The habitual suffix <i>-óš-</i> .....	74
6	Affix extraction: adverbs .....	76
6.1	Morphological categories and lexical borrowing .....	76
6.1.1	Manner adverbs.....	76
6.1.2	Adverb comparison.....	79
6.1.3	De-nominal adverbs.....	80
6.1.4	De-verbal adverbs .....	81
6.2	Extracted affixes .....	81
6.2.1	The similative manner suffixes <i>-šon</i> and <i>-ijaššan</i> .....	81
6.2.2	The temporal-ordinal suffix <i>-dikán</i> .....	83
7	Affix extraction and ‘gap filling’ .....	86
7.1	The concept of ‘gap filling’ and its application .....	86
7.2	‘Gap filling’ as a constraint on lexical borrowing .....	89
7.3	‘Gap filling’ as a constraint on affix extraction .....	91
7.4	Discussion .....	93
8	Conclusions.....	96
	Symbols and abbreviations .....	99
	References.....	101

# 1 Introduction

In his seminal study on language contact, before setting off to exemplify borrowing of bound morphology, Weinreich (1953: 31–32; footnotes omitted) remarks:

What appears at first blush to be a transfer of highly bound morphemes often turns out, upon a fuller analysis, to be something else. It sometimes happens that free forms are transferred into a language in pairs, with and without an affix. The presence of the pair in the recipient language enables even its unilingual user to analyze the two-morpheme compound into a base and affix, and to extend the affix to other, indigenous bases. [...] After such items are discounted, however, there remains a residue of cases which can be explained in no other way than by the outright transfer of a highly bound morpheme.

In this passage Weinreich distinguishes very clearly between two different routes ‘foreign’ (L2) affixes may make it into a language (L1): either through borrowing of free forms followed by L1-internal analysis and extension of bound morphemes these free forms contain; or through an ‘outright’ borrowing of bound morphemes from L2 into L1. In other words, he distinguishes between the introduction of L2 affixes through mediation of lexical borrowing and without it. I will term the former mechanism AFFIX EXTRACTION, the latter mechanism AFFIX COPYING, and employ the label AFFIX BORROWING as a cover term for both mechanisms.

Some authors, including Weinreich it appears, do not consider, or hesitate to consider, the mechanism of affix extraction to represent affix borrowing at all. The reason is quite obvious: with affix extraction there is no ‘outright transfer’, in Weinreich’s terms, of L2 affixes (bound form–function units) into L1, a development that would parallel the ‘outright transfer’ of L2 word forms (free form–function units) into L1 — which is what is usually called *borrowing* in the domain of lexicon. Instead, with affix extraction the ‘outright transfer’ into L1 only affects L2 free forms, and the fact that this results in the presence of ‘foreign’ bound forms in the morphology of L1 is, in a sense, an internal matter of the L1 linguistic system. The clearest evidence for the L1-internal nature of affix extraction is, as Weinreich observes, the fact that even monolingual

speakers may innovate the morphological structure of their L1 by extracting new affixes from lexical borrowings from a language they do not speak or understand.

Yet, affix extraction is a type of contact-induced change, in the sense of Thomason's (2001: 62) definition: "any linguistic change that would have been less likely to occur outside a particular contact situation is due at least in part to language contact." It is a type of contact-induced change that would have been not only less likely, as Thomason requires, but plainly impossible to occur without language contact and, in our case, lexical borrowing. In addition, on encountering an obviously 'foreign' affix in a language, it is often difficult if not impossible to find out in what way the affix made it into the language. The role of lexical mediation remains unclear, for example, in the many instances of demonstrably borrowed affixes in the languages of Arnhem Land, explored in detail in Heath (1978). Finally, though affix extraction is clearly distinct from affix copying, i.e. affix borrowing in the strict sense, both mechanisms share certain properties and outcomes (cf. Winford 2003: 91–97).

Weinreich (1953), who seems to be touching, allow me to exaggerate a little bit, on every thinkable issue of the agenda of modern contact linguistics, devotes merely the two sentences in the above quote, to the issue of affix extraction, plus a few examples, such as the Hebrew-origin plural suffixes in Yiddish, e.g. *doktórj-**im*** 'doctors', or the French-origin diminutive suffix in English, e.g. *kitchen-**ette*** (p. 31). The mechanism appears to be clear (or irrelevant). Affix extraction, nevertheless, may get much more interesting than Weinreich's oft-cited examples suggest.

Although affix extraction presupposes lexical borrowing, it is not merely epiphenomenal to it. There are CONDITIONS that must be met in order for affix extraction to be applicable at all and there are, presumably, also several FACTORS (or CONSTRAINTS, if formulated negatively and in absolute terms) that (co-)determine whether affix extraction will actually take place, or not, when it *is* applicable. While the necessary conditions are, in a way, part of the definition of the mechanism of affix extraction, and can be thus formulated *a priori*, the identification of the determining factors or constraints may, in principle, only result from empirical investigation of actual instances of affix extraction across a variety of contact situations (though hypothesized factors based on some *a priori* expectations may of course be subjected to testing).

The general condition on affix extraction is that of IDENTIFIABILITY: speakers must be able to identify, within their L1, the ‘foreign’ (L2-origin) affix that is to be extracted, i.e. to identify it as a morphemic, form–function, unit. The identifiability of an L2-origin affix presupposes not only that there are, in the L1, morphologically complex lexical borrowings that contain a reflex of the source L2 morpheme, the DERIVATIONS in an extended sense (i.e. including morphologically complex inflectional forms), but also that there are paradigmatically related lexical borrowings that do *not* contain a reflex of the source L2 morpheme, which serve as BASES for the derivations.

Finally, affix extraction assumes not only adoption of an affix within loanwords from a certain L2 and its (potential) paradigmatic identification, but also its analogical, language-internal, EXTENSION to other L1 lexemes and other etymological compartments within the L1 lexicon. While identifiable affixes may (but need not!) cease to be recognized as morphemes with the loss of L1 speakers’ bilingualism in the source L2 of the affixes (in case of significant lexical replacement of L2 loanwords), extended affixes become more stable, though not necessarily permanent, part of the L1 morphological structure.

The present thesis is a CASE STUDY on affix extraction in a variety of Romani, SELICE ROMANI, that has been and continues to be strongly influenced by Hungarian. The asymmetrical contact situation is one of native or near-native active bilingualism in Hungarian of all native speakers of Selice Romani, and of extensive lexical and grammatical borrowing from Hungarian into Selice Romani. The general objective of this thesis is to explore the mechanism of affix extraction, with special attention to its various complexities that might not be encountered in situations of less intimate or less asymmetrical bilingualism.

Somewhat paradoxically, bilingualism — and especially the situation of general and near-balanced bilingualism in a speech community (as is the case of the Selice Romani L1 speakers and their generally high competence in Hungarian) — appears to complicate the conceptual apparatus required to capture the mechanism of affix extraction in an adequate manner. Above all, the analyst needs to distinguish at least three levels of affixal morphology (and of morphemics in general) in the bilingual contact situation. First, the bilingual L1 speakers are able to analyze (in the sense of emergent

grammar, roughly) the morphemic structure of L2 word forms: these are the L2 affixes. Second, the bilingual L1 speakers are able to identify affixes within lexical borrowings from an L2 into their L1: these I will call IMPORTED (L2-origin) affixes. And, finally, those imported affixes that actually get extended to other L1 lexemes are the EXTRACTED affixes. Though one might be tempted to assume formal and functional identity between the L2 affixes in a word form and the imported affixes in the L1 lexical borrowing of this word form, several instances of affix extraction in Selice Romani suggest that this is not the case: there may be mismatches both in form (due to re-analysis of boundaries and allomorphic selection, for example) and in function.

All the data on SR presented in this thesis come from my own linguistic research on SR that was carried out during short but numerous fieldtrips to Selice, Slovakia, between 1997 and 2007.<sup>1</sup> Parts of the thesis have been published in Elšík (in press). The descriptive sources on Hungarian that I have consulted include Abandolo (1988), Kenesei *et al.* (1998), Siptár & Törkenczy (2000), Samu (1971), and Tompa (1968).

The structure of the thesis is as follows: Section 2 is a brief information on the L1 variety, Hungarian Romani of Selice, including an outline of grammatical borrowings from Hungarian, viz. other than those discussed in the following sections. Sections 3–6 present an overview of SR affixes that clearly or very likely result from extraction from Hungarian loanwords. The individual sections are devoted to lexical borrowing and affix extraction in the four major word classes that both Hungarian and SR possess: verbs (Section 3), nouns (Section 4), adjectives (Section 5), and adverbs (Section 6).<sup>2</sup> Subsections on individual extracted affixes describe the synchronic properties of these affixes (their shape, function, productivity, derivational bases, allomorphy etc.); deal with the morphological category they express and with markers competing with the extracted affixes, if any; and discuss the Hungarian sources of the extracted affixes. Section 7

---

<sup>1</sup> I wish to thank the late Milena Hübschmannová for introducing me to the Selice Rumungro community; Július Lakatoš and Alena Krézsová for their hospitality and native speaker expertise; and the Roma Culture Initiative of the Open Society Institute, Budapest, for their financial support of my SR research in 2001–2002.

<sup>2</sup> There are only instances of affix copying in SR pro-words. See Section 2.6 and Elšík (in press) for some discussion.

addresses the issue of predictability of lexical borrowing of morphologically complex forms and of affix extraction, and evaluates the SR data from the perspective of the so-called 'gap filling' hypothesis. Section 8 concludes.

## 2 Hungarian Romani of Selice

### 2.1 *The language and its speakers*

The language under description is a variety of Romani (Indo-Aryan, Indo-European) spoken by long-settled Roms (Gypsies) of southern Slovakia and northern Hungary, which is classified in Romani dialectology as the Northern (non-Vendic) subgroup of the South Central group of Romani dialects (cf. Boretzky 1999, Elšík *et al.* 1999) and usually referred to as ‘Rumungro’ in Romani linguistics. The variety I chose to describe, Selice Romani (SR), is one of the few Northern South Central (NSC) Romani varieties whose speakers are Hungarian bilinguals. Although all NSC Romani varieties have been influenced by Hungarian, most NSC Romani speakers presently live in ethnically Slovak parts of Slovakia and are Slovak bilinguals. An overwhelming majority of NSC Romani communities in Hungary and in the Hungarian parts of Slovakia have undergone language shift to Hungarian.<sup>3</sup> For more details on the dialectological and sociolinguistic situation of Romani in Slovakia see Elšík (2003).

The NSC Romani dialects are seriously underdescribed. As for grammar, there is no book-format grammatical description of any NSC dialect. An article co-authored by the present writer (Elšík *et al.* 1999) describes a selection of salient phonological and morphological features in several NSC varieties. As for lexicon, the NSC varieties of Hungary have been included in Vekerdi’s (1983) multidialectal dictionary, and a variety of Nógrád (Hungary) is described in Rác (1994). There is no lexical description of any NSC variety of Slovakia. The NSC variety of Nógrád is documented in Görög’s (1985) text collection, and another NSC variety (of unclear provenance) is documented in Müller (1869).

SR is the language of some 1,350 Rom (Gypsy) inhabitants of the Hungarian village of Selice (Hungarian *Sók*, SR *Šóka*) in southwestern Slovakia. In addition, there are about 150 Roms in the village who speak a different (a North Vlax) dialect of

---

<sup>3</sup> There are numerous further dialects of Romani in current contact with Hungarian, some of them only distantly related to SR, e.g. Lovari (Hutterer & Mészáros 1967), Cerhari (Mészáros 1976), Sinti (Mészáros 1980), Gurvari (Vekerdi 1971), and the more closely related Vend (Vekerdi 1984).

Romani. The former Roms are referred to as *Rumungri* (originally ‘Gypsy-Hungarians’) by the latter group, who are called *Pojáki* (originally ‘Poles’) by the Rumungri. Both groups use the ethnonym *Rom* for their own group and both are called *cigányok* ‘Gypsies’ by Hungarians, although the Hungarian villagers clearly differentiate between *magyar cigányok* ‘Hungarian Gypsies’ (i.e. the Rumungri) and *oláh cigányok* ‘Romanian Gypsies’ (i.e. the Pojáki). At present, both Rom groups taken together slightly outnumber the Hungarian population of the village. Until recently, however, the Hungarians were in a demographic majority and they remain to be the socially, economically, and politically dominant group in the village.

SR is prevalently an oral language. Some Rumungri are able to write letters or text messages in Romani but the language is not used for regular written communication. Nor is it used in massmedia or in formal education. Although Romani in general is an officially recognized language in Slovakia, there is no recognition of the South Central (Rumungro) dialect specifically and, so far, there have been no attempts at its standardization. SR is the language of family and in-group communication among the local Rumungri and the language of inter-group communication between the Rumungri and the local Pojáki. While the latter learn SR as their second dialect of Romani (and speak a distinct ethnolect of it), the Rumungri usually do not learn the dialect of the Pojáki. Many Hungarian villagers understand SR well, although only a few have some active competence in it and they are rarely fluent speakers. While all Selice Rumungri born before 1975 or so are native speakers of SR, in some families children are presently spoken to only in Hungarian or Slovak, and left to acquire some competence in SR in adolescent and adult peer groups, if at all. Thus, SR is not a safe language, though it is not seriously endangered yet.

All school-age or older L1 speakers of SR are multilingual. First of all, they are fluent and highly competent in Hungarian, which they use especially in their everyday communication with the Hungarian villagers. Some very young children may be monolingual in SR, although early acquisition of Hungarian appears to be the prevailing pattern nowadays. In addition, most Rumungri are fluent in Slovak, the official and dominant language of Slovakia, which they use outside of the village. Also, most have acquired at least passive competence in Czech through their exposure to Czech

massmedia and employment-related stays in the Czech part of the former Czechoslovakia (in 1960–1980’s almost all families of the Selice Rumungri community spent ten to thirty years there). Though both Hungarian and Slovak (and to some extent Czech as well) may be classified as current L2s of SR, it is clear that Hungarian enjoys a special sociolinguistic status: *inter alia* it is the language of the secondary ethnic identity of the Selice Rumungri, who frequently refer to themselves as ‘Hungarian’ Roms, accepting the attribute ascribed to them by Hungarians.

As evidenced by lexical borrowings, SR shares with other Romani dialects previous contact with West Iranian (Persian and/or Kurdish), Ossetic, Armenian, and especially Greek; the latter language also had an enormous impact on Romani grammar. On the other hand, most South Slavic loanwords in SR are dialect-specific within Romani. Some of them can be identified as Serbian/Croatian or even Ikavian Serbian/Croatian (Elšik *et al.* 1999); for the sake of convenience, I will refer to the South Slavic contact layer of SR simply as ‘Ikavian’, although some of the South Slavic borrowings may be of different (e.g. East South Slavic) origin. Linguistic contact of SR with Hungarian is likely to have lasted for at least two centuries. Widespread multilingualism of the Selice Rumungri in Slovak and Czech did not develop before the 1920’s and 1950’s, respectively. While these secondary current L2s have contributed only a few marginal established loanwords, Hungarian has exerted, and continues to exert, a strong lexical and grammatical influence on SR.

## **2.2 Typology**

The typological profile of Asian (Proto-)Romani was altered rather significantly already before the arrival of its speakers to Europe. Matras (2002: 196) argues that, for example, the development of interrogative-based relativizers or the reduction of non-finite constructions could have taken place in a western Asian convergence area, i.e. before the contact of Romani with Greek in Asia Minor. The latter language, nevertheless, remains the major source of typological innovations that are shared by Romani as a whole: the development of a proclitic definite article, the emergence of prepositions (or a significant

expansion of their inventory), the shift to a basic predicate–object order, and more (cf. Matras 1994, 2002: 198–199).

Post-Greek L2s have had a less significant impact on major typological parameters of SR. Two developments in morphological typology deserve a mention. In its Greek period, Romani possessed a single prefix: the deprivative *bi-* ‘un-’ of Indo-Aryan or West Iranian origin. Matter borrowing of several pronominal prefixes from Ikavian and Hungarian, of a superlative prefix from Hungarian, and a grammaticalization of another pronominal prefix due to pattern replication from Hungarian, has increased the number of prefixes in SR by eight. Second, there is some marginal evidence that separatist exponence, which prevails in the largely agglutinative Hungarian, has been gaining ground in SR at the expense of fusion, although it is difficult to argue for contact-induced innovations here.<sup>4</sup> Outstanding syntactic developments due to contact with Hungarian include the creation of a class of preverbs, the ‘re-introduction’ of non-finite subordinate constructions, and various modifications in word order patterns.

### 2.3 *Lexicon*

Out of a much larger inventory of early loanwords into Romani (as attested in different Romani dialects; cf. e.g. Boretzky 1995, Boretzky & Iglá 1994, 2004), SR retains ca. 20 loanwords from Iranian languages, ca. 10 loanwords from Armenian, and ca. 30 loanwords from Greek. In addition, there are over 40 loanwords from South Slavic, some of which can be identified as (Ikavian) Serbian/Croatian and which are mostly not shared with other dialects of Romani. Most of the pre-Hungarian loanwords are nouns, while verbs and adjectives are less numerous. Only relatively few pre-Hungarian function loanwords have been retained. While there are a few stable noun loanwords from the secondary current L2s of SR speakers (e.g. *obrazovka* ‘screen’ from Slovak, *pepšo* ‘black

---

<sup>4</sup> To mention one example: Most Romani dialects possess a small class of nouns that fuse their roots with an oblique suffix due to a phonological contraction: cf. *mos-* < \**muj-es-*, an oblique stem of *muj* ‘face, mouth’. In SR, the oblique stem of the above noun is non-fusional (*muj-es-*), although the noun *mos-tar* ‘slap in the face’ (a lexicalized ablative) suggests that the contraction had affected SR as well. While the non-fusional inflection of *muj* clearly results from a secondary, morphological, development in SR, it is impossible to prove that this instance of morphological decomposition is due to Hungarian influence.

pepper’ from Czech), and while nonce borrowing of nouns and verbs from these languages is rather common, the by far most important current source of loanwords is Hungarian. Hungarian loanwords include basic vocabulary in domains such as body parts, bodily functions, kinship, or physical properties (e.g. ‘knee’, ‘to breathe’, ‘son-in-law’, ‘weak’).

Unlike some Romani varieties that employ internal word-formation processes to create a layer of secret vocabulary in certain semantic domains (cf. Matras 2002: 223), SR does not seem to avoid loanwords (such as *čendéri* ‘policeman’ < Hungarian) in these domains. Instances of pattern replication without matter borrowing in complex referring expressions are exceptional, e.g. *sobota-kurko* [Saturday-Sunday] ‘weekend’ calquing local Hungarian *szombat-vasárnap*. An overwhelming majority of Hungarian compounds are borrowed rather than translated, e.g. *fog-orvoš-i* ‘dentist’ < *fog-orvos* [tooth-doctor], though translations of lexicalized preverb–verb collocations are common. Some Hungarian compounds may be decomposed into adjective–noun collocations, e.g. *világ-ik-o háború* [world-ADJ-NOM.SG.M war] ‘world war’ < *világ-háború* [world-war].

Phraseological idioms are commonly translated from Hungarian. As several Hungarian types of greetings and similar expressions are missing in the traditional Rumungri culture, some speakers have started to fill in the ‘gap’ by using Hungarian expressions, e.g. *szia* ‘hi; bye’, *jó étvágyat* ‘bon appetit’. Some indigenous politeness expressions are used in wider contexts due to cultural contact. For example, *palikerav* ‘I greet; I thank’ is not traditionally used after being served a meal or coffee at home, but some Rumungri would now use it in this context, as the local Hungarians do.

## **2.4 Phonology**

The inventory of SR phonemes is identical to that of Hungarian, with two exceptions. First, SR retains distinctive aspiration in voiceless stops and affricates, e.g. *čór-* [tʃ<sup>h</sup>o:r] ‘to steal’ vs *čhor-* [tʃ<sup>h</sup>or] ‘to pour’, which is absent from Hungarian. Second, Hungarian rounded front vowels are usually replaced with their unrounded counterparts in

loanwords, e.g. *csütörtökön* [tʃytörtökön] > *čiterteken* [tʃiterteken] ‘on Thursday’, although some speakers now tend to retain them in certain loanwords. Both vowel and consonant inventories of Romani have been enlarged due to contact with Hungarian. Instances of contact-induced phoneme loss are rare: they include the merger of the voiceless uvular fricative [χ] with the glottal fricative /h/ [h] and the merger of the palatal lateral [lʲ] with the palatal approximant /j/ [j], e.g. \*[χalʲam] > *hájam* [hɔ:jam] ‘we ate’. On the other hand, contact with Hungarian has given rise to several phonemic distinctions and numerous new phonemes in SR.

A major contact-induced change has been the development of distinctive phonological quantity: vowel length, e.g. *phirav-* [pʰirav] ‘to wear’ vs *phírav-* [pʰi:rav] ‘to make [so.] walk’, and consonant gemination, e.g. *čuča* [tʃutʃa] ‘empty’ (an inflectional form) vs *čučča* [tʃutʃ:a] ‘breasts’. Both types of quantity have spread to the pre-Hungarian lexical component, although some individual geminates remain restricted to the Hungarian component. The inventory of vocalic qualities, too, has been enlarged due to contact. Although the open-mid front vowels – the short [æ] and the long [æ:] – are mostly restricted to Hungarian loanwords, they are phonologically distinct from their closed-mid counterparts, e.g. [dæ] ‘but’ vs [de] ‘give!’. In addition to the phonological quantity difference, the long /á/ [ɔ:] is distinguished through phonetic rounding from the short /a/ [a], as it is in the local Hungarian dialect. Contact with Hungarian has also triggered the development of a series of palatal consonants from palatalized dentals or palatalized velars, e.g. \*[tatʲar] > *taʹar-* [tacaɾ] ‘make warm’, \*[kʰil] > *tʰil* [cʰil] ‘butter’.

The Hungarian-origin phonemes play an important role in morpho-phonological alternations. In addition, several morpho-phonological rules are borrowed. For example, a morpheme-initial palatal approximant triggers gemination and a shift to a palatal of a preceding morpheme-final dental stop, as it does in Hungarian, e.g. *kafid-i* [kafidi] ‘table’ → {kafid-ja} *kafidʹd-a* [kafiʃ:a] ‘tables’. SR also borrows vowel harmony from H, although it remains restricted to a single type of alternation that affects only a few indigenous affixes, e.g. *farkašš-a* [farkaʃ:a] ‘wolves’ vs *kemívešš-e* [kæ:mi:væʃ:æ] ‘bricklayers’, *bika-ha* [bikaha] ‘with a bull’ vs *kečke-he* [kætʃkæhæ] ‘with a goat’. Apart

from the development of long vowels and geminate consonants, the syllable structure of the pre-Hungarian component has remained unaffected by contact with Hungarian. On the other hand, there is no adaptation of Hungarian loanwords in terms of their syllable structure. The distribution of long vowels in SR suggests that they developed before the Hungarian-induced general shift of stress to word-initial position, e.g. \*[barval'o] > \*[barva:l'o] > *barválo* [b'arvɔ:lo] 'rich'. Intonation patterns are largely identical to those of the local Hungarian dialect.

## 2.5 Syntax

A number of clause-level syntactic features that SR shares with Hungarian is due to a typological or areal similarity between the two languages, rather than due to immediate borrowing from Hungarian into SR. For example, both languages have uninflected pre-verbal negators, allow pro-drop, and use a copula verb in non-verbal predication (though, unlike Hungarian, SR does not allow copula deletion in the third-person present affirmative). SR also shares with Hungarian negative agreement of the predicate with negative pro-words; this is clearly a post-Greek pattern in SR, though Ikavian is a more likely source than Hungarian.

The major structural domain of syntactic borrowing from Hungarian into SR is clause combining and phrase combining. SR borrows all of its coordinating conjunctions with the exception of conjunctive coordinators, which are pre-Hungarian: plain disjunctive *vad'* 'or', contrastive disjunctive *vad' – vad'* 'either – or', free-choice alternative *ha – ha* 'whether – or' (1), and several connectors with adversative and contrastive functions, e.g. *de* 'but', *azomba* 'however', *mégiš* 'still, even so', *hanem* 'but rather', and *meg* and *pedig* 'but, in turn' (2). Borrowed adverbial subordinators include the causal *mert* and *mivel* 'since, because' (3), and several non-simultaneous temporal subordinators: the posterior *mire* and *mielét* 'before' (4), the posterior–durative *még* 'until' (5), and the anterior durative *mióta* 'since' (6).

- (1) *Beršeste čak trinval bašavlahi vad' bijav, vad' bállo.*  
 year.LOC.SG only three.times play.3SG.REM CONJ weddingCONJ ball

‘In a year he just played three times, either at a wedding, or at a ball.’

- (2) *De ón na džan ánglal, hanem téle džan.*  
CONJ they NEG go.3PL to.the.front CONJ downward go.3PL  
‘But they are not progressing, they are rather sinking.’

- (3) *Mivel čoháni ssa,*  
CONJ witch COP.3SG.PRET  
*na tromalahi and-i khangéri te džan.*  
NEG dare.3SG.REM in-DEF.F church(F) COMP go.3PL.SUBJ  
‘Since she was a witch, she did not dare to go to the church.’

- (4) *Mielét hasa, thov tre vasta*  
CONJ eat.2SG.FUT wash.IMP2SG.GEN:PL hand.PL  
‘Before you are going to eat, wash your hands.’

- (5) *Addig phírd’om, még le n- alakjom.*  
to.that.extent walk.PFV.1SG CONJ 3SG.M.ACCNEG find.PFV.1SG  
‘I did not stop walking until I found him.’

- (6) *Mióta džukela hi amen, náne amen mačka.*  
CONJ dog.PL COP.3.PRES we.ACC COP.NEG.3.PRES we.ACC cat  
‘Since we have kept dogs, we do not keep a cat.’

Clausal complements of predicates of utterance, propositional attitude, (acquisition of) knowledge, immediate perception and the like, are introduced by the Hungarian-origin general subordinator *hod’* (7a). Like in Hungarian, this subordinator is also employed to introduce several types of adverbial clauses (7b: reason clause) and, optionally, embedded interrogative clauses (7c) and embedded polar questions (7d). The latter are – obligatorily, unless an alternative construction is used – marked by the question enclitic *-i*, which is also borrowed from Hungarian. The subordinator *hod’* may

also precede various pre-Hungarian subordinators that introduce embedded commands and other clausal complements of manipulative predicates (7e), and purpose clauses (7f). Unlike in Hungarian, however, the subordinator *hod'* cannot introduce such clauses by itself.

(7)a. *Halijom, hod' má n- aná le*  
 understand.PFV.1SG COMP already NEG bring.1SG.FUT 3SG.M.ACC  
*uppe gódi.*  
 on brain  
 'I understood that I will not persuade him any more.'

