

The importance of not belonging: Paradigmaticity and Latinate nominalizations in Serbo-Croatian

Marko Simonović, Utrecht University

Boban Arsenijević, University of Niš

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, which display relatively high values of productivity and frequency. 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 Latinate stratum of the S-C lexicon. In other words, we argue that members of these two classes of nominalizations were first borrowed as simplex words. After demonstrating that the sequences -cija and -itet display quantitative patterns characteristic of the productive native suffixes, we argue that these quantitative effects, in conspiracy with the shared semantic properties of the nouns ending in these sequences, have led to a reanalysis of the initial simplex borrowed forms as derived nouns involving a stem and the suffix -cija or -itet, respectively. Pragmatic, phonotactic 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.

1. Abstract morphology typically is not subject to borrowing

It is a common observation in the literature on language contact that content words are much more easily borrowed than function words, and that free morphemes are much more easily borrowed than bound morphemes (Backus & Versschik 2012). This is related to another (formulation of the same) observation, namely that paradigms are generally not borrowable: in most language contact situations, there is a single form, termed the *initial surface form* in Simonović (forthcoming), which serves as the base of the whole new lexical item in the borrowing language.¹

This general observation also matches the finding that the initial stage of borrowing – code-switching – tends to target semantically specific items (Backus 2001). Since bound morphology typically has general and abstract meaning, it is expected not to be targeted by code-switching. Finally, the general, abstract meanings expressed by bound morphology and their grammatical effects (e.g. of nominalizing or adjectivizing affixes) are typically covered by some native

¹ 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.

mechanism, so borrowing such bound morphemes would lead to multiple realizations of the same semantic content, the same set of syntactic features, taking place in the same general context. Such situation is clearly dispreferred both on grounds of economy and of introducing optionality into the system.

2. ... except everywhere around

However, many if not all European languages seem to present blatant cases of exactly the situation we just described as excluded. Modern Slavic and Germanic languages have developed Latinate lexical strata, with derivational patterns which hold within the stratum and which yield semantic effects as abstract as that of nominalization, often in apparent competition with the native derivational pattern.

(1) Latinate and native nominalization patterns in (apparent) competition

	Native	Latinate
English	absurd-ness socialis-ing	absurd-ity socialis-ation
Dutch	absurd-heid socialiser-ing	absurd-iteit socialis-atie

The main goal of this paper is to explore the status of Latinate nominalizations focusing on their behavior in Serbo-Croatian (S-C). The S-C reflexes of the Latin nominalizing suffixes *-itas* (~ English *-ity*) and *-tio* (~ English *-tion*) are *-itet* and *-cija* respectively. Importantly, S-C nouns in *-cija* do not only come from the Latin suffix *-tio*, but also from the later Latin deadjectival nominalizing suffixes *-tia*, *-cia*, and *-ia* which were combined with stems ending in *-nt*.

(2) Later Latin nominalizations in *-aacia*, *-aatia* and *-(t)ia* (Miller 2006: 36-37)

konspira-cija, from Lat. conspira-acia	arogan-cija, from Lat. arrogant-(t)ia
conspire- <i>cija</i>	arrogant- <i>cija</i>
‘conspiracy’	‘arrogance’
(cf. konspirativ-n-a, conspative-Adj-FemSg,	arrogant-n-a, arrogant-Adj-FemSg)

Our main question is whether S-C Latinate nominalizations instantiate the process of borrowing derivational suffixes as abstract morphological items and structural patterns, and, relatedly, whether, or to what extent, Latinate derived nouns can be considered part of the paradigms of related adjectival and verbal stems.

We argue that, although Latinate nominalizations are ‘real’, the Latinate patterns are not borrowed but emergent in the recipient lexicon. As we will show, the lexicon prevents the emergent Latinate patterns from competing with the productive native nominalizing suffixes by integrating the former as markedly non-paradigmatic.

Note that our use of the term paradigmatic is somewhat 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 do not identify paradigms with inflection, but we do define them using similar properties. We are also leaving out the discussion of the usefulness of the distinction between inflection and derivation altogether due to space limitations. In a nutshell, our view is that the traditional boundary between inflection and derivation does not coincide with that between paradigmatic and non-paradigmatic relations: the paradigmatic space takes up all that is considered inflection and a good part of what is considered derivation morphology.

We borrow the definition of the paradigm from Arsenijević & Simonović (2013), where it is the set of all the words productively and compositionally derived from a certain stem, irrespective of the divide between inflection and derivation. This means that in order to be part of a certain paradigm (i.e. to be paradigmatic), an item needs to be formed from the exact stem the paradigm is based on, and to be semantically transparent in respect of the contribution of the stem, the other morphemes present (typically affixes) and the structure of the derivation. In this view, for instance, a deverbal nominalization pattern produces nominalizations which are members of the paradigms of a class of verbs if these nominalizations are productively and compositionally derived from that class of verbs. On the other hand, a deverbal nominalization is not member of any verbal paradigm if it idiosyncratically targets arbitrary verbs or arbitrary members of a verb class, or if its semantics tends to undergo lexicalization and appears idiomatic (a sketch is presented in section 3, and Arsenijević & Simonović 2013 give a detailed overview of deadjectival nominalization in this perspective). This view yields a hierarchical organization of morphology in terms of paradigms: the entire set of declension forms of a deverbal nominalization (inflection) is naturally part of the paradigm of the nominalization itself, but it is also true that, in case the nominalization satisfies the above conditions (paradigmatic derivation), this same set is part of the paradigm of the verb (for illustration, see (8) below).

