

## РОДОВЫЕ ПОНЯТИЯ В БЫТОВОЙ ЛЕКСИКЕ КАК ОБЛАСТЬ ТОНКИХ РАЗЛИЧИЙ МЕЖДУ СЕРБСКИМ И ХОРВАТСКИМ ЯЗЫКОМ

**А. Ч. Пиперски** (apiperski@gmail.com)

МГУ, Москва, Россия

В бытовой лексике сербского и хорватского языков есть заметные различия. Различия видовых наименований (слов с такими значениями, как 'ложка', 'очки', 'паспорт') хорошо осознаются носителями, а различия родовых наименований ('посуда', 'столовые приборы', 'канцелярские принадлежности') менее заметны.

**Ключевые слова:** быт, бытовая лексика, род, родовые понятия, родовые наименования.

## GENERIC TERMS IN EVERYDAY VOCABULARY AS A SPHERE OF SUBTLE DIFFERENCES BETWEEN SERBIAN AND CROATIAN

**A. Ch. Piperski** (apiperski@gmail.com)

Lomonosov Moscow State University, Moscow,  
Russian Federation

There are significant differences in the everyday vocabulary of Serbian and Croatian. The speakers are aware of diverging specific terms (e. g., words for 'spoon', 'glasses', 'passport'), but they fail to notice some diverging generic terms (words for 'kitchenware', 'cutlery', 'writing supplies'). This is explained by the fact that generic terms show considerable amount of variation even within one language and cannot serve as markers of identity.

**Key words:** everyday items, everyday vocabulary, generic terms, identity marker.

It is typical for everyday vocabulary to exhibit considerable variation. To study this variation, a group of researchers under the lead of Boris Iomdin organized a survey for speakers of various languages (see Iomdin et al., this volume). The participants were shown 33 pictures and were requested to provide a name for each object in the picture (a specific, or subordinate term) and a word denoting the group to which this object belongs (a generic, or superordinate term). For example, a picture of the chair could be described with the words *chair* and *furniture*.

There were 5 speakers of Croatian and 6 speakers of Serbian who participated in the survey. Though this sample cannot count as representative, their responses were extremely instructive and made it possible to put forward some hypotheses about the lexical differences between Serbian and Croatian. These hypotheses were verified using other methods.

There is a number of well-described divergences between Serbian and Croatian literary languages (for references see Greenberg 2004, Tošović 2009)<sup>1</sup>. They do not impede mutual understanding, even though they are present on all levels of the language system — in phonetics, morphology, word formation, vocabulary, and syntax. But lexical differences are the most striking ones, and no wonder that they receive the greatest attention.

Among the responses to our survey there were some well-known pairs distinguishing the two languages:

*S naočare* (5) / *naočari* (1) vs. Cr *naočale* (4) / *naočali* (0) ‘glasses’<sup>2</sup>

*S kašika* (6) vs. Cr *žlica* (5) ‘spoon’

*S karmin* (5) vs. Cr *ruž (za usne)* (5) ‘lipstick’

*S lenjir* (6) vs. Cr *ravnalo* (5) ‘ruler’.

*S pasoš* (6) vs. Cr *putovnica* (3), *pasoš* (2) ‘passport’

Known phonological and morphological differences between Serbian and Croatian could also be observed:

*S odeća* (11) vs. Cr *odjeća* (8) ‘clothes’ (about socks, gloves and high-heeled shoes)

*S prtljag* (1) (masc.) vs. Cr *prtljaga* (4) (fem.) ‘luggage’ (about trunks and cosmetics bags).

More important is that the survey showed some new facts that have never been considered as markers of differentiation between Serbian and Croatian.

High-heeled shoes were described by all 6 Serbian respondents as *cipele*, but 4 speakers of Croatian called them *štikle* (only 1 Croatian-speaking person wrote *ženske cipele* ‘ladies’ shoes’). The word *cipele* actually has the meaning ‘shoes’ according to all the dictionaries of Serbian and Croatian, but the word *štikla* is described as having only the meaning ‘heel’ even in the voluminous and up-to-date Croatian

---

<sup>1</sup> The other two languages closely related to Serbian and Croatian, namely Bosnian and Montenegrin, will not be discussed in this paper because no speakers of these languages took part in the survey and I lack reliable data about everyday vocabulary in these languages.