(7)b. *Daráhi, hod' našlija o lóvo.*  
 fear.1SG.REM COMP get.lost.PFV.3SG DEF.M money(M)  
 'I was afraid that the money had gotten lost.'

(7)c. *Na šunde láche, (hod') ko vičinel taj so.*  
 NEG hear.PFV.3PL well COMP who shout.3SG and what  
 'They did not hear well who was shouting and what.'

(7)d. *Na džanav, (hod') muká -i man tutar*  
 NEG know.1SG COMP let.1SG.FUT-Q 1SG.ACC 2SG.ABL  
*te čumiden.*  
 COMP kiss.3PL.SUBJ  
 'I do not know whether I will let you kiss me.'

(7)e. *Phend'a mange, (hod') khére nek áčovav.*  
 say.PFV.3SG 1SG.DAT COMP at.home OPT stay.1SG.SUBJ  
 'S/he told me to stay at home.'

(7)f. *Site le papaleg uppe alakhes,*  
 must 3SG.M.ACC again upward find.2SG

(*hod'*) *káj nek džanesahi le te phenen.*  
 COMP where OPT know.2SG.REM 3SG.M.ACC COMP say.3PL.SUBJ  
 ‘You have to discover it again, in order to be able to say it.’

Due to pattern replication from West Iranian or Greek, complement clauses of modal predicates were finite in the early European stages of Romani: the subordinate verb was introduced by an indigenous non-factual complementizer and showed subject person–number agreement with the matrix verb (Matras 2002: 161). Pattern replication from Hungarian has resulted in a development of a non-finite complement form in SR, through fossilization of a frequent finite form of the subordinate verb: the subordinate verb now invariably shows third plural subjunctive inflections, irrespective of the person–number of the matrix verb. This non-finite construction, which may be termed the *subjunctive infinitive* (or the ‘new’ infinitive, Boretzky 1996), encodes not only clausal complements of modal predicates but also clausal complements of some manipulative verbs and tightly integrated same-subject purpose clauses. See 3.3.3 for more details.

Pattern replication from Hungarian has also occurred in relative clauses. Although SR relativizers are formally identical to interrogatives, whereas Hungarian relativizers are not, the former partly copy the ‘ontological’ restrictions of the latter: human head nouns usually select a person pro-word (‘who’) as a relativizer in SR, while non-human head nouns mostly select a thing pro-word (‘what’).<sup>5</sup>

Linear order of the predicate, its arguments and adverbial adjuncts is flexible in Romani, being largely determined by pragmatic factors (cf. Matras 1995, 2002: 167–174). While syntactic non-configurationality is also characteristic of SR, numerous aspects of SR clause-level order appear to have been borrowed from Hungarian, likewise a non-configurational language. A prominent example is the tendency to position

---

<sup>5</sup> The ontological match is not complete, however, as the thing relativizer is not ungrammatical with human head nouns, although it is much rarer than, say, in Slovak-influenced varieties of South Central Romani. In addition, the indigenous local interrogative (‘where’), which was, due to pattern borrowing from Greek, the general relativizer in earlier stages of Romani (Matras 2002: 177), is still rarely attested with non-local head nouns in SR.

focussed constituents immediately before the finite verb; this frequently results in clause-final position of the copula in non-verbal predications (8; second line).

- (8) *Odá hi gadžikano sokáši.*  
 that.M COP.PRES.3 nonGypsy(a) habit  
*Romano sokáši tista áver hi.*  
 Gypsy(a) habit sheer other COP.PRES.3  
 ‘That’s a non-Gypsy habit. The Gypsy habit is completely different.’

On the other hand, linear order at the noun phrase level is syntactically determined in SR: all types of adjectival modifiers, including descriptive adjectives, adnominal possessors, demonstratives, and numerals, always precede their head nouns. While the modifier–noun order prevails in all Romani dialects (cf. Matras 2002: 165–167), it has been fully grammaticalized in SR due to contact with Hungarian. The alternative noun–modifier order is simply ungrammatical, except in cases of afterthought whereby the postposed modifier is a nominalized apposition. SR exhibits an etymological split in the order of adpositions: while those borrowed from Hungarian are postposed to their object noun phrases, adpositions of pre-Hungarian origin always remain preposed.<sup>6</sup> An analogical split occurs with focus particles meaning ‘also, too’: the indigenous *te* is preposed to the focused element, while the Hungarian-origin *iš* is postposed.

## 2.6 *Function words*

In addition to lexical verbs, nouns, adjectives, and manner adverbs, SR has borrowed numerous function, or less lexical, words from its different L2s. The modal particle of possibility *šaj* ‘can’ is likely to be of West Iranian origin (Matras 2002: 196). Greek is the source of the cardinal numerals *efta* ‘seven’, *ofto* ‘eight’, *eňňa* ‘nine’, and *trianda* ‘thirty’ and the ordinal *trito* ‘third’; the quantifier *buka* ‘a little, a piece of’; the address particle

---

<sup>6</sup> This contrasts with the contact-induced postpositioning of inherited prepositions in some Romani dialects influenced by postpositional languages such as Turkish or Finnish (cf. Matras 2002: 206).

*more* ‘hey, man!’; the temporal deictic particle *paleg* ‘then, after that’ (< ‘again’);<sup>7</sup> and the temporal adverb *táha* ‘tomorrow’. Ikavian provided the quantifiers *dosta* ‘enough’, *sako* ‘every’, and *cilo* ‘whole’; the distributive particle *po*; the optative/permissive particle *nek* ‘let’, which has also been grammaticalized into a subordinator; the focus particle *ni* ‘not even, neither’ and the related coordinator *ni – ni* ‘neither – nor’; the negative pronoun *ništa* ‘nothing’; and the preverb *préku* ‘through; across, over’, which has been grammaticalized within SR from a borrowed spatial adverb. Only some elderly speakers of Selice Rumungro use the proximative preposition *uze* ‘at/to the vicinity of’ of Ikavian origin; others use an indigenous proximative preposition.

Most function words have been borrowed from Hungarian, the current L2. Hungarian is the source of numerals (see below), the quantifier *čepo* ‘few, little; a few, a little’ (< ‘a drop of’), the degree words *igen* ‘very, very much’ and *túl* ‘too, too much’, the generic obligative particle *musaj* ‘one has to’, numerous preverbs (e.g. *át* ‘through; across, over’ or *sít* ‘apart’), and a few marginal postpositions (e.g. *serint* ‘according to’ or *felé* ‘in the direction of’). Borrowing from Hungarian is extensive in discourse-related function words, such as repetition adverbs (*újbú* or *újra* ‘again, anew’), utterance-level adverbs (*talán* ‘perhaps’, *bistoš* ‘certainly’, *perse* ‘of course, sure’, *bizoň* ‘indeed’), phasal adverbs (*még* ‘still’ and *má* ‘already’), focus particles (*iš* ‘also, too’, *čak* ‘only’, *ippen* ‘just’, *pont* ‘exactly’, *egís* ‘entirely’), affirmative answer particles (the regular *hát* ‘yes’, and the contrary-to-expectation *de* ‘but yes’), interjections (*ehe*), fillers (*hát*), sequential discourse markers (*no*), and more.

In addition to function words, Rumungro has borrowed several function-word affixes. The Greek-origin suffix *-t-* derives regular ordinals from cardinal numerals, e.g. *dúj* ‘two’ → *dúj-t-o* ‘second’. The Ikavian-origin prefix *ni-* and the Hungarian-origin prefixes *vala-*, *akár-*, and *minden-* apply to interrogative pro-words, e.g. *káj* ‘where’ → *ni-kháj* ‘nowhere’ (negative), *vala-káj* ‘somewhere’ (specific indefinite), *akár-káj* ‘anywhere whatsoever’ (free-choice), and *minden-káj* ‘everywhere’ (universal quantification). The Hungarian-origin prefixes *am-* and *ud’an-* apply to deictic pro-words, e.g. *asso* ‘such’ → *am-asso* ‘such like the other’ (deictic contrast) and *ud’an-asso* ‘just

---

<sup>7</sup> The repetition particle *papaleg* ‘again’ (< \**pal-pale*) has developed through reduplication of the Greek loanword.

such like this/that one' (deictic identity). All of the pronominal prefixes must have been borrowed without the mediation of lexical borrowing.

There are also several instances of pattern replication from Hungarian in function words. The genderless Hungarian is the source of gender neutralization in the nominative of the SR third person singular pronoun: the original feminine form *ó* 'she' has replaced the original masculine form *\*ó* 'he', assuming a gender-neutral function 's/he' (cf. H *ő* 's/he').<sup>8</sup> On the other hand, the development of a distinction between local pro-words of stative location and direction, e.g. *káj* 'where' vs *kija* 'whither', is likely to have been modelled on an identical distinction in Hungarian. Due to a complex interplay of pattern replication and internal re-analysis, the universal-quantification prefix *sa-* has developed as an alternative to the borrowed universal-quantification prefix *minden-* (see above), e.g. *sa-káj* 'everywhere'. Pattern replication has also been involved in the grammaticalization of the reciprocal pronoun *jékh-ávr-* [one-(an)other-] 'each other', which is a compound of an identical structure as the Hungarian reciprocal pronoun *egy-más*. The expression of the phasal expression 'no longer' as a negation of 'already' is clearly modelled on Hungarian.<sup>9</sup> In syntax, adnominal cardinal numerals (optionally in case of 'one') have lost case agreement with their head nouns due to Hungarian influence, e.g. *dúj* (*\*dúj-e*) *muršenca* [two (*\*two-OBL*) man.PL.SOC] 'with two men'.

A final note concerns borrowing of Hungarian numerals. Two types of loans must be distinguished: morphologically integrated loanwords, which have no inherited, pre-Hungarian alternative (the cardinals *nulla* 'zero', *ezeri* 'thousand', and *miliomo* 'million', the ordinal *éšéno* 'first', and most fraction numerals), and morphologically unintegrated loanwords, which alternate with inherited numerals. The unintegrated numerals allow or require, due to Hungarian influence, the singular of some of their head nouns, viz. of some Hungarian-origin nouns denoting currency units: contrast *pándžvárdeš hallér-ja* 'fifty hellers' (indigenous numeral, plural noun) with *etven hallér-i* 'fifty hellers' (Hungarian numeral, singular noun) < H *ötven hallér*. Note that the latter

---

<sup>8</sup> However, oblique case forms of the pronoun have remained differentiated for gender, e.g. the accusative *le* 'him' vs *la* 'her' (cf. Hungarian *őt* 'him, her').

<sup>9</sup> The expression of 'not yet' as a negation of 'still' is congruent with Hungarian, but is likely to be pre-Hungarian.

construction is not necessarily a code-switch, as the singular noun is morphologically adapted in SR. The alternation between inherited and borrowed expressions also concerns various de-numeral derivations and compounds, e.g. *eftavardeš beršiko* or *hetven ívešno* < H 'seventy-year-old'.

### 3 Affix extraction: verbs

#### 3.1 Inflectional adaptation

Verbs are commonly borrowed into SR, as they are into any other Romani dialect. Pre-Greek and (presumably) early Greek loan-verbs show full morphological integration and are structurally indistinguishable from indigenous verbs. Post-Greek loan-verbs, on the other hand, are marked out by a specific adaptation marker, the Greek-origin suffix *-in-*, which is added to an inflectional stem of the L2 verb, e.g. *vič-in-* ‘to shout’  $\triangleleft$  Ikavian *vič-*, and followed by regular indigenous inflections (with the exception of passive participle forms).<sup>10</sup> The suffix *-in-* was extracted from lexical borrowings of Greek verbs with the present stem in *-in-*. Through none of these have been retained in SR, the suffix has been extended to those Greek loan-verbs that originally contained a different suffix, e.g. *rum-in-* ‘to spoil TR’  $\triangleleft$  Greek *rim-az-*, *ir-in-* ‘to turn TR’  $\triangleleft$  Greek *jir-iz-*. Cross-dialect comparison suggests that the suffix *-in-* was originally specialized for non-perfective adaptation of some transitive loan-verbs in Romani (Matras 2002: 130). In SR, however, it has developed into a general, aspect- and valency-neutral, verb-adaptation marker.<sup>11</sup>

Nonce loan-verbs from the ‘post’-Hungarian L2s, Slovak and Czech, show a distinct pattern of morphological adaptation, which will be discussed in more detail in Section 3.3.4.

---

<sup>10</sup> The adaptation suffix *-in-* is absent in passive participles of morphologically adapted borrowed verbs. Instead, the participles contain the Greek-origin participle suffix *-ime*, e.g. *rum-ime* ‘spoiled’, *vič-ime* ‘shouted’. The suffix was extracted from Greek lexical borrowings and extended to all post-Greek loan-verbs and to several indigenous verb classes, e.g. *d-ime* ‘given’  $\leftarrow$  *d-* ‘give’ (cf. Elšík & Matras 2006: 331–332).

<sup>11</sup> The Greek-origin suffix *\*(V)s-*, which appears to have been the marker of perfective adaptation of all loan-verbs and of non-perfective adaptation of intransitive loan-verbs (Matras 2002: 130), has acquired novel functions in SR: it is now an integral part of the suffix *-(i)sal-*, which serves as a stem extension in several valency-changing or aktionsart derivations, e.g. *cid-* ‘to pull’  $\rightarrow$  *cid-isaj-ov-* ‘to stretch ITR’ (anticausative), *térn-o* ‘young’  $\rightarrow$  *térn-isaj-ár-* ‘to make young’ (factitive), *khand-* ‘to stink’  $\rightarrow$  *khand-isaj-ov-* ‘to stink intensively’ (intensive).

## 3.2 *Morphological categories and lexical borrowing*

### 3.2.1 **Verb inflection**

Hungarian verbs have a rich inflectional morphology (cf. Tompa 1968: 155–174, Abandolo 1988, Kenesei *et al.* 1998). The categories of tense and mood combine into six subparadigms: the synthetic present (indicative), past (indicative), present conditional, and imperative–subjunctive; and the periphrastic future (indicative) and past conditional. Intransitive verbs cross-reference person and number of the grammatical subject and transitive verbs, in addition, inflect for definiteness<sup>12</sup> and, marginally, person of the direct object. Non-finite forms are: the infinitive, a nominal verb form, which shows inflection for the person and number of the possessor in certain constructions; three participles, i.e. adjectival verb forms (active/present, passive/past, and future); and two converbs or gerunds, i.e. adverbial verb forms (simple/simultaneous and perfective).

SR verbs, too, are richly inflected. The categories of aspect, tense, and mood combine into six subparadigms: present–subjunctive, future, aorist or preterite, imperfect–conditional, counterfactual, and imperative. All TAM forms are synthetic. Unlike all other verbs, the copula and verb of existence (‘to be’) does not encode aspect in the past and has three additional distinctly encoded subparadigms: conditional, present subjunctive, and past subjunctive. All verbs cross-reference person and number of the grammatical subject. There are only two productive non-finite categories: the infinitive and the passive participle (participles of one of two structural types inflect for adjectival number and gender). A sample inflectional paradigm of a borrowed verb is shown in Table 1.

Rather than borrowing the inflectional forms of Hungarian verbs, SR adopts their INFLECTIONAL STEMS. Most Hungarian verbs are base-inflected, i.e. there is an inflectional form that is markerless with regard to the inflectional stem: it is the third-person singular indefinite present indicative form, presumably the most frequent inflectional form (cf. Bybee 1985, Haspelmath 2002). Here it is ambiguous whether SR

---

<sup>12</sup> This is an oversimplified label. In addition to definiteness of the direct object, person hierarchy between the subject and the object and further factors play a role in determining which of two sets of inflections, termed ‘centrifugal’ vs ‘centripetal’ by Abandolo (1988), is employed.

borrowes the inflectional stem, or the frequent and markerless inflectional form, e.g. *ír-in-* ‘to write’  $\triangleleft$  H *ír-* inflectional stem of the verb ‘to write’  $\approx$  *ír* ‘s/he writes’. However, one class of Hungarian verbs, the so-called *ikes igék* ‘verbs with *-ik*’, are stem-inflected, i.e. they possess no inflectional form that is markerless with regard to the inflectional stem, and so the inflectional stem is not a morphosyntactically free form. Here it is unambiguously the inflectional stem, rather than the third-person singular indefinite present indicative form, that is adopted in SR, e.g. *úsz-in-* ‘to swim’  $\triangleleft$  H cf. *úsz-* inflectional stem of the verb ‘to swim’, cf. *úsz-ik* ‘s/he swims’. The statement that SR adopts the inflectional stems of Hungarian verbs is thus a correct generalization over both types of instances.

Table 1: Inflectional paradigm of the verb *irin-* ‘to turn TR’

	present– subjunctive	future	imperfect– conditional	imperative
1SG	<i>ír-in-a-v</i>	<i>ír-in-á</i>	<i>ír-in-á-hi</i>	–
2SG	<i>ír-in-e-s</i>	<i>ír-in-e-h-a</i>	<i>ír-in-e-s-ahi</i>	<i>ír-in</i>
3SG	<i>ír-in-e-l</i>	<i>ír-in-l-a</i>	<i>ír-in-l-ahi</i>	–
1PL	<i>ír-in-a-s</i>	<i>ír-in-a-h-a</i>	<i>ír-in-a-s-ahi</i>	–
2PL	<i>ír-in-e-n</i>	<i>ír-in(-e)-n-a</i>	<i>ír-in(-e)-n-ahi</i>	<i>ír-in-e-n</i>
3PL	<i>ír-in-e-n</i>	<i>ír-in(-e)-n-a</i>	<i>ír-in(-e)-n-ahi</i>	–
INF	<i>ír-in-e-n</i>			
PTC	<i>ír-in-d-o/i/e ~ ír-ime</i>			
1SG	<i>ír-in-d'-om</i>		<i>ír-in-d'-om-ahi</i>	
2SG	<i>ír-in-d'-al</i>		<i>ír-in-d'-al-ahi</i>	
3SG	<i>ír-in-d'-a</i>		<i>ír-in-d'-á-hi</i>	
1PL	<i>ír-in-d'-am</i>		<i>ír-in-d'-am-ahi</i>	
2PL	<i>ír-in-d'-an</i>		<i>ír-in-d'-an-ahi</i>	
3PL	<i>ír-in-d-e</i>		<i>ír-in-d-é-hi</i>	
	aorist		counterfactual	

There is a single exception to the non-borrowability of inflectional forms of Hungarian verbs into SR,<sup>13</sup> viz. the lexical borrowing of INFINITIVE forms, which will be discussed in more detail in Section 3.3.3. The other non-finite forms may only be borrowed if they are lexicalized and converted into adjectives or adverbs in Hungarian, e.g. SR *forró-n-o* ‘hot, boiling hot’ < H *forr-ó* id. (adjective) ← ‘boiling’ (active participle) ← *forr* ‘to boil ITR’, SR *fordítva* ‘conversely’ < H *fordít-va* id. (adverb) ← ‘turning’ (simple converb) ← *fordít* ‘to turn TR’.

### 3.2.2 Possibility verbs

Hungarian has a morphological class of POSSIBILITY verbs marked by the suffix *-hat-* ~ *-het-*, which expresses any kind of possibility, including deontic possibility (permission) and epistemic possibility (probability). While possibility is considered to be a derivational category in more traditional descriptions of Hungarian (e.g. Tompa 1968: 115), Kenesei *et al.* (1998: 359) argue that the Hungarian possibility verbs are in fact inflectional verb forms, since they are fully productive and do not combine with certain derivational affixes.

SR has no morphological category of this kind. Possibility and impossibility of different kinds are expressed by means of the uninflected modal particles *šaj* ‘can’ (possibly of West Iranian origin) and *naštig* ‘cannot’, which are used in constructions with finite TAM-inflected verbs (cf. Elšík & Matras, in prep.).<sup>14</sup> SR does not allow lexical borrowing of the Hungarian possibility verb forms: they are always rendered by the analytic possibility construction (9).

(9) BASE

DERIVATION

---

<sup>13</sup> Non-borrowability of inflectional verb forms is, of course, a property of this particular L1, SR, not a general borrowing constraint. For example, numerous dialects of Romani in contact with Turkish, Crimean Tatar, Greek or East Slavic borrow verbs in their L2 inflected forms, sometimes including even their L2 negation and/or TAM auxiliaries (Elšík & Matras 2006, Elšík & Matras, in prep.).

<sup>14</sup> In addition, SR replicates the Hungarian participant-internal possibility (capability) construction with the personal modal verb *tud* ‘to know’ > SR *džan-* ‘to know’ plus an infinitive complement (cf. Elšík & Matras, in prep.).

H	<i>úsz-ik</i> 's/he swims' ▽	<i>úsz-hat(-ik)</i> 's/he can (may, is able to) swim'
SR	<i>ús-in-el</i> id.	* <i>úshat-in-el</i> <i>šaj ús-in-el</i> id.

### 3.2.3 Frequentative verbs

Hungarian possesses a number of FREQUENTATIVE or ITERATIVE de-verbal derivations (cf. Tompa 1968: 109–111, Kenesei *et al.* 1998: 360), though only the suffix *-(V)gat-* ~ *-(V)get-* is fully productive (10). The meanings of the derived verbs are often strongly lexicalized (10b).

(10)	BASE	DERIVATION
a.	<i>lép</i> 'to step' <i>olvas</i> 'to read'	<i>lép-eget</i> 'to step repeatedly' <i>olvas-gat</i> 'to read from time to time'
b.	<i>hall</i> 'to hear' <i>mos</i> 'to wash'	<i>hall-gat</i> 'to listen, to be silent' <i>mos-ogat</i> 'to wash the dishes'

The category of frequentatives or iteratives has also developed in SR, most likely due to replication from Hungarian. SR derives frequentatives by means of the indigenous suffix *-ker-* ~ *-ger-* (11), whose original function was to mark transitivity or valency increase (cf. Matras 2002). The first part of the extended allomorph *-in-ger-* must be cognate with the Greek-origin loan-verb adaptation marker *-in-* (see Section 3.1), and has probably resulted from analogical extension from frequentatives of borrowed verbs. The frequentative derivation is fully productive and may apply to loan-verbs and to all kinds of derived verbs, including primary frequentatives (11b). The meanings of the derivation include repetition or frequent occurrence of an action, or its distribution over several referents (subjects or objects). Often, one and the same derived verb may have several of these meanings, depending on context.

(11)	BASE	DERIVATION
a	<i>líp-in-</i> ‘to step’ $\triangleleft$ H <i>lép</i> <i>gen-</i> ‘to read’ <i>phuč-</i> ‘to ask’	<i>líp-in-ger-</i> ‘to step repeatedly’ <i>gen-in-ger-</i> ‘to read from time to time’ <i>phuč-in-ger-</i> ‘to ask many questions’
b.	<i>phuč-in-ger-</i> ‘to ask many questions’	<i>phuč-in-ger-ker-</i> ‘to ask many q. often’

Loanwords of Hungarian frequentatives into SR appear to be restricted to the more lexicalized ones. In addition, there are mostly no loanwords of the base verbs of borrowed Hungarian frequentatives, and so the imported Hungarian-origin frequentative marker is probably not identifiable as a morpheme within SR. This is illustrated in (12) and (13).

(12)	BASE	DERIVATION
H	<i>mos</i> ‘to wash’	<i>mos-ogat</i> ‘to wash the dishes’ $\nabla$
SR	* <i>moš-in-</i> <i>thov-</i> ‘to wash’	<i>mošogat-in-</i> ‘to wash the dishes’ <i>thov-ker-</i> ‘to wash repeatedly’
(13)	BASE	DERIVATION
H	<i>hall</i> ‘to hear’	<i>hall-gat</i> ‘to listen, to be silent’ $\nabla$
SR	* <i>hall-in-</i> <i>šun-</i> ‘to hear, to listen’	<i>halgat-in-</i> ‘to be silent’ (*‘to listen’) <i>šun-in-ger-</i> ‘to listen repeatedly’

### 3.2.4 Valency-changing verbs

On the valency-increasing side, Hungarian possesses the productive category of CAUSATIVE verbs, which are marked by means of the suffix *-(t)at-* ~ *-(t)et-* (cf. Kenesei *et al.* 1998: 359–360, Tompa 1968: 112–113).<sup>15</sup> The productive internal derivation of

<sup>15</sup> Some descriptions (e.g. Tompa 1968) appear to use the term ‘factitives’ for causatives derived from transitive verbs, reserving the term ‘causatives’ only for causatives from intransitive verbs.

causatives, borrowing of Hungarian causatives, and the extraction of an Hungarian-origin causative suffix in SR will be discussed in detail in Section 3.3.2.

There are two common valency-decreasing derivations in Hungarian, those marked by the productive suffix *-ód-* ~ *-őd-* and those marked by the common but “hardly productive” suffixes *-kod-* ~ *-ked-* ~ *-köd-* and *-koz-* ~ *-kez-* ~ *-köz-* (Kenesei *et al.* 1998: 360–361, but cf. Tompa 1968: 113–115). Both are usually termed ‘reflexive’ verbs in Hungarian grammarography, but if I understand the descriptions correctly, the former suffix appears to be best described as ANTICAUSATIVE, i.e. expressing spontaneous non-agentive events (cf. Haspelmath’s 1993 inchoatives). The latter two suffixes have REFLEXIVE, RECIPROCAL and apparently also various other MIDDLE (mediopassive) functions. In addition, there is an obsolete passive derivation in Hungarian (cf. Kenesei *et al.* 1998: 361), which need not concern us here, as it is likely to be absent in the Hungarian dialect of Selice.

SR has a common, but not completely productive, valency-decreasing ANTICAUSATIVE derivation, marked by the middle suffix *-(j)ov-* (14a–b) and its various extended allomorphs (this is also the marker of de-adjectival intransitive verbs or inchoatives, see Section 3.2.5). The meaning of the derived verb is often strongly lexicalized and unpredictable (14b). Middle verbs derived from intransitives remain intransitive, and express INTENSIVE actionality modification of the base verb (14c). There are no reflexive, reciprocal or passive derivations in SR; these functions are expressed through analytic constructions.<sup>16</sup>

(14)	BASE	DERIVATION
a.	<i>prav-</i> ‘to open TR’, PTC <i>prá-d-o</i>	<i>prá-d’-ov-</i> ‘to open ITR’
	<i>šun-</i> ‘to hear’, PTC <i>šun-d-o</i>	<i>šun-d’-ov-</i> ‘to be audible’
	<i>alakh-</i> ‘to find’	<i>alath’-ov-</i> ‘to be found’
b.	<i>dikh-</i> ‘to see, to look’	<i>dit’h’-ov-</i> ‘to seem, to look like’
	<i>muk-</i> ‘to leave, to let, to drop’	<i>muk-<u>isaj</u>-ov-</i> ‘to fall into bad ways’
	<i>phud-</i> ‘to blow’	<i>phud-<u>isaj</u>-ov-</i> ‘to get annoyed’

---

<sup>16</sup> Though the English translations of some of the anticausative derivations in (14a) seem to indicate passive functions, these derived verbs do not allow overt expression of agents and are conceived as non-agentive.