We draw our empirical material from Serbo-Croatian (S-C), a language which, already in the native stratum, displays various morphological and phonological reflexes of paradigmaticity. The main diagnostic is the prosody of the two types of nominalizations, the paradigmatic and the non-paradigmatic ones. Paradigmatic nominalizations surface with the elsewhere attested allomorphs of the base, whereas in the non-paradigmatic nominalizations, the base surfaces with a new allomorph, which is imposed by the prosody of the suffix. As we will show, the same type of prosodic distinction obtains between the native and the Latinate nominalizations. The Latinate nominalizing suffixes are always accented and dominant – they have a lexical tone which deletes whatever prosodic information the base may have contributed. The native nominalizing suffixes, on the other hand, present a continuum whose most common (and most paradigmatic) exponents illustrate the other extreme: they have no prosody to impose on the base (i.e. they are unaccented and recessive), so that even when the base is toneless, the nominalization displays the prosodic pattern with which the base surfaces on its own.

The rest of this article is organized as follows. Section 3 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 – thus we take them as paradigmatic suffixes – and the suffixes which display limited productivity, induce lexicalized meanings and have dominant prosody: the non-paradigmatic suffixes. Section 4 discusses the structural flattening concomitant with the non-paradigmatic prosodic pattern. In section 5, we are showing that Latinate nominalizations have expectedly joined the class of non-paradigmatic (idiomatic) derivations and making some initial speculations concerning the borrowing trajectory, arguing that the nominalizations in question got borrowed as semantically idiomatic,

morphologically *simplex* items. Their derivational relation with the base then emerged due to the fact that the number of borrowed event/property nouns sharing a final sequence had reached the critical mass, hence yielding a surface effect of a reconstructed (counterpart of the) source language derivational suffix in the recipient lexicon. Building on this idea, in section 6, we explore the borrowing patterns for the word classes which serve as base words in nominalizations – adjectives and verbs, comparing them to that for 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 Latinate nominalization, which effectively prevents the synchronic derivation of Latinate nominalizations from being transparent. As a consequence, Latinate nominalizations are prevented from joining the paradigms of S-C Latinate verbs and adjectives. In sum, Latinate items are allowed to establish some derivational relatedness, but that is never of the paradigmatic kind. Section 7 presents the recent dynamics, in which the suffix *-cija*, the S-C counterpart of the Latin *-tio/tion/cia*, is slowly approaching the native suffixes, taking on the split to paradigmatic and non-paradigmatic patterns, but also how it is kept at a safe distance from the core (native) patterns by its demanding prosody. In section 8, some consequences for the general theory of grammar and lexicon are drawn.

3. Native nominalizations

The specificity of S-C is that, in both deadjectival and deverbal nominalizations, 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.

Recent morphological literature has shown that the split to inflectional and derivational morphology, where all derivational morphology takes place before all inflection morphology, and surfaces closer to the stem, cannot be maintained (ever since Di Sciullo & Williams 1987: 69-71, see also Haspelmath 1995). This has also brought to question the notion of a paradigm. We use the notion of paradigm in a different sense, independent of the above distinction. As specified in section 2, we borrow the view from Arsenijević & Simonović (2013a), where 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.

The asymmetry between paradigmatic and non-paradigmatic derivations is reflected in different prosodic shapes in S-C. Observe the examples in (3), where the same base adjective and the same suffix combine to derive one paradigmatic and one idiomatic noun. In each of the examples, the paradigmatic nominalization surfaces with the prosodic pattern identical to that of the base adjective (the suffix is prosodically vacuous, i.e. comes without tone or stress), while the non-paradigmatic version has a prosodic pattern which indicates the involvement of a dominant suffix.

(3) Native deadjectival nominalizations

Base	Paradigmatic	Non-paradigmatic
prIvaatn-a ² privat-FSg 'private'	prIvaatn-oost privat-oost 'privateness'	privAAtn-OOst privat-oost 'privacy'
Opaasn-a dangerous-FSg 'dangerous'	Opaasn-oost dangerous-oost 'dangerousness'	opAAsn-OOst dangerous-oost 'danger'
sEksuaaln-a sexual-FSg 'sexual'	sEksuaaln-oost sexual-oost 'sexualness'	seksuAAln-OOst sexual-oost 'sexuality'

Somewhat different behavior is observed with verbs, where each verb derives only one nominalization. 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*, more precisely phonologically represented as *-VV.je* 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 *-VV.je*-nominalizations, but these always have a prosodic pattern unattested elsewhere in the paradigm of the verb and typically come with a somewhat shifted or idiomatic semantics.

(4) Native deverbal nominalizations

Verb	Paradigmatic	Non-paradigmatic
pridružIIvA-ti ^{IMP} conjoin-Inf 'to conjoin' (imperf.)	pridružIIvAA-nje conjoin-N 'conjoining'	/
pridrUUŽI-ti ^{PERF} conjoin-Inf 'to conjoin' (perf.)	/	pridružEE-njE conjoin- <i>nje</i> 'accession'
rešAAvA-ti ^{IMP} solve-Inf 'to solve' (imperf.)	rešAAvAA-nje solve- <i>nje</i> 'solving'	/
rEEŠI-ti ^{PERF}	/	rešEE-njE

2 The following format is used to represent S-C prosody: capital letters mark all the syllables which get surface prominence (stress and tone) and double vowels mark long vowels. Single capitalized vowels 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 with the stress falling on the leftmost syllable of the span. In sum, the capital letters can be read as indicators of tone-bearing syllables, whereas the distribution of stress is predictable from that of tone (e.g. *tatAtA* stands for *ta'tAtA*, whereas *tAtata* stands for *'tAtata*).

solve-Inf
'to solve' (perf.)

solve-*nje*
'solution'

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

- imperfective verbs derive paradigmatic nominalizations with gerund semantics (*pridružIIvAAnje* 'conjoining'), whereas
- perfective verbs derive non-paradigmatic nominalizations (if any) with lexicalized semantics (*pridružEEEnjE* 'accession').

