

Research Article

Marko Simonović, Boban Arsenijević*

The importance of not belonging: Paradigmaticity and loan nominalizations in Serbo-Croatian¹

<https://doi.org/10.1515/opli-2018-0021>

Received March 5, 2018; accepted July 9, 2018

Abstract: In a number of Slavic and Germanic languages, various derivational affixes and morphological patterns of Latin origin are relatively common, and bear effects as abstract as deriving event nouns from verbs and property nouns from adjectives. This seems to contradict the general observation that abstract morphology typically is not subject to borrowing. We discuss the status of two Serbo-Croatian (S-C) nominalizing Latinate suffixes, *-cija* and *-itet*, complemented by one Germanic suffix, *-er*. On our analysis, these are not borrowed suffixes and derivational patterns, in the sense that they were present in another language and got copied into S-C, but rather suffixes and patterns which emerged within S-C, more specifically in the borrowed stratum of the S-C lexicon. Crucial factors in their emergence were the shared semantic properties of the nouns ending in the respective sequences (*-cija*, *-itet* and *-er*), and the quantitative properties of these sequences closely matching those of native derivational suffixes. Pragmatic, phonological and prosodic constraints apply to these derivations to the effect that the suffixes that have emerged in the borrowed domain of the lexicon never enter a competition with the native nominalization patterns.

Keywords: loan suffixes; paradigms; nominalizations, Serbo-Croatian

1 Introduction: Are Latinate suffixes borrowed suffixes?

Modern Slavic and Germanic languages are generally analysed as having a contact-induced Latinate stratum in their lexicon which is characterized by specific morphological processes. Serbo-Croatian (S-C) *konzervirati* ‘to conserve’, *konzervacija* ‘conservation’, *konzervatorij(um)* ‘conservatory’ and *konzervativan* ‘conservative’ are generally considered morphologically complex, as are their respective Dutch counterparts *conserveren*, *conservatie*, *conservatorium* and *conservatief*. In each case, the suffixes which are at play here (e.g. *-irati*, *-acija*, *-atorij(um)*, *-ativan*) seem to be restricted to Latinate words, which makes it plausible to consider them borrowed suffixes, which entered the lexicon through language contact. However, there is another option: the words which nowadays contain the Latinate suffixes could have been borrowed as simplex items, and then the suffixes may have emerged in the recipient language once there were enough

¹ An earlier version of this article was presented at SinFonIJA 6 (Syntax, Phonology and Language Analysis) in Niš, in September 2013. An early draft appeared as Chapter 11 of the doctoral thesis Simonović (2015). We would like to thank the audience of SinFonIJA 6, René Kager, Wim Zonneveld, Daniel Nikolić and two anonymous reviewers for their comments. All the remaining mistakes are ours. We acknowledge financial support from the Slovenian Research Agency (program No. P6-0382).

*Corresponding author: Boban Arsenijević, University of Graz, Graz, Österreich, Austria, E-mail: b.arsenijevic@gmail.com
Marko Simonović, University of Nova Gorica

words sharing the same sequence. In this article we argue that the latter analysis is correct.

The emergence analysis matches a number of generalisations from the language-contact literature which seem to be hard to reconcile with massive borrowing of suffixes. First, borrowing is observed to target content words rather than functional items (Backus & Versschik 2012, Backus 2001). Second, the initial stage of borrowing – code-switching – tends to target semantically specific items (Backus 2001). Finally, paradigms are generally observed not to be borrowable: in most language contact situations, there is a single form, termed the *initial surface form* in Simonović (2015), which serves as the base of the whole new lexical item in the borrowing language.²

This all makes borrowed suffixes unexpected. Since bound morphology typically has general and abstract meaning, it is expected not to be targeted by code-switching. The general, abstract meanings expressed by bound morphology and their grammatical effects (e.g. the ontological restriction to events and properties and the lexical- and syntactic-category shift effected by nominalizing or adjectivizing affixes, respectively) are typically covered by some native mechanism, so borrowing such bound morphemes would lead to multiple realizations of the same semantic content, and of the same set of syntactic features, taking place in the same lexicon and in the same set of grammatical, semantic and pragmatic contexts. Such situation is clearly dispreferred both on grounds of economy and of introducing computational optionality into the system.

In this paper we discuss three loan-origin suffixes in S-C: the deverbal nominalizing suffix *-cija* for event nouns (a cognate of the Latin *-tion*, *-aacia*, *-aatia* and *-(t)ia*), the deadjectival nominalizing suffix *-itet* for property nouns (a cognate of the Latin *-itas*) and the deverbal nominalizing suffix *-er* for nomina agentis (a cognate of the Germanic *-er*). All these suffixes have abstract and flexible semantics, they are all bound morphemes, and hence seem to contradict the generalization above if analysed as borrowed. On our emergence analysis, these data actually lend additional evidence to the generalizations on borrowing.

We present a detailed analysis of the different ways in which these three suffixes have been integrated in the recipient language, S-C. On the one hand, as described in more detail in section 4, suffixes *-cija* and *-itet* may share their stems/bases with their native counterparts *-nje* and *-ost*, respectively, and establish a systematic relation with them (they complement each other both in terms of the nominalizing semantics which they express and in terms of their paradigmatic status). On the other hand, the suffix *-er* simply competes with the native agentive nominalizing suffixes.

- | | | | |
|--------|-----------------------|-----|------------------------|
| (1) a. | manipul-acija | vs. | manipul-ir-anje |
| | manipul- <i>acija</i> | | manipul-V- <i>anje</i> |
| | ‘manipulation’ | | ‘manipulating’ |
| b. | debil-itet | vs. | debil-n-ost |
| | debil- <i>itet</i> | | debil-Adj- <i>ost</i> |
| | ‘debility’ | | ‘retardedness’ |
| c. | stop-er | vs. | *stop-ir-ač |
| | hitchhike- <i>er</i> | | hitchhike-V- <i>er</i> |
| | ‘hitchhiker’ | | |

A different pattern is observed with respect to generality: suffixes *-cija* and *-er* have both extended their domains from borrowed stems only – into the domain of native bases, while suffix *-itet* resists combination with native stems (only one example attested in the corpus).

² A famous instance of whole paradigms from two different languages being apparently present within the same code is Romani in Bulgaria, in which many Turkish verbs are conjugated using exclusively Turkish affixes (Matras 2009, 182–184). However, the boundary between the two languages seems to be somewhat of a linguists’ construct. As Matras reports, “[w]hen asked in the majority language, Bulgarian, to answer a question in *ciganski* (i.e. ‘the Gypsy language’), consultants often responded in Turkish, rather than in Romani” (184). For these speakers, there is no real distinction between Turkish and Romani items, since both are part of the in-group code opposed to the out-group code of Bulgarian. In such cases, it is hard to claim that there is strict separation between the Romani and the Turkish lexicon, which would be a necessary condition for speaking of a transfer.

- | | | |
|-----------------------|-------------------------|---|
| (2) a. zeza-an-cija | b. selj-ob-er | c. propa-l-itet (the only attested example) |
| joke-Adj- <i>cija</i> | peasant-Suff- <i>er</i> | fail-Pcpl- <i>itet</i> |
| 'fun' | 'tacky person' | 'loser' |

In a nutshell, our emergence analysis runs as follows. Loan-origin suffixes are not really borrowed: only entire nouns treated as morphological simplexes were subject to borrowing. The suffixes then reemerged in the target language from a surface generalization over a large number of items sharing the same final sequence and a set of semantic and syntactic properties. We identify three quantitatively measurable properties as crucial in this process: productivity, frequency and generality. Final sequences manifesting these properties to a degree that matches native suffixes were likely to emerge as suffixes as they provided sufficient ground for the acquisition as suffixes. Those with a relatively high productivity and a relatively high generality were even able to expand to the native domain.

Finally, having (re-)emerged in S-C, all these suffixes derived nominals that fit in the existing system of the recipient language. Event nominals and property nominals in S-C come in two types, one of which derives from syntactically complex expressions (more precisely, from predications, see Arsenijević 2010, Arsenijević & Simonović 2013, Simonović & Arsenijević 2014), and the other from structurally simplex bases. In congruence with the reemergence analysis and as confirmed by their prosodic properties, the emerging loan-origin suffixes only derive nominals of the latter type. Nomina agentis in S-C do not exhibit this duality. This has resulted in the different behavior of the event-denoting *-cija* and property-denoting *-itet* nouns on the one hand, and the agent-denoting *-er* nouns on the other in the recipient derivational system in which they emerged. The fact that the loan-origin suffixes exhibit the same morphological and semantic behavior as their native counterparts and not that of the corresponding suffixes from the source-language argues in favor of our reemergence analysis.