- c. *asa-* ‘to laugh, to smile’                      *asa-saj-ov-* ‘to guffaw’  
*khand-* ‘to stink’                                      *khand-isaj-ov-* ‘to stink intensively’

Loanwords of the Hungarian anticausative verbs in *-ód-* ~ *-őd-* are unattested in SR. Loanwords of the Hungarian verbs in *-kVd-* are attested, although their base verbs are usually not borrowed (15–16). Finally, loanwords of the Hungarian verbs in *-kVz-* are attested together with loanwords of their base verbs (17–18), but the imported suffix does not get extracted in SR.

- |      |  |   |
|------|--|---|
| (15) | BASE   | DERIVATION                                  |
| H    | <i>visel</i> ‘to carry, to bear’                               | <i>visel-ked-ik</i> ‘to behave, to act’     |
|      |  | ▽   |
| SR   | * <i>višel-in-</i><br>[cf. <i>ledž-</i> ‘to carry’]            | <i>višéked-in-</i> id.                      |
|      |  |   |
| (16) | BASE   | DERIVATION                                  |
| H    | <i>keres</i> ‘to look for, to earn’                            | <i>keres-ked-ik</i> ‘to trade, to barter’   |
|      | ▽  | ▽   |
| SR   | * <i>kereš-in-</i><br>[cf. <i>rod-</i> ‘to look for, to earn’] | <i>kerešked-in-</i> id.                     |
|      |  |   |
| (17) | BASE   | DERIVATION                                  |
| H    | <i>gondol</i> ‘to think’                                       | <i>gondol-koz-ik</i> ‘to think, to reflect’ |
|      | ▽  | ▽   |
| SR   | <i>gondul-in-</i> id.  | <i>gondú(-)koz-in-</i> id.                  |
|      |  |   |
| (18) | BASE   | DERIVATION                                  |
| H    | <i>talál</i> ‘to find, to hit the aim’                         | <i>talál-koz-ik</i> ‘to meet’               |
|      | ▽  | ▽   |
| SR   | <i>talál-in-</i> ‘to guess’                                    | <i>talá(-)koz-in-</i> id.                   |

### 3.2.5 Factitive and inchoative verbs

There are two classes of productive de-adjectival verb derivations in Hungarian that, in a sense, parallel the valency-changing de-verbal derivations (cf. Tompa 1968: 117–119, Kenesei *et al.* 1998: 361–362). First, the suffix *-ít*, rarely *-sít*, derives transitive verbs meaning ‘to make [so/sth.] ADJ’, which may be termed FACTITIVE verbs (19a). Second, intransitive verbs meaning ‘to become or get ADJ(er)’ or ‘to make oneself ADJ’, which may be termed INCHOATIVE verbs, are derived by means of the suffix *-od-* ~ *-ed-* ~ *-öd-* or the less productive, but common, suffix *-ul* ~ *-ül* (19b).

(19)	BASE	DERIVATION
a.	<i>erős</i> ‘strong, powerful’	<i>erős-ít</i> ‘to strengthen’
	<i>fiatal</i> ‘young’	<i>fiatal-ít</i> ‘to make young(er)’
	<i>mély</i> ‘deep’	<i>mély-ít</i> ‘to deepen’
	<i>gyenge</i> ‘weak’	<i>gyeng-ít</i> ‘to weaken’
	<i>kész</i> ‘ready, prepared’	<i>kész-ít</i> ‘to prepare’
b.	<i>erős</i> ‘strong, powerful’	<i>erős-öd-ik</i> ‘to become strong’
	<i>fiatal</i> ‘young’	<i>fiatal-od-ik</i> ‘to become young(er)’
	<i>mély</i> ‘deep’	<i>mély-ed</i> ‘to get deep(er)’
	<i>gyenge</i> ‘weak’	<i>gyeng-ül</i> ‘to become weak(er)’
	<i>kész</i> ‘ready, prepared’	<i>kész-ül</i> ‘to prepare oneself’

Both de-adjectival categories are also encountered in SR. Factitives are derived by means of the suffix *-(j)ar-* ~ *-(j)ár-* (20a), and inchoatives are derived by means of the suffix *-(j)ov-* (20b).<sup>17</sup> Both derivations may apply to all pre-Hungarian qualitative adjectives and to several classes of internally derived qualitative adjectives. It remains to be investigated, however, how productive they are with Hungarian-origin adjectives, though they do apply at least to some of them.

<sup>17</sup> The complex allomorphy of the derivations is not well understood yet: both suffixes possess various extended variants and may, but need not, trigger deletion of the adaptation marker of their borrowed adjective bases.

(20)	BASE	DERIVATION
a.	<i>zor-ál-o</i> ‘strong, powerful’	<i>zor-aj-ár-</i> ‘to strengthen’
	<i>térn-o</i> ‘young’	<i>térn-isaj-ár-</i> ‘to make young(er)’
	<i>míl-n-o</i> ‘deep’ ◁ H <i>mély</i>	<i>míl-ň-ar-</i> ‘to deepen’
	<i>d’eng-av-o</i> ‘weak’ ◁ H <i>gyenge</i>	<i>d’eng-isaj-ár-</i> ‘to weaken’
b.	<i>zor-ál-o</i> ‘strong, powerful’	<i>zor-aj-ov-</i> ‘to become strong’
	<i>térn-o</i> ‘young’	<i>térn-isaj-ov-</i> ‘to become young(er)’
	<i>míl-n-o</i> ‘deep’ ◁ H <i>mély</i>	<i>míl-ň-ov-</i> ‘to become deep(er)’
	<i>d’eng-av-o</i> ‘weak’ ◁ H <i>gyenge</i>	<i>d’eng-isaj-ov-</i> ‘to become weak(er)’

Loanwords of Hungarian factitive verbs in SR are mostly accompanied by loanwords of their base adjectives (21–22). However, the opposite does not hold: most loanwords of Hungarian adjectives are not accompanied by loanwords of corresponding factitive derivations. Despite its morphemic identifiability, the imported factitive suffix *-ít-* does not undergo any lexical extension within SR.

(21)	BASE	DERIVATION
H	<i>kész</i> ‘ready, prepared’	<i>kész-ít</i> ‘to prepare’
	▽	▽
SR	<i>kís-n-o</i> id.	<i>kís(-)ít-in-</i> id.
(22)	BASE	DERIVATION
H	<i>szabad</i> ‘free’	<i>szabad-ít</i> ‘to liberate’
	▽	▽
SR	<i>sabad-n-o</i> id.	<i>sabad(-)ít-in-</i> id.

While lexical borrowing of Hungarian factitives appears to be fairly rare, loanwords of Hungarian inchoatives are simply unattested, and may be missing in SR. Intransitive counterparts to SR loanwords of Hungarian factitives are usually formed through analytic reflexivization, e.g. *kísít-in-av* ‘I prepare [sth.]’ → *kísít-in-av man* ‘I

prepare myself, I am getting ready’ (there is no SR \**kísil-in-* < H *kész-ül* ‘to prepare oneself, to get ready’).

### 3.2.6 De-nominal verbs

There are a number of de-nominal verb derivations in Hungarian (cf. Tompa 1968: 117–118, Kenesei *et al.* 1998: 357–358). Perhaps most productive are the derivations in *-(V)z-* and *-(V)l-*, which have a wide range of functions. Both semantically and formally related to the de-adjectival factitive verbs are de-nominal transitive verbs that are derived by means of the factitive suffix, though the allomorph *-(V)sít*, which is rare with de-adjectival factitives, appears to prevail with the de-nominal derivations. These derivations may be termed PSEUDO-FACTITIVES. Finally, there are also de-nominal (and de-adjectival) verbs in the suffix *-(s)kod-* ~ *-(s)köd-*, meaning ‘to behave as, to work as’.

De-nominal verb derivation is much less developed in SR, and there is no productive pre-Hungarian derivation. Although SR borrows different kinds of de-nominal verbs (23–24), and although some of the imported verb-deriving suffixes are identifiable within SR (24), only a reflex of the Hungarian suffix *-(V)z-* got actually extracted (see Section 3.3.1).

(23)	BASE	DERIVATION
H	<i>semmi</i> ‘nothing’	<i>semmi-sít</i> ‘to destroy’
		▽
SR	* <i>šemmi</i> [cf. <i>ništa</i> ‘nothing’ < Ik]	<i>šemmiššít-in-</i> id.
(24)	BASE	DERIVATION
H	<i>kasza</i> ‘scythe’	<i>kaszá-l</i> ‘to scythe’
	▽	▽
SR	<i>kas-a</i> id.	<i>kas(-)ál-in-</i> id.

### 3.3 *Extracted affixes*

#### 3.3.1 The de-nominal suffix *-áz-*

The SR suffix *-áz-* is a productive means to derive verbs from nouns. The meanings of the derived verbs are varied. The verbs derived from non-human nouns may have the following meanings: ‘to produce N’, ‘to use N’, ‘to work on N’, ‘to collect N’, ‘to catch N’, ‘to consume N constantly, to be fond of consuming N’, and more (25). Most of these derived verbs have lexicalized meanings, some of which may be rather idiosyncratic, e.g. *máčh-o* ‘fish’ → *máčh-áz-in-* ‘to catch bait for fishing’.<sup>18</sup> Verbs derived from human nouns, on the other hand, are mostly occasionalisms that are assigned the pragmatically most relevant meaning from the following semantic domain: ‘to be in touch with N constantly, to visit N frequently’, ‘to be after N’, ‘to speak about N constantly’, ‘to speak like N’ or similar (26).<sup>19</sup>

(25) BASE	DERIVATION
<i>ásva</i> ‘teardrop’	<i>ásv-áz-in-</i> ‘to shed tears’
<i>paramis-i</i> ‘fairy-tale’	<i>paramis-áz-in-</i> ‘to tell fairy-tales’
<i>d’íl-i</i> ‘song’	<i>d’ij-áz-in-</i> ‘to sing’
<i>perhas</i> ‘joke’	<i>perhas-áz-in-</i> ‘to make jokes’
<i>mirikl-i</i> ‘bead’	<i>mirikj-áz-in-</i> ‘to string beads’
<i>čhúr-i</i> ‘knife’	<i>čhurj-áz-in-</i> ‘to fight with knives’
<i>čuňik</i> ‘whip’	<i>čuňik-áz-in-</i> ‘to whip’
<i>džukel</i> ‘dog’	<i>džukl-áz-in-</i> ‘to play with dogs’
<i>phuv-ja</i> ‘field(s)’ <sup>20</sup>	<i>phuvj-áz-in-</i> ‘to work on field(s)’
<i>huhur</i> ‘mushroom’	<i>huhur-áz-in-</i> ‘to collect mushrooms’
<i>trast</i> ‘iron’	<i>trast-áz-in-</i> ‘to collect iron’
<i>čirikl-i</i> ‘bird’	<i>čirikj-áz-in-</i> ‘to catch birds’

<sup>18</sup> Semantically different from the loanword *halás-in-* ‘to fish’ < H *halász-ik*.

<sup>19</sup> The translations in (4) indicate the most common meanings and do not exhaust the potential polysemy of the derivations.

<sup>20</sup> Onomasiological plural of the noun *phú* ‘earth, land, country’.

*máčh-o* ‘fish’

*ákhor* ‘nut’

*mas* ‘meat’

*mol* ‘wine’

*máčh-áz-in-* ‘to catch bait for fishing’

*ákhor-áz-in-* ‘to eat/collect nuts’

*mas-áz-in-* ‘to eat meat constantly’

*mol-áz-in-* ‘to drink wine constantly’

(26) BASE

*dél* ‘God’

*rašaj* ‘priest’

*čhaj* ‘Gypsy girl’

*rakl-i* ‘non-Gypsy girl’

*gadž-i* ‘non-Gypsy woman’

*pojáckiň-a* ‘Vlax G. woman’

*phen* ‘sister’

DERIVATION

*dévl-áz-in-* ‘to speak about God often’

*rašaj-áz-in-* ‘to talk like a priest’

*čhaj-áz-in-* ‘to be after Gypsy girls’

*rakj-áz-in-* ‘to be after non-Gypsy girls’

*gadž-áz-in-* ‘to visit non-G. women’

*pojáckiň-áz-in-* ‘to visit Vlax G. women’

*pheň-áz-in-* ‘to visit sister(s) often’

The derived verbs are mostly intransitive, though a few may be also used transitively, e.g. *čuňik-áz-in-* ‘to whip’, *d’ij-áz-in-* ‘to sing’. The base nouns are mostly indigenous, though a few are pre-Hungarian loanwords, e.g. *paramisi* ‘fairy-tale’ < Greek, or derivations based on pre-Hungarian loanwords, e.g. *pojác-kiň-a* ‘Vlax Gypsy woman’ ← *poják-o* ‘Vlax Gypsy (man)’ < Ikavian. The suffix *-áz-* itself does not show any allomorphy, though it does trigger complex allomorphic variation in the base. In most instances, the derivational stem of the verb inherits the irregularities of a PRE-OBLIQUE stem of the noun, on which it is based, e.g. *dél*, oblique stem *dévl-es-*, pre-oblique stem *dévl-* → *dévl-áz-in-*. The suffix *-áz-* must be followed by the loan-verb adaptation suffix *-in-* (see Section 3.1).

The category of de-nominal verbs had been present in SR before the introduction of the derivation in *-áz-*, but the competing derivations apply only to a few nouns and are not productive, e.g. *dand* ‘tooth’ → *dand-er-* ‘to bite’, *hól-i* ‘anger’ [< Greek] → *hoj-ar-* ‘to make angry’, *jár-o* ‘flour’ → *já-jar-* ‘to put too much flour’, *kan* ‘ear’ → *kan-d-* ‘to obey’, *khaň* ‘fart’ → *khaň-ar-* ‘to fart’, *likh* ‘louse’ → *likh-ajár-* ‘to delouse’.

The source morpheme of the SR suffix *-áz-*, the Hungarian suffix *-(V)z-*, is a productive marker of de-nominal verbs, which are commonly borrowed into SR.

Importantly, the extracted SR suffix does not correspond in shape to any allomorph of the Hungarian suffix, and so the extracted suffix must have resulted from re-analysis of morpheme boundaries (27).

(27)	BASE	DERIVATION
H	<i>cigaretta</i> ‘cigarette’	<i>cigaretta-z-ik</i> ‘to smoke cigarettes’
	▽	▽
SR	<i>cigaretta-a</i> id.	<i>cigaretta(-)áz-in-</i> id.

### 3.3.2 The causative suffix *-atat-*

The SR suffix *-atat-* is a productive marker of causativity. It derives causative verbs from two synchronic classes of verbs. First, it applies to all verbs that are derived by the productive de-substantival suffix *-áz-* (see Section 3.3.1), e.g. *huhur-áz-atat-in-av-* ‘to make [so.] collect mushrooms’ ← *huhur-áz-in-* ‘to collect mushrooms’ (← *huhur* ‘mushroom’). Second, it applies to three verbs that contain the loan-verb adaptation suffix *-in-* on synchronic analysis but are not derived by the de-substantival suffix *-áz-*, viz. *pot’-atat-in-av-* (alongside *pot’-in-av-*) ‘to make [so.] pay’ ← *pot’-in-* ‘to pay’, *prut’-atat-in-av-* (alongside *prut’-in-av-*) ‘to make [so.] kick’ ← *prut’-in-* ‘to kick’, and *uš-atat-in-av-* (alongside *uš-in-av-*) ‘to make [so.] roll’ ← *uš-in-* ‘to roll’.<sup>21</sup> The use of the suffix *-atat-* is obligatory in causatives derived from the first class of verbs but optional, though common, in those derived from the second class of verbs. While the suffix does not trigger any allomorphy in the morphemes of its derivational base, it is unique among derivational suffixes in that it is not the final suffix of (the inflectional stem of) the base

---

<sup>21</sup> These are indigenous verbs that have been re-analyzed as containing the loan-verb adaptation suffix *-in-*. Curiously, the suffix *-(a)tat-* does not apply to pre-Hungarian (Greek or Ikavian) loan-verbs that contain the adaptation suffix, e.g. *vič-in-av-* ‘to make [so.] shout’ (\**vič-atat-in-av-*) ← *vič-in-* ‘to shout’. My native speaker consultants accepted the causative *mol-atat-in-av-* ‘to make [so.] pray’, which is derived from the Ikavian-origin verb *mol-in-* ‘to pray’, as grammatical but they appear to use only the causative *mol-in-av-*, i.e. the one without the suffix *-(a)tat-*, in spontaneous discourse.

verb. Instead, it is inserted immediately before the adaptation suffix *-in*,<sup>22</sup> which is the final suffix of the derivational base.

While the category of morphological causatives has been inherited from Indo-Aryan, dialect comparison within Romani suggests that its retention and productivity in SR is due to pattern replication from Hungarian (cf. Hübschmannová & Bubeník 1997, Matras 2002: 120, Elšík & Matras 2006: 432).<sup>23</sup> The Hungarian-origin suffix *-atat-* competes with two indigenous causative suffixes in SR, viz. the unproductive *-jar-* ~ *-ajár-*<sup>24</sup> and the productive *-av-*. The former applies to several closed classes of indigenous verbs, e.g. *beš-ajár-* ‘to seat’ ← *beš-* ‘to sit’. The latter can apply to almost any verb, including some of those that usually take the unproductive indigenous suffix, e.g. *beš-av-* ‘to seat’ ← *beš-* ‘to sit’. It may also apply to causative verbs based on intransitives, thus deriving the so-called second causatives, e.g. *beš-ajár-av-* or *beš-av-av-* ‘to make [so.] seat’.

As the reader may have noticed above, the causatives derived by the suffix *-atat-* contain in addition the productive causative suffix *-av-* as the final morpheme of the derivation’s inflectional stem, e.g. *huhur-áz-atat-in-av-* ‘to make [so.] collect mushrooms’. This pattern, however, does not represent the second causatives (which would contain one more instance of the suffix *-av-*, e.g. *huhur-áz-atat-in-av-av-* ‘to make [so.] make [so.] collect mushrooms’), but double marking of the first causatives. Double causative marking is also typical of borrowings from Hungarian: although causatives of Hungarian loan-verbs may be a) morphologically adapted lexical borrowings of Hungarian causatives, e.g. *dógoz-tat-in-* ‘to make [so.] work’ < Hungarian *dolgoz-tat*, or b) internal derivations from a non-causative loan-verb by the productive indigenous causative marker, e.g. *dógoz-in-av-* ‘to make [so.] work’ ← *dógoz-in-* ‘to work’ <

---

<sup>22</sup> The unusual linear order of the causative suffix *-(a)tat-* still does not make it an infix, as there is a (synchronic) morphemic boundary between the morphemes that surround it.

<sup>23</sup> Morphological causatives are much less productive, for example, in those closely related South Central dialects whose speakers are no longer bilingual in Hungarian. Still, causative morphology appears to be more developed in Romani than in most of its European contact languages, which show preference for anticausative derivation (cf. Haspelmath 1993: 102).

<sup>24</sup> The suffix *-jar-* ~ *-(j)ajár-* is productive as a marker of factitive verbs, i.e. transitive verbs derived from adjectives, e.g. *mát-ar-* ‘make drunk’ ← *máto* ‘drunk’.

Hungarian *dolgoz-ik*, the by far most common strategy is c) the combination of both mechanisms, e.g. *dolgoz-tat-in-av-* ‘to make [so.] work’. This pattern of double causative marking also elucidates the unusual linear order of the suffix *-atat-* in the internal causative derivations.

In Hungarian, causative verbs can be marked by several different suffixes, though only two of them are productive, viz. the ‘short’ *-at-* ~ *-et-* and especially the ‘long’ *-tat-* ~ *-tet-* (Kenesei *et al.* 1998: 359–360). While the selection of one or the other productive suffix is to some extent lexically determined, the number of syllables of the inflectional stem of the derivational base is the most important factor: most monosyllabic verbs take the ‘short’ suffix and most polysyllabic verbs take the ‘long’ one (cf. Tompa 1968: 112–113). Curiously, both suffixes also form part of passive derivations, cf. the causative *ír-at* ‘to make [so.] write’ vs the passive *ír-at-ik* ‘to be written’ (the suffix *-ik* is an inflectional marker) ← *ír* ‘to write’ (Kenesei *et al.* 1998: 361), and so their labelling as ‘causative’ markers is somewhat misleading. Lexical borrowings of Hungarian causatives into SR, which are abundant, retain the source language variation in causative marking, e.g. *ír-at-in-av-* < *ír-at* ‘to make [so.] write’, *dolgoz-tat-in-av-* < *dolgoz-tat* ‘to make [so.] work’, *mesé-tet-in-av-* ‘to make [so.] whiten’ < *meszel-tet* etc.<sup>25</sup> SR does not borrow the Hungarian passives, which, given the inflectional nature of the suffix *-ik*, would be indistinguishable from the borrowed causatives and which, moreover, are obsolete in current Hungarian (Kenesei *et al.* 1998: 361) and are likely to be lacking in the contact dialect.

Importantly, the extracted SR causative suffix *-atat-* does not correspond in shape to any of the Hungarian causative markers, although historically it must be based on their back allomorphs, i.e. *-at-* and/or *-tat-*. Should SR follow the Hungarian rules of allomorphy, causatives derived from the *-áz-* verbs, whose pre-causative stems are always polysyllabic (the suffix *-áz-* counts), would select the ‘long’ back suffix *-tat-*; and causatives derived from the other *-in-* verbs, whose pre-causative stems are monosyllabic (the suffix *-in-* does not count), would select the ‘short’ back suffix *-at-*. However, the

---

<sup>25</sup> The unusual allomorph *-tot-* in SR *keheg-tot-in-av-* ‘to make [so.] caught’ probably reflects Hungarian dialectal *köhög-töt* (cf. standard *köhög-tet*). Still, the phonological adaptation /ö > o/ (backing) rather than the usual /ö > e/ (unrounding) in SR remains puzzling.

employment of the ‘short’ *-at-* is ungrammatical with all causatives of the latter class, e.g. *\*pot’-at-in-av-*, and the employment of the ‘long’ *-tat-* is ungrammatical with most causatives of the former class, e.g. *\*huhur-áz-tat-in-av-*. The expected allomorph *-tat-* is attested, as a free variant, with a single causative of the former class: *d’ij-áz-tat-in-av-* (alongside *d’ij-áz-atat-in-av-*) ‘to make [so.] sing’ ← *d’ij-áz-in-* ‘to sing’ (← *d’il-i* ‘song’).

### 3.3.3 The infinitive suffix *-ñi*

The SR suffix *-ñi* is an infinitive marker that applies to de-nominal verbs derived by the suffix *-áz-*, e.g. stem *huhur-áz-in-* ‘to collect mushrooms’ → infinitive *huhur-áz-ñi*. Since the suffix *-ñi* may apply to any verb in *-áz-* and since the latter are productively derived, the infinitive suffix itself is productive. However, the suffix must be considered derivational in SR, as it applies neither to pre-Hungarian or post-Hungarian underived verbs, e.g. *mol-in-* ‘to pray’ → *\*mol-ñi* or *sledov-ál-in-* ‘to observe, to follow’ → *\*sledov-ál-ñi*, nor to other internally derived verbs, including derivations from the verbs in *-áz-*, e.g. *huhur-áz-in-* ‘to collect mushrooms’ → *huhur-áz-atat-in-* ‘to make [so.] collect mushrooms’ → *\*huhur-áz-atat-ñi*. The suffix *-ñi*, which does not show or trigger any allomorphy, is syntagmatically incompatible with the loan-verb adaptation marker *-in-*, which it replaces in the infinitive derivation.

The infinitive suffix *-ñi* competes with a fully productive morphosyntactic construction in SR that has been termed the ‘new’ infinitive<sup>26</sup> (Boretzky 1996) or the ‘subjunctive’ infinitive (Elšík, in press). All SR verbs, including those that may form the infinitive in *-ñi*, can be used in the subjunctive infinitive construction. Synchronically, the subjunctive infinitive is a non-finite subordinate verb form which is homonymous to the third-person plural subjunctive form, irrespective of the subject categories (viz. person and number) of the matrix verb. The subjunctive infinitive has developed from Early Romani finite subordinate constructions, through fossilization (obligatorification) of a

---

<sup>26</sup> The ‘old’, Indo-Aryan, infinitive has developed in Romani into the derivational category of de-verbal action nouns in *\*-iben* (SR *-ibe*) (Michael Beníšek, p.c.).

frequent finite form of the subordinate verb: the third-person plural form in case of SR.<sup>27</sup> The development of the subjunctive infinitive in SR is unambiguously due to pattern replication from Hungarian (Elšík *et al.* 1999, Elšík, in press). Both infinitives, the subjunctive one and the one marked by the suffix *-ñi*, are used in clausal complements of subject-inflected modal predicates (28), in clausal complements of some manipulative verbs (29), and in tightly integrated same-subject purpose clauses (30). Unlike the subjunctive infinitive, which must be introduced by the non-factual complementizer *te*, the infinitive in *-ñi* does not allow any complementizer.

- (28) a. *Kam-l-ahi te džukl-áz-in-en-Ø.*  
 want-3SG-REM COMP dog-VERB-LOAN-3PL-SUBJ
- b. *Kam-l-ahi džukl-áz-ñi.*  
 want-3SG-REM dog-VERB-INF  
 ‘S/he liked to play with dogs.’
- (29) a. *Muk-Ø man-Ø te d’ij-áz-in-en-Ø.*  
 leave-IMP.2SG I.OBL-ACC COMP song-VERB-LOAN-3PL-SUBJ
- b. *?uk-Ø man-Ø d’ij-áz-ñi.*  
 leave-IMP.2SG I.OBL-ACC song-VERB-INF  
 ‘Let me sing!’
- (30) a. *Dža-s-Ø te huhur-áz-in-en-Ø.*  
 go-1PL-PRES COMP mushroom-VERB-LOAN-3PL-SUBJ
- b. *Dža-s-Ø huhur-áz-ñi.*  
 go-1PL-PRES mushroom-VERB-LOAN-INF  
 ‘We are going to go and collect mushrooms.’

---

<sup>27</sup> Other Romani dialects that possess the subjunctive infinitive have fossilized the third-person singular or the second-person singular forms. Many Romani dialects, including all Romani dialects spoken in the Balkans, have not developed the subjunctive infinitive at all, thus retaining the Balkan-type finiteness in all kinds of subordinate clauses.

