

Lexeme Formation, Morphological Parsing and Constituency in Loanword Adaptation A Case Study from Slovenian and Serbo-Croatian

Marko Simonović
UiL-OTS, Utrecht University

I will present an analysis of loanword adaptation data from two relatively closely related Slavic languages – Slovene and Serbo-Croatian (henceforth SL and SC). The focus will be on nouns ending in mid vowels *-o* and *-e* (e.g. *metro*, *konto*, *matine*, *konte*), which coincide with the markers of neutral declensions in the native words of both languages¹. Curiously, these borrowed nouns practically never enter the neuter declensions. As I will show, the fashion in which novel items are assigned to inflectional classes and get paradigms calls for a model of adaptation which bears on lexeme formation and the structure of the lexicon in general.

The analysis is developed within a model which builds on OT-based research of synchronic stratified lexica (Ito & Mester 2001, 2002) – the *Portal model of loanword adaptation* (Simonović 2009). This model takes the lexicon to be inherently stratified, with onion-like layers inside-out. The outermost stratum, the most liberal one in terms of marked phonological structures and the most conservative one in terms of morphological combinability, is the natural host for new borrowings. The interactions on the outermost stratum are defined by the interaction of undominated markedness (which delimits what is possible in the language), faithfulness indexed to loanwords FAITH(LOAN) (see Ito & Mester 2001, 2002, Simonović 2009) and Lexical Conservatism (see below). Furthermore, words which are in a language as a result of ongoing language contact and which maintain the links to the original form in the source language in bilingual speakers, are considered to be in the *portal*. The concept of the *portal* is used to model the *uniformity* with which all the words containing a particular to-be-adapted structure are treated.

A prime example of portal-related uniformity comes from the present case study. Many loanwords which have entered Slovenian and Serbo-Croatian are structurally very similar to native neuter nouns in these languages (e.g. SL *metro* is similar to *srebro*_{NEUTRAL} “silver” in all relevant aspects). Other new nouns equally ending in an *o*, however, have structures alien to the pattern of neuter endings - e.g. *rešo* “hot plate” has a soft consonant followed by the “hard” ending *-o*, an unattested combination. From the perspective of the portal, as all the words are to be treated on a par, it is due to this latter group of nouns that in Slovenian (and Serbo-Croatian) the neuter declensions are “closed” for non-native items in *-o* altogether. Similar examples exist for the nouns ending in *-e*, the other potential neuters. In sum, the inflectional class of neuter nouns could not accommodate for some of the novel items, so it remained closed for all of them.

In the masculine paradigms, which these words have to enter as the only remaining option, the default nominative singular (= citation form) ending is null, but some exceptional items have the endings *-o* and *-e* (e.g. a proper noun as *Mark-o*). Consequently, the final *e/o* in borrowed words can either be parsed by the morphology as the nominative singular ending or left as stem-final. Generally speaking, the prosody of the initial form plays a crucial role - when stressed (SL) or long (SC), the final vowel is left unparsed, resulting in a paradigm as SL *metro-Ø*, *metro-ja*, *metro-ju*, *metro-jem* etc². In these nouns the entire segmental form of

¹ The loanword corpus consists of over 150 words borrowed from French, Italian, German, Spanish etc. into various Slavic languages, whose collection and analysis is a part of a larger PhD project on morphological adaptation.

² Due to space limitations, we present the paradigms of the relevant nouns using the following forms: nominative singular, genitive singular, dative singular and instrumental singular. All the other singular forms are homonymous to one of the provided forms.

the loanword is interpreted as the new stem and starts a new, loanword-only declensional type (with vowel-final stems) in both languages. On the other hand, if the final vowel is short and unstressed, it usually gets parsed, as in SL *konto*, *konta*, *kontu*, *kantom* etc. However, in both languages, some more recent words ending in prosodically “weak” vowels (*e* significantly more often than *o*) tend to remain unparsed: SL *sáke*, *sákeja*, *sákeju*, *sákejem*. Also, experiments on which we will report show that SC speakers³ have a tendency to keep foreign names unparsed regardless of other comparable parsable nouns (*Majlo-ø*, *Ludo-ø*, *Rilke-ø*, *Pele-ø*).

I will argue that FAITH(LOAN), although essential in an analysis of phonological importation, fails in capturing all the relevant aspects of paradigm formation. Instead, Lexical Conservatism (Steriade 1997), an independently required mechanism which requires all the inflected forms to replicate the initial surface form (listed allomorph) in as many respects as possible, should be used.

SC data provide us with a further insight as for what LC constraints should target. The two aforementioned nouns in *-e*, *Rilke* and *Pele*, are ambiguous - they can be interpreted either as native nicknames or as foreign names. In the former case the paradigm for *Rilke* is *Rilk-e*, *Rilk-et-a*, *Rilk-et-u*, *Rilk-et-om* etc. In the latter case, the name of the German writer declines as *Rilke*, *Rilke-a*, *Rilke-u*, *Rilke-om* etc. As can be seen, in both paradigms, the string /rilke/ is preserved in all forms. What makes the two paradigms different is the fact that the string in question is not a morphological constituent in the former case. This is why the paradigm in which the whole source language form functions as the new stem is selected (cf. also the adjective *rilke-ovski* “Rilkean”). In sum, the constraints responsible for the selection of the paradigm can demand the morphological *constituency* of the input segmental string. If this constraint is satisfied, the integrity of the source form is maintained within a lexical entry, a constellation which our model recognizes to be beneficial for the bilingual lexicon, due to the transparent resemblance to the source form. This matches the treatment of *o/e*-final nouns in other Slavic languages, where they are indeclinable altogether (e.g. Russian) or where they are always kept unparsed (recent developments in Bulgarian).

In conclusion, I will point out the directions for further research.

References

- Ito, Junko & Armin Mester. 2001. “Covert generalizations in Optimality Theory: the role of stratal faithfulness constraints”. *Studies in Phonetics, Phonology and Morphology* 7. 73–299.
- Ito, Junko & Armin Mester. 2002. “One phonology or many? Issues in stratal faithfulness theory”. *Phonological Studies* 5.121-126.
- Simonović, Marko. 2009. Immigrants Start on the Periphery – A Unified Approach to Loanword Phonology. MA thesis, Utrecht University.
- Steriade, Donca. 1997. “Lexical Conservatism”. *Linguistics in the Morning Calm*.157-179.

³ Comparable experiments with Slovenian speakers are obstructed by the fact that colloquial Slovenian has the super-unparsing procedure for male nouns in *-o*, *-e* and *-a*, so that most of the speakers retreat to the genitive forms *Ludo-t-a*, *Udo-t-a* etc. This declensional type also occurs in native names e.g. *Marko-t-a* (standard Slovenian *Mark-a*).