The rest of this article is organized as follows. Section 2 provides an analysis of the native nominalization patterns of S-C. A clear contrast is observed between the productive suffixes, which have a transparent meaning and unaccented, recessive prosody – a class we term paradigmatic suffixes – and the suffixes which display limited productivity, induce lexicalized meanings and have dominant prosody: the non-paradigmatic suffixes. Section 3 discusses the structural flattening and prosodic marking concomitant with the non-paradigmatic derivations. In section 4, we are showing that loan-stratum nominalizations have expectedly joined the class of non-paradigmatic (idiosyncratic) derivations, and we make some initial speculations concerning the trajectory of reemergence. Building on this view of reemergence, in section 5, we explore the borrowing patterns for the word classes which serve as base words in nominalizations – adjectives and verbs, comparing them to the borrowing pattern of nominalizations. We show that the way words are borrowed into S-C ensures that, on top of prosodic differences, there is a segmental mismatch between the base and the stem of the borrowed nominalization, which effectively prevents the synchronic derivation of borrowed nominalizations from being analyzable as transparent. As a consequence, borrowed nominalizations are prevented from joining the paradigms of S-C borrowed verbs and adjectives. Loan nominalizations are allowed to establish some derivational relatedness, but that is never of the purely paradigmatic kind. Section 6 reports on a quantitative corpus-based analysis of the three suffixes under scrutiny and their comparison with native suffixes. It shows that as predicted by the reemergence analysis, the three borrowed suffixes have the quantitative properties of native suffixes with a lower degree of frequency, productivity and generality. Section 7 presents the recent dynamics, in which the loan origin suffixes, especially the suffix *-cija*, are slowly approaching the native suffixes, taking on the split to more and less paradigmatic patterns, but also how they are kept at a safe distance from the core (native) patterns by their demanding prosody and specialized semantics. In section 8, some consequences for the general theory of grammar and lexicon are drawn.

2 Native nominalizations

A specificity of S-C derivational morphology is that, in deadjectival nominalizations denoting properties and deverbal nominalizations denoting events, it displays a native contrast between:

- (a) paradigmatic derivations, which strictly maintain the prosodic pattern of the base, and
- (b) non-paradigmatic derivations, which diverge from the prosodic pattern of the base.

Note that our use of the term paradigmatic is non-standard. Haspelmath (1995: 47) identifies the forms derived from one stem by means of inflection with the paradigm of the stem. He takes derivation to result in the emergence of a new stem, which, if subject to inflection, generates its own paradigm. The properties that he lists as criteria for the identification of inflectional, i.e. paradigm-generating morphology, are regularity, generality and productivity. He does not go much deeper into the characterization of these three properties: into how they are tested, measured, and what their precise definitions are. In this article, we define paradigms using a similar set of properties, yet we do not identify them with inflection. We are also leaving out the discussion of (the usefulness of) the distinction between inflection and derivation altogether due to space limitations. Yet, we take the distinction between paradigmatic and non-paradigmatic spaces to be directly determined by structural properties rather than by the traditional intuitive divide between inflection and derivation.

We adopt the view from Arsenijević & Simonović (2013a), where a paradigm is the set of all the forms which are compositionally derived by a productive morphological operation (typically affixation) from one and the same stem, and which target all stems of a certain class (hence semantically regular, class-general and productive).

The asymmetry between paradigmatic and non-paradigmatic derivations is reflected in S-C in different prosodic shapes. Observe the examples in (5), where the same base adjective and the same suffix combine to derive one paradigmatic and one idiosyncratic noun. In each of the examples, the paradigmatic nominalization surfaces with the prosodic pattern identical to that of the base adjective (suggesting a prosodically weak suffix, i.e. one which surfaces without tone or stress), while the non-paradigmatic version has a prosodic pattern which indicates the involvement of a dominant suffix.

(5) Native deadjectival nominalizations

Base	Paradigmatic	Non-paradigmatic
prIvaatn-a ³	prIvaatn-oost	privAAtn-OOst
privat-FSg	privat-oost	privat-oost
‘private’	‘privateness’ (being private)	‘privacy’ (e.g. in privacy policy)
Opaasn-a	Opaasn-oost	opAAsn-OOst
dangerous-FSg	dangerous-oost	dangerous-oost
‘dangerous’	‘dangerousness’	‘danger’
sEksuaaln-a	sEksuaaln-oost	seksuAAln-OOst
sexual-FSg	sexual-oost	sexual-oost
‘sexual’	‘sexualness’	‘sexuality’

Somewhat different behavior is observed with verbs, where each verb derives only one event-noun. S-C verbs are specified for grammatical aspect, bearing the values perfective or imperfective. The default nominalizing suffix, closely matching the English *-ing* and the German *-ung* (Ignjatović 2013), is the suffix *-anje/-enje* (composed from the passive participle suffix *-an/-en* and the mass nominalizing suffix *-je*,

³ The following format is used to represent S-C prosody: capital letters mark all the syllables which get surface prominence (stress and/or high tone) and double vowels mark long vowels. Single capitalized syllable nuclei in a word stand for the falling accent (where tone and stress co-occur) and will be encountered only on the first syllable (in native words). Any two adjacent capitalized syllable nuclei stand for a rising span, in which both syllables have a high tone, with the stress falling on the leftmost syllable of the span. In sum, the capital letters can be read as indicators of high-tone-bearing syllables, whereas the distribution of stress is predictable from that of tone (e.g. *tatAtA* stands for *ta'ta_Hta_H*, whereas *tAtata* stands for *'ta_Htata*).

more precisely phonologically represented as –*VVje* due to its lengthening of the final syllable of the stem – see Arsenijević 2007 for a semantic and Simonović & Arsenijević 2014 for a phonological analysis). All and only imperfective verbs in S-C productively derive deverbal nominalizations, which remain within the paradigm in the sense adopted in this paper. A small, arbitrary subset of perfective verbs also derive –*VVje*-nominalizations, but these always have a prosodic pattern unattested elsewhere in the paradigm of the verb and typically come with a somewhat shifted or idiosyncratic semantics.

(6) Native deverbal nominalizations

Verb	Paradigmatic	Non-paradigmatic
pridružIIvA-ti ^{IMP}	pridružIIvAAAn-je	/
conjoin-Inf	conjoin-ed-N	
‘to conjoin’ (imperf.)	‘conjoining’	
pridrUUŽI-ti ^{PERF}	/	pridružEEEn-jE
conjoin-Inf		conjoin-ed-N
‘to conjoin’ (perf.)		‘accession’
rešAAvA-ti ^{IMP}	rešAAvAAAn-je	/
solve-Inf	solve-ed-N	
‘to solve’ (imperf.)	‘solving’	
rEEŠI-ti ^{PERF}	/	rešEEEn-jE
solve-Inf		solve-ed-N
‘to solve’ (perf.)		‘solution’

We can thus draw a local generalization, elaborated more extensively in Simonović & Arsenijević (2013a, 2014), that:

- imperfective verbs derive paradigmatic nominalizations with gerund semantics (*pridružIIvAAAnje* ‘conjoining’), whereas
- perfective verbs derive non-paradigmatic nominalizations (if any), typically bearing factitive interpretations or referring to the point of phase-transition in a telic eventuality, but in a majority of cases with lexicalized semantics narrowed down to a special subcase of the compositional interpretation (*pridružEEEnjE* ‘accession’).

The ability to derive paradigmatic nominalizations is caused by the ontological conservativity of suffixes. Although they change the category of the base from verbs and adjectives into nouns – they do not affect their ontological status. Verbs derive event-nouns, adjectives derive property-nouns. Ontological conservativity implies compositionality, which is one of the central components of paradigmaticity in our sense.

It is important to point out that our view of the contrast between paradigmatic and non-paradigmatic morphological operations is not discrete – we take it rather as a continuum, where pairs of operations involving the same suffix or the same set of bases typically show different degrees of paradigmaticity, correlated with the prosodic behavior.