4. Forced Lexicalization

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

(5) /gledaan+VV.je / → glEdaanje, /kosovo/ → KOsovo, /ostrvo/ → Ostrvo

There is, however, a peculiar gap in the repertoire of post-lexical patterns, which, as far as we know, has not been spotted 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 forms which have the post-lexical pattern with an unstressed long syllable. Simplex nouns without the high tone in the lexicon are always like *KOsovo*, never like **KOsoovo*, as illustrated in (6).

(6) Simplex H-less neuter nouns

kOsovo ³	<i>no internal structure</i>
/kosovo/	

lAstovo	<i>no internal structure</i>
/lastovo/	

ostrvo 'island'	<i>no internal structure</i>
/ostrvo/	

All the nouns which have the surface prosodic configuration involving a high tone and stress on the initial syllable and a long syllable on any other syllable are morphologically derived and compositionally interpreted, as shown by the examples in (7).

³ *Kosovo* and *Lastovo* are toponyms, originally derived as possessive adjectives (from *kOOs* 'blackbird' and *lAAstA* 'swallow' + the possessive ending *-ov* in the neuter form), but have lost the connection to their source nouns in the intuition of a contemporary speaker – partially exactly through the prosodic dissociation, as explained below.

(7) Derived H-less neuter nouns

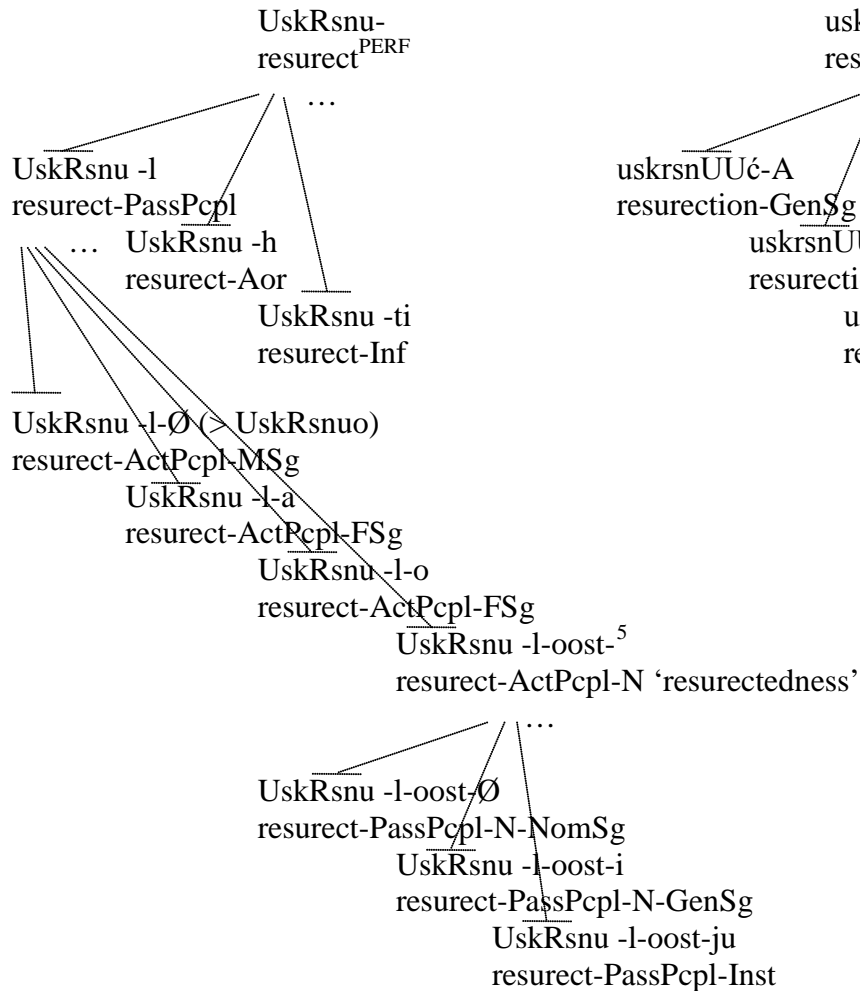
glEdaanje	prEdiivo	sklAdiište
glEda+an + VV.je	prEd + iiv + o	sklAd + iište
look+ed + N_{mass}	weave + able + $NeutSg$	order + N_{loc}
‘looking’	‘weft’	‘storage’

In sum, it is unclear what the destiny would be of the underlying representations containing no lexical tone/stress and a long vowel on a non-first syllable – all we know is that they would get neutralized with some other prosodic pattern. From the existing approaches (Zec 1999, Becker 2007), it follows that neutralization would consist in the deletion of the length. The fact that the right-aligned rising span occurs in non-paradigmatic deverbal nominalization, in which we reasonably expect loss of lexical stress due to the flattening of the structure, might indicate a need for a revision of the theory of post-lexical prosody for this type of nouns. This revision would be to the effect that the pattern in *pridružEEnje* (rising span over the two final syllables) would be post-lexical in cases where there is pre-final length, at least in the paradigms with no bare form. This would make the representation /pridružeenje/ indistinguishable from the representation /pridružeenje,H/. Importantly for our discussion – if *pridružEEnje* has the representation /pridružeenje (,H⁴)/, this means that non-paradigmatic nouns undergo *forced lexicalization*: they are forced to lose their internal structure and lexicalize as stems. This puts a prosodic stamp on them, marking that they sit at the top of the paradigm, i.e. that there is no bigger paradigm that they are part of. Effectively, this specifies their stems as separate simplex stems, and thus dissociates them from the paradigm of a semantically and derivationally related stem.

