

Severing affixes from categories and borrowed derivations

Approaches to derivational morphology that consider some derivational affixes as roots draw a significant part of their empirical basis from derivations containing affixes that are categorially flexible, affect prosody, and are unpredictable in terms of their semantic contribution. Examples of such affixes, profiled as root-selecting roots, are English *-ic* and *-al* and Dutch *-iek* and *-aal* (Lowenstamm 2014, Creemers et al. 2018). As observed in Creemers et al. (2018), these affixes are only flexible between two categories (nouns and adjectives) and originate from language contact, i.e. they are Romance/Latinate. Creemers et al. (2018) argue that in Dutch, in a certain sense, the latter may be an explanation of the former: Romance verbs feature the affix *-eer-*, which is 'greedy' and takes up all the space in which other flexible affixes might show up.

An account in which language contact strengthens (or even introduces) affixes that can be analysed as root-selecting roots has been proposed independently by Simonović (2015) and Simonović & Arsenijević (2018). These authors argue that loanwords enter the recipient language as morphologically simplex and carrying a specific meaning, whereas their segmentation into morphemes is a matter of subsequent reanalysis within the recipient language. Simonović & Arsenijević (2018) argue that in Bosnian/Croatian/Montenegrin/Serbian (BCMS) Latinate nominalisations join the class of nominalisations that they term 'non-paradigmatic', in certain cases distinct from their 'paradigmatic' counterparts only by affix-controlled prosody and a specialised meaning. The native contrast can be illustrated by the native verb *obrazovati* [ˈobrázouati] 'educate', that can derive the prosody-preserving nominalisation *obrazovanje* [ˈobrázouaːnje] 'educating' and the prosody-altering nominalisation *obrazovanje* [obrazoˈuaːné] 'education'. Both shifted prosody and specific meaning are arguably cues of a simplified structure, so that the learner is naturally guided to analysing the non-paradigmatic nominalisations as containing the same lexical material (roots) as their paradigmatic counterparts, but less intermediate (i.e. categorial) structure. Simonović & Arsenijević (2018) suggest that the same contrast is carried over to Latinate verbs, such as *educirati* [eduˈtsi:ráti], but now between the native nominalisation, e.g., *educiranje* [eduˈtsi:ráːnje] 'educating' and the Latinate nominalisation, e.g., *edukacija* [eduˈka:tsíja] 'education'.

Against this background, we focus on Latinate nominalisations in *-cija* and in BCMS, especially targeting their co-existence with corresponding native nominalisations derived from verbs which contain the formant *-ir-* (e.g. as in *educ-ir-anje* 'educating') and their internal complexity. Our goal was to compare *-cija* nominalizations with native *-nje* nominalizations in terms of the degrees to which they show internal verbal properties. To this end, we randomly selected 11 pairs of *-cija* nominalizations and native *-nje* nominalization derived from the same root, making sure that both members of the pair pass the frequency threshold of at least 3000 tokens in the CLASSLA-web corpus of Croatian (Ljubešić et al. 2024). Next, we conducted a corpus study drawing random samples of 50 concordance lines from the CLASSLA-web corpus for each of the 22 nominalizations, yielding a total of 1100 concordance lines¹. The concordance lines were manually annotated for overt signals of the presence of verbal structure within the noun phrase headed by the nominalization in question. Here we focus on two diagnostics of internal verbal structure per Grimshaw (1990), one positive — the presence of a genitive case-marked complement and one negative — the plural form of the head noun.

Based on these two diagnostics, our results indicate that native *-nje* nominals show more verbal properties than Latinate *-cija* nominals. Genitive case-marked complements were attested almost two times more frequently with *-nje* nominals (357/550 tokens) than with *-cija* nominals (190/550 tokens). At the level of specific pairs, in 10 pairs the *-nje* nominal has a genitive complement more often than its *-cija* counterpart, whereas one pair features a tie. The contrast

¹ In the course of the upcoming semester the authors intend to teach a course on the topic, so it is expected that by the time the conference is held, more data will be available.

is even stronger on the second diagnostic: around 20% of all *-cija* nominals were plural nouns (117/550 tokens), whereas only one instance of a plural noun was attested among the *-nje* nominals in the sample. Binomial logistic regression models predicting the probability of a genitive case-marked complement and plural marking based on the type of nominalization (*-cija* vs. *-nje*) and root indicate significant effects of both nominalization type and root for both dependent variables.

We account for these data by assuming that *-cija* nominalizations indeed lack verbal structure and replicate the native 'non-paradigmatic' nominalizations in displaying morphological complexity (e.g., *-a-c-ija*, *-i-c-ija*, *-k-c-ija*), whereby the specific elements do not have a clear semantic or syntactic import. This is further confirmed by the fact that the Latinate stratum (just like the native 'non-paradigmatic' nominals) also contains examples of zero-derived nominals, such as e.g. *kontrola* 'control' (cf. *kontroliranje* 'controlling').

As for the two diagnostics, the presence of a genitive case-marked complement with nominalizations of various types is not grammatically licensed, i.e. there is no grammatical feature/head that licenses genitive case-marked complements. However, the presence of verbal structure (e.g. *v*) with native *-nje* nominals makes its overt presence more likely, due to the fact that these nominalizations (unlike *-cija* nominalizations) always denote events, see the contrast in (1). In other words, with *-nje* nominals derived from transitive verbs the internal argument is always either overtly realised or implied. With *-cija* nominals the internal argument is often not present at all as they do not obligatorily refer to events (1a).

- (1) a. Jednostavan priključak na instalacije izveden u podu
simple connection on installations placed in floor
'A simple connection to the utilities placed on the floor'
b. Kontrola instaliranja drugoga softvera
control installing other software
'The control over installing another software'

As for plural marking, we assume, following lordăchioaia and Soare (2009), that native nominals derived from imperfective/biaspectual roots are atelic allowing both progressive and iterative readings, which makes nominal pluralization redundant. On the other hand, a *-cija* nominalization is either telic, denoting a single event, or it denotes an entity (an abstract or concrete result of the event described by the underlying verb), which means that pluralities of events or entities denoted by these nouns have to be overtly signalled in the form of a plural suffix.

References

- Creemers, A., Don, J., & Fenger, P. (2018). Some affixes are roots, others are heads. *Natural language & linguistic theory*, 36, 45-84.
- Grimshaw, J. (1990). *Argument structure*. MIT Press.
- lordăchioaia, G., & Soare, E. (2009). Structural patterns blocking plural in Romance nominalizations. In E. O. Aboh, E. van der Linden, J. Quer, & P. Sleeman (Eds.), *Romance languages and linguistic theory: Selected papers from 'Going Romance' Amsterdam 2007* (pp. 145-160). John Benjamins Publishing Company. <https://doi.org/10.1075/rllt.1>
- Ljubešić, N., Rupnik, P., & Kuzman, T. (2024). *Croatian web corpus CLASSLA-web.hr 1.0*. Slovenian language resource repository CLARIN.SI. <http://hdl.handle.net/11356/1929>
- Lowenstamm, Jean. 2014. Derivational affixes as roots. In *The syntax of roots and the roots of syntax*, ed. Artemis Alexiadou, Hagit Borer, and Florian Schäfer, 230–259. Oxford: Oxford University Press.
- Simonović, M. 2015. *Lexicon immigration service - Prolegomena to a theory of loanword integration*. Utrecht: LOT Dissertation Series 393.
- Simonović, M. & Arsenijević, B. (2018). The importance of not belonging: Paradigmaticity and loan nominalizations in Serbo-Croatian. *Open Linguistics*, 4(1), 418-437.