3 Properties of non-paradigmatic derivations

Non-paradigmatic derivations display prosodic patterns which point in the direction of deleting the prosody of the base and being affix-controlled. While this description seems plausible, we consider the option of the prosodic pattern surfacing in non-paradigmatic derivations being default (also referred to as post-lexical). Under that view, we do not need to stipulate one prosodically neutral and one prosodically prominent suffix: the prosodic asymmetries derive from the structure. In order to do this, we briefly review what is known on default prosody in S-C.

In most standard analyses (Zec 1999, Becker 2007), S-C words which bear no prosodic specification in the lexicon (no lexical tone), receive a default prosodic pattern whereby the initial syllable bears both high tone and stress. This pattern is illustrated in (7).

- (7) /gledaan+VV.je/ → glEdaanje, /kosovo/ → kOsovo, /ostrvo/ → Ostrvo
 ‘watching’ ‘Kosovo’ ‘island’

There is, however, a peculiar gap in the repertoire of prosodic patterns in simplex words, which, as far as we know, has not been discussed in the literature.

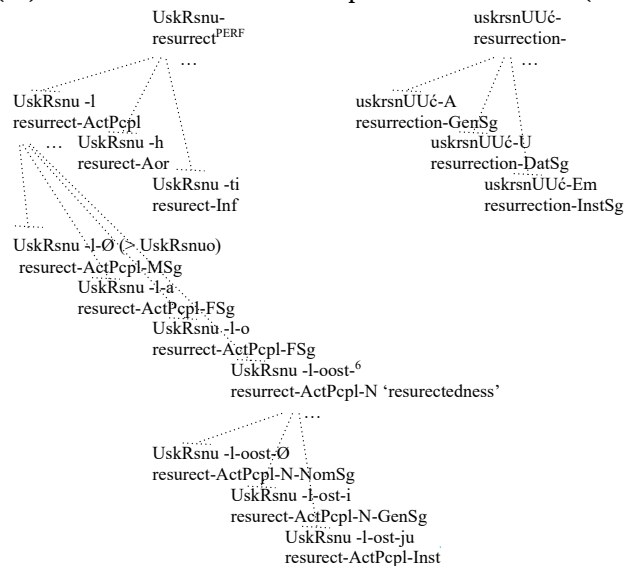
In the paradigms which have no bare, i.e. suffixless forms – the main feminine and neuter declensions in S-C – there are no morphologically simplex nouns which have the default pattern and an unstressed long syllable. Simplex nouns without a lexical prosodic specification (i.e. without a syllable specified for a high tone in the lexicon) are always like *kOsovo*, never like **kOsoovo*, as illustrated in (8). We argue that this gap supports an analysis according to which the pattern with a long rising penultima is a result of the default prosody assignment.

- (8) Simplex neuter nouns without an underlying H(igh tone)
- | | |
|---------------------|------------------------------|
| kOsovo ⁴ | <i>no internal structure</i> |
| /kosovo/ | |
| lAstovo | <i>no internal structure</i> |
| /lastovo/ | |
| Ostrvo ‘island’ | <i>no internal structure</i> |
| /ostrvo/ | |

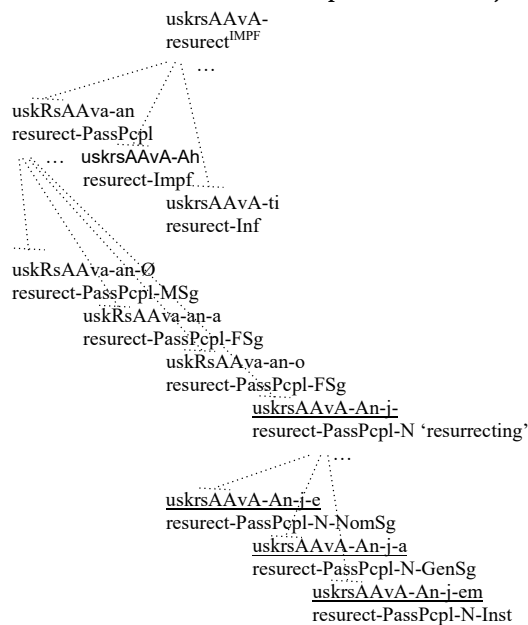
that they sit at the top of the paradigm, i.e. that there is no bigger paradigm that they are part of. Effectively, this dissociates them from the paradigm of a semantically and derivationally related stem.

In order to illustrate the hierarchical structure of paradigms, and the place of the two types of nominalization in this context, in (10) we are showing samples of the respective paradigms and their relations for the perfective and imperfective verbal stems from the base *uskrs-* 'resurrect'. In (10a), the stem of the nominalization from a perfective verb forms a paradigm of its own, independent of the paradigm of its base. In (10b), the stem of the imperfective nominalization joins, and further expands, the paradigm of the base, its paradigm is a sub-paradigm of that of the base).

(10) a. The nominalization of a perfective verb forms (i.e. sits at the top of) a distinct paradigm



b. The nominalization of an imperfective verb joins the verb's paradigm



⁶ Verbs of this class tend to drop the suffix *-nu* for a single occurrence of an event at this type of nominalization, yielding the form *UskRs-l-oost* instead of *UskRsnu-l-oost*, but both forms are attested in the corpus, and the form *UskRsnu-l-oost* is given in the interest of the uniformity of the example.

This view neatly matches the syntactic analysis of the asymmetry between paradigmatic and non-paradigmatic nominalizations in Arsenijević (2010), where paradigmatic nominalizations are nominalized predications, with the suffix merging with a complex structure and nominalizing it, as in (11a), while non-paradigmatic nominalizations have a flat structure: a structurally simplex stem is merged with a structurally simplex suffix, as in (11b).

- (11) a. [-oost [_{PredP} [_{DP} Jovan] ~~ljubazan~~ [_{AdjP} ljubazan]]] b. ljubazn-oost
 -oost J kind.Adj kind.Adj kind-oost
 Jovanova ljubazn-oost ljubAAzn-OOst
 Jovan's kind-ost kindness
 'Jovan's kindness' (trop) 'kindness' (generic)

4 Important properties of loan-nominalizations

In this section we highlight some properties characteristic of loan nominalizations, and show that they are expected given our analysis.

Linate strata in English, Dutch, German, S-C, and probably a number of other languages, have one property in common: they are all characterized by accented, dominant suffixes, as opposed to the native derivations which at least have the option of involving an unstressed, recessive suffix. This is illustrated by the Dutch examples in (12).

- (12) Unstressed native and stressed Linate suffixes
- | Base | Native | Linate | |
|--------------|---------------|--------------|--------------|
| absúrd | absúrdheid | absurdtéit | <i>Dutch</i> |
| socialiséren | socialiséring | socialisátie | |

A careful reader is probably coming up with an objection: *socialisátie* does not share the base of *socialiséren* (*socialisa-* vs. *socialiseer-*). Exactly this observation, applied to S-C and promoted to a generalization, makes the core evidence for a central argument of this paper. Distinct stems are ways for Linate nominalizations to remain at a safe distance from the paradigm that the semantics and the shared portion of segments (analysable as a base) associate it with. It is in this sense that patterns in Linate nominalizations uncover new dimensions of non-belonging. Such a constellation is very improbable to arise systematically in the native stratum.

Observe the examples in (13), from S-C.

- (13) S-C Nominalizations

Base	Paradigmatic	Non-paradigmatic
spEcijaaln-a	spEcijaaln-oost	specijAAln-OOst
special-FSg	special-oost	special-oost
'special'	'specialness' (being special)	'specialty' (expertise)
		special-ItEEt
		special-itet
		'culinary specialty'
elokvEntn-a	elokvEntn-oost	elokvEn-clja
eloquent-FSg	eloquent-oost	eloquent-cija
'eloquent'	'eloquentness' (being eloquent)	'eloquence'
defInIsa-ti	defInIsaa-nje	defInII-cija
devine-Inf	define-nje	define-cija
'to define'	'defining' (process)	'definition' (e.g. in science)
intonIIrA-ti	intonIIrAA-nje	intonAA-clja
intone-Inf	intone-nje	intone-cija
'to intone'	'intoning' (process)	'intonation' (in prosody)

As already briefly discussed in section 2, and as elaborated in e.g. Arsenijević & Simonović (2013), a special prosodic pattern for S-C non-paradigmatic derivations is no big news. Since Latinate nominalizations enter the lexicon from the outside (so, in a sense in addition to the always already present paradigmatic and non-paradigmatic derivations), it is fully expected for them to join the prosodic pattern of non-paradigmatic native derivations. However, Latinate nominalizations introduce a new level of non-belonging in a systematic way.