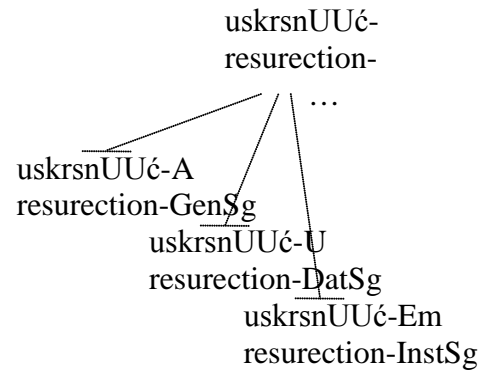
In order to illustrate the hierarchical structure of paradigms, and the place of the two types of nominalization in this context, in (8) we are showing samples of the respective paradigms and their relations for the perfective and imperfective verbal stems from the base *uskr̥s-* ‘resurrect’.

⁴ In this SBC representation, a floating high tone is associated to the stem. This representation is equivalent to the representation /violiin+a, H/ (surfacing as *violIIInA*) discussed by Zec (1999).

- (8) a. The nominalization of a perfective verb forms (i.e. sits at the top of) a distinct paradigm
paradigm of /usk^Hrsnu+ti/

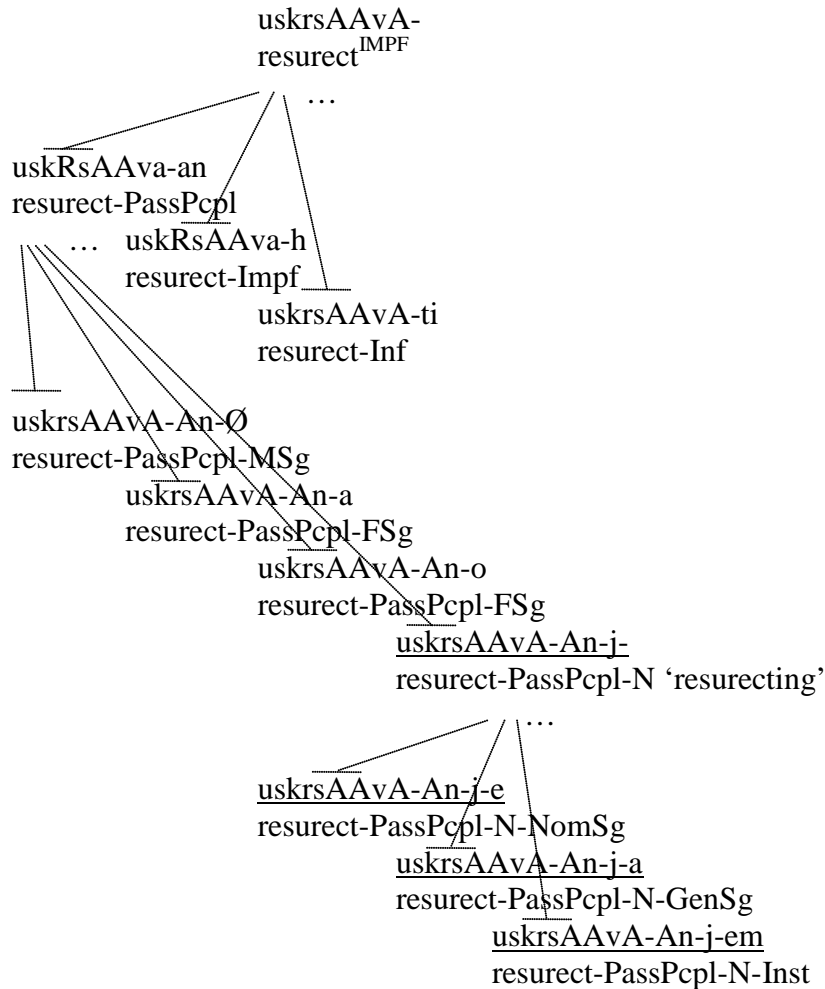


paradigm of /uskrsnUUć +e,H /



⁵ 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 uniformity of the example.

b. The nominalization of an imperfective verb joins its paradigm
paradigm of /uskrsaava^H+ti, uskrsaa^Hva+m/



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 (9a), while non-paradigmatic nominalizations have a flat structure: a structurally simplex stem is merged with a structurally simplex suffix, as in (9b).

(9) a. [-oost [PredP [DP Jovan] ljubazan [AdjP ljubazan]]]
-oost J kind.Adj kind.Adj

Jovanova ljubazn-oost
Jovan's kind-*ost*
'Jovan's kindness' (trop)

b. ljubazn-oost
kind-oost

ljubAAzn-OOst
kindness
'kindness' (generic)

5. Latinate Nominalizations

Latinate 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 (10).

(10) Unstressed native and stressed Latinate suffixes

Base	Native	Latinate	
absUrd	absUrdheid	absurditEIt	<i>Dutch</i>
socialisEren	socialisEring	socialisAtie	

A careful reader is probably coming up with an objection: *socialisAtie* does not share the base of *socialisEren* (*socialisa-* vs. *socialiseer-*). Exactly this observation, applied to S-C and promoted to a generalization, makes the base of our main argument in this paper. Distinct stems are ways for Latinate nominalizations to escape 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 Latinate nominalizations uncover new dimensions of non-belonging. Such a constellation is very improbable to arise systematically in the native stratum.

Observe the examples in (11), from S-C.

(11) S-C Nominalizations

Base	Paradigmatic	Non-paradigmatic
spEcijaaln-a special-FSg 'special'	spEcijaaln-oost special- <i>oost</i> 'specialness'	specijAAIn-OOst special- <i>oost</i> 'specialty' special-ItEEt special- <i>itet</i> 'culinary specialty'
elokvEntn-a eloquent-FSg 'eloquent'	elokvEntn-oost eloquent- <i>oost</i> 'eloquentness'	elokvEn-cIja eloquent- <i>cija</i> 'eloquence'
defInIsa-ti devine-Inf 'to define'	defInIsaa-nje define- <i>nje</i> 'defining'	definII-cija define- <i>cija</i> 'definition'
intonIIrA-ti intone-Inf 'to intone'	intonIIrAA-nje intone- <i>nje</i> 'intoning'	intonAA-cIja intone- <i>cija</i> 'intonation'

As already briefly discussed in section 3, 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 degree of

non-belonging to the paradigm. Generally speaking, paradigmatic deadjectival and deverbal nominalizations share both the stem and the prosody of the source 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 (12).