The source morpheme of the borrowed SR infinitive suffix, the Hungarian suffix *-ni* (dialectal *-nyi*), has the identical function of marking the infinitive. Unlike the Hungarian-origin SR infinitive in *-ňi*, however, the Hungarian infinitive inflects for person and number of the possessor in certain constructions (cf. Kenesei *et al.* 1998: 318). Like the SR infinitive in *-ňi*, the Hungarian infinitive does not allow a complementizer.

Lexical borrowings of Hungarian infinitives, which are employed in the same range of constructions as the SR infinitives internally derived by the suffix *-ňi*, are well attested in SR, e.g. *dógozňi* ‘to work’  $\triangleleft$  H *dolg-oz-ni*, cf. the inflectional stem *dógoz-in-* ‘to work’  $\triangleleft$  H *dolg-oz-*. Nevertheless, lexical borrowing of numerous Hungarian verbs is not matched by lexical borrowing of their infinitives, e.g. *ír-in-* ‘to write’  $\triangleleft$  H *ír*, but not *\*írňi*  $\triangleleft$  H *ír-ni*. The only option in such cases is to use the subjunctive infinitive of the morphologically adapted loan-verb, e.g. *(te) ír-in-en* ‘to write’. My corpus of SR actually only contains infinitive loanwords of those verbs that are, in Hungarian, derived by the de-nominal suffix *-(V)z-*. It is hardly an accident that the Hungarian-origin SR suffix *-ňi* is bound precisely to that SR suffix, viz. *-áz-*, that results from extraction of the Hungarian de-nominal suffix *-(V)z-*, and so we must conclude that the suffix *-ňi* has been extracted only in conjunction with the suffix *-áz-* (see Section 3.3.1). The constraint on lexical borrowing of Hungarian infinitives, however, remains to be explained.

### 3.3.4 The loan-verb adaptation suffix *-ál-*

The SR suffix *-ál-* is a productive loan-verb adaptation marker. It applies to nonce borrowings of ‘post’-Hungarian, i.e. Slovak or Czech, verbs, e.g. SR *klik-ál-in-* ‘to click’  $\triangleleft$  Slovak *klik-(a-t)* or *sledov-ál-in-* ‘to observe, to follow’  $\triangleleft$  Slovak *sled-ov-(a-t)*. The adopted stem may be termed the PRE-INFINITIVE stem: it corresponds to the infinitive stem of the Slovak or Czech verb without the ‘thematic’ (inflectional class) suffix. However, loanwords of those Slovak or Czech verbs that employ the complex ‘thematic’ marker *-ov-a-* in their infinitive stem retain the extension suffix *-ov-* in the adopted stem. Interestingly, the Slovak or Czech verbs that get borrowed into SR are all imperfective: there appears to be no mechanism to borrow Slovak or Czech perfective verbs, e.g.

\**klik-n-ál-* < Slovak *klik-n-(ú-t)*. The suffix *-ál-*, which is specialized for adopting ‘post’-Hungarian verbs, is obligatorily followed by the Greek-origin suffix *-in-*, which is the general loan-verb adaptation suffix in SR. The suffix *-ál-* does not have any allomorphs.

The source morpheme of the SR suffix is the Hungarian suffix *-ál-*. Though noting that “several seemingly derived forms have no nominal or any other bases”, Kenesei *et al.* (1998: 358) describe the Hungarian suffix as a de-nominal verb-deriving marker with two allomorphs, *-ál-* ~ *-izál-*. They assume, for example, the following derivations, some of them involving truncation of the base stem: *kodifikáció* ‘codification’ → *kodifik-ál* ‘to codify’, *politika* ‘politics’ → *polit-izál* ‘to be engaged in politics’, *bagatell* ‘trifle’ → *bagatell-izál* ‘to make look like a trifle’, *modern* ‘modern’ (adjective!) → *modern-izál* ‘to modernize’. This attempt to find a synchronic derivational relationship between pairs of nominal and verbal loanwords is mistaken, in my view. Arguably, the Hungarian suffix has a single allomorph, *-ál-*, and the same function as the SR suffix, viz. to adapt borrowed verbs, cf. *kodifik-ál*, *politiz-ál*, *bagatelliz-ál*, *moderniz-ál*.<sup>28</sup> Note that Hungarian adaptation suffix *-ál-* is synchronically distinct from the de-nominal verb-deriving suffix *-(V)l* (cf. Kenesei *et al.* 1998: 357), from which it has probably developed through re-analysis.<sup>29</sup>

---

<sup>28</sup> This appears to confirm Bakker’s (1997) point that loanword adaptation markers in an L1 (Romani) may result from borrowing of L2 (Greek) adaptation markers.

<sup>29</sup> In Elšík (in press) I have wrongly assumed that the SR adaptation suffix *-ál-* has developed through morphemic re-analysis of the de-nominal Hungarian suffix *-(V)l-*, none of whose allomorphs has the shape *-ál-*, and that the functional shift in SR from a de-nominal marker to a loan-verb adaptation marker was probably triggered by the formal similarity of the stems in *-ál-* to the Slovak past participles in *-l*, e.g. SR *sledov-ál-* ≈ Slovak *sledova-l*.

## 4 Affix extraction: nouns

### 4.1 *Inflectional adaptation*

Nouns are commonly borrowed into SR. Pre-Greek and some (presumably early) Greek noun loans show full morphological integration and are structurally indistinguishable from indigenous nouns; they have a so-called *oikoclitic* morphology. Some (presumably late) Greek and post-Greek noun loans, on the other hand, have so-called *xenoclitic* morphology (Elšík & Matras 2006: 324–333), which is characterized, above all, by borrowed nominative inflections, mostly of Greek origin. These inflections were extracted from lexical loans of Greek nominative noun forms, and extended to later loanwords as well. For example, the xenoclitic nominative singular feminine suffix *-a* was extracted from Greek-origin nouns, e.g. *cip-a* ‘skin’ < *tsip-a*, and extended to nouns borrowed from Ikavian, e.g. *pét-a* ‘oven’ < *pěc*, and Hungarian, e.g. *virág-a* ‘flower’ < *virág*. Gender assignment of loan-nouns from Hungarian, a genderless language, is determined by inflectional class assignment, which is a complex issue that will not be addressed here.

### 4.2 *Morphological categories and lexical borrowing*

#### 4.2.1 **Noun inflection**

Hungarian nouns (cf. Tompa 1968, Kenesei *et al.* 1998, Siptár & Törkenczy 2000: 38–43) inflect for the categories of number, associativity, and case and, when used in various possessive constructions, also for person, number, and anaphoricity of the possessor. The category of number, both that of the noun’s referent and that of the possessor, has two values: singular and plural. The category of associativity, traditionally considered to be derivational, is encoded in human nouns by the separatist suffix *-ék-*. The number of case values is a matter of dispute, ranging between 16 and 23 in various descriptions of the language. Some traditional case forms of nouns are better considered to be de-nominal adverbial derivations. There is no grammatical gender in Hungarian.

SR nouns have an inherent gender and inflect for the categories of number and case. Human nouns, in addition, inflect for the category of associativity in the plural. Nouns may be of masculine or feminine gender. The category of number has two values: singular and plural. The category of nominal case has eight values (nominative, accusative, dative, locative, ablative, sociative, genitive, and vocative), but inanimate nouns mostly lack the accusative case. SR (and Romani in general) shows differential object marking, whereby inanimate (and, in some dialects, also non-referential animate) direct objects are marked by the nominative. The genitive case exhibits *Suffixaufnahme*, viz. inflection for gender, number, and adjectival case in agreement with its head noun. Lexical classification into inflectional classes is manifested in the selection of the so-called ‘internal’ case suffixes, which also cumulate gender and number. They are the nominative, oblique, and vocative suffixes. The so-called ‘external’ case suffixes, which are separatist and insensitive to lexical classification, are suffixed to the oblique stem; the accusative is the markerless oblique case form. A sample inflectional paradigm of a SR human noun is displayed in Table 2.

Table 2: Inflectional paradigm of the noun *lavutár-i* ‘Gypsy musician; violin player’

	SG	PL	PL.ASS
NOM	<i>lavutár-i</i>	<i>lavutár-ja</i>	<i>lavutár-i-n-ger-e</i>
ACC	<i>lavutár-i</i>	<i>lavutár-e-n[-Ø]</i>	<i>lavutár-i-n-ger-e-n[-Ø]</i>
DAT	<i>lavutár-i-s-ke</i>	<i>lavutár-e-n-ge</i>	<i>lavutár-i-n-ger-e-n-ge</i>
LOC	<i>lavutár-i-s-te</i>	<i>lavutár-e-n-de</i>	<i>lavutár-i-n-ger-e-n-de</i>
ABL	<i>lavutár-i-s-tar</i>	<i>lavutár-e-n-dar</i>	<i>lavutár-i-n-ger-e-n-dar</i>
SOC	<i>lavutár-i-ha</i>	<i>lavutár-e-n-ca</i>	<i>lavutár-i-n-ger-e-n-ca</i>
GEN	<i>lavutár-i-s-ker-o/i/e/a</i>	<i>lavutár-e-n-ger-o/i/e/a</i>	<i>lavutár-i-n-ger-o/i/e/a</i>
VOC	<i>lavutár-i-na</i>	<i>lavutár-alen</i>	<i>lavutár-i-n-ger-alen</i>

Hungarian nouns that get borrowed into SR are mostly adopted either in their nominative form or stem, which is markerless, or in their possessive stem, especially if it is irregular, e.g. SR *vej-o* ‘son-in-law’ < H *vő*, possessive stem *vej-*. Nevertheless, SR does not borrow POSSESSIVE forms of Hungarian nouns, expressing possessivity through free indigenous genitive forms of nouns and pro-nouns, cf. (31) where L stands for the nominal lexeme.

(31)	‘my son-in-law’	‘my sons-in-law’	‘our son-in-law’	‘our sons-in-law’
H	<i>vej-e-m</i>	<i>vej-e-i-m</i>	<i>vej-ünk</i>	<i>vej-e-i-nk</i>
	L-POSS-1SG	L-POSS-PL-1SG	L-POSS.1PL	L-POSS-PL-1PL
SR	<i>m-r-o</i>	<i>m-r-e</i>	<i>am-ar-o</i>	<i>am-ar-e</i>
	1SG-GEN-ADJ	1SG-GEN-ADJ	1PL-GEN-ADJ	1PL-GEN-ADJ
	<i>vej-o</i>	<i>vej-i</i>	<i>vej-o</i>	<i>vej-i</i>
	L(M)-NOM.SG	L(M)-NOM.PL	L(M)-NOM.SG	L(M)-NOM.PL

Though the development of the category of ASSOCIATIVITY in SR is clearly due to replication from Hungarian (Elšík 2004, Elšík & Matras 2006: 322–323), SR does not borrow Hungarian associative forms in their material form. In SR the category of associativity combines with the plural: there is a regular (non-associative) plural and an associative plural. The associative plural of a noun denotes the referent of the corresponding singular together with a group of associated persons, e.g. *lavutár-ingere* ‘the violin player and his group’, rather than plurality of referents denoted by the singular. The latter is expressed by a regular plural, e.g. *lavutár-ja* ‘violin players’.

The singular oblique marker *-is-* of the noun in Table 2 may be segmented into a classification suffix *-i-* (which is shared by nouns of a single inflectional class) and the masculine singular oblique suffix *-s-* (which is shared by all masculines). Similarly, the regular plural oblique marker *-en-* consists of the classification suffix *-e-* and the plural oblique suffix *-n-*. The associative forms inflect almost like nominalized genitive forms of the regular plural. The significant difference is in the quality of the classification suffix. The classification suffix of the associative plural is that of the singular (*-i-*), not of the regular plural (*-e-*). On the other hand, the associative is clearly marked as non-

singular by the plural oblique suffix *-n-*. The functions of markers involved in the formation of the associative forms may be summarized as follows: The oblique marker *-n-* indicates non-singular reference: accordingly, it occurs in the regular plural and in the associative plural, but not in the singular. The classification marker *-i-* indicates that a single referent is in focus: accordingly, it occurs in the singular and in the associative plural, but not in the regular plural. The associative forms are in fact nominalizations of genitive forms in *-C-n-ger-* (where *-C-* represents a classification suffix).<sup>30</sup>

Lexical borrowing of Hungarian PLURAL noun forms is non-existent in SR. However, there is an exceptional instance of adoption of the plural form of a pro-word. In some registers of Hungarian the universal quantifier *minden*, which translates as ‘every’ in adnominal uses and as ‘everything’ in nominal uses, can form the plural *minden-ek* ‘everything’ by means of the regular plural suffix. This Hungarian plural form is the source of the universal-quantificational pro-noun *mindenek-o* ‘everything’ in SR.<sup>31</sup> Interestingly, the Hungarian plural has been reinterpreted as singular: the SR pro-noun inflects like a singular xenoclitic noun, requires singular agreement, and cannot form plural forms. This is unlike the other instances of inflected-form adoption, where the function of the Hungarian inflectional category value is retained in SR.

SR commonly borrows CASE forms of Hungarian nouns and pro-nouns, including nominalized adjectives and pro-adjectives. The loanwords of these inflected nominal forms, however, function as (potentially derived) adverbs rather than (inflectional) noun forms in SR. In many cases, especially when based on pro-words, there is no lexical borrowing of the base forms into SR (32–35), and so no potential for identification of the imported case markers.

(32)	BASE	DERIVATION
H	<i>maga</i> ‘self’	<i>magá-tól</i> ‘by oneself’ (ablative)
		▽
SR	<i>*maga</i>	<i>magátú</i> id.

<sup>30</sup> Genitive associative forms result from a morphological haplology of the expected forms (e.g. *lavutár-i-n-ger-* < *\*lavutár-i-n-ger-e-n-ger-*).

<sup>31</sup> The SR pro-noun *mindeneko* is synonymous to a less-frequent, internally derived, pro-noun *minden-so*.

(33)	BASE	DERIVATION
H	<i>annyi</i> ‘that much’	<i>annyi-ra</i> ‘to that extent’ (sublative)
		▽
SR	* <i>añni</i> [cf. <i>at’i</i> ‘that much’]	<i>aññira</i> id.
(34)	BASE	DERIVATION
H	<i>új</i> ‘new’	<i>új-ból</i> ‘anew, again’ (elative)
		▽
SR	* <i>új-n-o</i> [cf. <i>név-o</i> ‘new’]	<i>újbu</i> id.
(35)	BASE	DERIVATION
H	<i>má</i> ‘today’	<i>má-ig</i> ‘until today’ (terminative)
		▽
SR	* <i>má</i> [cf. <i>adádi</i> ‘today’]	<i>májig</i> id.

Nevertheless, there are also many instances where the base forms are borrowed as well, and the imported case markers may be identified as morphemes within SR. This is especially the case with the Hungarian illative/inessive (inessive in all non-standard varieties) case in *-ba* ~ *-be* and the superessive case in *-(V)n*. The former is imported, for example, as the regular marker of temporal adverbs of month (36–37), and the latter, for example, as the regular marker of temporal adverbs of day of week (38–39). None the less, only the Hungarian ablative suffix in *-tól* ~ *-től* got actually extracted within SR (see Section 4.3.1).

(36)	BASE	DERIVATION
H	<i>január</i> ‘January’	<i>január-ba</i> ‘in January’ (inessive)
	▽	▽
SR	<i>januvár-i</i> id.	<i>januvár(-)bo</i> id.
(37)	BASE	DERIVATION

<i>januvár-i</i>	<i>januvár(-)bo</i>
<i>februvár-i</i>	<i>februvár(-)bo</i>
<i>márciuš-i</i>	<i>márciuš(-)ba</i>
<i>ápriliš-i</i>	<i>ápriliš(-)ba</i>
<i>májuš-i</i>	<i>májuš(-)ba</i>
<i>júniuš-i</i>	<i>júniuš(-)ba</i>
<i>júliuš-i</i>	<i>júliuš(-)ba</i>
<i>águstuš-i</i>	<i>águstuš(-)ba</i>
<i>september-i</i>	<i>september(-)be</i>
<i>október-i</i>	<i>október(-)be</i>
<i>november-i</i>	<i>november(-)be</i>
<i>december-i</i>	<i>december(-)be</i>

(38)	BASE	DERIVATION
H	<i>kedd</i> ‘Tuesday’	<i>kedd-en</i> ‘on Tuesday’ (superessive)
	▽	▽
SR	<i>kedd-o</i> id.	<i>kedd(-)en</i> id.
(39)	BASE	DERIVATION
	<i>hetfen-o</i> ‘Monday’	<i>hetfen</i> ‘on Monday’
	<i>kedd-o</i> ‘Tuesday’	<i>kedd(-)en</i> ‘on Tuesday’
	<i>serd-a</i> ‘Wednesday’	<i>serd(-)án</i> ‘on Wednesday’
	<i>čitertek-o</i> ‘Thursday’	<i>čitertek(-)en</i> ‘on Thursday’
	<i>pintek-o</i> ‘Friday’	<i>pintek(-)en</i> ‘on Friday’

#### 4.2.2 Diminutive nouns

Diminutive nouns represent a fully productive derivational category in both languages. While Hungarian makes use of a variety of diminutive suffixes, e.g. *-cska* ~ *-cske* or *-ka* ~ *-ke*, SR diminutives are all derived by means of a single, fully productive marker, the indigenous suffix *-ór-* < ER *\*-oř-*, e.g. *gra* ‘horse’ → *grast-ór-o* ‘little horse’, *čhúr-i*

‘knife’ → *čhurj-ór-i* ‘little knife’. This suffix also applies to loanwords from Hungarian (40).

(40)	BASE	DERIVATION
H	<i>könyv</i> ‘book’	<i>könyv-ecske</i> ‘little book’
	▽	
SR	<i>keňv-o</i> id.	<i>keňv-ór-o</i> id.

Lexical borrowings of Hungarian diminutives are well attested in SR,<sup>32</sup> but many of them lack non-diminutive counterparts, and so their diminutive meaning (and consequently the significance of their imported derivational structure) is weakened or lost altogether, e.g. SR *mihečk-a* ‘bee’ < H *méh-ecske* ‘little bee’ ← *méh* ‘bee’,<sup>33</sup> SR *varjúk-a* ‘crow’ < H *varjú-ka* ‘little crow’ ← *varjú* ‘crow’, SR *meňečk-a* ‘daughter-in-law’ < H *meny-ecske* ‘dear daughter-in-law, young woman’ ← *meny* ‘daughter-in-law’. Even if Hungarian diminutives retain their diminutive meaning when borrowed into SR, their non-diminutive bases need not be borrowed (41). This must have contributed to the lack of extraction of the imported Hungarian diminutive suffixes in SR.

(41)	BASE	DERIVATION
H	<i>hold</i> ‘moon’	<i>hold-acska</i> ‘dear moon’
		▽
SR	* <i>hód-o/a</i> <i>čhon</i> ‘moon’	<i>hód(-)ačk-a</i> id. <i>čhon-ór-o</i>

<sup>32</sup> SR also freely borrows diminutive or familiar forms of personal names (first names or nicknames) from Hungarian, e.g. SR *Adrik-a* < H *Adri-ke* = familiar form of *Adriana*, or SR *Zoli* < H *Zol-i* = familiar form of *Zoltán*.

<sup>33</sup> SR has a non-diminutive noun *mih-a* ‘womb’ < H *méh* id., which however results from borrowing of a merely homonymous lexeme.

### 4.3 *Extracted affixes*

#### 4.3.1 *The separative suffix -tú*

The SR suffix *-tú* is an unproductive ‘ablative’ marker that derives spatial or temporal adverbs from several other spatial or temporal adverbs. In spatial adverbs, the suffix marks *separative* orientation (‘motion from’), a category that expresses (actual or potential) abandonment of the localization (spatial configuration) encoded by the adverbial root or another localization marker (cf. Elšík & Matras 2006). Though the ‘item’ semantics of the suffix may be described in this unified way, the ‘processual’ semantic effect of the derivation varies with different classes of derivational bases.

Three classes of spatial adverbs may be distinguished in this context, each consisting of a pair of derivational chains of adverbs. Each derivational chain, in turn, consists of three members: a) a synchronically underived adverb, whose form is either marked by the so-called old locative suffix *-e ~ -i* or, more rarely, markerless; b) an adverb synchronically derived from the underived adverb by means of the so-called old ablative suffix *-al*; and c) an adverb synchronically derived from the old ablative adverb by means of the suffix *-tú*. The old ablative adverbs thus function as the immediate derivational bases of the derivations in *-tú* and the old locative or markerless adverbs function as their ultimate derivational bases. Now, the bases have differing orientation functions in the four different classes of adverbs.

First, a well-behaved orientation paradigm is encountered in the class of vertical spatial adverbs, which are based on the roots *upr-* ‘up’ and *tél-* ‘down’ (42). Here the separative derivations in *-tú* clearly contrast with their old ablative bases, which encode the stative orientation (‘at rest’). This category expresses actuality of the localization encoded by the adverbial root (or another localization marker), as in ‘My brother lives *up* [upstairs, up the hill etc.]’. The old locative vertical adverbs encode the directive orientation (‘motion to’), a category that expresses (actual or potential) acquisition of the localization encoded by the localization marker, as in ‘I will go *up* to see my brother.’

(42)	DERIVATION	IMMEDIATE BASE	ULTIMATE BASE
	<i>upr-al-tú</i> ‘from up’	← <i>upr-al</i> ‘up, above’	← <i>upr-e</i> ‘up(wards)’

<i>tél-al-tú</i> ‘from down’ separative	← <i>tél-al</i> ‘down, below’ stative	← <i>tél-e</i> ‘down(wards)’ directive
--	--	---

The orientation paradigm is less differentiated in the class of horizontal adverbs, which are based on the roots *ángl-* ‘in the front’ and *pál-* ‘in the back’ (43). Here, too, the separative derivations in *-tú* contrast with their old ablative bases. The latter however encode an undifferentiated stative–directive orientation. (The old locative adverbs actually do not belong to the class of horizontal spatial adverbs as far as their function is concerned: *ángl-e* ‘fore-’ is only used as a non-spatial aktionsart pre-verb and *pál-e* ‘back, backwards’ expresses spatial or metaphorical reversal.)

(43)	DERIVATION	IMMEDIATE BASE	ULTIMATE BASE
	<i>ángl-al-tú</i> ‘from the front’ separative	← <i>ángl-al</i> ‘in/to the front’ stative–directive	← <i>ángl-e</i> ‘fore-’ specific functions
	<i>pál-al-tú</i> ‘from the back’ separative	← <i>pál-al</i> ‘in/to the back’ stative–directive	← <i>pál-e</i> ‘back(wards)’ specific functions

The containment adverbs, which are based on the roots *ánd(r)-* ‘in’ and *á(v)r-* ‘out’ (44a), and the distance adverbs, which are based on the root *dúr-* ‘far’ (44b), present yet a different functional division between the forms. Here the functional distinction between the derivations in *-tú* and their old ablative bases is somewhat less clear. While the former encode the typical separative orientation (‘motion from’), the function of the latter is not well understood.<sup>34</sup> The old locative containment adverbs encode the directive orientation and the markerless distance adverb encodes an undifferentiated stative–directive orientation. Containment adverbs of the stative orientation are not part of the relevant derivational chain.<sup>35</sup>

<sup>34</sup> In a previous publication (Elšik & Matras 2006) a somewhat premature terminological distinction between ‘dynamic’ and ‘static’ separatives was used.

<sup>35</sup> The stative containment adverbs are compounds consisting of the directive (sic!) remote deictic adverb *ond’a* ‘thither, to there’ plus the directive (sic!) containment adverbs, viz. *ond’ánd-e* ‘inside’ and *ond’ár-i* ‘outside’. The structure of these adverbs (including the quirky phenomenon of a directive+directive compound rendering a stative adverb) has been calqued on Hungarian.

(44)	DERIVATION	IMMEDIATE BASE	ULTIMATE BASE
a.	<i>ánder-al-tú</i> ‘from inside’	← <i>ánder-al</i> ‘(from) inside’	← <i>ánd-e</i> ‘in(wards)’
	<i>ávr-al-tú</i> ‘from outside’	← <i>ávr-al</i> ‘(from) outside’	← <i>ár-i</i> ‘out(wards)’
	separative	separative perspective	directive
b.	<i>dúr-al-tú</i> ‘from far’	← <i>dúr-al</i> ‘(from) far’	← <i>dúr</i> ‘(to) far’
	separative	separative perspective	stative–directive

In temporal adverbs, the SR suffix *-tú* marks posterior–durative relation (45), and this must be its original function in SR. While the suffix’s source morpheme, the Hungarian ablative suffix *-tól ~ -től*, does not possess the functions of the SR spatial *-tú*, it may mark the posterior–durative relation. Consequently, we must assume an extraction from loanwords of Hungarian temporal ablatives, such as (46), and a later, L1-internal, extension of the extracted suffix *-tú* into the separative spatial functions.

(45)	BASE	DERIVATION
	<i>idž</i> ‘yesterday’	<i>idžal-tú</i> ‘since yesterday’
(46)	BASE	DERIVATION
H	<i>tavaly</i> ‘last year’	<i>tavaly-tól</i> ‘since the last year’
		▽
SR	<i>tavval id.</i>	<i>tavval(-)tú id.</i>

### 4.3.2 The abstract/collective suffix *-(a)šág-*

The SR suffix *-(a)šág-* is an unproductive marker that derives ABSTRACT nouns (of xenoclitic masculine inflection) from just two nominal lexemes. First, the collective abstract noun *lukest-ašág-o* ‘army’, lit. ‘soldier-ness’, is derived from the masculine noun *lukest-o* ‘soldier’. Second, the abstract noun *ništa-šág-o* ‘unimportant matter’, lit. ‘nothing-ness’, is derived from (the nominative form of) the negative pro-noun *ništa* ‘nothing’. The suffix *-(a)šág-* does not trigger any alternation in its derivational bases but

has itself two allomorphs. The allomorph *-ašág-*, present in the derivation *lukest-ašág-o*, may be perhaps considered to be basic, as the negative pro-noun *ništa* ‘nothing’ has a variant byform *ništ*, and so *ništa-šág-o* can be alternatively segmented as *ništ-ašág-o* (47).