Due to the stratification of the lexicon, S-C derived words can be ordered in respect of their degree of non-belonging to the paradigm. Generally speaking, paradigmatic deadjectival and deverbal nominalizations share both the stem and the prosody with the related adjective or verb. Non-paradigmatic native nominalizations share the stem, but differ in the prosodic pattern. In Latinate nominalizations, we encounter a further degree: these deadjectival and deverbal derivations differ both in the stem and in the prosodic pattern from their related adjectives and verbs. This latter observation is illustrated by the underlined divergent stems in (14).

(14) Different stems in native and loan nominalizations

Base	Paradigmatic	Non-paradigmatic/native	Non-paradigmatic/loan
<u>spEcijaaln</u> -a 'special'	spEcijaaln-oost 'specialness' (one's being special)(expertise)	specijAAIn-OOst 'specialty'	specijal-ItEEt 'culinary specialty'
<u>elokvEntn</u> -a 'eloquent'	elokvEntn-oost 'eloquentness' (one's being eloquent)		elokvEn-clJa 'eloquence' (can be generic)
<u>defInIsa</u> -ti 'to define'	defInIsaa-nje 'defining' (process)		definII-cija 'definition' (in science etc.)
<u>intonIIrA</u> -ti 'to intone'	intonIIrAA-nje 'intoning' (process)		intonAA-clJa 'intonation' (the prosodic phenomenon)

The ubiquitous difference in the stems between the native and the borrowed derivation patterns originates from the way borrowing into S-C takes place, i.e. from an obligatory morphological integration of the borrowed verbal and adjectival stems by means of integration suffixes. Borrowed adjectives are normally integrated into the S-C grammar by the adjectivizing suffix *-n* (Simonović 2012), and borrowed verbs are integrated using the verbal suffixes *-isa*, *-ova* or *-ira* (Simonović & Samardžić 2013).

(15) Borrowed adjectives receive the native adjectivizing suffix *-n*

German	Suffix	S-C	
steril	-n	steril-n-a	(the ending <i>-a</i> in all 3 examples is for NomFSg)
privat	-n	privat-n-a	
elementar	-n	elementar-n-a	

(16) Borrowed verbs receive the (native) verbalizing suffixes *-isa*, *-ova* or *-ira*.

German	Suffix	S-C	
engag-ieren	-ova	angaž-ova-ti	(the ending <i>-ti</i> in all 3 examples is for Inf)
jogg-en	-ira	džog-ira-ti	
defin-ieren	-isa	defin-isa-ti	

The borrowing mechanism eliminates the possibility that a foreign adjectival/verbal stem directly becomes a S-C adjectival/verbal stem: the mechanism always adds native material first. As a consequence, Latinate nominalizations, consisting of a Latinate stem and a Latinate suffix, never share the stem of the native derivation from a Latinate base.

(17) Different stems in nominalizations derived by native and Latinate suffixes

Base	S-C nominalization	Latinate nominalization
spEcijaaln-a	spEcijaaln-oost, specijAaln-OOst	special-ItEEt
‘special’ (being special)	‘specialness’, ‘specialty’ (expertise)	‘culinary specialty’
elokvEntn-a	elokvEntn-oost	elokvEn-clja
‘eloquent’	‘eloquentness’ (being eloquent)	‘eloquence’ (rather generic)

Simonović & Samardžić (2013) show that loan verbs typically enter S-C lexicon as biaspectual, and only later have the option of getting profiled as perfective or imperfective forms. Since only imperfective verbs productively derive deverbal event-nouns, it is predicted that borrowed deverbal nominalizations (including the Latinate ones) will be able to take on the semantic pattern typical of the perfective (non-paradigmatic) deverbal nominalizations, and will hence block the emergence of native perfective (non-paradigmatic) nominalizations from borrowed verbs. This prediction is confirmed: perfective borrowed verbs never derive nominalizations using the native pattern (stress attracting –*VV,je*).

(18) Borrowed verbs do not derive non-paradigmatic native nominalizations

Verb	Paradigmatic Nominalization	Non-paradigmatic nominalization
kopIIrAti ^{BI}	kopIIrAAnje	/
‘to copy’	‘copying’	
iskopIIrAti ^{PERF}	*iskopiranje	kopija
‘to copy’		‘copy’
karikIIrAti ^{IMP}	karikIIrAAnje	/
‘to caricature’	‘caricaturing’	
iskarikIIrAti ^{PERF}	*iskarikiranje	karikatura
‘to caricature’		‘caricature’

The Germanic suffix *-er* does not show this behavior of filling the non-paradigmatic slot while leaving the paradigmatic slot to a native counterpart. This suffix simply derives the single available agentive noun from its verbal base. Just like with event- and property nouns, the agentive noun derived is always non-paradigmatic, and the suffix imposes its own prosody – but there are no paradigmatic native counterparts.

- (20) kompajl + Er → kompAjlEr adapt + Er → adAptEr blink + Er → blInkEr
 compile compiler adapt adapter blink blinker
 (cf. kompajl-ovati ‘to compile’) (cf. adapt-irati ‘to adapt’) (cf. blink-ati ‘to blink’)

This difference cannot be explained by the different source languages of these suffixes, as the property described in this section is not attested in the respective suffixes in Latin. This behavior is rather determined by the recipient lexicon. Among the native S-C nominalizations, the ontologically conservative ones (those where the derived noun refers to the same ontological class as the base, e.g. deverbal nouns referring to events and deadjectival nouns referring to properties) generally display the duality between a paradigmatic and a non-paradigmatic form, whereas non-conservative ones (those that do not refer to the same ontological class as their bases, e.g. agentive deverbal nouns – which refer to individuals, not to events like verbs they are derived from)⁷ generally display no such duality. Loan origin suffixes hence only reproduce this regularity, they simply display the properties imposed by the morphological system of S-C.

⁷ To explain the notion of ontological conservativity: this means that the deverbal nominalizations denoting events, such as gerunds, will be ontologically conservative: *to read* describes an eventuality, and *reading* does too. The category has changed, but the ontological class remained the same. Similar with deadjectival nominalizations denoting properties and tropes. The adjective *red* denotes a property, and the nominalization *redness* does too – the category changed, the ontological class did not. But agent nominalizations are not ontologically conservative. The verb *to read* describes an event, yet the noun *reader* denotes an individual, a participant in the event. Both the category and the ontological class have changed.

This is exactly what is predicted if these suffixes emerge in S-C, and the opposite of what is predicted if they are borrowed (in which case they would be expected to be borrowed with the properties they have in the source language).

5 The emergence mechanism

Our central claim is that in the initial stage of integration, loan-origin nominalizations were borrowed as simplex words, with idiosyncratic semantics (the only kind of semantics simplex words can have) and without a systematic grammatically encoded relation to their verbal or adjectival counterparts, even when these had also been borrowed. However, the rise in the number of simplex items sharing a final sequence and a number of semantic and syntactic properties (e.g. all *-tion* nouns being event nominals), while at the same time sharing a segmental sequence with a verb or adjective, led to a surface generalization which eventually triggered the emergence of a set of suffixes. These suffixes were limited in distribution to the domain of stems which do not participate in any native paradigm: they were separated from the paradigms of the related borrowed verbs and adjectives as a consequence of the requirement that a borrowed adjectival or verbal stem be integrated by a native suffix in order to derive a paradigm with native affixes. Consider the systematic lack of identity between the verbal and the nominal stem in (21).

- (21) Latinate nominalizations not part of the paradigm of the borrowed stem

<u>defIn</u> Isa-ti	<u>defIn</u> II-cija	<u>frustr</u> IIrA-ti	<u>frUstr</u> AA-cija
define-Inf	defini-cija	frustrate-Inf	frustra-cija
‘to define’	‘definition’	‘to frustrate’	‘frustration’
<u>inton</u> IIrA-ti	<u>inton</u> AA-clja	<u>evol</u> IIrA-ti	<u>evol</u> UU-clja
intone-Inf	intona-cija	evolve-Inf	evolu-cija
‘to intone’	‘intonation’	‘to evolve’	‘evolution’

That borrowed nominalizations were typically stressed on the suffix (i.e. that in Romance the respective endings are all stress-bearing) matched conveniently with the fact that in S-C non-paradigmatic derivations involve (what looks on the surface like) stressed suffixes. This has additionally supported the emergence of suffixes effectively (nearly) matching the suffixes from the language of origin. The related emergence of loan verbal and adjectival stems, as we have already described, involved the additional step of attaching a native verbal or adjectival suffix, respectively.

