

Paradigm Analogy and the Illusion of Phonological Regularity in Serbo-Croatian Borrowings

The borrowing of foreign vocabulary into a recipient language is rarely phonologically neutral: loanwords frequently undergo systematic restructuring that conforms to patterns of the target grammar (Kenstowicz 2005). This talk examines the $k \rightarrow \check{c}$ ([k] \rightarrow [tʃ]) alternation in loanwords of Serbo-Croatian that, on the surface, resemble the well-known process of Velar Palatalization—whereby stem-final velars [k, g, x] surface as postalveolars [tʃ, ʒ, ʝ] before palatal-initial suffixes—but which, I argue, are better understood as paradigmatically organized analogical patterns.

The data come from agentive formations with the suffix *-ar*, which attaches to common nouns and typically denotes profession. In the native vocabulary, the $k \rightarrow \check{c}$ alternation applies inconsistently: compare *mleko* \rightarrow *mlek-ar* ‘milkman’ and *slika* \rightarrow *slik-ar* ‘painter’ with *jezik* \rightarrow *jezič-ar* ‘philologist’ and *stoka* \rightarrow *stoč-ar* ‘cattle breeder’. No principled phonological generalization captures this distribution. Turning to borrowed stems of Greco-Latin origin, however, a strikingly regular pattern emerges. Four stem types can be distinguished: stems in which *i* directly precedes the velar ($\sim i\mathbf{k}$) exhibit palatalization, while stems that lack *i* do not ($\sim \emptyset\mathbf{k}$); stems containing *i* in the penultimate syllable but lacking a final *k* ($\sim i\emptyset$); and stems lacking both ($\sim \emptyset\emptyset$). The distribution across these types is summarized in Table 1.

Vocabulary	Stem Type	Example	Derived Form (NOM.MASC.SG)
NATIVE	$k > k$	<i>mleko</i> ‘milk’	<i>mlek-ar</i> ‘milkman’
		<i>slika</i> ‘painting’	<i>slik-ar</i> ‘painter’
	$k > \check{c}$	<i>jezik</i> ‘language’	<i>jezič-ar</i> ‘philologist’
		<i>stoka</i> ‘cattle’	<i>stoč-ar</i> ‘cattle breeder’
LOAN	$\sim \emptyset\mathbf{k}$	<i>biblioteka</i> ‘library’	<i>bibliotek-ar</i> ‘librarian’
		<i>apoteka</i> ‘pharmacy’	<i>apotek-ar</i> ‘pharmacist’
	$\sim i\mathbf{k}$	<i>lingvistika</i> ‘linguistics’	<i>lingvistič-ar</i> ‘linguist’
		<i>matematika</i> ‘mathematics’	<i>matematič-ar</i> ‘mathematician’
	$\sim i\emptyset$	<i>istorija</i> ‘history’	<i>istori-č-ar</i> ‘historian’
		<i>biblija</i> ‘Bible’	<i>bibli-č-ar</i> ‘biblicist’
$\sim \emptyset\emptyset$	<i>astma</i> ‘asthma’	<i>astma-tič-ar</i> ‘asthmatic’	
	<i>psorijaza</i> ‘psoriasis’	<i>psorija-tič-ar</i> ‘psoriatic’	

Table 1: Patterns of *-ar* formations in the native and loan lexicon.

The loan data reveal that the conditioning of the alternation extends well beyond what a feature-spreading type of account for velar palatalization can capture (cf. Morén 2007). In the $\sim i\emptyset$ and $\sim \emptyset\emptyset$ types, the segments *-č-* and *-tič-* appear where no velar is present in the base to palatalize. Moreover, the suffix *-ar* carries no palatal segment that could independently trigger assimilation. Any representation-based analysis is therefore forced to either (i) posit abstract latent material in the underlying stem (e.g. *-ik* or *-tik* extensions), or (ii) treat the suffix as having multiple phonologically-conditioned allomorphs. Both moves are costly, requiring abstract representational distinctions not recoverable from surface forms. I propose instead that the distribution reflects a paradigm-level analogical pattern, captured within a word-based framework

(Blevins 2006). In this approach, morphological structure is defined over relations among whole word forms rather than over morpheme concatenation or feature operations. High-frequency $\sim ik$ forms such as *matematičar* and *lingvističar* constitute an entrenched structural template. Forms of the $\sim i\emptyset$ and $\sim \emptyset\emptyset$ types align themselves with this template, by producing $-č-$ or $-tič-$ in the same pre-*ar* slot so as to conform to the shape of the dominant paradigm (Bybee 1985; Dawdy-Hesterberg–Pierrehumbert 2014). The same configuration extends to adjectival formations with the suffix *-n*. Here, vowel-final bases ($\sim \emptyset V$) surface with $-tič-$ (e.g. *limfa* ‘lymph’ → *limfa-tič-ni* ‘lymphatic’; *trauma* → *trauma-tič-ni* ‘traumatic’ etc.), while consonant-final bases ($\sim \emptyset C$) surface with $-ič-$ (e.g. *egoist* → *egoist-ič-ni* ‘egoistic’; *cilindar* ‘cylinder’ → *cilindr-ič-ni* ‘cylindrical’ etc.), eventually extending the $\sim \emptyset\emptyset$ pattern. Table 2 illustrates the analogical structure shared across both paradigms. The shaded cells indicate the material inserted in each form to bring it into conformity with the dominant $\sim ik$ template.

Analogical Relations		Pattern Type	Suffix
m a t e m a t i č a r		$\sim ik$	<i>-ar / -n</i>
i s t o r i č a r		$\sim i\emptyset$	
a s t m a t i č a r		$\sim \emptyset V$	$\sim \emptyset\emptyset$
a s t m a t i č n i		$\sim \emptyset C$	
e g o i s t i č n i			<i>-n</i>

Table 2: Analogical relations across agentive and adjectival paradigms.

The split between $\sim \emptyset V$ and $\sim \emptyset C$ in the adjectival domain does not require underlying representations distinguishing vowel-final from consonant-final stems. Rather, the two patterns are distinguished by phonotactic constraints on well-formed consonant clusters (e.g. **cilindrtični*): where such a constraint is violated, $-ič-$ is preferred.

The broader implication is that what appears to be a phonological process in the borrowed lexicon is in fact a paradigmatically structured phenomenon. The regularization observed in loanwords reflects adaptation to an entrenched analogical template rather than the productive application of a purely phonological process.

References

- Blevins, James P. (2006). Word-based morphology. *Journal of Linguistics* 42(3). 531–573.
- Bybee, Joan (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins Publishing.
- Dawdy-Hesterberg, Lisa–Janet B. Pierrehumbert (2014). Learnability and generalisation of Arabic broken plural nouns. *Language, Cognition and Neuroscience* 29(10). 1268–1282.
- Morén, Bruce (2007). Consonant–vowel interactions in Serbian: Features, representations and constraint interactions. *Lingua* 116(8), 1198–1244.
- Kenstowicz, Michael (2005). *The phonetics and phonology of Korean loanword adaptation*. Paper presented at First European Conference on Korean Linguistics, Leiden University, February 2005.