(12) Different stems in native and loan nominalizations

Base	Paradigmatic	Non-paradigmatic/native	Non-paradigmatic/loan
<u>spEcijaal</u> n-a 'special'	<u>spEcijaal</u> n-oost 'specialness'	<u>specijAA</u> n-OOst 'specialty'	<u>specijal</u> -ItEEt 'culinary specialty'
<u>elokvEntn</u> -a 'eloquent'	<u>elokvEntn</u> -oost 'eloquentness'		<u>elokvEn</u> -cIja 'eloquence'
<u>defInIsa</u> -ti 'to define'	<u>defInIsaa</u> -nje 'defining'		<u>definII</u> -cija 'definition'
<u>intonIIrA</u> -ti 'to intone'	<u>intonIIrAA</u> -nje 'intoning'		<u>intonAA</u> -cIja 'intonation'

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).

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

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

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

German	Suffix	S-C
engag-ieren	-ova	angažovati
jogg-en	-ira	džogirati
defin-ieren	-isa	definisati

The borrowing mechanism eliminates the possibility that a foreign adjectival/verbal stem directly becomes a S-C adjectival/verbal stem: it 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.

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

Base	S-C	Latinate
spEcijaaln-a 'special'	spEcijaaln-oost, specijAAIn-OOst 'specialness', 'specialty'	special-ItEEt 'culinary specialty'
elokvEntn-a 'eloquent'	elokvEntn-oost 'eloquentness'	elokvEn-cIJa 'eloquence'

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 nominalizations, 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 non-paradigmatic nominalizations using the native pattern (stress attracting -VV.*je*).

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

Verb	Paradigmatic Nominalization	Non-paradigmatic nominalization
šutIIrAti ^{IMP} 'to kick, to shoot'	šutIIrAAnje 'kicking, shooting'	
šUtnuti ^{PERF} 'to kick, to shoot'		šUt *šutnUUĆE 'kick, shoot'
štAAmpati ^{BI} 'to print'	štAAmpaanje 'printing'	
OdštAAmpati ^{PERF} 'to print out'		štAAmpa *odštampAAnje 'press'

6. The borrowing mechanism

So the question now is how does grammar plan how to integrate the borrowed nominalisations, in such a neat way that paradigms are separated and memberships are marked? And the answer is: of course it does not. Rather than being result of an actual computation, this optimal outcome emerges as a consequence of a conspiracy of a number of independently justified processes.

We argue that in the initial stage of integration, Latinate nominalizations were borrowed as simplex words, with idiomatic semantics (as simplex words are supposed to have) and unrelated to their verbal or adjectival counterparts, even when these had also been subject to borrowing. However, the rise in the number of simplex items sharing a final sequence and a number of semantic and syntactic properties (e.g. *-tion* nominalizations 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 limited in distribution to the domain of stems which do not participate in any native paradigm. The isolation from the native paradigms came as a consequence of the requirement that a borrowed stem be integrated by a native suffix in order to derive a paradigm with native affixes.

(17) Latinate nominalizations not part of the paradigm of the borrowed stem

<u>defInIsa</u> -ti	<u>definII</u> -cija	frustrIIrA-ti	frUstrAA-cija
define-Inf	defini- <i>cija</i>	frustrate-Inf	frustra- <i>cija</i>
‘to define’	‘definition’	‘to frustrate’	‘frustration’
<u>intonIIrA</u> -ti	<u>intonAA</u> -cIja	evoluIIrA-ti	evolUU-cIja
intone-Inf	intona- <i>cija</i>	evolve-Inf	evolu- <i>cija</i>
‘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 stressed suffixes. This has additionally supported the emergence of suffixes effectively (nearly) matching the suffixes from the language of origin.

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

borrowed		setting the stem		nominalizations mismatch
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 –*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.

(19) Latinate suffixes with native stems

pretumba-ti	pretumba-cija	uživ-ati	uživ-an-cija
scramble-Inf	scramble- <i>cija</i>	enjoy-Inf	enjoy- <i>an-cija</i>
‘to scramble’	‘scrambling’	‘to enjoy’	‘joy’
švaler(is)a-ti	švalera-cija	za-jeb-ati	zajeb-an-cija
have_affair-Inf	have_affair- <i>cija</i>	for-fuck-Inf	for-fuck- <i>an-cija</i>
‘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- <i>en-cija</i>
‘to entertain’	‘entertainment’	‘to relax’	‘relaxation’

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

7. Immigrants: interbreeding and productivity

This section presents arguments in support of our analysis in which the suffixes –*cija* and –*itet* are not loan suffixes (borrowed Latin suffixes –*tio/tia/cia* and –*itas*, respectively), but rather new endings emerging within the Latinate stratum of the lexicon, in the interaction of the fact that a sufficiently large number of items share certain syntactic and semantic properties while at the same time having the same final sequence, and the fact that the derivational morphology of S-C is predominantly suffixal.

We crucially assume that the process of emergence was originally mediated by acquisition, and that all suffixes, native and ‘borrowed’ in fact need to emerge and survive in the process of acquisition, under a set of quantitative conditions. Once these quantitative conditions – most of which relate to frequency and productivity – are fulfilled by a shared segment of a set of loanwords, it is fully expected that they will be reanalyzed from simplex words into words derived by suffixes like *-cija* and *-itet*.