- (22) Borrowing by adding a native suffix and borrowing followed by suffix emergence

borrowed	setting the stem	stem mismatch in nominalizations
defin _v	→ defin + isa _v	→ definisa+nje ‘defining’
definicija _N	→ defini< >cija _N	→ defin+icija ‘definition’

Recent decades have brought about a further change in the status of the suffixes of the borrowed domain –*cija* and –*er* and to a marginal extent also –*itet* – they have started to occur with native stems, in a limited, pragmatically restricted domain of the colloquial language, in nominalizations marked for a comical, vulgar or intimate character.

- (23) Latinate suffixes with native stems

pretumba-ti	pretumba-cija	uživ-ati	uživ-an-cija
scramble-Inf	scramble-cija	enjoy-Inf	enjoy-an-cija
‘to scramble’	‘scrambling’	‘to enjoy’	‘joy’
švaler(is)a-ti	švalera-cija	za-jeb-ati	zajeb-an-cija
have_affair-Inf	have_affair-cija	for-fuck-Inf	for-fuck-an-cija
‘to have an affair’	‘having affairs’	‘to trick’	‘fun’
zanima-ti	zanima-cija	opusti-ti	opušt-en-cija

entertain-Inf	entertain- <i>cija</i>	relax-Inf	relax-en- <i>cija</i>
‘to entertain’	‘entertainment’	‘to relax’	‘relaxation’
sel-j-ak	sel-j-ob-er	propal-a	propal-itet
village-Adj-N	village-Adj-Suff- <i>er</i>	ruinous-FSg	ruinous- <i>itet</i>
‘villager’	‘tacky person’	‘ruinous’	‘looser’
balav	balav-and-er		
slobbery	slobbery-Suff- <i>er</i>		
‘slobbery’	‘stripling’		
ljepot-an	ljepot-an-er		
beauty-N	beauty-N- <i>er</i>		
‘beautiful guy’	‘beautiful guy (with a comical nuance)’		

The following section presents an analysis of these cases of “interbreeding”.

6 Productivity in the loan stratum and in general

In this section we present an additional argument for the emergence analysis of the loan-origin suffixes which is quantitative in nature, and relies on the degree of productivity and generality of the suffixes under discussion. We postulate a coconut-structured lexicon (see also Simonović & Arsenijević 2014), in which the outer layer is hard and strict, unproductive and idiosyncratic, the layer in between is somewhat softer, but still solid, with limited productivity and stratal restrictions, while the inner volume is fluid, flexible, strongly productive, interconnected, and compositional. Each layer is quantitatively characterized by certain typical measures of frequency, productivity and combinatorial selectiveness (i.e. generality, Haspelmath 1995), which generally increase from outer to inner layers. We take the layers to be formed by clusters of shared quantitative properties of these items (such as productivity and frequency), but also some interdependencies (the likelihood that if a stem occurs with one, say with the suffix *-cija*, it is likely to occur with another one, e.g. *-nt*). The layers in this structure therefore exhibit a degree of autonomy. Items combine within their layers, but not between them, and their transition from one layer to another is slow, and requires that the item reaches the quantitative properties of the target layer.

As we argue that the quasi-borrowed suffixes and derivational patterns are actually newly emerged items, occurring in the medial stratum, at the edge of the inner space, we predict that the suffixes in the focus of this paper occur with quantitative properties comparable with the lower section of the range characteristic of native suffixes. Moreover, we predict that the loan-origin suffixes will predominantly occur with the bases that they are borrowed with. Those sufficiently productive and promiscuous among them will first start to combine with other bases in their own domain of the coconut’s hard tissue – i.e. with other loan bases, and then possibly also with native stems – but the native stems they combine with will be significantly fewer in number.

An account based on real borrowing of suffixes needs to make the following two predictions:

1. borrowed suffixes will show grammatical properties from the source language – this prediction has been rejected in section 5 by the strict non-paradigmatic behavior of the conservative loan suffixes and by their properties determined by the recipient grammar;
2. borrowed suffixes will not be subject to constraints regarding the origin of the stem: once a suffix is borrowed, it is equally likely to combine with the stem it is borrowed with, with another loan base, or with a native base – this prediction we tested on a corpus.

Our corpus research is pursued on our own morphologically annotated corpus of contemporary Serbo-Croatian.⁸ We are aware that our corpus of contemporary S-C reflects the quantitative properties of the

⁸ We worked on our own Morphologically Annotated Corpus of Serbo-Croatian (MASC) with 11716712 words, involving 228702 different lemmas. The corpus is mainly formed from a selection of representative texts from the internet, with a close to equal participation of different registers and thematic fields. The representativeness of the corpus has been tested by comparing its frequency data with the data excerpted from the more established Corpus of Contemporary Serbian (CCS) of the University of Belgrade, for the commensurable domains that could be retrieved in the latter. Our corpus was used in the research as due to its morphological annotation, it offered quantitative insights which would be inaccessible for automatic retrieval in CCS.

already emerged suffixes, not the situation from which they have emerged. However, since in our quantitative reports for the loan origin suffixes we count only the borrowed nominalizations (without those in which the suffix attaches to a native stem or to a borrowed stem with which it does not combine in any of the potential languages of origin), we consider it safe to assume that such reports do not significantly depart from the situation during emergence. Finally, the process of emergence repeats at the ontogenetic level for every new acquirer of S-C, and the data we examined realistically represent a contemporary acquirer's input.

The research shows that *-cija* is by far the most frequent borrowed suffix in S-C. The frequency of occurrence of words from this class is 4.0176 occurrences per 1000 words (0.53829 for *-er*), but the average number of occurrences per member of the class in 1000 words is 0.00239 (second to the suffix *-itet*, with the average frequency of a member of the class reaching 0.00508 occurrences per 1000 words). The table in (24) shows a comparison between the three borrowed suffixes, three native suffixes of a lower productivity, and three native suffixes of a higher productivity (quantitative measures of productivity are discussed below), where the borrowed suffixes clearly pattern with the native suffixes of a lower degree of productivity.

(24) Examples for the three classes compared in respect of frequency

		Frequency of the suffix	Average frequency of a class member ⁹
borrowed	<i>-itet</i>	0.70557	0.00508
	<i>-er</i>	0.48532	0.00100
	<i>-cija</i>	4.0176	0.00239
native, low to average productivity	<i>-aj</i>	0.44518	0.01764
	<i>-ište</i>	0.72256	0.0019
	<i>-telj</i>	0.28831	0.00067
native, average to high productivity	<i>-ac</i>	8.00242	0.00294
	<i>-ica</i>	17.2147	0.00299
	<i>-ski</i>	18.4216	0.00116

We consider average frequency of a class member (which is proportional to the frequency of the suffix and counter-proportional to its generality, i.e. to the number of stems with which it combines) to be at least as important for the reemergence of these suffixes in the recipient language as absolute frequency. This measure expresses the frequency of the suffix relative to the domain of its possible application (the number of bases in the language which it may target), which determines its likelihood of occurring when the conditions for its occurrence are met.

Another important parameter is productivity. In order to emerge as a suffix from a set of simplex borrowings (in order to trigger the acquisition pattern of a productive suffix), a sequence needs to match the quantitative measure of productivity of a productive native suffix. Potential suffixes with a quantitative behavior below a productivity threshold fail to trigger the (re)analysis of a set of simplex borrowed words as words derived via a suffix. Only sequences displaying productivity patterns higher than the relevant threshold have a chance to survive as suffixes, or, in the perspective of acquisition – to emerge as such.

In the literature exploring derivational morphology from the quantitative perspective, productivity is taken to be directly linearly proportional to the number of hapaxes involving the respective suffix, and inversely linearly proportional to the number of tokens involving the suffix. The reasoning is that a productive suffix is available to appropriate stems irrespective of their frequency, and infrequent combinations are unlikely to be memorized – which is the strategy that may preserve the non-productive items. Baayen & Lieber (1991) and Plag (2003) propose the formula in (25):

⁹ Average frequency of the lexemes derived by the suffix, i.e. the frequency of the suffix divided by the number of stems it has been attested to combine with.