(47)	BASE	DERIVATION
a.	<i>lukest-o</i> ‘soldier’	<i>lukest-ašág-o</i> ‘army’
b.	<i>ništa</i> ‘nothing’	<i>ništa-šág-o</i> ‘unimportant matter’
	<i>ništ</i>	<i>ništ-ašág-o</i>

The de-substantival derivation of abstract nouns is in general uproductive in SR. The suffix *-(a)šág-* competes with the synonymous indigenous suffix *-ipe*, which derives abstract nouns (of oikoclitic masculine inflection) from only a few oikoclitic nouns, e.g. *čohan-ipe* ‘witchcraft, wizardry’ from *čohán-i* ‘witch’ and *čohán-o* ‘wizard’.<sup>36</sup> One of the two bases of the *-(a)šág-* derivation, the noun *lukest-o* ‘soldier’, has xenoclitic inflection despite of its Indo-Aryan etymology:<sup>37</sup> cf. the xenoclitic nominative plural *lukest-i*, oblique singular *lukest-os-*, oblique plural *lukest-on-*, and vocative singular *lukest-o-na* (alongside the oikoclitic *lukest-eja*). The other base, the pro-noun *ništa* ‘nothing’, is a lexical borrowing of Ikavian *ništa*. Although, strictly speaking, the distinction between oikocclisis and xenocclisis is irrelevant for pro-nouns, the nominative form *ništa* can be synchronically recognized as borrowed due to a contrast with non-nominative forms based on the oblique stem *ni-so-s-*, which are internally derived. Thus, the selection between the two de-substantival abstract suffixes appears to be determined by the oiko/xeno status of the derivational base.

<sup>36</sup> However, the suffix *-ipe* is fully productive as a means of deriving abstract nouns from adjective bases, e.g. *nasvál-o* ‘ill’ > *nasval-ipe* ‘disease’. The suffix *-(a)šág-* cannot derive de-adjectival abstract nouns.

<sup>37</sup> The noun *lukesto* ‘soldier’ has resulted from onomasiological nominalization of the passive participle *luk-est-o* < *\*unk(h)l-ist-o* of a verb meaning ‘to move out, to move up’ and secondarily also ‘to mount, to ride [a horse]’. The verb itself, which must have an Indo-Aryan etymology (cf. Boretzky & Igla 1994: 123, 291, Mānušs *et al.* 1997: 127), has been lost in SR and in the Central Romani dialects in general. Several other Romani dialects show similar semantic developments of the participle, e.g. Latvian Romani *ukhl’isto* ‘rider’ or German Sinti *klisto* ‘policeman’.

### 4.3.3 The action suffix *-áš<sup>1</sup> ~ -íš-*

The SR suffix *-áš<sup>1</sup> ~ -íš-* is a productive marker deriving ACTION nouns from verbs. The ‘back’ allomorph *-áš<sup>1</sup>* applies productively to de-nominal verbs derived by the suffix *-áz-* (48a) and to a few underived indigenous verbs (48b); the precise lexical membership of the latter group remains to be investigated. The ‘front’ allomorph *-íš-* appears to apply to a single verb (48c). The suffix is incompatible with the loan-verb adaptation suffix *-in-*.

(48)	BASE	DERIVATION
a.	<i>ásv-áz-in-</i> ‘to shed tears’ <i>čhurj-áz-in-</i> ‘to fight with knives’ <i>d’ij-áz-in-</i> ‘to sing’ <i>huhur-áz-in-</i> ‘to collect mushrooms’ <i>máčh-áz-in-</i> ‘to catch bait for fishing’ <i>moj-áz-in-</i> ‘to drink wine constantly’ <i>paramis-áz-in-</i> ‘to tell fairy-tales’ etc.	<i>ásv-áz-áš-i</i> ‘shedding tears’ <i>čhurj-áz-áš-i</i> ‘fighting with knives’ <i>d’ij-áz-áš-i</i> ‘singing’ <i>huhur-áz-áš-i</i> ‘collecting mushrooms’ <i>máčh-áz-áš-i</i> ‘catching bait for fishing’ <i>moj-áz-áš-i</i> ‘constant drinking of wine’ <i>paramis-áz-áš-i</i> ‘telling fairy-tales’ etc.
b.	<i>muk-</i> ‘to leave, to let, to drop’	<i>muk-áš-i</i> ‘divorce, separation’
c.	<i>pel-in-</i> ‘to quarrel’	<i>pel-íš-i</i> ‘quarrelling, quarrel’

### 4.3.4 The agentive suffix *-áš<sup>2</sup>*

The SR suffix *-áš<sup>2</sup>* is an unproductive marker deriving AGENTIVE nouns from a few nouns (49a). In addition, one de-verbal derivation is attested in my corpus (49b). The meanings of the individual derivations are lexicalized, though most share the semantic component ‘the one who likes N (to V)’. The suffix, which does not show any allomorphy, applies to the so-called PRE-OBLIQUE inflectional stem<sup>38</sup> of the base nouns

---

<sup>38</sup> Pre-obliques stems are inflectional stems of nouns that precede the oblique suffix. While there are identical to nominative (base) stems with most classes of nouns, they may also show various morpho-

and to the non-perfective inflectional stem of the base verb. The bases are prevalingly oikoclitic, but some are xenoclitic: *šliva*  $\triangleleft$  Ikavian. The derived nouns are masculine and show xenoclitic inflection.

(49)	BASE	DERIVATION
a.	<i>džuvl-i</i> ‘woman’	<i>džuvj-áš-i</i> ‘womanizer’
	<i>gra</i> ‘horse’	<i>grast-áš-i</i> ‘coach-man, horse rider, whip’
	<i>mas</i> ‘meat’	<i>mas-áš-i</i> ‘meat lover’
	<i>šlív-a</i> ‘plumb; plumb brandy’	<i>šlív-áš-i</i> ‘plumb brandy lover’
b.	<i>mát-ov-</i> ‘to get drunk’	<i>mát-ov-áš-i</i> ‘drunkard’

The derivation of agentive nouns is in general uproductive in SR. The agentive suffix *-áš<sup>-2</sup>* does not compete with any other morphological marker of agentive nouns, with a single exception of an imported but unextracted marker in a pair of loanwords, perhaps from Aromanian (50). The overwhelming majority of agentive nouns in SR are lexical borrowings, especially from Hungarian.

(50)	BASE	DERIVATION
	<i>lavut-a</i> ‘violin’ $\triangleleft$ Arom	<i>lavut-ár-i</i> ‘violin player’ $\triangleleft$ Arom

The SR suffix *-áš<sup>-2</sup>* must have resulted from extraction of the Hungarian suffix *-(V)s*, where V stands for the vowel-harmony conditioned vowels /a e ö o/. This suffix is one of the most multifunctional affixes in Hungarian. It derives several semantic types of de-nominal, de-numeral and de-adjectival adjectives as well as several semantic types of de-nominal and de-numeral nouns (cf. Kenesei *et al.* 1998: 344–345, 352, 356, 362–363, 367, 370, Tompa 1968: 126–127, 130–131). Relevant for the SR derivation in *-áš<sup>-2</sup>* is the Hungarian de-nominal noun derivation that is usually described as marking OCCUPATION (Kenesei *et al.* 1998: 352). Tompa’s (1968: 126–127) description makes it clear that the

---

phonological alternations, e.g. base stem *gra* ‘horse’ → oblique stem *grast-es-* → pre-oblique stem *grast-*; base stem *džuvl-* ‘woman’ → oblique stem *džuvj-a-* → pre-oblique stem *džuvj-*.

meanings of this derivation are in fact fairly varied.<sup>39</sup> They may be perhaps described more generally as denoting socially relevant roles of human referents relative to objects of their activity (the referents of the base nouns). Some examples of this Hungarian derivation are displayed in (51). As many other Hungarian affixes, the suffix *-(V)s* triggers lengthening of the stem-final /a/ (51b).

(51)	BASE	DERIVATION
a.	<i>lakat</i> ‘padlock’	<i>lakat-os</i> ‘locksmith’
	<i>bolt</i> ‘shop’	<i>bolt-os</i> ‘shop assistant’
b.	<i>kártya</i> ‘card’	<i>kártyá-s</i> ‘card player’
	<i>kurva</i> ‘whore’	<i>kurvá-s</i> ‘womanizer’

Numerous Hungarian occupation (etc.) nouns in *-(V)s* are borrowed into SR, and morphologically adapted through integration into the masculine xenoclitic class in *-i*, e.g. *lakatoš-i*, *bótoš-i*, *kártáš-i*, *kurváš-i* and many more. Often, though not always, the derivational bases of the Hungarian occupation nouns are also borrowed, e.g. *lakat-o*, *bót-a*, *kárt-a*, though not *\*kurv-a*.<sup>40</sup> Borrowing of both the derivations in *-(V)s* and their bases enables an identification of the imported suffix within SR.

(52)	BASE	DERIVATION
	H <i>bolt</i> ‘shop’	H <i>bolt-os</i> ‘shop assistant’
	▽	▽
	SR <i>bót-a</i> id.	SR <i>bót(-)oš-i</i> id.

<sup>39</sup> Tompa (1968: 126–127) distinguishes the following subcategories of the so-called occupation nouns: ‘der mit dem im Gruntword genannten Gerät, Stoff usw. Arbeitende’, ‘der Fahrer eines Verkehrsmittels’, ‘der mit einem Sportgerät Sport Treibende’, ‘der etwas Genießende, mit etwas Spielende’, ‘der ein Instrument Spielende’, ‘der Verkäufer einer Ware’, ‘der zu einer bestimmten Zeit Arbeitende’, and more.

<sup>40</sup> The indigenous noun *lubn-i* ‘whore’ is used instead and appears to block the borrowing of Hungarian *kurva*.

The extracted SR suffix *-áš<sup>-2</sup>*, however, does not correspond in shape to any of the Hungarian allomorphs of the suffix *-(V)s*. Re-analysis of boundaries must have taken place in SR, whereby the lengthened /a:/ of the Hungarian stems ending in /a/ has been re-interpreted as part of the derivational suffix. This has been possible due to the re-analysis of the base inflectional stem (53).

(53)	BASE	DERIVATION
	H <i>kártya</i> ‘card’	H <i>kártyá-s</i> ‘card-player’
	▽	▽
	SR <i>kárt’-a</i> id.	SR <i>kárt’-áš-i</i> id.

#### 4.3.5 The pecunial suffix *-eš-*

The SR suffix *-eš-* marks PECUNIAL nouns derived from numerals. The derivation only applies to three indigenous numerals, all of which are morphologically simple. The derived nouns are either feminine, in which case they denote a ‘banknote’ of the value denoted by the base numeral (54a), or they are masculine, in which case they denote any ‘piece of money’ of the relevant value (54b). The general meaning ‘piece of money’ is interpreted as ‘coin’ or ‘banknote’, according to the most pragmatically relevant reading, and the same holds for the interpretation of the currency. Thus, the most common interpretation for *deš-eš-n-o* ‘10-unit piece of money’ is ‘10-crown coin’ etc.

(54)	BASE	DERIVATION
a.	<i>deš</i> ‘ten’	<i>deš-eš-n-i</i> ‘10-unit banknote’
	<i>biš</i> ‘twenty’	<i>biš-eš-n-i</i> ‘20-unit banknote’
	<i>šel</i> ‘hundred’	<i>šel-eš-n-i</i> ‘100-unit banknote’
b.	<i>deš</i> ‘ten’	<i>deš-eš-n-o</i> ‘10-unit piece of money’
	<i>biš</i> ‘twenty’	<i>biš-eš-n-o</i> ‘20-unit piece of money’
	<i>šel</i> ‘hundred’	<i>šel-eš-n-o</i> ‘100-unit piece of money’

The derived nouns are, in fact, nominalized adjectives, and so two morphological processes can be recognized to be involved in this derivation: de-numeral adjective derivation and nominalizing conversion. The suffix *-eš-* is obligatorily followed by the loan-adjective adaptation suffix *-n-* and the derived nouns inflect like nominalized adjectives, showing oikoclitic inflection: these structural properties testify to the underlying adjectival nature of the derived forms.<sup>41</sup> The derived masculines probably reflect the gender of the masculine noun *lővo* ‘money’, while it is unclear what determines the feminine gender of the ‘banknote’ derivations.<sup>42</sup> It should be emphasized that the de-numeral adjectives that underlie the converted nouns are not used as modifiers with an overt head noun, and so the obligatory nominalizing conversion is onomasiological rather than syntactic, i.e. determined by elision of the head noun. The suffix *-eš-* does not show or trigger any allomorphy.

The SR suffix *-eš-* must have resulted from extraction of the Hungarian suffix *-(V)s*, where V stands for the vowel-harmony conditioned vowels */a e ö o/*. This suffix is, in addition to its further functions (see Section 4.3.4), a fully productive means to derive de-numeral nouns that refer to the number denoted by the base cardinal numeral or relational adjectives thereof (cf. Kenesei *et al.* 1998: 344–345, 356, 370).<sup>43</sup> The de-numeral derivation (55) is vague between nominal and adjectival readings, which are specified by syntactic context only. While naming the number is the general function of this Hungarian derivation, the individual derivations are often used as names of numbered bus or tram lines, rooms etc., and also banknotes and coins (Tompa 1968: 127).

(55)	<p>BASE</p> <p><i>nulla</i> ‘zero’</p> <p><i>két ~ kettő</i> ‘two’</p>	<p>DERIVATION</p> <p><i>nullá-s</i> ‘(of) the number 0’</p> <p><i>kett-es</i> ‘(of) the number 2’</p>
------	--	---

---

<sup>41</sup> An alternative, directly nominal, de-numeral derivation of the type *\*deš-eš-i* etc. is unattested.

<sup>42</sup> There is no noun feminine noun meaning ‘banknote’ or the like. The meaning ‘banknote’ is rendered by the collocation *papírko lővo* ‘paper money’, which is not strictly lexicalized.

<sup>43</sup> Kenesei *et al.* (1998: 344–345) describe the de-numeral adjectives as “attributive adjectival forms of numerals”.

<i>tíz</i> ‘ten’	<i>tíz-es</i> ‘(of) the number 10’
<i>husz</i> ‘twenty’	<i>husz-os</i> ‘(of) the number 20’
<i>ötven</i> ‘fifty’	<i>ötven-es</i> ‘(of) the number 50’
<i>száz</i> ‘hundred’	<i>száz-as</i> ‘(of) the number 100’
<i>ezer</i> ‘thousand’	<i>ezer-es</i> ‘(of) the number 1000’

The extracted SR suffix *-eš-* corresponds in shape to the ‘front unrounded’ allomorph *-es* of the Hungarian suffix *-(V)s*. Its selection with the indigenous bases *deš*, *biš* and *šel* conforms to the rules of Hungarian vowel harmony.

#### 4.3.6 The tool suffix *-ó*

The SR suffix *-ó* is an unproductive marker of TOOL or instrument, deriving a single noun from a verb base (56). The derived noun is masculine and has xenoclitic inflection. The base verb *l-av-av-* is an indigenous causative of the indigenous verb *l-* ‘to take’.<sup>44</sup> In addition to its semantically regular causative meaning ‘to make [so/sth] take [so/sth], to have [so/sth] taken [by so/sth]’, the causative verb has several lexicalized meanings, the most common ones being ‘to undergo artificial abortion’ (i.e. ‘to have the fetus taken’), and ‘to photograph’ (i.e. ‘to have the image of so./sth. taken’). It is of course the latter lexicalized meaning that is the semantic base of the noun in question: the meaning of the noun ‘camera’ is construed as ‘the tool that photographs’. The suffix *-ó* does not show or trigger any allomorphy.

(56)	BASE	DERIVATION
	<i>l-av-av-</i> ‘to photograph’	<i>l-av-av-ó</i> ‘camera’

<sup>44</sup> Note that the causative verb *l-av-av-* ‘to make [so/sth] take’ is irregular in that it is derived by a double application of the indigenous causative suffix *-av-*. The irregular double application of the causative suffix is likely to be connected to the non-syllabic shape of the base verb’s inflectional stem *l-* ‘to take’, though the ungrammatical regular form *\*l-av-* does not violate any phonological constraint in SR.

There is no competing tool derivation in SR. Nouns denoting tools are either underived lexemes (e.g. indigenous *čhúr-i* ‘knife’, borrowed *tover* ‘ax’ < Kurdish *tavar*, *kalapáč-i* ‘hammer’ < H *kalapács*) or loanwords of Hungarian tool derivations.

The SR suffix *-ó* must have resulted from extraction of the Hungarian suffix *-ó ~ -ő*. This suffix is a productive means to form the so-called ‘present’ participles of verbs (cf. Tompa 1968: 59, 126, Laczkó 2001). The present participles are noun modifiers that can retain much of the argument structure of their base verb and that encode ‘simultaneous’ or ‘imperfective’ secondary predications. The present participles of most verbs are active, or agent-oriented, though those derived from potential verbs are passive, or patient-oriented. Numerous present participles are converted into adjectives, as evidenced by the possibility of their predicative use and by the possibility to derive degree forms and manner adverbs from them. Many de-participial adjectives are then converted into nouns.

In grammars of Hungarian these ultimately de-participial nouns are considered to represent a productive de-verbal noun derivation (cf. Kenesei *et al.* 1998: 355, Tompa 1968: 122). Most nominal derivations in *-ó ~ -ő*, many of which have lexicalized meanings, can be classified into three semantic classes: nouns denoting human agents, including occupations (57a); nouns denoting tools or instruments (57b); and nouns denoting regular location of the action (57c).<sup>45</sup> Some of the derived tool nouns alternate with compounds consisting of a de-participial adjective in *-ó ~ -ő* and the noun *gép* ‘machine’, which are preferred in more technical or less colloquial registers, e.g. *fényképez-ő+gép* lit. ‘photographic machine’ ≈ *fényképez-ő* ‘camera’, *repül-ő+gép* lit. ‘flying machine’ ≈ *repül-ő* ‘airplane’.

(57)	BASE	DERIVATION
a.	<i>el+árul</i> ‘to betray’	<i>árul-ó</i> ‘traitor’
	<i>bír</i> ‘to manage, to bear’	<i>bír-ó</i> ‘judge’
	<i>fest</i> ‘to paint’	<i>fest-ő</i> ‘painter’
	<i>igazgat</i> ‘to control, to direct’	<i>igazgat-ó</i> ‘director’

<sup>45</sup> Kenesei *et al.* (1998: 355) only describe the agent and location meanings of the derivation and do not mention or exemplify the tool meaning.

	<i>keresked-ik</i> ‘to trade, to barter’	<i>keresked-ő</i> ‘merchant’
	<i>nyomoz</i> ‘to detect’	<i>nyomoz-ó</i> ‘detective’
	<i>tanít</i> ‘to teach’	<i>tanít-ó</i> ‘teacher’
	<i>tanul</i> ‘to learn, to study’	<i>tanul-ó</i> ‘pupil’
	<i>vezet</i> ‘to lead, to guide, to drive’	<i>vezet-ő</i> ‘boss, leader, guide, driver’
b.	<i>fényképez</i> ‘to photograph’	<i>fényképez-ő</i> ‘camera’
	<i>fog</i> ‘to hold, to grasp’	<i>fog-ó</i> ‘tongs’
	<i>hűt</i> ‘to refrigerate’	<i>hűt-ő</i> ‘fridge’
	<i>kapcsol</i> ‘to connect, to switch’	<i>kapcsol-ó</i> ‘switch’
	<i>legyez</i> ‘to fan’	<i>legyez-ő</i> ‘fan’
	<i>ragaszt</i> ‘to glue’	<i>ragaszt-ó</i> ‘glue’
	<i>reszel</i> ‘to file, to rasp, to grate’	<i>reszel-ő</i> ‘file, rasp, grater’
	<i>repül</i> ‘to fly’	<i>repül-ő</i> ‘airplane’
	<i>törülköz-ik</i> ‘to wipe/dry oneself’	<i>terülköz-ő</i> ‘towel’
	<i>vés</i> ‘to hollow out’	<i>vés-ő</i> ‘chisel’
c.	<i>ebédel</i> ‘to have dinner’	<i>ebédl-ő</i> ‘dining room’
	<i>foly-ik</i> ‘to flow’	<i>foly-ó</i> ‘river’
	<i>mosd-ik</i> ‘to wash ITR’	<i>mosd-ó</i> ‘sink, washbasin’

Lexical borrowings of the Hungarian nouns in *-ó* ~ *-ő* into SR are abundant, e.g. agent *áruló* ‘traitor’, *bíró* ‘judge’, *festő* ‘painter’, *igazgató* ‘director’, *kereskedő* ‘merchant’, *nyomozó* ‘detective’, *tanító* ‘teacher’, *tanuló* ‘pupil’, *vezető* ‘boss, leader, guide, driver’; tool *fogó* ‘tongs’, *hűtő* ‘fridge’, *kapcsoló* ‘switch’, *ledéző* ‘fan’, *ragasztó* ‘glue’, *reselő* ‘file, rasp, grater’, *repülő* ‘airplane’, *terítkező* ‘towel’, *víűő* ‘chisel’; location *fojó* ‘river’, *moűő* ‘sink, washbasin’, and more. Note that the Hungarian phoneme /ű/ is phonologically adapted as /ó/ in SR when it is word-final,<sup>46</sup> so that all loanwords of the

<sup>46</sup> The ‘backing’ of word-final /ű/ also affects loanwords of (synchronically) underived Hungarian words, e.g. H *tűű* ‘lungs’ > SR *tűű* id. However, the Hungarian phoneme /ű/ undergoes an ‘unrounding’ adaptation to /é/ in word-medial positions, e.g. H *tűű+baj* ‘tuberculosis’ > SR *tűű-baj-a* id. There is a single exception to this distribution of adaptational processes affecting the Hungarian phoneme /ű/, viz. H *fű* ‘head’ > SR *fű* ‘boss’ (\*‘head’).

Hungarian nouns in *-ó ~ -ő* end in /ó/. SR also borrows many of the base verbs of these nouns, e.g. *árul-in-* ‘to betray’, *bír-in-* ‘to manage, to bear’, *fešt-in-* ‘to paint’, *kerešked-in-* ‘to trade’, *ňomoz-in-* ‘to detect’, *vezet-in-* ‘to guide, to drive’ (but not \*‘to lead’), *kapčol-in-* ‘to connect, to switch’, *led’ez-in-* ‘to fan’, *ragast-in-* ‘to glue’, *resel-in-* ‘to file, to rasp, to grate’, *repil-in-* ‘to fly’, *viš-in-* ‘to hollow out’, though not all of them, \**tanít-in-* ‘to teach’ (cf. indigenous *sikav-* or *sikjar-*), \**tanul-in-* ‘to learn, to study’ (cf. *sikjov-*), \**fog-in-* ‘to hold, to grasp’ (cf. *astar-*), \**foj-in-* ‘to flow’ (cf. *čujov-*) etc.<sup>47</sup> The simultaneous lexical borrowing of the Hungarian nouns in *-ó ~ -ő* and of their base verbs thus enables the identification of the suffix *-ó* within SR as an imported marker of agents (58) or tools (59).

(58)	BASE H <i>fest</i> ‘to paint’ ▽ SR <i>fešt-in-</i> id.	DERIVATION H <i>fest-ő</i> ‘painter’ ▽ SR <i>fešt(-)ó</i> id.
------	---	--

(59)	BASE H <i>reszel</i> ‘to file [etc.]’ ▽ SR <i>resel-in-</i> id.	DERIVATION H <i>reszel-ő</i> ‘file [etc.]’ ▽ SR <i>resel(-)ó</i> id.
------	--	---

Despite of this the imported agent and tool suffix *-ó* has been extended to a single internal derivation, the noun *lavav-ó* ‘camera’,<sup>48</sup> where it is to be interpreted as a tool marker. In other words, there has been no extraction of the suffix in its agent-marking function.

<sup>47</sup> Loanwords of verbs that serve as bases for the location nouns in *-ó ~ -ő* are unattested in my SR corpus, though there appears to be no structural reason why such verbs should not be borrowable.

<sup>48</sup> The internal derivation of this noun blocks lexical borrowing of its Hungarian semantic counterpart: there is no \**fěňkípezó* ‘camera’ in SR.

### 4.3.7 The artificiality prefix *mí-*

The SR prefix *mí-* is a marker of ARTIFICIALITY. It derives nouns denoting artificial objects from nouns denoting natural objects of the same type. The examples attested in my corpus include numerous derivations from certain body part nouns (60a), but also a few derivations from nouns denoting other kinds of natural objects (60b). The derivation is presumably productive, being applicable whenever it pragmatically makes sense to distinguish between a natural and an artificial object of the same type.<sup>49</sup> The prefix does not exhibit or trigger any allomorphy.

(60)	BASE	DERIVATION
a.	<i>angušt</i> ‘finger/toe’	<i>mí-angušt</i> ‘artificial finger/toe’
	<i>bal</i> ‘hair’	<i>mí-bal</i> ‘artificial hair’
	<i>čúči</i> ‘breast’	<i>mí-čúči</i> ‘artificial breast, fake breast’
	<i>dand</i> ‘tooth’	<i>mí-dand</i> ‘artificial tooth’
	<i>jakh</i> ‘eye’	<i>mí-jakh</i> ‘artificial eye’
	<i>kan</i> ‘ear’	<i>mí-kan</i> ‘artificial ear’
	<i>pro</i> ‘leg/foot’	<i>mí-pro</i> ‘artificial leg/foot’
	<i>va</i> ‘hand/arm’	<i>mí-va</i> ‘artificial hand/arm’
b.	<i>kašt</i> ‘tree; wood’	<i>mí-kašt</i> ‘artificial tree’

The source morpheme, the Hungarian noun *mű* ‘creation, artificial thing, artefact [etc.]’, is a free morpheme that frequently forms the first, modifying, part of noun–noun compounds. Some of these Hungarian compounds have been lexically borrowed into SR, e.g. *mí-keňek-a* < H *mű-könyök* ‘artificial elbow’, *mí-virág-a* < H *mű-virág* ‘artificial

---

<sup>49</sup> My native speaker consultants have refused, on pragmatic grounds, several artificial derivations I suggested, e.g. *mí-šéro* ‘artificial head’ ← *šéro* ‘head’ (using comments such as ‘you can say that word, it’s Romani, but it does not make sense’). One derivation whose refusal surprised me was *mí-músi*. Until that session I wrongly assumed that *músi* means ‘upper limb’, i.e. the arm including the hand. However, the following discussion of the item made it clear to me that it actually means ‘upper arm’, i.e. the part of the arm between the shoulder and the elbow (the upper limb can be unambiguously referred to by the expression *cilo va*, lit. ‘the whole hand/arm’). I must have agreed that it is difficult to imagine someone with an artificial upper arm but a natural lower arm and hand.

flower'. The noun *mŭ* itself, however, has not been borrowed as a free form, and so loanwords such as *mí-keňek-a* cannot be considered to be transparent compounds in SR and the SR morpheme *mí-* is best analyzed as a prefix. The derivational category of artificiality is an innovation within SR that has developed due to contact with Hungarian, a language that lacks such a derivational category.

## 5 Affix extraction: adjectives

### 5.1 Inflectional adaptation

Adjectives are commonly borrowed in SR. Since the distinction between xenoclitic and oikoclitic adjective inflection has been lost (see Section 5.2.1), borrowed adjectives now inflect like indigenous adjectives, showing Indo-Aryan inflections, e.g. *žut-o* ‘yellow’ < Ikavian *žut*. Adjectives borrowed from Hungarian are integrated into SR morphology by means of overt morphological adaptation and inflectional integration. A borrowed SR adjective, e.g. *keňní-n-o* ‘easy’, consists of a phonologically adapted form of the source adjective, e.g. *keňní-* < *könnyű*; an adaptation suffix, e.g. *-n-*; and an inflection, e.g. the base form (nominative singular masculine) suffix *-o*.