Our first argument for the emergence analysis comes from the nature of correspondence between the Latin suffixes and their S-C counterparts. 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 (Arsenijević & Simonović 2013), in the domain of native stems, S-C nominalizing suffix *-cija* has one (emerging) increasingly paradigmatic and one non-paradigmatic version, with distinct properties in respect of the degree of productivity, semantic contribution, selection of the stem and relation with other suffixes, such as the Latinate suffix *-(a/e)nt*. Crucially, the split between *-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 (which we dub *-cija* and *-ancija* for the time being, i.e. until a unifying analysis is formulated). Interestingly, the development of *-cija* goes in the direction of mimicking the twofold behavior of the native deverbal nominalizing suffix *-VVje* (Simonović & Arsenijević 2014). The split is between the more strongly non-paradigmatic version *-cija*, which takes stems of any category, including verbal participles, and the version *-ancija*, which takes root-like disyllabic items turned into adjectival stems by the adjectival suffix *-an* (Simonović & Arsenijević 2013), and which shows signs of paradigmatic behavior (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 morpho-syntax, which makes further evidence that the suffixes *-cija* and *-ancija* have (re-)emerged in S-C.

As already hinted at the end of section 6, in present day S-C, the process of the integration of the ending *-cija* into the lexicon has entered a phase in which it is displaying limited productivity with native stems, but the domain of its productivity is pragmatically restricted to the strongly colloquial lexicon, i.e. to stems which are inappropriate for use in higher registers (swearwords, strong negative evaluative terms, morally disqualified phenomena etc.).

At some levels, suffix *-ancija* behaves as a rather 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 paradigm with some hierarchical depth (not a flat paradigm like those of typical non-members) and 2) it picks only imperfective verbal stems and derives agentive process-denoting nominalizations.⁶ In the latter property, it patterns with paradigmatic S-C deverbal

⁶ 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 (*zavitlavati*^{IMPF} ‘to joke’: *zavitlati*^{PERF} ‘to hurl’), and some have perfective variants with a marked perfectivizing

nominalization suffixes. At another level, 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 (20)). While the paradigmatic properties are part of its becoming productive, we take the non-paradigmatic behavior to be result of a mechanism that keeps the suffixes 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.

(20) Phonotactic restrictions of suffixes *-ancija* and *-ant* on the stem

udvar-AncIja, court- <i>ancija</i> 'courting'	udvAAr <u>A</u> -ti se court-Inf Refl 'to court'	udvar-Ant court- <i>ant</i> the one who courts (pejorative)
zafrk-AncIja, jokeV- <i>ancija</i> 'joking'	zafrkAA- <u>vA</u> -ti se joke-Inf Refl 'to joke'	zafrk-Ant joke- <i>ant</i> 'joker'
zavitl-AncIja, jokeV- <i>ancija</i> 'joking'	zavitlAA- <u>vA</u> -ti se joke-Inf Refl 'to joke'	zavitl-Ant joke- <i>ant</i> 'joker'
zajeb-AncIja, jokeV- <i>ancija</i> 'joking'	zajebAA <u>vA</u> -ti se joke-Inf Refl 'to joke'	zajeb-Ant joke- <i>ant</i> 'joker'

affix (*-nu* in the perfective variant in pairs like *zafrkavati*^{IMPF} 'to joke' : *zafrknuti*^{PERF} 'to joke'). Finally, the last verb in (20), *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 disyllabic 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*.

prevar-AncIja, cheat- <i>ancija</i> 'cheating'	(prE)vAra-ti cheat-Inf 'to cheat'	prevar-Ant cheat- <i>ant</i> 'cheater'
govor-AncIja talk- <i>ancija</i> 'talking/speech'	govOrI-ti talk-Inf 'to talk'	govor-Ant talk- <i>ant</i> 'rhetor' (pejorative)

The truly non-paradigmatic suffix *-cija* is added to 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.

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

a. Non-verbal stems:

sigur-a-cija certain-a- <i>cija</i> 'certainty'	??sigur-a-nt certain-a- <i>nt</i>	šklopo-cija ONOMAT- <i>cija</i> 'run down machine'	*šklopo-nt ONOMAT-a- <i>nt</i>
oskud-a-cija scarce-a- <i>cija</i> 'poverty'	*oskud-a-nt scarce-a- <i>nt</i>	švaler-a-cija womenizer-a- <i>cija</i> 'affair'	*švaler-a-nt womenizer-a- <i>nt</i>

b. Perfective verbal stems:

opušten-cija relaxed- <i>cija</i> 'relaxedness'	*opušten-nt relaxed- <i>nt</i>	raspušten-cija disbanded- <i>cija</i> 'disorganization'	*raspušten-nt disbanded- <i>nt</i>
zaljubljen-cija fallen_in_love- <i>cija</i> 'being in love'	*zaljubljen-nt fallen_in_love- <i>nt</i>	izgubljen-cija lost- <i>cija</i> 'confusion'	*izgubljen-nt lost- <i>nt</i>

Our second argument for the emergence analysis is quantitative in nature, and relies on the degree of productivity of the suffixes observed. We postulate a coconut-structured lexicon (see also Simonović & Arsenijević 2014), in which the outer layer is hard and strict, unproductive and idiomatic, the layer in between is somewhat softer, but still solid, with limited productivity, while the inner volume is flexible, strongly productive, interconnected, and compositional. Each layer is quantitatively characterized by certain typical measures of frequency, productivity and combinatorial selectiveness. As we argue that the quasi-borrowed suffixes and derivational patterns are actually newly emerged items, we predict that the suffixes in the focus of this paper occur with quantitative properties between the level of marginally productive suffixes (for Latinate suffixes less likely to combine with native stems) and those of average productivity (Latinate suffixes more liberal in respect of native stems). Accounts involving real borrowing of suffixes (borrowing accounts) make no quantitative predictions.