(25) Formula for calculating the index of productivity of an affix

$$P = n_1^{\text{aff}} / N^{\text{aff}},$$

P stands for the productivity index, n_1^{aff} is the number of hapaxes, in our case of words involving the respective suffix which appear in the corpus only once, and N^{aff} is the aggregate number of occurrences of the suffix in the corpus (aggregate number of tokens of all the words involving the suffix in the corpus). The values of P for the suffixes in (24) are as follows:

(26) Productivity indices per suffix

Loan-origin:

$P_{\text{-itet}} = 0.01113$	(rar-itet,	rar-itet,	‘rarity’)
$P_{\text{-cija}} = 0.02666$	(tradi-cija,	tradi-cija,	‘tradition’)
$P_{\text{-er}} = 0.05771$	(stop-er,	stop-er,	‘hitchhiker’)

Native derivational suffixes of low to mid P-value:

$P_{\text{-aj}} = 0.01764$	(pokuš-aj,	try-aj,	‘attempt’)
$P_{\text{-ište}} = 0.03496$	(ognj-ište,	fire-ište,	‘fire-place’)
$P_{\text{-ski}} = 0.06285$	(mor-ski,	sea-ski,	‘maritime’)

Native derivational suffixes of mid to high P-value:

$P_{\text{-n}} = 0.08935$	(led-n-o,	ice-n-NomNSg,	‘icy’)
$P_{\text{-telj}} = 0.09976$	(uči-telj,	teach-telj,	‘teacher’)
$P_{\text{-ost}} = 0.11991$	(mlad-ost,	young-ost,	‘youth’)

The observed suffixes from the domain of loanwords are of a similar degree of productivity as low to mid productivity native suffixes. In combination with their high average frequency of a class member, this has resulted in their emergence as suffixes in the recipient language. Suffix *-itet* has a low index of productivity, but a very high average frequency of a class-member, which suggests that it is unlikely to derive new nouns, but survives due to the high functional load among the existing ones. Suffixes *-cija* and *-er* have a somewhat higher productivity, which turns out to be sufficient to derive new nominalizations (see section 7).

A final relevant quantitative property is what Haspelmath (1995) terms generality. Generality is the opposite of selectivity: a derivational pattern that targets a larger number of bases is more general than the one that targets a smaller number. This property, which also relates to the intuitive notion of productivity (though not directly to productivity estimated in terms of hapaxes), can be measured by the number of stems that the suffix combines with. The table in (27) gives the number of different stems that each of the suffixes above combines with as an estimate of the degree of generality (G).

(27)

Loan-origin:

$G_{\text{-itet}} = 138$	(rar-itet,	rar-itet,	‘rarity’)
$G_{\text{-er}} = 721$	(stop-er,	stop-er,	‘hitchhiker’)
$G_{\text{-cija}} = 1681$	(tradi-cija,	tradi-cija,	‘tradition’)

Native derivational suffixes of low to mid G-value:

$G_{\text{-aj}} = 181$	(pokuš-aj,	try-aj,	‘attempt’)
$G_{\text{-ište}} = 380$	(ognj-ište,	fire-ište,	‘fire-place’)
$G_{\text{-telj}} = 429$	(uči-telj,	teach-telj,	‘teacher’)

Native derivational suffixes of a high G-value:

$G_{\text{-ost}} = 15614$	(mlad-ost,	young-ost,	‘youth’)
$G_{\text{-ski}} = 15886$	(mor-ski,	sea-ski,	‘maritime’)
$G_{\text{-n}} = 16957$	(led-n-o,	ice-n-NomNSg,	‘icy’)

Together with productivity, generality determines how likely a suffix is to expand beyond its original domain. A high value of generality means that the suffix combines with a large number of bases, which are

likely to display a higher degree of diversity. It is easier for such a set to expand than for a narrower set of rigidly constrained bases.

Suffixes *-er* and *-cija* are in the range of productivity of average to low productivity native suffixes, while *-itet* remains at the bottom of this layer. Similar holds for their generality: suffixes *-er* and *-cija* have a G value somewhat higher than that displayed by native suffixes of low to average generality, and suffix *-itet* again sits in the lower portion of this range.

The quantitative data thus support our view that suffixes *-er*, *-itet* and *-cija* have emerged in the domain of loanwords once the respective lexeme-final sequences had reached the quantitative properties of native suffixes, which has resulted in a cognitive advantage of treating them as suffixes of the loan lexical domain as opposed to memorizing the entire lexemes as simplexes. The measures we obtained indicate further that while suffixes *-cija* and *-er* behave like productive derivational mechanisms, fully active in S-C (uneven distribution of frequency over the members of the class, with a relatively large number of hapaxes, high generality), the suffix *-itet* gives quantitative effects closer to a barely productive suffix, result of a derivational mechanism that used to be more productive in the past (relatively few hapaxes and a relatively high average frequency, low level of generality).

It is expected that if any loan origin suffixes combine with native stems, those will be loan domain suffixes of higher productivity and generality. Our quantitative data support a two-step expansion. The suffix is first extended to apply also to loan domain bases which have not been borrowed with this suffix (either because they do not combine in the language of origin, or because they come from different languages), as in (28a, b). Then it expands further to native stems of the target language, as in (28c, d). Nominalizations of the former type, a combination of the suffix with a new borrowed base, are about twice as numerous as those of the latter type, where the suffix takes a native base.

- (28) a. tajkun-iza-cija c. sera-n-cija
 tycoon-V-*cija* shit-Adj-*cija*
 ‘~tycoonisation’ ‘~shitisation’
 b. cirkuz-an-er d. selj-omb-er
 circus-Suff-*er* village-Suff-*er*
 ‘circus-artist’ ‘tacky person’

Due to its relatively low levels of productivity and generality, the suffix *-itet* fails to significantly expand over the limits of the loan stems with which it has been borrowed and from which it emerged. Its existence is more due to the emergence of the corresponding adjectives and the consequent emergence of the respective bases, than to the suffix itself.

After showing in sections 6 and 7 that the loan-origin suffixes behave as determined by the recipient language grammar and semantics, in this section we showed that they show quantitative patterns predicted by the reemergence analysis. They tend to have quantitative properties in the lower portion of the range of native suffixes and to be restricted to the bases with which they were borrowed, or to gradually expand via other borrowed bases towards the native lexicon. While our analysis directly predicts this quantitative distribution, direct borrowing theories would need additional means to explain it. Perhaps the best, and simplest way to illustrate this is to compare the nouns derived by the suffixes under discussion for the number of bases borrowed with the respective suffixes versus the number of native bases with which they combine.

(29)

	Number of stems borrowed with the suffix	Number of native stems
-itet	134	1
-er	460	13
-cija	1629	22

Exactly as predicted by our reemergence hypothesis, the observed loan-origin suffixes are all primarily combined with the stems with which they have been borrowed, and marginally also with native stems.

7 Dynamics of combinations with native bases

We crucially assume that the process of emergence is mediated by acquisition, and that all suffixes, native and loan-origin in fact need to emerge and survive in the process of acquisition, under a set of quantitative conditions. Once these quantitative conditions are fulfilled by a shared final sequence in a set of loanwords with a shared semantic component and certain morpho-syntactic properties, the sequence will be reanalyzed into a suffix, i.e. simplex words will be reanalyzed into words derived by suffixes like *-er*, *-cija* and *-itet*.

The first argument for the emergence analysis in this section comes from a further degree of assimilation of the loan origin suffixes to the native system. As already specified in section 2, it is not only one Latin suffix that corresponds to the S-C *-cija*, but actually three: *-tio*, *-tia* and *-cia*. In what follows, we present data showing that just like it is the case with several native S-C suffixes (see section 2, and for more details Arsenijević & Simonović 2013), S-C nominalizing suffix *-cija* has one (emerging) increasingly paradigmatic and one non-paradigmatic version, with distinct properties in respect of the semantic contribution, selection of the stem and relation with other suffixes, such as the Latinate suffix *-(a/e)nt*. Importantly, the split between the Latin suffixes *-tio*, *-tia* and *-cia* (*-tio* for process/result denoting deverbal nominalizations, *-cia* and *-tia* mostly for deadjectival and denominal nominalizations, e.g. Miller 2006: 36, 97) does not syntactically or semantically match the split between the two versions of *-cija* in the native S-C domain. While at least a partial match is expected under a direct borrowing analysis, the re-emergence analysis predicts that the division in the recipient language will be determined only by the derivational morphology of that language.