There are two suffixes in SR, *-n-* and *-av-*, that serve as markers of overt morphological adaptation of adjectives borrowed from Hungarian. They are both dedicated to this function, as they do not have any other functions in the language. They are pre-inflectional, since they are part of the inflectional stem of borrowed SR adjectives. However, they are not derivational, since they do not derive words within SR. The distribution of the two adjective-adapting suffixes is not random: it is conditioned by phonological properties of the source adjective, more specifically by its final syllable rhyme. Hungarian adjectives ending in a consonant or a long vowel are adapted by the suffix *-n-*, while Hungarian adjectives ending in a short vowel are adapted by the suffix *-av-*, which in addition triggers a deletion of the final vowel of the source form.

### 5.2 Morphological categories and lexical borrowing

#### 5.2.1 Adjective inflection

Hungarian adjectives (and, indeed, all kinds of adjectival words with the exception of demonstratives) are uninflected when used as ATTRIBUTIVE modifiers (cf. Kenesei et al. 1998: 330–331). This contrasts sharply with SR adjectives (and also most other kinds of adjectival modifiers), which inflect for the agreement categories of number, gender, and case when used attributively. However, the category of adjectival case in SR (and in

Romani in general) is distinct from the category of nominal case (see Section 4.2.1) in that attributive adjectives show partial or ‘deflected’ case agreement with their head nouns, distinguishing merely two cases: the nominative and the oblique. The latter is used in agreement with any non-nominative case of the head noun.

When used in PREDICATIVE positions, Hungarian adjectives inflect for number, agreeing in it with the subject noun phrase. The adjectival plural marker *-(V)k* is identical to the nominal one. SR predicative adjectives, on the other hand, inflect for the agreement categories of number and gender, and so they are less differentiated than SR attributive adjectives. Both Hungarian and SR adjectives are NOMINALIZED when used as heads of noun phrases, and show full nominal inflection (see Section 4.2.1). In SR there is an additional inflectional category — I term it STATUS —, which is restricted to a productive class of adjectives with an underlying long vowel in the final syllable of their inflectional stem. These adjectives distinguish between predicative and nominalized forms, which retain the underlying long vowel, and attributive forms, where the vowel undergoes shortening.<sup>50</sup> A sample inflectional paradigm of a status-encoding adjective is shown in Table 3.

Table 3: Inflectional paradigm of the adjective *bár-o* ‘big’

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

Inflectional classes of Hungarian adjectives are determined solely by rules of vowel harmony, i.e. phonotactically. In SR there are three productive inflectional classes and a few minor and unproductive ones. The distinction between xenoclitic and oikoclitic adjective inflection, which is reconstructable for earlier post-Greek stages of Romani, has

<sup>50</sup> The SR category of status is unique within Romani (cf. Elšík 2002). It has developed as a by-product of the development of phonologically distinctive vowel length, which was triggered by contact with Hungarian. Hungarian is indirectly responsible for the creation of an inflectional category that it itself does not possess.

been lost in SR due to internal analogical developments in favour of the once oikoclitic, indigenous, inflections (Elšík & Matras 2006: 329).

On the whole, SR adjective inflection is more differentiated than Hungarian adjective inflection, both in terms of inflectional categories and in terms of lexical classification. Hungarian possess no categories that SR lacks.

### 5.2.2 Adjective comparison

The category of degree is isomorphous in the two languages in contact. The comparative is derived from the positive form by means of a comparative (or rather: non-positive) suffix, and the superlative is, in turn, derived from the comparative form by means of a superlative prefix. While the non-positive suffix is different in Hungarian (61) and in SR (62), the superlative prefix is identical, i.e. clearly of Hungarian origin in SR. Since, however, there is no lexical borrowing of the Hungarian comparative or superlative degree forms into SR, the SR superlative prefix *leg-* must have resulted from affix copying rather than affix extraction.

(61)	POSITIVE	COMPARATIVE	SUPERLATIVE
ADJ	<i>olcsó</i> → 'cheap'	<i>olcsó-bb</i> → 'cheaper'	<i>leg-olcsó-bb</i> '(the) cheapest'
(62)	POSITIVE	COMPARATIVE	SUPERLATIVE
ADJ	<i>ócsó-n-o</i> <math>\triangleleft</math> H → 'cheap'	<i>ócsó-n-eder</i> → 'cheaper'	<i>leg-ócsó-n-eder</i> '(the) cheapest'

### 5.2.3 Attenuative and similative adjectives

There are several semantic classes of de-adjectival evaluative derivations of adjectives in both Hungarian (cf. Tompa 1968: 130–134, Kenesei *et al.* 1998: 367) and SR. In Hungarian, ATTENUATIVE adjectives are derived from qualitative adjectives by means of the suffix *-(V)s* (63a). The same suffix derives SIMILATIVE adjectives from de-nominal relational adjectives in *-i* (63b).

(63)	BASE	DERIVATION
a.	<i>ősz</i> ‘grey (of hair)’	<i>ősz-es</i> ‘greyish (of hair)’
b.	<i>üzlet-i</i> ‘relating to business’	<i>üzlet-i-as</i> ‘business-like’

SR attenuative adjectives are derived by means of one of two suffixes. The Ikavian-origin suffix *-ast-* is completely productive and can apply to any qualitative adjective whatsoever, including derived adjectives and loanwords from Hungarian (64a). When re-iterated, the resulting double or multiple attenuative emphasizes the attenuative modification of the base adjective (64b). The second attenuative suffix, the indigenous *-ikán-*, is restricted to (some) underived indigenous bases (64c). One interesting exception is the adjective *rom-an-ikán-o* ‘Gypsy-like’, which — unlike the other de-adjectival derivations in *-ikán-* — is based on a de-nominal relational adjective and has similitive rather than attenuative meaning (64d).<sup>51</sup>

(64)	BASE	DERIVATION
a.	<i>gull-o</i> ‘sweet’ <i>somnak-un-o</i> ‘golden’ <i>és-n-o</i> ‘grey (of hair)’ < H <i>ősz</i>	<i>gull-ast-o</i> ‘rather sweet’ <i>sumnak-un-ast-o</i> ‘goldish’ <i>és-n-ast-o</i> ‘greyish (of hair)’
b.	<i>gull-ast-o</i> ‘rather sweet’ <i>gull-ikán-o</i> ‘rather sweet’	<i>gull-ast-ast-o</i> ‘slightly sweet’ <i>gull-ikán-ast-o</i> ‘slightly sweet’
c.	<i>gull-o</i> ‘sweet’ <i>somnak-un-o</i> ‘golden’ <i>és-n-o</i> ‘grey (of hair)’ < H <i>ősz</i> <i>gull-ast-o</i> ‘rather sweet’ <i>gull-ikán-o</i> ‘rather sweet’	<i>gull-ikán-o</i> ‘rather sweet’ * <i>sumnak-un-ikán-o</i> * <i>és-n-ikán-o</i> * <i>gull-ast-ikán-o</i> * <i>gull-ikan-ikán-o</i>
d.	<i>rom-án-o</i> ‘related to Gypsies’	<i>rom-an-ikán-o</i> ‘Gypsy-like’

<sup>51</sup> Further similitive ethnic adjectives in *-ikán-* are derived from nouns rather than from adjectives.

#### 5.2.4 Diminutive adjectives

Diminutivization of adjectives is marginal in Hungarian (Tompa 1968: 129) and productive though rare in SR. SR diminutive adjectives, which are derived by the indigenous suffix *-ór-*, mark polite, emotional, or otherwise evaluative relation of the speaker towards the referent of the modified noun (65), rather than lessening the quality denoted by the base adjective, which is the function of attenuatives.

(65)	BASE	DERIVATION
	<i>kal-o [rom]</i> ‘black [G. man]’	<i>kal-ór-o [rom]</i> ‘sweet black [G. man]’

#### 5.2.5 Deprivative adjectives

Hungarian possesses a productive derivational category of deprivative adjectives, marked by means of the suffix *-(a)t(a)lan ~ -(e)t(a)len* (cf. Tompa 1968: 123, 133, Kenesei *et al.* 1998: 365–367). The derivation applies to nouns (66a), (mostly transitive) verbs, including those inflected for possibility (66b), relational adjectives in *-i*, and similitive adjectives in *-szer-ű* (66c). On the other hand, the SR derivation of deprivative adjectives, marked by the indigenous prefix *bi-*, is completely unproductive. There are only deprivative derivations from a few qualitative adjectives (67), while relational adjectives, nouns, or verbs cannot serve as bases for the derivation.

(66)	BASE	DERIVATION
a.	<i>könyv</i> ‘book’ <i>só</i> ‘salt’	<i>könyv-telen</i> ‘bookless’ <i>só-t(a)lan</i> ‘saltless’
b.	<i>olvas</i> ‘read’ <i>olvas-hat</i> ‘to be able to read’	<i>olvas-atlan</i> ‘unread’ <i>olvas-hat-atlan</i> ‘unreadable’
c.	<i>üzlet-i</i> ‘relating to business’ <i>könyv-szer-ű</i> ‘like a book’	<i>üzlet-i-etlen</i> ‘unbusiness-like’ <i>könyv-szer-ű-tlen</i> ‘unlike a book’

(67)	BASE	DERIVATION
------	------	------------

*lon-d-o* ‘salty’

*bi-lon-d-o* ‘saltless’<sup>52</sup>

Despite the lack of productivity of the category in SR, lexical borrowing of Hungarian deprivative adjectives does not appear to be particularly common. The attested examples are mostly loanwords of de-nominal deprivative adjectives; their bases may but need not be borrowed (68–69). Loanwords of de-verbal deprivative adjectives are very rare and those of de-adjectival deprivative adjectives are unattested. There is no extraction of the Hungarian origin deprivative suffix in SR.<sup>53</sup>

(68)	BASE	DERIVATION
H	<i>bűn</i> ‘sin, crime’	<i>bűn-telen</i> ‘sinless’
	▽	▽
SR	<i>bín-o</i> id.	<i>bín(-)telen-n-o</i> ‘innocent’
(69)	BASE	DERIVATION
H	<i>szerencse</i> ‘luck’	<i>szerencsé-tlen</i> ‘unlucky’
		▽
SR	* <i>serenč-e</i> [cf. <i>bast</i> ‘luck’ < WIr]	<i>serenčétlen-n-o</i> id.

### 5.2.6 Negative adjectives

Kenesei *et al.* (1998: 368) argue that negative adjectives should be considered to represent a derivational category, despite the conventional separate spelling of the negator (70), which is homonymous to the regular predicate negator. SR negative adjectives by means of the prefix *na-* (71), which is likewise homonymous to the regular

---

<sup>52</sup> Note that the deprivative adjective meaning ‘saltless’ is a de-nominal derivation in Hungarian, but a de-adjectival one in SR.

<sup>53</sup> The Hungarian deprivative suffix has been extracted in the closely related Romani variety of the neighbouring village of Vlčany. Nevertheless, the deprivative adjectives internally derived by means of the Hungarian-origin suffix in the Vlčany variety, e.g. *bast* ‘luck’ → *bas-talan-n-o* ‘unlucky’, are not even acceptable to my SR consultants.

predicate negator, appear to be rare, though several lexemes are well attested. There is no lexical borrowing of Hungarian negative adjectives into SR.

- |      |                                      |   |
|------|--------------------------------------|---|
| (70) | BASE<br><i>érdekes</i> ‘interesting’ | DERIVATION<br><i>nem érdekes</i> ‘not interesting’  |
| (71) | BASE<br><i>lách-o</i> ‘good, nice’   | DERIVATION<br><i>na-lách-o</i> ‘not good, not nice’ |

### 5.2.7 De-nominal adjectives

Hungarian de-nominal derivations of adjectives are numerous (cf. Tompa 1968: 130–134, Kenesei *et al.* 1998: 362–365). Several suffixes are phrasal affixes, as they may apply to complex noun phrases (72). For example, the complex marker *-bel-i* may be suffixed to spatial noun phrases, including those inflected for the person and number of the possessor.

- |      |   |   |
|------|---|---|
| (72) | BASE<br><i>ház-unk</i> ‘our house’<br><i>három óra</i> ‘three hours’<br><i>rövid függöny</i> ‘short curtain’<br><i>rövid haj</i> ‘short hair’ | DERIVATION<br><i>ház-unk-bel-i</i> ‘of the loc. in our house’<br><i>három óra-s</i> ‘three-hour-long’<br><i>rövid függöny-ös</i> ‘with short curtains’<br><i>rövid haj-ú</i> ‘short-haired’ |
|------|---|---|

In SR, adjectives derived from nouns are mostly relational, though some are qualitative. De-nominal adjectives are derived by means of one of the following productive suffixes: the indigenous *-ál-*, *-án-*, *-ikán-*, *-un-*, *-utn-*, or the Greek-origin *-(i/c)k-*.<sup>54</sup> The choice of the suffix is mostly determined by semantics of the derivational base, though there are various kinds of exceptions. To name just one of several examples,

---

<sup>54</sup> In addition to the productive de-nominal adjective-deriving suffixes, there are also several unproductive suffixes, which are restricted to qualitative derivations. Conversion, too, only creates qualitative adjectives, e.g. *béng* ‘devil’ → ‘naughty, mischievous’.

the suffix *-án-* is productively applied to nouns denoting domestic animals (73a), but it is also found in one ethnic adjective (73b).

(73)	BASE	DERIVATION
a.	<i>bakr-o</i> ‘sheep, ram’	<i>bakr-án-o</i> ‘relating to sheeps, mutton’
	<i>bál-o</i> ‘pig’	<i>bal-án-o</i> ‘relating to pigs, pork’
	<i>čibók-a</i> ‘chicken’ < H <i>csibóka</i>	<i>čibók-án-o</i> ‘relating to chickens’
	<i>džukel</i> ‘dog’	<i>džukl-án-o</i> ‘relating to dogs’
	<i>gra</i> ‘horse’ < Arm <i>grast</i>	<i>grast-án-o</i> ‘relating to horses’
	<i>kečk-e</i> ‘goat’ < H <i>kecske</i>	<i>kečk-án-o</i> ‘relating to goats’
b.	<i>rom</i> ‘Gypsy’	<i>rom-án-o</i> ‘Romani, relating to Gypsies’

There is often a mismatch between the origin of the base and that of the suffix. For example, the Hungarian-origin xenoclitic noun in (74a) takes an indigenous suffix, while the indigenous oikoclitic noun in (74b) takes the Greek-origin suffix. Nevertheless, within certain semantic subclasses of derivations, origin of the base may co-determine the choice of the suffix (74c).<sup>55</sup>

(74)	BASE	DERIVATION
a.	<i>sarv-a</i> ‘horn’ < H <i>szarv</i>	<i>sarv-ál-o</i> ‘having horns’
b.	<i>čór</i> ‘thief’	<i>čór-k-o</i> ‘relating to thieves’
c.	<i>somnakaj</i> ‘gold’	<i>somnak-un-o</i> ‘golden’
	<i>araň-o</i> ‘gold’ < H <i>arany</i>	<i>araň-ik-o</i> ‘golden’

Sometimes a single base takes alternative suffixes and then the suffix may co-determine the semantics of the derivation. For example, within derivations based on

---

<sup>55</sup> Both internal derivations, *somnak-un-o* and *araň-ik-o*, are used alongside the loanword *araňoš-n-o* ‘golden’ < H *aranyos*.

ethnic nouns there is a distinction between relational (75a) and similitive (qualitative) adjectives (75b).<sup>56</sup>

(75)	BASE	DERIVATION
a.	<i>próst-o</i> ‘non-Gypsy’	<i>próst-ik-o</i> ‘relating to non-Gypsies’
	<i>serv-o</i> ‘Slovak’	<i>serv-ik-o</i> ‘relating to Slovaks’
	<i>ungr-o</i> ‘Hungarian’	<i>ungr-ik-o</i> ‘relating to Hungarians’
b.	<i>próst-o</i> ‘non-Gypsy’	<i>próst-ikán-o</i> ‘non-Gypsy-like’
	<i>serv-o</i> ‘Slovak’	<i>serv-ikán-o</i> ‘Slovak-like’
	<i>ungr-o</i> ‘Hungarian’	<i>ungr-ikán-o</i> ‘Hungarian-like’

SR borrows numerous Hungarian de-nominal adjectives in the suffix *-V(s)*, but only those that are qualitative. The Hungarian-origin imported suffix *-V(s)*, however, does not get extracted in SR. Relational Hungarian de-nominal adjectives, e.g. those in *-i*, *-ú ~ -ű* and *-nyi*, are not borrowed at all.

### 5.2.8 De-adverbial adjectives

Hungarian de-adverbial adjectives (Kenesei *et al.* 1998: 369–370) are derived from spatial or temporal adverbs by means of the relational suffix *-i* (76).

(76)	BASE	DERIVATION
	<i>bent</i> ‘inside’	<i>bent-i</i> ‘inner’
	<i>ma</i> ‘today’	<i>ma-i</i> ‘of today’

SR de-adverbial adjectives, which are derived by means of the indigenous suffixes *-un-*, *-utn-* or *-ukn-* (and their various irregularly extended allomorphs), are likewise prevalently based on spatial or temporal adverbs and are relational. The suffix

---

<sup>56</sup> The semantically parallel distinction between *rom-án-o* ‘relating to Gypsies’ vs *rom-an-ikán-o* ‘Gypsy-like’ is structurally different, in that the similitive derivation is de-adjectival, rather than de-substantival, and so structurally identical to the attenuative adjectives (see above).

*-un-*, shared with some de-nominal adjective derivations, applies to spatial adverbs of vertical and horizontal localization (77a). The suffixes *-utn-* or *-ukn-* are specific to de-adverbial derivations, and there is a tendency for the former to apply to spatial adverbs (77b) and the latter to temporal adverbs (77c). The qualitative adjective *polók-un-o* ‘slow’ is the only adjective derived from a manner adverb (77d).<sup>57</sup>

(77)	BASE	DERIVATION
a.	<i>upr-</i> ‘up’	<i>upr-un-o</i> ‘upper’
	<i>tél-</i> ‘down’	<i>tél-un-o</i> ‘lower’
	<i>ángl-</i> ‘in the front’	<i>ángl-un-o</i> ‘front’
	<i>pál-</i> ‘in the back’	<i>pál-un-o</i> ‘back’
b.	<i>ándr-</i> ‘inside’	<i>ándr-utn-o</i> ‘inner’
	<i>ávr-</i> ‘outside’	<i>ávr-utn-o</i> ‘outer’
	<i>táha</i> ‘tomorrow’ < Greek <i>taxia</i>	<i>táha-t-utn-o</i> ‘tomorrow’s’
c.	<i>linaj-e</i> ‘in the summer’	<i>linaj-ukn-o</i> ‘relating to summer time’
	<i>tavval</i> ‘last year’ < H <i>tavaly</i>	<i>tavval-ukn-o</i> ‘last year’s’
	<i>korán</i> ‘early in t. morning’ < H <i>korán</i>	<i>korán-ukn-o</i> ‘early morning’s’
	<i>ra’aha</i> ‘in the morning’	<i>ra’a-l-ukn-o</i> ‘morning’s’
	<i>idž</i> ‘yesterday’	<i>idž-al-ukn-o</i> ‘yesterday’s’
	<i>dúr</i> ‘far’	<i>dúr-all-ukn-o</i> ‘far, faraway’
d.	<i>polók-e</i> ‘slowly; in a low voice’	<i>polók-un-o</i> ‘slow’

### 5.2.9 De-verbal adjectives

Hungarian possesses several productive classes of de-verbal adjectives (cf. Tompa 1968: 123–124, Kenesei *et al.* 1998: 365–367), in addition to the de-verbal deprivative adjectives discussed in Section 5.2.5. While adjectival conversion of verbal participles is a matter of lexicalization, the suffix *-hat-ó* ~ *-het-ő*, which consists of the possibility suffix and the active participial suffix, is considered to represent a productive derivation.

<sup>57</sup> The dominant pattern is to derive manner adverbs from adjectives.

SR de-verbal adjectives, on the other hand, are rare and unproductive. Restrictions on the formation of perfective participles and their conversion into adjectives are not well understood, though lexicalized participles are attested especially with transitive and ‘unaccusative’ verbs. Some verbs of emotion and bodily states appear to be the only semantic class of verbs that allow derivation of non-participial de-verbal adjectives (79). See Section 5.3.1 for the derivation in *-óš-*.

(78)	<p>BASE</p> <p><i>asa-</i> ‘to laugh, to smile’, PTC <i>-n-</i></p> <p><i>ladža-</i> ‘to be shy’, PTC <i>-n-</i></p> <p><i>khand-</i> ‘to stink’, PTC <i>-l-</i></p>	<p>DERIVATION</p> <p><i>asa-nd-o</i> ‘smily’</p> <p><i>ladža-nd-o</i> ‘shy’</p> <p><i>khand-in-o</i> ‘bloody, fucking’</p>
------	--	--

### 5.3 *Extracted affixes*

#### 5.3.1 *The habitual suffix -óš-*

The SR suffix *-óš-* is an unproductive de-verbal adjective marker. It applies to indigenous causative verbs in *-(a)v-* derived from emotion verbs (79a) or, rarely, from a few other verbs (79b). The meanings of the derivations are lexicalized and some of the derivations appear to be based on causatives only as far as their form is concerned. As several of the derivations appear to encode a constant or usual property, the suffix may be termed the habitual suffix for convenience. The suffix *-óš-*, which does not exhibit or trigger any allomorphy, is obligatory followed by the loan-adjective adaptation suffix *-n-*.

(79)	<p>BASE</p> <p>a. <i>asa-v-</i> ‘to make [so.] laugh/smile’</p> <p><i>dara-v-</i> ‘to frighten = make [so.] fear’</p> <p><i>ladža-v-</i> ‘to make [so.] shy/ashamed’</p> <p>b. <i>khand-av-</i> ‘to make [so/sth.] stink’</p>	<p>DERIVATION</p> <p><i>asa-v-óš-n-o</i> ‘funny, ridiculous’</p> <p><i>dara-v-óš-n-o</i> ‘frightening’</p> <p><i>ladža-v-óš-n-o</i> ‘usually shy/ashamed’</p> <p><i>khand-av-óš-n-o</i> ‘constantly stinky’</p>
------	---	---

De-verbal adjectives are otherwise very rare and unproductive in SR and there is no specialized derivation of habitual or similar adjectives. With a few exceptions, they

are lexicalized passive participles. Interestingly, other SR de-verbal adjectives are mostly based on the same class of verbs, as are — ultimately, i.e. via the (partly formal) causative derivation — the adjectives in *-óš-*.

The source morpheme of the SR suffix *-óš-*, the Hungarian suffix *-ós-* ~ *-ős-*, is a productive marker of de-verbal adjectives, whose meaning can be described as ‘usually V-ing’ (Kenesei *et al.* 1998: 367) or as denoting constancy or inclination. Although no lexical borrowings of Hungarian adjectives in *-ós-* ~ *-ős-* are attested in my corpus of SR, the obligatory presence of the adaptation suffix *-n-* after *-óš-* strongly suggests that the latter has been extracted from adjectives borrowed from Hungarian. SR selects the ‘back’ allomorph of the Hungarian suffix, which conforms to the rules of Hungarian vowel harmony: the last vowel of all the bases of the SR adjectives in *-óš-* is the ‘back’ vowel /a/.

## 6 Affix extraction: adverbs

### 6.1 Morphological categories and lexical borrowing

#### 6.1.1 Manner adverbs

Hungarian has three productive suffixes deriving MANNER adverbs from adjectives (cf. Kenesei *et al.* 1998: 371–372).<sup>58</sup> The selection among the suffixes is, to a great extent, determined lexically and by the derivational structure of their adjective bases; semantic contrast between the derivations is rare. The suffix *-(V)n* applies to most underived and derived adjectives, including comparative and superlative degree forms (see Section 6.1.2). The suffix *-ul ~ -ül* applies to several underived adjectives and, productively, to derived deprivative adjectives and to, derived or underived, ethnic adjectives. The least frequent, though productive, suffix *-lag ~ -leg* applies to several types of derived adjectives, including lexicalized active participles.

In SR, too, there are several productive suffixal markers deriving de-adjectival manner adverbs. First, the indigenous suffix *-e* < ER *\*-es*<sup>59</sup> applies to most indigenous adjectives (80a). Its extended version *-ón-e* used to apply only to borrowed adjectives, but with the collapse of xenoclis in adjective inflection (see Sections 5.2.1) it has also extended to numerous indigenous bases (80b). Some of these latter derivations alternate with derivations by means of the mysterious suffix *-óvat* (80c).<sup>60</sup> Fourth, the suffix *-a* of Greek origin is specialized for deriving manner adverbs from most ethnic adjectives, viz. those that are derived by means of the Greek-origin relational suffix *-(i/c)k-*. Finally,

---

<sup>58</sup> In more traditional descriptions (e.g. Tompa 1968: 201, 205–206, cf. also Kenesei *et al.* 1998: 192), these suffixes are treated as case markers: the ‘modal-essive 1’ *-(V)n*, the ‘essive-modal’ *-ul ~ -ül*, and the ‘modal-essive 2’ *-lag ~ -leg*.

<sup>59</sup> The adverbial suffix *-e* is etymologically identical to the oblique singular masculine suffix *-e* < ER *\*-es*, which reflects the Old Indo-Aryan genitive masculine suffix *-asya*.

<sup>60</sup> The manner suffix *-óvat* is likely to be borrowed from Hungarian, as it is only attested in the speech of younger speakers of SR and nowhere else in Romani. While the phoneme /ó/ of *-óvat* probably reflects the initial phoneme of the extended manner marker *-ón-e*, the remaining part of the suffix, viz. /vat/, remains obscure. What comes closest as a potential source morpheme is the Hungarian converb suffix *-va ~ -ve*, which, however, does not fit in several respects.

there is the extracted Hungarian-origin suffix *-šon*, which will be discussed in Section 6.2.1.