Our corpus research, pursued on our own morphologically annotated corpus of contemporary Serbo-Croatian shows that *-cija* is by far the most frequent borrowed suffix in S-C.⁷ We have found 1681 different nouns derived by this suffix (the next, second loan suffix in this respect is *-er*, found on 724 different agentive nouns). The frequency of occurrence of words from this class is 4.0176 occurrences per 1000 words (1.04432 for *-er*), and 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 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 four classes compared in respect of frequency

		Number of stems selected	Frequency of the suffix	Average frequency of a class member
borrowed	-itet	139	0.70557	0.00508
	-er	722	1.04415	0.00145
	-cija	1681	4.0176	0.00239
native, low to average productivity	-aj	182	0.44518	0.01764
	-ište	381	0.72256	0.0019
	-telj	430	0.28831	0.00067
native, average to high productivity	-ac	2718	8.00242	0.00294
	-ica	5756	17.2147	0.00299
	-ski	15886	18.4216	0.00116

However, while general information about frequency gives a rough general insight, the crucial parameter here is productivity. In order to emerge as a suffix from a set of simplex borrowings (especially in order to trigger the acquisition pattern of a productive suffix), a sequence needs to display the quantitative properties of a productive suffix. This is because suffixes with a quantitative behavior typical of low productivity items may disappear in favor of storing their derivations as simplex words, while suffixes displaying high productivity patterns have a strong tendency to survive as suffixes, or, in the perspective of acquisition – to emerge as such.

In the literature on derivational morphology in quantitative perspective, productivity is taken to

⁷ We worked on our own Morphologically Annotated Corpus of Serbo-Croatian (MASC) 11716712 words, involving 228702 different lemmas. The corpus is mainly formed from 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.

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 behind this is that a productive suffix is likely to be used with highly infrequent stems, while such combinations are unlikely to be stored – which is the strategy at work with non-productive items. Thus, Baayen & Lieber (1991) and Plag (2003) propose the formula in (25)

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

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

where 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 (productivity index) for the suffixes above (plus several additional native nominalization suffixes with different degrees of productivity, for comparison) are as follows:⁸

(26) Productivity indices per suffix (3 borrowed and 3 native suffixes)

$P_{\text{-itet}} = 0.01113$

$P_{\text{-cija}} = 0.02666$

$P_{\text{-er}} = 0.04495$

$P_{\text{-aj}} = 0.01764$ (*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’)

$P_{\text{-n}} = 0.08935$ (*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’)

In addition to giving a surprising measure of productivity for the given suffixes (e.g. *-telj*, which is marginally productive, receives a higher value than the highly productive *-ski* and *-n*), this formula is not sufficiently well fitted to measure productivity in the border cases such as the emerging and disappearing suffixes, where the number of tokens and the number of hapaxes are very small. Consider for instance the suffix *-itet* in the domain of native stems. Only one token has been found in the corpus, *propal-itet* failed-*itet* ‘punk’. Its productivity in the native domain, calculated using the formula above, reaches the maximal value of productivity of 1.0 (one hapax divided by one token), which is incompatible with the fact that this suffix derives only one word in the domain of native stems.

The formula above is insensitive to one important factor, which Haspelmath (1995) terms generality, and which we measure by the number of stems that the suffix combines with. The index of productivity has to be directly proportional to the aggregate number of stems that the affix combines with, as its ability to combine with a larger number of stems clearly verifies a higher productivity of an affix. Moreover, the number of stems selected by an affix only becomes relevant when it exceeds the level up to which storing simplex words is more economical than productive derivations. We stipulate a corrective measure at the level of five

8 The estimations of productivity of suffixes given in this sentence are based on the introspective method.

derived words (the final rankings and relative differences would not be significantly affected if this number were somewhat larger or smaller). Hence we reach the formula:

(27) Index of productivity formula modified

$$P = (N_{\text{aff}}^{\text{stem}} - 5) * n_{\text{l}}^{\text{aff}} / N^{\text{aff}}.$$

(number of stems selected by the affix lowered by five, times the number of hapaxes, divided by the number of tokens)

Productivity indices for the loan domain suffixes calculated by this formula are more realistic:

(28) Productivity indices per suffix recalculated

Borrowed	Low/average productivity native	High productivity native
$P_{\text{-itet}} = 1.49123$	$P_{\text{-aj}} = 3.12193$	$P_{\text{-ski}} = 998.0762$
$P_{\text{-er}} = 32.25981$	$P_{\text{-ište}} = 13.14623$	$P_{\text{-n}} = 1514.67956$
$P_{\text{-cija}} = 44.64776$	$P_{\text{-telj}} = 42.39935$	$P_{\text{-ost}} = 1871.682$

Suffixes *-er* and *-cija* are in the range of productivity of average to low productivity native suffixes (20 to 30 times lower than the most productive suffixes like *-ski*), while *-itet* remains at (or below) the productivity level of native suffixes with marginal productivity.

And if we look only into the spreading of the suffixes from the Latinate domain into the domain of native stems, by measuring their productivity with native stems only, we get results indicating that *-itet* is unproductive in the native domain, *-cija* (with an aggregate of 8 hapax instances and with 1 of each moderately and highly frequent derived nominals found in the corpus) is close to non-productive, *-ancija* (17 hapaxes, 2 moderately frequent nouns and two nouns of a relatively high frequency) is at the level of the marginally productive native suffixes, and the suffix *-er* (18 hapaxes and 4 moderately frequent nouns) reaches closer to the domain of moderate productivity, mainly due to the low frequency of its derived nouns.

(28) Productivity indices per suffix in the domain of native stems

$P_{\text{-itet}} = -4 \rightarrow 0$	$P_{\text{-ancija}} = 1.10122$
$P_{\text{-cija}} = 0.27211$	$P_{\text{-er}} = 5.33333$

The formula probably involves imprecision in regard of the relevance of each of the factors: size of the domain of stems, number of hapaxes and number of tokens. Fine-tuning might be due in terms of coefficients or non-linear factoring. Still, the measures we obtained indicate quite reliably that while suffixes *-cija* and *-er* show the quantitative distribution typical of average to low productivity suffixes (uneven distribution of frequency over the members of the class, with a relatively large number of hapaxes), the suffix *-itet* gives quantitative effects closer to the patterns of lower productivity (not many hapaxes and a relatively high average frequency) – but still within the range delimited by the quantitative properties of comparable native affixes.