Interestingly, the development of *-cija* in the domain of native bases of S-C mimics the twofold behavior of the native deverbal nominalizing suffix *-VVje* (Simonović & Arsenijević 2014), but with an inverse quantitative distribution. The split is between the more strongly non-paradigmatic version *-cija*, which takes stems of any category, including verbal participles, and the version *-(a)n-cija*, where *-cija* takes root-like disyllabic items turned into adjectival stems by the adjectival suffix *-an* (Simonović & Arsenijević 2013) and shows properties of paradigmaticity (i.e. of deriving items which belong to a paradigm broader than the narrow paradigm of the derived word). In the interest of simplicity, we keep referring to the two suffixes as *-cija* and *-ancija*, until we reach the arguments for unification. The split between *-cija* and *-ancija* is strongly influenced by the native phonology and morphosyntax, which makes further evidence that the suffixes *-cija* and *-ancija* have (re-)emerged in S-C.

Suffix *-ancija* mostly behaves as a regular S-C deverbal nominalizer: 1) it shares its stem with another deverbal suffix – the Latinate agentive nominalization suffix *-ant*, hence being part of a broader paradigm with some hierarchical depth and 2) it picks only imperfective verbal stems and derives process-denoting nominalizations (which may then metaphorically extend their semantics). In the latter property, it patterns with paradigmatic S-C deverbal nominalization suffixes. At another level, however – at the level of the exponence of morphemes – it shows non-paradigmatic behavior: it forces verbal stems into a disyllabic pattern, which leads to the deletion of material in stems with more than two syllables (material targeted

by deletion is underlined in the middle column in (30)¹⁰). In addition, together with loan domain nominalizations with this suffix, it shows a non-paradigmatic prosodic pattern with the accent on the antepenultimate syllable. We take the paradigmatic properties and access to the native bases to signal that the suffix is reaching the status of a native suffix – but only with a specialized semantics which prevents it from entering competition with the native morphology. The migration of the suffix towards the core of the lexicon is here again delimited by the mechanisms preventing that the domains of stems targeted by the native and of those targeted by the integrated suffixes overlap.

(30) Phonological restrictions of suffixes –*ancija* and –*ant* on the stem

udvar-AncIja,	udvAArA-ti se	udvar-Ant
court- <i>ancija</i>	court-Inf Refl	court- <i>ant</i>
‘courting’	‘to court’	‘the one who courts (pejorative)’
zafrk-AncIja,	zafrkAA-vA-ti se	zafrk-Ant
jokeV- <i>ancija</i>	joke-Inf Refl	joke- <i>ant</i>
‘joking’	‘to joke’	‘joker’
zaviti-AncIja,	zavitiAA-vA-ti se	zaviti-Ant
jokeV- <i>ancija</i>	joke-Inf Refl	joke- <i>ant</i>
‘joking’	‘to joke’	‘joker’
zajeb-AncIja,	zajebAAvA-ti se	zajeb-Ant
jokeV- <i>ancija</i>	joke-Inf Refl	joke- <i>ant</i>
‘joking’	‘to joke’	‘joker’
prevar-AncIja,	(prE)vArA-ti	prevar-Ant
cheat- <i>ancija</i>	cheat-Inf	cheat- <i>ant</i>
‘cheating’	‘to cheat’	‘cheater’
govor-AncIja	govOrI-ti	govor-Ant
talk- <i>ancija</i>	talk-Inf	talk- <i>ant</i>
‘talking/speech’	‘to talk’	‘rhetor’ (pejorative)

¹⁰ It may appear that some of the stems are actually perfective, but there is evidence that this is not the case. The perfective analysis is possible with *zajebancija*, where the imperfective verb *zajebavati* is as plausible as the perfective verb *zajebati*. Moreover, the examples *prevarancija* and *prevarant* involve as a base the infrequent colloquial secondary imperfective *prevarati* (from the perfective verb *prevariti*). Here as well, the alternative is analyzing the two as derived from the much more frequent and less restricted perfective *prevariti*. However, other examples evidence that it is not the perfective verb that acts as the base, as some of them do not even have a perfective variant (e.g. *udvarati*^{IMPF} ‘to court’: **udvoriti*^{PERF}), some have perfective variants with unrelated semantics (*zavitalavati*^{IMPF} ‘to joke’: *zavitalati*^{PERF} ‘to hurl’), and some have perfective variants with a marked perfectivizing affix (–*nu* in the perfective variant in pairs like *zafrkavati*^{IMPF} ‘to joke’: *zafrknuti*^{PERF} ‘to joke’). Finally, the last verb in (30), *govoriti*, brings crucial evidence that the contribution of the verb is limited to a bisyllabic unit ending in a consonant, since this verb has no stem ending in an *a*. We conclude that in all these cases, the deviation is from the imperfective base, but sometimes obscured by the fact that the pattern in question is highly templatic and paradigm-avoiding. From the perspective of paradigm-avoiding behavior, the dominant suffix –*AncIja* is a well-behaved non-paradigmatic suffix, as it guarantees, in two ways, not to combine with any existing allomorph of the verb: (1) the stem always surfaces without any prosodic prominence and (2) the stem does not correspond to any verbal base (due to the modification and due to the fact that verbal bases typically end in vowels).

Further evidence that only imperfective semantics underlies the bases of the –*ancija*-nouns comes from the paraphrases of these nouns in terms of relative clauses (related with the habitual nature of the agentivity).

- (i) Paraphrased agentive nominalizations
 zajebant je neko ko se zajebava/#zajebe
 teaser is someone who teases^{IMPERF/PERF}
 zafrkant je neko ko se zafrkava/#zafrkne
 joker is someone who jokes^{IMPERF/PERF}
 prevarant je neko ko vara/#prevari
 cheater is someone who cheats^{IMPERF/PERF}

The tension between the template which allows only for dysyllabic stems and the semantic pressure to select imperfective verb stems is also attested by the less frequent and strongly marked version of *zajebancija* – *zajebavancija*, extended by the imperfectivizing suffix –*va*.

In the native domain, the truly non-paradigmatic suffix *-cija* combines with non-verbal native stems (or borrowed stems which do not combine with *-cija* in the source language, as in *siguracija*) and to the stems of perfective verbs (which, as presented above, do not productively derive nominalizations). Nominalizations derived in such a way never denote processes, especially not agentive ones, but either refer to stative situations, or to objects related in one way or other to the meaning of the stem. This ending is added to adjectival and nominal stems, to onomatopoeic words or to the passive participle form of a perfective verb. These nominalizations do not have corresponding *-ant/ent* or other agentive nominals, which verifies their lack of agentive entailment and their extreme non-paradigmaticity.

(31) Stems selected by the ending *-cija*

a. Non-verbal stems:

sigur-a-cija	??sigur-a-nt	šklopo-cija	*šklopo-nt
certain-a-cija	certain-a-nt	ONOMAT-cija	ONOMAT-a-nt
‘certainty’	‘run down machine’		
oskud-a-cija	*oskud-a-nt	švaler-a-cija	*švaler-a-nt
scarce-a-cija	scarce-a-nt	womanizer-a-cija	womanizer-a-nt
‘poverty’	‘affair’		

b. Perfective verbal stems:

opušten-cija	*opušten-nt	raspušten-cija	*raspušten-nt
relaxed-cija	relaxed-nt	disbanded-cija	disbanded-nt
‘relaxedness’	‘disorganization’		
zaljubljen-cija	*zaljubljen-nt	izgubljen-cija	*izgubljen-nt
fallen_in_love-cija	fallen_in_love-nt	lost-cija	lost-nt
‘being in love’	‘confusion’		

We analyze the sequence *-an* in the suffix *-ancija* as the S-C adjectivizing suffix *-an* (the same suffix that derives the Passive Participle, Simonović & Arsenijević 2013), added to a modified verbal stem. Recall that the *-ancija* class involved a restriction to disyllabic stems and a paradigmatic kind of relation with the *-ant* class of agentive nouns.