(80)	BASE	DERIVATION
a.	<i>šukár</i> ‘nice, beautiful’ <i>gadž-ikán-o</i> ‘non-Gypsy’	<i>šukár-e</i> ‘nicely, beautifully’ <i>gadž-ikán-e</i> ‘in a non-Gypsy way’
b.	<i>erd'av-o</i> ‘bad’ ◁ <i>Ik rđav</i> <i>harn-o</i> ‘short’ <i>mát-o</i> ‘drunk’	<i>erd'av-ón-e</i> ‘badly, in a sick state’ <i>harn-ón-e</i> ‘shortly, in a short way’ <i>mát-ón-e</i> ‘in a drunk way or state’
c.	<i>harn-o</i> ‘short’ <i>mát-o</i> ‘drunk’	<i>harn-óvat</i> ‘shortly, in a short way’ <i>mát-óvat</i> ‘in a drunk way or state’
d.	<i>pojác-k-o</i> ‘relating to Vlax Gypsies’	<i>pojác-k-a</i> ‘in the Vlax Gypsy way’

None of the SR adverb-deriving suffixes applies to adjectives borrowed from Hungarian. Instead, all manner adverbs semantically corresponding to Hungarian-origin adjectives are themselves loanwords from Hungarian. A few examples are shown in (81–83). Note that the formal relationships between the borrowed adjectives and the borrowed adverbs are varied, both due to allomorphy imported from Hungarian and due to the variation in the SR adjective-adaptation suffixes (see Section 5.1).

(81)	BASE	DERIVATION
H	<i>olcsó</i> ‘cheap’ ▽	<i>olcsó-n</i> ‘cheaply’ ▽
SR	<i>óčo-n-o</i> id.	<i>óčo(-)n</i> id.

(82)	BASE	DERIVATION
H	<i>bátor</i> ‘brave’ ▽	<i>bátr-an</i> ‘bravely’ ▽
SR	<i>bátor-n-o</i> id.	<i>bátr(-)an</i> id.

(83)	BASE	DERIVATION
------	------	------------

H	<i>gyenge</i> ‘weak’	<i>gyengé-n</i> ‘in a weird way’
	▽	▽
SR	<i>d’eng-av-o</i> id.	<i>d’eng(-)én</i> id.

Interestingly, all of the numerous examples of borrowed de-adjectival adverbs in my corpus of SR are loanwords of a single derivational type of Hungarian adverbs, viz. those in *-(V)n*. Lexical borrowings of the Hungarian de-adjectival adverbs in *-ul ~ -ül* and *-lag ~ -leg* are unattested. This is most likely a plain consequence of the fact that there are no loanwords of those Hungarian adjectives that take these adverb-deriving suffixes; at least they are unattested in my corpus. Though the lack of such borrowed adjectives is partly accidental,<sup>61</sup> a structural reason may be invoked at least for the lack of loanwords of Hungarian ethnic adjectives (which derive manner adverbs by means of *-ul ~ -ül*). Most ethnic adjectives in SR are internal de-nominal derivations by means of the suffix *-(i/c)k-*, and this also applies to all ethnic adjectives that are based on Hungarian-origin nouns (84). The internal SR derivation of ethnic adjectives from Hungarian base nouns ‘blocks’ lexical borrowing of the (homonymous) Hungarian ethnic adjectives.

(84)	N	ADJ	ADV
	‘Pole’	‘Polish’	‘in Polish’
H	<i>lengyel</i> =	<i>lengyel</i> →	<i>lengyel-ül</i>
	▽		
SR	<i>lend’el-i</i> →	<i>lend’el-k-o</i> →	<i>lend’el-k-a</i>

Generally, it appears to be the case that internal derivation of manner adverbs from Hungarian-origin adjectives is ‘blocked’ by systematic lexical borrowing of the category of manner adverbs from Hungarian.

---

<sup>61</sup> For example, the SR adjective *erd’av-o* ‘bad’ < Ikavian *rđav* has not been replaced by a loanword of its Hungarian semantic equivalent *rossz*, which takes the suffix *-ul ~ -ül* to derive its manner adverb.

### 6.1.2 Adverb comparison

Unlike degree forms of Hungarian adjectives, the degree forms of Hungarian adverbs (cf. Kenesei *et al.* 1998: 348–349) do not derive from one another (i.e. the superlative from the comparative and the comparative from the positive, see Section 5.2.2), and so the de-adjectival adverb derivation can be described as having scope over the degree derivation (85). In other words, though there are adverbs of different degrees, there is no derivational category of adverbial degree.

(85)	POSITIVE	COMPARATIVE	SUPERLATIVE
	‘cheap’	‘cheaper’	‘(the) cheapest’
ADJ	<i>olcsó</i> →	<i>olcsó-bb</i> →	<b><i>leg-olcsó-bb</i></b>
	↓	↓	↓
ADV	<i>olcsó-n</i>	<i>olcsó-bb-an</i>	<i>leg-olcsó-bb-an</i>
	‘cheaply’	‘more cheaply’	‘(the) most cheaply’

In SR, there are no specifically adverbial degree forms of internally derived manner adverbs. Instead, comparative and superlative forms of semantically corresponding adjectives are also used adverbially (86).

(86)	POSITIVE	COMPARATIVE	SUPERLATIVE
	‘beautiful’	‘more beautiful’	‘(the) most beautiful’
ADJ	<i>šukár</i> →	<i>šukar-eder</i> →	<b><i>leg-šukar-eder</i></b>
	↓		
ADV	<i>šukár-e</i>	<i>šukar-eder</i>	<i>leg-šukar-eder</i>
	‘beautifully’	‘more beautifully’	‘(the) most beautifully’

Hungarian-origin manner adverbs show a complex pattern of degree marking in SR. There are two structural options, as displayed in (87). First, internally derived comparative and superlative forms of the semantically corresponding Hungarian-origin adjectives may be used adverbially, as above. Or, alternatively, loanwords of the Hungarian adverbial comparative and superlative forms are employed. The latter option,

i.e. lexical borrowing of adverbial degree forms, appears to prevail in the varieties of younger speakers of SR.<sup>62</sup>

(87)	POSITIVE	COMPARATIVE	SUPERLATIVE
	‘cheap’	‘cheaper’	‘(the) cheapest’
ADJ	<i>óció-n-o</i> < H →	<i>óció-n-eder</i> →	<i>leg-óció-n-eder</i>
ADV		<i>óció-n-eder</i>	<i>leg-óció-n-eder</i>
	<i>óción</i> < H	<i>ócióbban</i> < H	<i>legócióbban</i> < H
	‘cheaply’	‘more cheaply’	‘(the) most cheaply’

### 6.1.3 De-nominal adverbs

According to Kenesei *et al.* (1998: 370–371) there is no productive de-substantival derivation of adverbs in Hungarian. However, they hasten to add, some of the traditional ‘case’ suffixes (e.g. modal *-lag* ~ *-leg*, temporal *-kor*, and essive-formal *-ként*, cf. Tompa 1968: 200–205, Kenesei *et al.* 1998: 192) are better considered to be adverb-deriving suffixes on some structural criteria.

Lexical borrowings of the adverbs derived by the modal suffix are attested in Selice Romani, e.g. SR *ešetleg* < H *eset-leg* ‘eventually’ ← *eset* ‘case, story, event’, and loanwords of the adverbs derived by the temporal suffix are actually abundant. They are the regular means to express several kinds of temporal adverbials: of clock time, e.g. SR [pándž] *órakkor* < H [öt] *óra-kor* ‘at [five] o’clock’ ← *óra* ‘hour; o’clock; clock’; of seasons, viz. SR *tavaskor* < H *tavaszkor* ‘in the spring’ ← *tavas* ‘spring’ and *éskor* < H *ősz-kor* ‘in the autumn’ ← *ősz* ‘autumn’; of feasts, e.g. SR *húsvítkor* < H *húsvét-kor* ‘at Easter’ ← *húsvét* ‘Easter’, SR *karáčoňkor* < H *karácsony-kor* ‘at Christmas’ ← *karácsony* ‘Christmas’; and several others, e.g. SR *aratáškor* < H *aratás-kor* ‘during the

<sup>62</sup> Lexical borrowing of Hungarian degree forms is the only option with those Hungarian-origin adverbs that lack regular adjective bases, e.g. positive *d’akran* ‘often’ < H *gyakran*, comparative *d’akrabban* ‘more often’ < H *gyakrabban*, and superlative *legd’akrabban* ‘(the) most often’ < H *leggyakrabban*.

harvest’ ← *aratás* ‘harvest’. Despite of this, there is no extraction of the imported suffix *-kor* in SR, cf. (88).

(88)	‘11’	‘at 11’	‘at 11 o’clock’
H	<i>tizenegy</i>	<i>tizenegy-kor</i>	<i>tizenegy óra-kor</i>
		▽	▽
SR		<i>tizened’(-)kor</i>	<i>tizened’ óra(-)kkor</i>
	<i>dešujék</i>	* <i>dešujék-kor</i>	<i>dešujék óra(-)kkor</i>

#### 6.1.4 De-verbal adverbs

According to Kenesei *et al.* (1998: 371) a number of converbs or adverbial participles, which are marked by the suffix *-va ~ -ve*, “behave as adverbs, although they are highly suspect of being lexicalized.” Though the diachronic process is lexicalization of an inflectional verb form, on synchronic analysis the lexicalized adverbs in *-va ~ -ve* can be perhaps considered to be adverbs derived from verbs. Only lexicalized converbs get borrowed into SR.

## 6.2 Extracted affixes

### 6.2.1 The similative manner suffixes *-šon* and *-ijaššan*

The SR suffix *-šon* derives manner adverbs from two classes of adjectives. First, it productively applies to similative ethnic adjectives, deriving similative ethnic adverbs. The base adjectives of this derivation are de-nominal or, rarely, de-adjectival, e.g. *ungr-o* ‘Hungarian (man)’ → *ungr-ikán-o* ‘Hungarian-like’ → *ungr-iká-šon* ‘in a Hungarian-like way’, and *rom* ‘Gypsy (man)’ → *rom-án-o* ‘relating to Gypsies’ → *rom-an-ikán-o* ‘Gypsy-like’ → *rom-an-iká-šon* ‘in a Gypsy-like way’. Second, the suffix *-šon* derives an ‘expletive’ pro-adverb from an ‘expletive’ pro-adjective, which is in turn derived from an

‘expletive’ pro-noun: *kov-a* → *kov-ál-o* → *kov-á-šon*.<sup>63</sup> The suffix does not have any allomorphs but triggers truncation of the final consonant of the inflectional stem of its adjective base. The SR suffix *-ijaššan*, which does not show or trigger any allomorphy, derives similative manner adverbs directly from ethnic nouns, without the mediation of similative ethnic adjectives. Though only one derivation is attested in my corpus of SR, viz. *ungr-o* ‘Hungarian’ → *ungr-ijaššan* ‘in a Hungarian-like way’, it is possible that the suffix applies to further xenoclitic ethnic nouns; it certainly does not apply to the pronominal ‘expletive’, cf. *\*kov-ijaššan*.

The de-nominal derivation of similative manner adverbs in *-ijaššan* competes with the de-adjectival derivation by means of the suffix *-šon*. In fact, the single attested de-nominal adverb *ungr-ijaššan* has a fully synonymous de-adjectival counterpart, viz. *ungr-iká-šon*. The two derivations differ in that the former is a de-nominal adverb derivation with inherently similative meaning, while the derivations in *-šon* are derived from SIMILATIVE adjectives, and so they cannot be considered to mark similativity themselves.

The suffix *-šon* probably results from extraction of the imported Hungarian-origin manner suffix *-(V)n*. While the initial /š/ of the suffix may be explained by re-analysis of morpheme boundaries in the loanwords of the manner adverbs derived from adjectives in *-(V)s* (89), the origin of the vowel /o/ remains obscure. In standard Hungarian, it is only found in a few irregular manner adverbs, such as *nagy-on* ‘very’, which, however, are not borrowed into SR. Perhaps the /o/ here is a dialectal reflex of standard Hungarian /a/ in the allomorph *-an*, though presently SR freely borrows adverbs in *-an*, and so, presumably, the local dialect does not possess a regular allomorph *-on*.

---

<sup>63</sup> The ‘expletive’ pro-words are used in situations when the speaker cannot remember a certain word or does not want to pronounce it, for whatever reason, cf. the example:

*Ó-j hi kov-al-i romn-i ...*  
 3.NOM-SG COP.3.PRES EXPL-ADJ-NOM.SG.F Gypsy\_woman(F)-NOM.SG  
*no čulo furč-av-i, kov-á-šon vaker-el, džan-es?*  
 well little weird-LOAN-NOM.F.SG EXPL-ADJ-ADV speak-3SG know-2SG  
 ‘She is this kind of woman, how shall I say it ... well, a little bit weird, she talks like ...  
 you know what I mean?’

(89)	BASE	DERIVATION
H	<i>érdekes</i> ‘interesting’	<i>érdekes-en</i> ‘interestingly’
	▽	▽
SR	<i>érdekeš-n-o</i> id.	<i>érdekeš(-)šen</i> id.

The SR suffix *-ijaššan* corresponds to a sequence of several Hungarian suffixes, as found in manner adverbs derived from those similitive adjectives in *-as* that are, in turn, derived from relational adjectives in *-i*, e.g. *város* ‘town’ → *város-i* ‘of the/a town’ → *város-i-as* ‘like town people’ (adjective) → *város-i-as-an* ‘like town people’ (adverb). The SR suffix corresponds to the ‘back’ variants of the morphemes involved, and its application to the ‘back’ noun *ungr-o* conforms to the rules of Hungarian vowel harmony.

### 6.2.2 The temporal-ordinal suffix *-dikán*

The SR suffix *-dikán* derives temporal–ordinal adverbs (meaning ‘on a certain day of a/the month’) from pronominal quantifiers, i.e. pro-words expressing the ontological categories of number, amount, or temporal duration. Two of the bases are indigenous, viz. the deictic quantifier *at'i* ‘this many, this much, this long’ < OIA *ati* and the interrogative quantifier *kit'i* ‘how many, how much, how long’ < OIA *kati*, and further bases are derived by means of indefinite prefixes: cf. the specific indefinite *vala-kit'i*, the free-choice indefinite *akár-kit'i*, and the negative *ni-kit'i*. The derivation of the temporal–ordinal adverbs from the indigenous bases is shown in (90).<sup>64</sup>

(90)	BASE	DERIVATION
a.	<i>at'i</i> ‘this many’	<i>at'i-dikán</i> ‘on this (day of a/the month)’
	<i>kit'i</i> ‘how many’	<i>kit'i-dikán</i> ‘on which (d. of a/the month)’
b.	<i>vala-kit'i</i> ‘some number of’	<i>vala-kit'i-dikán</i>

<sup>64</sup> Like most quantifiers, including cardinal numerals, the base quantifiers do not inflect for agreement categories when used as attributive modifiers, although they do inflect for case when they are used in head positions and nominalized, e.g. nominative *kit'-i* vs oblique stem *kit'-en-*. The marker *-dikán* is suffixed to the modifier and nominative head forms, rather than to the inflectional stem of the base lexemes.

<i>akár-kiti</i> ‘any number of’	<i>akár-kiti-dikán</i>
<i>ni-kiti</i> ‘no number of’	<i>ni-kiti-dikán</i>

The morphological category of temporal–ordinal adverbs is productive in SR. The suffix *-dikán* competes with temporal uses of internally derived locative (masculine singular) forms of nominalized ordinal numerals, e.g. *pándž* ‘five’ → *pándž-t-o* [five-ORD-NOM.SG.M] ‘fifth’ → *pándž-t-ón-es-te* [five-ORD-OBL-OBL.SG.M-LOC] ‘on the fifth (day of a/the month)’.<sup>65</sup> In fact, the temporal locative form may also derive from the interrogative ordinal quantifier, which in turn is derived from the cardinal interrogative quantifier, e.g. *kiti* ‘how many (etc.)’ → *kiti-t-o* [how\_many-ORD-NOM.SG.M] → *kiti-t-ón-es-te* [how\_many-ORD-OBL-OBL.SG.M-LOC] ‘on which (day of a/the month)’. The forms *kiti-dikán* and *kiti-t-ón-es-te* are synonymous, though the latter appears to be restricted to the speech of older SR speakers. In a parallel fashion, the older speakers’ temporal locative forms of ordinal numerals are presently giving way to lexical borrowings of Hungarian temporal–ordinal adverbs, e.g. *etedikén* ‘on the fifth (day of a/the month)’ < H *ötödikén*. Nevertheless, Romani numerals do not allow the internal derivation of temporal–ordinal adverbs by means of the suffix *-dikán* and the pronominal quantifiers do not compete with loanwords, cf. \* < H *hányadikán* ‘on which (day of a/the month)’. This rather complex pattern is summarized in (91):

(91)	PRON. QUANTIFIER	NUMERAL
LOC	+ <i>kiti-t-ón-es-te</i>	+ <i>pándž-t-ón-es-te</i>
<i>-dikán</i>	+ <i>kiti-dikán</i>	–
LOAN	–	+ <i>etedikén</i> < H

Like the SR locatives, the Hungarian temporal–ordinal adverbs are actually temporal uses of a case form, viz. of the superessive case in *-(V)n*. The Hungarian

<sup>65</sup> The masculine singular form of the nominalized ordinal reflects, in all likelihood, the categories of the implied head noun, viz. *dí(ve)* ‘day’. However, constructions with the head noun overtly expressed, e.g. *pándž-t-e díve-s-te* [five-ORD-OBL.SG.M day-OBL.SG.M-LOC] ‘on the fifth day (of the month)’, are unattested.

adverbs in addition involve third-person singular possessor marking, which cross-references the (elidable) noun expressing the name of month (cf. Kenesei *et al.* 1998: 344). The following derivational chain (92) should elucidate the morphological structure of the Hungarian temporal–ordinal adverbs:

(92)	<i>öt</i> ‘five’	cardinal numeral
	↓	
	<i>öt-öd</i> ‘a/the fifth’	fraction numeral
	five-ORD	
	↓	
	<i>öt-öd-ik</i> ‘fifth’	ordinal numeral
	five-ORD-IDENT	
	↓	
	<i>öt-öd-ik-e</i> ‘the fifth [day of a/the month]’	temporal–ordinal NP
	five-ORD-IDENT-POSS:3SG	
	↓	
	<i>öt-öd-ik-é-n</i> ‘on the fifth [day of a/the month]’	temporal–ordinal adverbial
	five-ORD-IDENT-P:3SG-SUPERESS	

The extracted SR marker *-dikán* thus corresponds to a sequence of four Hungarian suffixes. Although lexical borrowing of ordinal numerals is attested in SR, e.g. *husadik-o* ‘twentieth’  $\triangleleft$  H *husz-ad-ik*, there is no independent extraction of the Hungarian ordinal marker *-Vd-ik-*, and so no *\*kiti-dik-o*, for instance. The marker *-dikán* thus appears to be unsegmented within SR. Next, it should be observed that all surface allomorphs of the Hungarian ordinal marker *-Vd-ik-* begin in a linking vowel, viz. *a ~ e ~ ö*, and so the consonant-initial shape of the extracted suffix *-dikán* must result from internal SR re-analysis. Finally, there is no ‘front’ variant *\*-dikén* of the extracted suffix in SR. Nevertheless, the selection of the ‘back’ variant *-dikán* with the ‘back’ base *at'i* and the neutral base *kit'i* conforms to the rules of Hungarian vowel harmony.

## 7 Affix extraction and ‘gap filling’

### 7.1 The concept of ‘gap filling’ and its application

The concept of gap filling has often been discussed as a criterion of contact-induced change. The basic idea behind the notion is that the existence of functional and/or structural ‘gaps’ in an L1 *vis-à-vis* an L2 in borrowing situations, or in an L2 *vis-à-vis* an L1 in situations of acquisitional interference, triggers or facilitates contact-induced developments of new functions and/or structures in the influenced language. The concept has been applied to a number of different types of contact-induced developments: to replica grammaticalization of functional categories such as evidentiality, tense, or possessive reflexivity, both in borrowing situations and in situations of acquisitional interference (Heine & Kuteva 2005: 125–130), to matter borrowing and pattern replication of function words (Campbell 1987, Harris & Campbell 1995: 128–130), to L1 productivity of extracted derivational affixes (Dalton-Puffer 1996: 221), or to matter borrowing of inflectional and derivational affixes (Heath 1978: 115–116). In a section on functionally based constraints on morphological borrowing, Winford (2003: 96) summarizes some of the previous views as follows: “The existence of gaps in the morphemic inventory of a recipient language facilitates the importation of new morphemes and functional categories from a source language.”

According to Heath (1978: 115), the concept of gap filling implies a THERAPEUTIC — and so, one may add, teleological — view of language change: “[b]orrowings are interpreted as devices to fill functional gaps, to create or renew useful distinctions, to establish a more symmetrical system, and so forth.” The therapeutic view is evident, for example, in Harris & Campbell’s (1995: 129–130) discussion of borrowing and replication of conjunctions from Spanish into Pipil, a language whose pre-contact devices for coordination and subordination were perceptually non-salient, and so “not as *efficient* for the hearer to process” (p. 129; emphasis mine) as the more overt and more differentiated borrowed or replicated devices. Heine & Kuteva (2005: 129), on the other hand, stress that the languages with presumed gaps “were used quite effectively prior to acquiring a new category as a result of language contact”, and so that the concept of gap

filling can be meaningful only with reference to a **TYPOLOGICAL CONTRAST** between the languages in contact.

It follows that the therapeutic and the contrastive views of gap filling also involve different conceptualizations of what the term *gap* stands for. On the former view, gap is a **DEFICIENCY** of the language being influenced, though perhaps one that can be realized as such only in contrast to the influencing language. On the latter view, gap is merely a **typological DIFFERENCE** or mismatch between the two languages in contact, though presumably only differences of a certain categorial abstractness and/or structural importance or generality would qualify as gaps in practice. This issue is rarely addressed more explicitly. I will adopt here a broad contrastive view of gaps as **CATEGORIAL** differences between languages in contact, and of gap filling as any contact-induced development that results in the reduction of these differences in such a way that the influenced language undergoes **CONVERGENCE** or assimilation to the influencing language with regard to the linguistic category in question.

As far as affix extraction is concerned, the gap filling hypothesis needs to be tested, in my view, on two hierarchical levels. Gap filling may be expected to be operative, first, with regard to lexical borrowability of morphologically complex L2 forms and, second, with regard to the actual extractability of identifiable L2-origin affixes (and, ideally, also with regard to the extent of their lexical extension). The expectations are, perhaps, that lexical borrowing of morphologically complex L2 forms expressing certain inflectional or derivational category or extraction of L2-origin affixes of a certain category will be facilitated if there is no grammatical category, ideally a regular or productive one, in the L1 that is synonymous or near-synonymous to the L2(-origin) morphological category. Note that the previous wording is asymmetrical in that an L2(-origin) morphological category is matched with a **GRAMMATICAL**, i.e. not necessarily morphological, category in the L1.<sup>66</sup>

---

<sup>66</sup> An alternative is, of course, to assume a gap in the L1 even if an L2(-origin) morphological category corresponds to an L1 *syntactic* category. This alternative is less felicitous, in my view, since what we are investigating are *categorial* mismatches and the fact that we are looking at L2(-origin) forms of a certain *morphological* category is derivative of the fact that we are exploring a morphological mechanism.

The above formulation also takes into account regularity or PRODUCTIVITY of the L1 category.<sup>67</sup> Though this might represent mixing of several criteria, it might also turn out to be a more realistic approach to the concept of linguistic category. It seems clear that the existence of merely lexical or circumlocutional, i.e. non-grammatical, means to encode a certain SEMANTIC category in the original L1 system does not disqualify the introduction, via lexical borrowing or affix extraction, of a novel morphological category of identical or similar semantics from being an instance of gap filling. Situations where the pre-borrowing or pre-extraction L1 system had some highly irregular, completely unproductive or otherwise severely restricted, nevertheless grammatical, means to encode the relevant category come close to this, although, strictly speaking, the relevant category did exist as a grammatical category prior to its expansion or renewal through borrowing or extraction. It might be useful to allow for a more flexible approach here.

Confronted with the data they studied or reviewed, most authors have grown, or remained, rather skeptical about the predictive power of the concept of gap filling, should it be formulated as a strict constraint on contact-induced language change. After reviewing numerous instances of direct diffusion (i.e. matter borrowing) of affixes between the languages of Arnhem Land, Heath (1978: 117) concludes that although gap filling can play a role in particular instances, it does “not account for very much of the morphemic diffusion” he describes. In addition to their impression that the number of uncontroversial instances of gap filling is fairly small, Heine & Kuteva (2005: 130) also draw attention to the “many cases where the model language has a grammatical category for which there is no equivalent in the replica language but where — nevertheless — no gap filling took place.”

In spite of the established observation that gap filling can only be one of many factors in contact-induced change, as a hypothesis to be tested the concept should, none the less, be formulated as a strict constraint. The strongest possible formulation involves logical equivalence, and I will stick to it for the moment (but see Section 7.4 for a revision).

---

<sup>67</sup> The productivity of the L2 category appears not to be at stake, though perhaps taking into account relative productivity of the relevant L2 and L1 categories would be a viable model.

## **7.2 'Gap filling' as a constraint on lexical borrowing**

A strong version of the gap filling hypothesis predicts that morphologically complex L2 forms expressing a certain category will be lexically borrowed into an L1 if and only if there is no regular or productive grammatical category in that L1 that is synonymous or near-synonymous to the L2 morphological category. Equivalently, it predicts that morphologically complex L2 forms expressing a certain category will not be lexically borrowed into an L1 if and only if there is a regular or productive grammatical category in that L1 that is synonymous or near-synonymous to the L2 morphological category. On other hand, the predictions of the hypothesis can be shown to be wrong if lexical borrowing of morphologically complex L2 forms takes place despite the presence of a synonymous grammatical category in the L1, or if lexical borrowing of morphologically complex L2 forms does not takes place despite the absence of a synonymous grammatical category in the L1.

There are several instances of PREDICTED lexical borrowing of morphologically complex Hungarian forms into SR. SR lacks, for example, productive means to derive denominal verbs, names of occupations and other agentive nouns, tool and place nouns, and habitual and deprivative adjectives. Lexical borrowing of these categories or, to be more precise, of their individual instances, i.e. of individual derived lexemes, can thus be viewed as supplying referring expressions that cannot be formed internally in SR (unless circumlocution etc. is used). Needless to say, the need to possess such referring expressions in an L1 is a matter of cultural (extralinguistic) factors and/or of linguistic contrast with the L2, rather than of any inherent need of the L1 grammatical system.