Crucially, it is confirmed that borrowed suffixes show quantitative properties of average to low productivity native suffixes, and there is a sharp difference between the quantitative patterns these suffixes display in the Latinate stratum and with native stems, supporting the emergence analysis of the borrowed suffixes.

These observations open several interesting questions, the most important for our discussion

being the question what triggered the split inside the *-cija*-nouns with native stems, which we labeled as the *-cija* and *-ancija* nouns, and why it is exactly the context of the sequence *-an* that hosts the subclass with a special behavior. In quantitative perspective, this could be a consequence of the sequence *-an-* preceding the suffix *-cija* with a significantly higher frequency than any other sequence. But this is not the case. Several sequences significantly exceed *-an* in this respect, on all the possibly relevant parameters. Observe the data comparing the sequences *-za* and *-an* in (29).

(29) Comparing the quantitative parameters for the sequences *-an-* and *-za-* before *-cija*

	<i>-an(-cija)</i>	<i>-za(-cija)</i>
Number of different stems	50	423
Number of tokens	643	6435
Tokens per 1000 words	0.054879	0.549216
Index of productivity	2.729393	21.04615
Average freq. per stem	0.001098	0.001298

Another possible explanation may be coming from the fact that a very large number among the Latinate nominalizations originating from the late Latin *-(t)ia* class are preceded by *-n-*, while a significant number of these are in fact derivationally related to the late Latin *-ant/ent* nouns and hence have an *ant/ent* counterpart. It could thus be the case that the distinction between *-tio* and other nominalizations is preserved in S-C, based on their different semantics (preserving the information about the different category of their stem) and their relation with the *-ant/ent* class. The situation in S-C does not support this explanation either, because not only is the number of *-cija* nominalizations preceded by *-(a)n-* relatively small, as already discussed, but of all the borrowed *-n-cija* items found in the corpus (N=171), only 34 (i.e. less than 20%) do have a *-nt* counterpart. Finally, the suffix *-ancija* in S-C is sensitive to the aspectual content of the stem, which signals that it does not select for a stem deprived of verbal properties – something that differentiates the late Latin suffixes *-(t)ia* and *-cia* from *-tio* (Miller 2006).

We argue that the reason for the emergence of the *-ancija* class, with a more productive behavior than the (rest of the) *-cija* class lies in other aspects of S-C morphology. In particular, we analyze the sequence *-an* as the default 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.

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

udvar-an-cija, court- <i>an-cija</i> 'courting'	udvara-ti court-Inf 'to court'	udvar-an-t court- <i>an-t</i> the one who courts (pejorative)
zafrk-an-cija, jokeV- <i>an-cija</i> 'joking'	zafrka- <u>va</u> -ti joke-Inf 'to joke'	zafrk-an-t joke- <i>an-t</i> 'joker'
zavitl-an-cija, jokeV- <i>an-cija</i> 'joking'	zavitla- <u>va</u> -ti joke-Inf 'to joke'	zavitl-an-t joke- <i>an-t</i> 'joker'

zajeb-an-cija, jokeV- <i>an-cija</i> 'joking'	zajeba- <u>va</u> -ti joke-Inf 'to joke'	zajeb-an-t joke- <i>an-t</i> 'joker'
prevar-an-cija, cheat- <i>an-cija</i> 'cheating'	prevara-ti cheat-Inf 'to cheat'	prevar-an-t cheat- <i>an-t</i> 'cheater'
govor-an-cija talk- <i>an-cija</i> 'talking/speech'	govori-ti talk-Inf 'to talk'	govor-an-t ⁹ talk- <i>an-t</i> '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 default 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, idiomatic semantics, idiosyncratic choice of suffixes) variant (see Arsenijević and Simonović 2013 for a detailed discussion).

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 *-VV.je* nominalizations (imperfective stems, as briefly presented in respect of the example (4)), but it also imposes phonotactic constraints which truncate these stems.

Native morphology playing a role in the behavior of borrowed suffixes, in terms of factors such as native suffixes (*-an*) and native tendencies (the formation of paradigmatic and non-paradigmatic domains), with native grammatical and semantic oppositions realized by this behavior, make a strong argument in support of the view that the borrowed suffixes are actually suffixes that have emerged in the Latinate domain of the target language rather than suffixes readily borrowed from the source language.

8. Conclusions

The architecture of the lexicon is coconut-like in respect of productivity. Its core contains 'soft' stems / patterns / constructions which form domains of paradigmatic derivations, characterized by being: (a) maximally productive, (b) semantically transparent and (c) phonologically non-demanding. Next, 'medium' layer involves the domains of constrained productivity, with idiosyncratic semantics and own patterns in the domain of prosody. Borrowed derivations form 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

⁹ 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.

this movement is slow and limited in extent.

Latinate nominalizations belong to the softer side of the outer, hard layer. They have not only special prosodic shapes, but also stems which differ from those of the related inner layers verbs/adjectives. However, they do get reanalyzed into stems and emerging suffixes, unlike the more recently borrowed English derivations of the type *rijaliti* “reality show” *sekjuriti* “security personnel”, *sebniti* “celebrity”, *sešn* “session” (only in fashion) and *fešn* “fashion”. 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 idiomatic semantics and limited possibilities of morphological relating with other items in the lexicon, and tend to remain that way. 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 7). 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, H. R. & R. Lieber. 1991. Productivity and English derivation: a corpus based study. *Linguistics* 29: 801-844.
- 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 (in preparation). *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. Submitted to *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.