(32) Suffixes *-(an)cija* and *-(an)t* reanalyzed

udvar-an-cija,	udvar <u>a</u> -ti	udvar-an-t
court-an-cija	court-Inf	court-an-t
‘courting’	‘to court’	the one who courts (pejorative)
zaf ^r rk-an-cija,	zaf ^r rk <u>a</u> -v <u>a</u> -ti	zaf ^r rk-an-t
jokeV-an-cija	joke-Inf	joke-an-t
‘joking’	‘to joke’	‘joker’
zav ⁱ tl-an-cija,	zav ⁱ tl <u>a</u> -v <u>a</u> -ti	zav ⁱ tl-an-t
jokeV-an-cija	joke-Inf	joke-an-t
‘joking’	‘to joke’	‘joker’
zajeb-an-cija,	zajeb <u>a</u> -v <u>a</u> -ti	zajeb-an-t
jokeV-an-cija	joke-Inf	joke-an-t
‘joking’	‘to joke’	‘joker’
prevar-an-cija,	prevar <u>a</u> -ti	prevar-an-t
cheat-an-cija	cheat-Inf	cheat-an-t
‘cheating’	‘to cheat’	‘cheater’
govor-an-cija	govor <u>i</u> -ti	govor-an-t ¹¹

¹¹ Examples like this one, where the closest vowel to the stem to be deleted is not *-a* (here *i*, in *govor-i-ti*) show that the *a* in *-an* is really part of the adjectival suffix and not an undeleted part of the stem.

talk- <i>an-cija</i>	talk-Inf	talk- <i>an-t</i>
‘talking/speech’	‘to talk’	‘rhetor’ (pejorative)

The reanalysis that licensed the extension of the suffix *-an-cija* into the native domain of stems involved the reanalysis of the sequence *-an-* in *-ant* as the native adjectival suffix, added to roots with a lexical-semantic action component. The sequence *-t* was reanalyzed as a suffix specifying the bearer of the active property. The difference between the *-cija*-class and the *-ancija*-class is not in the suffix but in the stem: what we referred to as the *-ancija* class is the suffix *-cija* added to a combination of a root and the adjectivizing suffix *-an*, while what we referred to as the *-cija* class is the same suffix in combination with stems which can be structurally more complex (e.g. nomina agentis or passive participles). This neatly fits the tendency of S-C grammar to have its suffixes split into a paradigmatic (structurally complex stems, compositionally derived, regular morphology) and non-paradigmatic (flat stems, idiosyncratic semantics, idiosyncratic choice of suffixes) variant (see Arsenijević and Simonović 2013 for a detailed discussion). In this way, suffix *-cija* acquires properties of a native suffix: it requires a native adjectivizing suffix to be added to the base before it nominalizes them, and it joins a paradigm with another suffix (in this case with the suffix *-t* from the complex *-ant*).

As discussed above, the suffix we referred to as *-ancija* forms its own new paradigms and is systematically separated from the native paradigm of the stem it selects by the mechanism of flattening instantiated in the disyllabic template. It selectively targets only (some among) those native stems which do productively derive the native deverbal *-VVje* nominalizations (imperfective stems, as briefly presented in respect of the example (6)), but it also imposes phonological constraints which lead to a truncation of these stems.

The role of native morphology in the behavior of borrowed suffixes, in terms of inserted native suffixes (e.g. *-an*) and native tendencies (the formation of paradigmatic and non-paradigmatic domains), with native grammatical and semantic oppositions realized, make a strong argument in support of the view that the loan-origin suffixes have emerged in the borrowed domain of the target language rather than being readily borrowed from the source language.

8 Conclusions

The emergence of the three loan-origin suffixes has informed our theory of architecture of the lexicon. This architecture is coconut-like in respect of productivity. Its core contains ‘soft’ stems / patterns / constructions which form domains of paradigmatic derivations, characterized by being: (a) highly productive, (b) semantically transparent, (c) general and (d) phonologically non-demanding. Next, ‘medium’ layer involves the domains of constrained productivity, with idiosyncratic semantics and own prosodic patterns. Borrowed material forms the ‘hardened’ outer layer of extreme non-belongers, with restricted compositionality and possible heavy phonological restrictions. Items and classes move between these three layers, but this movement is slow and limited in extent.

Loan nominalizations belong to a middle layer, neither fluid, nor completely hard. They have not only special prosodic shapes, but also stems which differ from those of the related verbs/adjectives from inner layers. However, they do get reanalyzed into stems and emerging suffixes, unlike the more recently borrowed English derivations of the type *reality/rijaliti* ‘reality show’ *security/sekjuriti* ‘security personnel’, *celebrity/selebriti* ‘celebrity’, *session/sešn* ‘(photo) session’, *fashion/fešn* ‘fashion’, and *vision/vižn* ‘vision’. All these items are integrated as simplex stems (e.g. *rijaliti-ja* ‘reality-GenSg’, *rijaliti-ju* ‘reality-DatSg’) and no analysis into stems and suffixes is performed.

A relatively high productivity of what looks like borrowed Latinate nominalizing suffixes is no blatant case of massive borrowing of bound morphology for general purposes. Loan nominalizations enter with idiosyncratic semantics and limited possibilities of morphological relating with other items in the lexicon, and tend to remain that way – at least until they reach certain quantitative thresholds. Derivational patterns are not really imported – they rather reemerge in the recipient language. A number of constraints stand in

the way of these patterns entering competition with the native ones, since they remain isolated by their different stems (forming separate paradigms), non-paradigmatic prosody (which specifies that these items should remain in isolation) and other restrictions (such as the restriction to disyllabic stems or the pragmatic effects, as discussed in section 5). Thus, even when they start developing productive behavior, their migration towards the soft core of the lexicon is slowed by properties reflecting their loan past.

References

- Arsenijević, Boban. 2007. Značenje srpskog imenskog sufiksa *-je* (The meaning of the Serbian nominalizing suffix *-je*). *Jezik, književnost, umetnost*, vol. 1: Jezik i društvena kretanja, Kragujevac, Srbija: FILUM, 171-176.
- Arsenijević, Boban. 2010. On two types of deadjectival nominals in Serbian. *Suvremena lingvistika* 70, 129-145.
- Arsenijević, Boban & Marko Simonović. 2013. The role of syntax in stress assignment in Serbo-Croatian. In Folli, Raffaella, Christina Sevdali, and Robert Truswell (eds.). *Syntax and its Limits*. Oxford: Oxford University Press.
- Baayen, Harald R. & Rochelle Lieber. 1991. Productivity and English derivation: a corpus based study. *Linguistics* 29: 801-844.
- Backus, Ad. 2001. The role of semantic specificity in insertional codeswitching: evidence from Dutch Turkish. In Jacobson, R. (ed.). *Codeswitching Worldwide II*. Berlin: Mouton de Gruyter:125-154.
- Backus, Ad and Anna Versschik. 2012. Copyability of (bound) morphology. In J. Johanson and M Robbeets (eds.). *Copies Versus Cognates in Bound Morphology*:123-149. Leiden: Brill.
- Becker, Michael. 2007. Tone licensing and categorical alignment in Serbo-Croatian. In Leah Bateman et al. (eds.) *University of Massachusetts Occasional Papers in Linguistics 32: Papers in Optimality Theory III*. Amherst: GLSA, pp. 1–19.
- Haspelmath, Martin. 1995. Word-class-changing inflection and morphological theory. In: Booij, Geert & van Marle, Jaap (eds) *Yearbook of Morphology 1995*, 43–66. Dordrecht: Kluwer.
- Miller, Gary. D. 2006. Latin suffixal derivatives in English and their Indo-European ancestry. Oxford: Oxford University Press.
- Plag, Ingo 2003. *Word-Formation in English*. Cambridge: Cambridge University Press.
- Simonović, Marko. 2012. The Emergence of Post-cyclic Prosody in Loanword In-tegration: Toneless Latinate Adjectives in Serbo-Croatian. *Acta Linguistica Hungarica* 59.
- Simonović, Marko. 2015. *Lexicon Immigration Service: Prolegomena to a Theory of Loanword Integration*. Ph.D thesis, Universiteit Utrecht.
- Simonović, Marko & Boban Arsenijević. 2013. In and out of paradigms: How to do everything with very few affixes. *JeNom* 6. Universitat Pompeu Fabra, Barcelona, June 20-22nd 2013.
- Simonović, Marko & Boban Arsenijević. 2014. Regular and honorary membership: On two kinds of deverbal nouns in Serbo-Croatian. *Lingua e Linguaggio*.
- Simonović, Marko & Tanja Samardžić. 2013. Exponence, productivity and default pattern - A study of verb aspect in Serbo-Croatian. Handout from the presentation at 19 ICL Geneva, 23 July 2013.
- Zec, Draga. 1999. Footed tones and tonal feet: rhythmic constituency in a pitch-accent language. *Phonology* 16. 225–264.