However, there are also numerous instances of UNPREDICTED lexical borrowing of morphologically complex forms from Hungarian into SR. To name the clearest examples: SR freely borrows Hungarian causative and frequentative verbs, de-verbal action nouns, de-adjectival abstract nouns, de-nominal qualitative adjectives, and de-adjectival manner adverbs, although it possesses fully productive means to derive these categories internally. It also borrows Hungarian infinitives of some verbs, although the subjunctive infinitive construction is fully productive. The fact that the productivity, or even the creation, of some of the above SR categories (viz. causativity, frequentativity, and the subjunctive infinitive) is due to pattern replication from Hungarian does not

invalidate the evidence, since the replicated morphological categories were not introduced via lexical borrowing (there is some comparative evidence that they had developed before the influx of lexical borrowings of identical functions).<sup>68</sup>

In other instances, however, the existence of a productive grammatical category in SR does appear to BLOCK lexical borrowing of morphologically complex Hungarian forms. SR has regular syntactic means to express possibility (viz. the pre-European modal particles *šaj* ‘can’ and *našti* ‘cannot’, cf. Elšík & Matras, in prep.) and adnominal possessors (viz. nominal and pronominal genitives) and, indeed, it does not borrow the possibility forms of Hungarian verbs and possessor-inflected forms of Hungarian nouns. SR also possesses productive internal derivation of factitive and inchoative verbs, associative forms of nouns, comparative and superlative forms of adjectives, and attenuative, similitive and relational adjectives and, indeed, morphologically complex Hungarian forms that express these categories are not borrowed. Again, some of the above SR categories, viz. associativity, degree, and possibly also the attenuative and similitive adjectives, have developed or remained productive due to pattern replication (and in case of the superlative due to pattern replication accompanied by affix copying) from Hungarian. All of the above instances of apparent blocking are predicted by the gap filling hypothesis.

Finally, there are only few not so clear instances of UNPREDICTED LACK of lexical borrowing. SR does not borrow the Hungarian wife derivations in *-né*, although it does not possess any grammatical category specialized for this function. Here it might be argued, however, that at least in some, though by no means all, instances the SR feminine (motion) derivation may cover the semantics of the Hungarian wife derivations.

---

<sup>68</sup> To take the example of the infinitive: while the subjunctive infinitive is present and obligatory in all South Central (and, indeed, in all Central) dialects of Romani, the Hungarian-marked infinitive is restricted to SR and a few neighbouring Romani varieties (cf. Elšík *et al.* 1999). This clearly suggests a recent local innovation.

### 7.3 'Gap filling' as a constraint on affix extraction

When applied to affix extraction proper, a strong version of the gap filling hypothesis predicts that a lexically imported L2 affix will be extracted if and only if its extraction would create a novel morphological category in L1. Equivalently, it predicts that an imported L2 affix will not be extracted if and only if its potential extraction would not create a novel morphological category in L2. On the other hand, the predictions of the hypothesis are wrong if affix extraction takes place without introducing a novel morphological category, or if affix extraction does not take place even if it would have introduced a novel morphological category.

There are quite a few instances of PREDICTED extraction of lexically imported Hungarian-origin affixes in SR, i.e. such instances that introduce a novel morphological category into the language. First, the categories of tool nouns, pecunial nouns, habitual adjectives, and inherently similitive manner adverbs can all be shown to have been absent from SR before the extraction of the Hungarian-origin suffixes *-ó*, *-eš-*, *-óš-*, and *-ijaššan*, respectively, even though some of the latter suffixes might be unproductive.<sup>69</sup> Second, although there had been morphological means in SR to form de-nominal verbs and agentive nouns before the extraction of the productive Hungarian-origin suffixes *-áz-* and *-áš-*, respectively, the pre-extraction means were completely unproductive. A curious instance of a novel SR category is the category of artificial nouns marked by the extracted prefix *mí-*. This is only a category of compounding in the source language Hungarian, yet it has ended up as a derivational, i.e. certainly more grammatical, category in SR.

Instances of UNPREDICTED extraction are, however, as numerous as those of predicted extraction. Several lexically imported Hungarian-origin affixes have been extracted in SR, although they did not introduce a novel morphological category into the language. The Hungarian-origin causative suffix *-atat-* competes with several causative suffixes of Indo-Aryan origin, including the fully productive suffix *-av-*, which is, in addition, syntagmatically compatible with the extracted causative suffix. The Hungarian-

---

<sup>69</sup> However, the productivity of even severely lexically restricted affixes such as the tool *-ó* is difficult to evaluate, given they have been extended after all.

origin loan-verb adaptation suffix *-ál-* is syntagmatically compatible with the pre-Hungarian, Greek-origin, loan-verb adaptation suffix *-in-*. The extracted infinitive suffix *-ñi* competes with a non-finite subjunctive construction. Although the subjunctive infinitive results from pattern borrowing from Hungarian, i.e. from the source language of the extracted affix, there is comparative and language-internal evidence that the extraction of the infinitive suffix must have taken place *after* the development of the subjunctive infinitive.<sup>70</sup> The Hungarian-origin de-verbal action suffix *-áš- ~ -iš-* competes with the synonymous productive suffix *-ibe* of Indo-Aryan origin. The Hungarian-origin suffix *-(a)šág-* competes with the suffix *-ipe* of Indo-Aryan origin in deriving abstract nouns from nouns. Though lexically restricted and perhaps unproductive, the derivational category of de-nominal abstract nouns can be shown to have been present in SR before the extraction. Though mostly applying to similitive adjectives, the Hungarian-origin manner suffix *-šon* is not inherently similitive, and so it competes with several pre-Hungarian suffixes deriving manner adverbs. The Hungarian-origin suffix *-dikán*, which marks temporal–ordinal adverbs, competes with internal adverbial locatives in this function. Finally, although the extraction of the Hungarian-origin separative and posterior-durative suffix *-tú* helped to restore certain formal distinctions in the spatial category of orientation (and in the temporal relations modelled on it), it did not introduce the category of separative orientation or posterior-durativity into SR.

Of the test cases of extractable but unextracted Hungarian-origin affixes in SR most appear to conform to the predictions of the gap filling hypothesis. Above all, extraction of the de-adjectival abstract suffix *-šág- ~ -šíg-*, which is freely imported within lexical borrowings, appears to have been BLOCKED by the presence of the productive internal derivation of de-adjectival abstract nouns by means of the Indo-Aryan suffix *-ipe*. The Hungarian ordinal and multiplicative suffixes remain restricted to Hungarian numerals, while Romani numerals employ the productive pre-Hungarian

---

<sup>70</sup> To take the example of the infinitive: while the subjunctive infinitive is an absolutely productive morphosyntactic construction, the infinitive suffix is lexically restricted to a class of verbs derived by another extracted affix, which is likewise absent in the closely related South Central dialects. It is quite safe to conclude that the extraction of the Hungarian-origin infinitive suffix did not introduce the infinitive category into SR.

ordinal and multiplicative suffixes. There are several productive adjective-deriving suffixes in SR, and (so) the adjective-deriving suffix *-(V)š-* is not extracted from the numerous Hungarian loanwords.

Finally, there appear to be few, if any, unambiguous instances of UNPREDICTED LACK of extraction in SR. Good examples of this class of instances should be lexically imported Hungarian-origin affixes that *have not* been extracted, even though they would have introduced a novel morphological category into the language if they *had* been extracted. What comes closest to this delimitation are perhaps the temporal suffix *-kor* and the inessive and superessive suffixes in the numerous temporal and spatial adverbs borrowed from Hungarian. Although they compete with pre-Hungarian morphological means of encoding temporal and spatial relations, viz. with the derivations in *-e ~ -i* and the adverbial locative, and with adpositional encoding, they appear to have introduced categorial distinctions into the domain of temporal and spatial adverbs that were absent previously.

## **7.4 Discussion**

At first sight, the criterion of gap filling appears to fail as a predictor, or strict constraint, when applied to the SR data. Both with regard to lexical borrowing of morphologically complex Hungarian forms into SR and with regard to extraction of Hungarian-origin affixes within SR all four logically possible outcomes are attested: not only the predicted instances of borrowing or extraction and of lack of borrowing or extraction, but also the unpredicted instances of borrowing or extraction and of lack of borrowing or extraction.

On closer inspection, however, it turns out that, despite of this, there is a (non-trivial) empirical generalization to be made. Note that the instances of unpredicted lack of lexical borrowing and, more importantly, also the instances of unpredicted lack of affix extraction of extractable affixes are relatively few (definitely less numerous than the instances of each of the other three types of outcome), and that these few instances are, moreover, not among the ones that are readily interpretable. This means that, if we give up the logically stronger equivalence-based version of the gap filling hypothesis, we may formulate a weaker, viz. IMPLICATIONAL, version of it, which is a much more successful

generalization.<sup>71</sup> The implicational formulations of the gap filling hypothesis that appear to hold for the SR data, at least as statistical tendencies, are:

- (93) IF SR lacks a (productive) grammatical category that Hungarian has, THEN morphologically complex L2 word forms expressing this category will be borrowed.
- (94) IF SR lacks a (productive) grammatical category that Hungarian has, THEN extractable L2-origin affixes expressing this category will be extracted.

Or, to put it more plainly: if there is a gap, it will be filled. Importantly, being implications rather than equivalences, the generalizations in (93) and (94) say nothing about the instances where there is no gap in SR: morphologically complex Hungarian word forms may be borrowed and extractable Hungarian-origin affixes may be extracted even if they express categories that had been present in SR before. In other words, should I continue to use the metaphors of gaps and filling, OVERFILLING (as well as proper filling) is allowed in SR, but UNDERFILLING is not (or is only rare: these are the mostly ambiguous instances of unpredicted lack of borrowing or extraction). Importantly, the implicational re-formulation of the gap filling hypothesis completely invalidates the metaphor of BLOCKING (i.e. lack of borrowing or extraction predicted under the equivalence formulation), which I have employed as a shorthand in Sections 7.2–3. Even under the equivalence version of the gap filling hypothesis, the term ‘blocking’ is merely an economical (but misleading) way of saying that the hypothesis works with certain examples. Unless independent motivation is encountered, the concept is useless and should be abandoned.

I should stress that the generalizations in (93) and (94) are generalizations about the CURRENT OUTCOMES of the relevant borrowing mechanisms in SR *vis-à-vis* Hungarian, not attempts at formulating any generally valid constraints on borrowing. That generalizations such as these are specific to a concrete contact situation should also be

---

<sup>71</sup> That I have been inspired by implicational formulations of empirical generalizations as practiced in linguistic typology (e.g. Croft 2003) should be obvious.

clear from the metaphors of underfilling and overfilling. It is reasonable to assume that an L1 that only recently started to borrow or extract, or whatever, from an L2 will first show underfilling (categorial mismatch *vis-à-vis* that L2), and only after a prolonged contact overfilling (categorially redundant borrowing etc.) will start taking place.<sup>72</sup> In other words, it is reasonable to assume that the amount of borrowing etc. is a function of length and intensity of contact, which is, of course, a commonplace in contact linguistics (compare, for instance, Thomason & Kaufman's [1988: 65–109] 'borrowing scales').

Finally, it should be noted in connection with the concept of categorially redundant extraction (overfilling) that no instance of REPLACIVE extraction is attested in SR. All extracted Hungarian-origin affixes that do not introduce a novel morphological category add to the complexity of SR in that they are just additional allomorphs, lexically conditioned for the most part, of the category in question.

---

<sup>72</sup> The metaphors I have used here are misleading in that they might suggest that languages moving, diachronically, from the underfilling to the overfilling stage will have to pass through an in-between stage, where all gaps are filled already but categorially redundant borrowing does not take place yet. This, of course, need not be the case. Different categories are likely to behave in different ways.

## 8 Conclusions

Affix extraction is one of two contact-induced mechanisms that result in the introduction of ‘foreign’ (L2-origin) affixes into an L1. The distinction between affix extraction and affix copying (affix borrowing in the narrow sense) concerns the role of lexical borrowing: the former mechanism is mediated by lexical borrowing, while the latter mechanism is not. Three aspects of affix extraction may be distinguished: lexical borrowing of L2 word forms, paradigmatic identification of L2-origin L1 affixes within the borrowed word forms, and lexical extension of the L2-origin affixes in the L1 lexicon. Analogically, one may distinguish between L2 morphemes, IMPORTED L2-origin morphemes, and EXTRACTED L2-origin morphemes. While lexical borrowing and paradigmatic identifiability are the necessary conditions for affix extractions, lexical extension is the actualization of affix extraction.

I have discussed numerous instances of extracted Hungarian-origin affixes in SR, a variety of Romani spoken by fluent active Hungarian bilinguals. Several of the instances illustrate that the mechanism of affix extraction is not as straightforward as the examples that are usually cited in contact linguistic literature might suggest.

First of all, several of the extracted Hungarian-origin affixes (viz. *-áš<sup>-2</sup>*, *-ašág-*, *-atat-*, *-áz-*, *-eš-*, and *-šon*) do not correspond in form to any allomorph of their Hungarian source morpheme. RE-ANALYSIS OF MORPHEME BOUNDARIES must have taken place here. While re-analysis of this type is evident only after lexical extension of the L2-origin affix takes place, it strongly suggests, nevertheless, that the segmental identity of L2-origin morphemes imported within lexical borrowings might be different from the segmental identity of their L2 counterparts in the source word forms. Unanalyzed extraction of segments that are polymorphemic in the L2 (viz. *-dikán* and *-ijaššan*) appears to confirm this hypothesis. In fact, it has been illustrated that the mechanism of affix extraction rarely targets single morphemes. The relevant affixes are extracted from loanwords *together* with their pre-inflectional loanword adaptation markers (cf. the loan-adjective adaptation marker in *-óš-n-*, and the loan-verb adaptation marker in *-áz-in-*, *-al-in-*, or *-atat-in-av-*), and, it can be argued, also together with their L1 inflections. In fact, sometimes, when lexical borrowings of the relevant category are unattested, the presence

of L1 adaptational morphology can be used as indirect evidence for affix extraction, rather than affix copying. Re-analysis of morpheme boundaries might also result in a MORPHEMIC SPLIT. This is the case of the SR agentive suffix *-ás-* vs. the pecunial suffix *-eš-*, both of which have the same source morpheme in Hungarian.

Second, affix extraction may be SELECTIVE in that it affects only certain ALLOMORPHS of a poly-allomorphous L2 morpheme. This may happen irrespective of whether re-analysis of boundaries takes place, or not. In most such instances, the Hungarian-origin affixes in SR are based on the ‘back’ allomorphs of Hungarian morphemes, e.g. H *-ó ~ -ő* > SR *-ó*, H *-ós ~ -ős* > SR *-óš-*, H *-tól ~ -től* > SR *-tú*, H *-(t)at- ~ -(t)et-* > SR *-atat-*, H *-i-as-an ~ -i-es-en* > SR *-ijaššan*, H *-ság ~ -ség* > SR *-(a)šág-*, H *-Vdik-é-n ~ -Vdik-á-n* > SR *-dikán*. Exceptional is the extraction of both Hungarian allomorphs in the case of the SR action suffix *-ás-* ~ *-iš-*. It has been shown that the SR allomorph selected through extraction is always harmonic with its base, conforming to the rules of Hungarian vowel harmony. This in itself may be taken as a piece of evidence for affix extraction, rather than copying.

Significant FUNCTIONAL SHIFTS are rare: most extracted SR affixes serve identical or similar functions as their source Hungarian morphemes. The only outstanding exception is the SR separative suffix *-tú*, whose source morpheme’s original function is semantically similar, but clearly distinct (see Section 4.3.1 for details). Here the functional shift might be connected to the fact that the affix operates within a relatively self-contained, closed class of items (non-lexical spatial and temporal adverbs). The issue, however, clearly needs further investigation. The final curiosity is the RE-ANALYSIS OF THE MORPHOLOGICAL STATUS of the SR artificiality prefix *mí-*, which results from extraction of a Hungarian lexical morpheme.

Whether affix extraction will, or will not, take place is presumably determined by various FACTORS. I have only discussed one criterion that has been employed in studies on language contact, viz. the concept and hypothesis of gap filling. Other possible factors include various TRANSPARENCY criteria: semantic transparency, categorial clarity, and unifunctionality on the semantic side, and morphotactic transparency, segmentability (e.g. sharpness of boundaries), and saliency (e.g. pronouncability) on the formal side (cf. Weinreich 1953, Heath 1978, Dressler 1985, Dressler *et al.* 1987, Dalton-Puffer 1996,

Winford 2003: 94–96). Also viable with regard to affix extraction might be the concept and potential criterion of DIFFERENTIAL IDENTIFIABILITY, as determined by differing type frequencies of paradigmatically related sets of lexical borrowings that may serve as the source of extraction: presumably, the greater the number of lexical borrowings that are paradigmatically related with regard to an L2-origin (or, indeed, indigenous) affix, the more readily that affix can be identified. These ideas, however, have not been addressed in this thesis and I will not pursue them any further here.

It should be evident from the previous discussion that any attempt at explaining, or even predicting, which L2-origin affixes will, or will not, be extracted in an L1, has to work with two TIERS of constraining factors: one concerning lexical borrowing and one concerning lexical extension of affixes. In other words, the lack of extraction of an L2 affix might be due to the fact that the affix is simply *not extractable*, since there is either a constraint on lexical borrowing of morphologically complex L2 word forms that contain that affix, in which case the L2 affix does not make it into the L1 in any form, or such morphologically complex word forms are borrowed without simultaneous borrowing of their L2 morphological bases, in which case the L2 affix is only imported as a non-morphemic segment. Only the lack of extraction of an L2 affix that is in principle *extractable*, i.e. identifiable as a morpheme within L1, needs to be explained by the actual, proper, CONSTRAINTS on affix extraction. Constraints on lexical borrowing thus function, in a sense, as a secondary FILTER with regard to affix extraction. This double-tiered constraining mechanism is something that distinguishes affix extraction from affix copying (where lexical mediation does not play any role whatsoever).

As discussed in Section 7.4, the concept of constraint itself should be employed with caution. Although it has been shown that an implicational version of the gap filling hypothesis works with regard to the SR data, it also became obvious that the neat generalization (SR allows overfilling but not underfilling with regard to lexical borrowing of morphologically complex forms and affix extraction) is a statement about the current state of affairs in a particular language contact situation, rather than a prediction concerning language change.

## Symbols and abbreviations

$X < Y$	historical relation: X has developed from Y
$X \triangleleft Y$	borrowing relation: X is a lexical borrowing of Y
$X \leftarrow Y$	morphological relation: X is derived from (based on) Y
1	1st person
2	2nd person
3	3rd person
ACC	accusative
ADJ	adjective-deriving marker, adjectivizer, adjectival agreement
ADV	adverb-deriving marker
Arm	Armenian
Arom	Aromanian
ASS	associative
ATTR	attributive, attribute
CONJ	conjunction
COMP	complementizer
COP	copula / verb of existence
DEF	definite article
ER	Early Romani
EXPL	expletive
F	feminine
FUT	future
G.	Gypsy
GEN	genitive
H	Hungarian
Ik	(Ikavian) Serbian/Croatian
IMP	imperative
INF	infinitive
ITR	intransitive
LOAN	loanword adaptation marker

LOC	locative
M	masculine
NSC	Northern (non-Vedic) South Central Romani dialects
NEG	negator, negative form
NOM	nominative
OBL	oblique case, oblique stem
OPT	optative particle
ORD	ordinal marker
POSS	possessor
PRED	predicative, predicate
PRES	present
PFV	perfective
PL	plural
PRET	preterite
PTC	participle
Q	question particle
REM	remote = imperfect tense or conditional mood
SG	singular
SOC	sociative (instrumental/comitative) case
SR	Selice Romani
SUBJ	subjunctive
SUPERESS	superessive (case)
TR	transitive
VERB	verb-deriving marker, verbalizer

## References

- Abandolo, Daniel. 1988. *Hungarian inflectional morphology*. Budapest: Akadémiai Kiadó.
- Bakker, Peter. 1997. Athematic morphology in Romani: the borrowing of a borrowing pattern. In: Matras, Yaron, Peter Bakker, and Hristo Kyuchukov (eds.) *The typology and dialectology of Romani*. Amsterdam: John Benjamins. 1–21.
- Boretzky, Norbert, and Birgit Iгла. 1994. *Wörterbuch Romani-Deutsch-Englisch für den südosteuropäischen Raum: mit einer Grammatik der Dialektvarianten*. Wiesbaden: Harrassowitz.
- Boretzky, Norbert. 1995. Armenisches im Zigeunerischen (Romani und Lomavren). *Indogermanische Forschungen* 100, 137–155.
- Boretzky, Norbert. 1996. The ‘new infinitive’ in Romani. *Journal of the Gypsy Lore Society*, fifth series, 6, 1–51.
- Boretzky, Norbert. 1999. Die Gliederung der Zentralen Dialekte und die Beziehungen zwischen Südlichen Zentralen Dialekten (Romungro) und Südbalkanischen Romani-Dialekten. In: Halwachs & Menz 1999, 210–276.
- Boretzky, Norbert. 1999. Grammatical interference in Romani: Loan formations for foreign categories. *Acta Linguistica Hungarica* 46, 3–4, 169–200.
- Boretzky, Norbert, and Birgit Iгла. 2004. *Kommentierter Dialektatlas des Romani*. Wiesbaden: Harrassowitz.
- Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.
- Campbell, Lyle. 1987. Syntactic change in Pipil. *International Journal of Americal Linguistics* 53, 3, 253–280.
- Croft, William. 2003. *Typology and universals. Second Edition*. Cambridge: Cambridge University Press.
- Dalton-Puffer, Christiane. 1996. *The French influence on Middle English morphology*. Berlin & New York: Mouton de Gruyter.
- Dressler, Wolfgang U. 1985. On the predictiveness of natural morphology. *Journal of Linguistics* 21, 321–327.

- Dressler, Wolfgang U, Willi Mayerthaler, Oswald Panagl, and Wolfgang Wurzel. 1987. *Leitmotifs in Natural Morphology*. Amsterdam: Benjamins.
- Elšík, Viktor. 2002. The synchrony and diachrony of vowel length in Šóka Romani. *The 6th International Conference on Romani Linguistics*. Graz, 12–14 September 2002. Conference presentation.
- Elšík, Viktor. 2003. Interdialect contact of Czech (and Slovak) Romani varieties. *International Journal of the Sociology of Language* 2003, 162, 41–62.
- Elšík, Viktor. 2004. Unusual features in Šóka Romani: Associative and ostensives. *International Colloquial on Romani Linguistics*. Graz, 11–13 November 2004.
- Elšík, Viktor, Milena Hübschmannová, and Hana Šebková. 1999. The Southern Central (*ahi-imperfect*) Romani dialects of Slovakia and northern Hungary. In: Halwachs & Menz 1999, 277–390.
- Elšík, Viktor and Yaron Matras. 2006. *Markedness and language change: the Romani sample*. Berlin: Mouton de Gruyter.
- Elšík, Viktor and Yaron Matras. In preparation. Modals in Romani. In: Björn Hansen, Ferdinand de Haan, and Johan van der Auwera (eds.) *Modals in the languages of Europe*.
- Görög, Veronika (ed.) 1985. *Berki János mesél cigány és magyar nyelven*. Budapest.
- Halwachs, Dieter W. and Florian Menz (eds.) 1999. *Die Sprache der Roma. Perspektiven der Romani-Forschung in Österreich im interdisziplinären und internationalen Kontext*. Klagenfurt: Drava.
- Harris, Alice C. and Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Haspelmath, Martin. 1993. More on the typology of inchoative/causative verb alternations. In: Comrie, Bernard and Maria Polinsky (eds.) *Causatives and transitivity*. Amsterdam: John Benjamins. 87–120.
- Haspelmath, Martin. 2002. *Understanding morphology*. London: Arnold.
- Heath, Jeffrey. 1978. *Linguistic diffusion in Arnhem Land*. (= Australian Aboriginal Studies: Research and Regional Studies 13). Canberra: Australian Institute of Aboriginal Studies.

- Heine, Bernd and Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.
- 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*. Amsterdam: John Benjamins. 133–145.
- Hutterer, Miklós and György Mészáros. 1967. *A lovári cigány dialektus leíró nyelvtana. Hangtan, szóképzés, alaktan, szótár*. Budapest: Magyar nyelvudományi társaság.
- Kenesei, István, Vago Robert M., and Anna Fenyvesi. 1998. *Hungarian*. London & New York: Routledge.
- Laczkó, Tibor. 2001. A comprehensive analysis of *-o/-ő*, a multifunctional deverbal nominalizer in Hungarian. In: Kenesei, István (ed.) *Argument structure in Hungarian*. Budapest: Akadémiai Kiadó. 13–49.
- Matras, Yaron. 1994. Structural Balkanisms in Romani. In: Reiter, Norbert, Uwe Hinrichs & Jiřina van Leeuwen-Turnovcová (ed.) *Sprachlicher Standard und Substandard in Südosteuropa und Osteuropa*. Wiesbaden: Harrassowitz. 195–210.
- Matras, Yaron. 1995. Connective (VS) word order in Romani. *Sprachtypologie und Universalienforschung* 48, 1, 189–203.
- Matras, Yaron. 2002. *Romani: a linguistic introduction*. Cambridge: Cambridge University Press.
- Mészáros, György. 1976. The Cerhári Gipsy dialect. *Acta Orientalia Academiae Scientiarum Hungaricae* 30, 351–367.
- Mészáros, György. 1980. *A Magyarországi szinto cigányok (történetük és nyelvük)*. Budapest: Magyar nyelvudományi társaság.
- Müller, Friedrich. 1869. Beiträge zur Kenntniss der Rom-Sprache. *Sitzungsberichte der philosophisch-historischen Classe der Kaiserlichen Akademie der Wissenschaften zu Wien* 61, 1, 149–206.
- Rácz, Sándor Romano. 1994. *Kárpáti cigány – magyar, magyar – kárpáti cigány szótár és nyelvtan*. Budapest.
- Samu, Imre. 1971. *A mai magyar nyelvjárások rendszere*. Budapest: Akadémiai Kiadó.
- Schrammel, Barbara. 2005. Borrowed verbal particles and prefixes: a comparative approach. In: Schrammel, Barbara, Dieter W. Halwachs, and Gerd Ambrosch (eds.)

- General and applied Romani linguistics: Proceedings from the 6th International Conference on Romani Linguistics*. Munich: Lincom Europa. 99–113.
- Siptár, Péter and Miklós Törkenczy. 2000. *The phonology of Hungarian*. Oxford: Oxford University Press.
- Thomason, Sarah Grey. 2001. *Language contact: an introduction*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah Grey and Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Tompa, József. 1968. *Ungarische Grammatik*. The Hague: Mouton.
- Vekerdi, József. 1971. The Gurvari Gypsy dialect in Hungary. *Acta Orientalia Academiae Scientiarum Hungaricae* 24, 381–389.
- Vekerdi, József. 1983. *A magyarországi cigány nyelvjárások szótára*. Pécs.
- Vekerdi, József. 1984. The Vend Gypsy dialect in Hungary. *Acta Linguistica Academiae Scientiarum Hungaricae* 34, 65–86.
- Weinreich, Uriel. 1953. *Languages in contact: Findings and problems*. New York: Linguistic Circle of New York.
- Winford, Donald. 2003. *An introduction to contact linguistics*. Oxford: Blackwell.