<sup>2</sup> The number of relevant responses obtained in the survey is given in parentheses. If the total of Croatian responses lies under 5 or the total of Serbian responses lies under 6, this means that the other responses were irrelevant for the topic under discussion (e. g., they contain absolutely different words). The Serbian words are rendered in Roman script.

dictionary by Vladimir Anić (2003). Thus, a metonymy ‘heels’ > ‘high-heeled shoes’ has occurred in Croatian, but not in Serbian, and this fact has not yet been reflected in the dictionaries.

Another semantic difference could be observed among the generic terms. The word *vaza* ‘vase’ was categorized as *nameštaj* by 3 Serbian respondents out of 6, while none of the Croatians used the words *namještaj* or its typical Croatian synonym *pokućstvo* (two of them could not even think of a suitable category for vases and left this field blank). Furthermore, *stolnjak* ‘tablecloth’ is categorized as *nameštaj* by 2 Serbians, and 1 Serbian regards  *pokrivač* ‘blanket, comforter’ as *nameštaj*, too. It can be inferred that the S *nameštaj* has a more general meaning ‘furniture, home décor or accessories’, while the meaning of Cr *namještaj/pokućstvo* is limited to ‘furniture’. In a somewhat outdated Serbo-Croatian-Russian dictionary by I. Tolstoy (1957) the word *nameštaj* is translated as ‘мебель, обстановка’. It seems that the first meaning applies to Croatian and the second one to Serbian.

When examining the differences between Serbian and Croatian, scholars are usually concerned with absolute differences of the type “the word *X* exists in one of the languages but not in the other one” and pay less attention to statistical differences of the type “the word *X* exists in both languages, but it occurs significantly more frequently in one of them”<sup>3</sup>. The results of our survey allowed me to posit some hypotheses about differences of the second type, which could be then verified using greater amount of data from the Google search engine (www.google.com; the results were retrieved on 31.01.2011). For each word or phrase 3 queries were made: in Roman script in the domain *.rs*, in Roman script in the domain *.hr* and in Cyrillic script in the domain *.rs*<sup>4</sup>. The results were presented in form of tables.

	site:.rs	site:.hr
Roman script	$s_{cyr}$	$h$
Cyrillic script	$s_{lat}$	
Total ( $\Sigma$ )	$s = s_{cyr} + s_{lat}$	$h$

A coefficient  $k = s : h$  was calculated for each word/phrase. This coefficient reflects the frequency of a word/phrase in Serbian texts relative to its frequency in Croatian texts.

Given the proximity of the two languages, I assume that the words that are not specifically Serbian or specifically Croatian (which is the case with the most words)

<sup>3</sup> There are in fact some studies of lexical divergences between Serbian and Croatian that use simple statistical methods (cf. Grčević 2002), but most scholars tend to rely on their own impressions and not on statistical data. However, highly elaborated statistical methods have been proven useful for analyzing lexical divergences between closely related language systems (cf. Berdicevskis *forthcoming* for comparison of the Russian language in Russia and Latvia).

<sup>4</sup> Surely there are texts written in Serbian in the domain *.hr* and texts written in Croatian in the domain *.rs*, but their amount is negligible.

have similar frequencies in both languages. This means that  $k$  does not vary significantly for such words, and its mean value ( $\bar{k}$ ) reflects only the size of the corpus for each language. It turns out that  $\bar{k} \approx 0.2$  (in other words, the Google database contains 5 times more Croatian texts than Serbian). For words/phrases which are characteristic only for one of the languages,  $k$  will greatly diverge from the mean value. If there is a pair of synonymous words/constructions  $w_1$  and  $w_2$ , for which  $k_1$  is significantly less than  $\bar{k}$  and  $k_2$  is significantly greater than  $\bar{k}$  (or vice versa), the usage of the members of this pair constitutes a statistical difference between Serbian and Croatian.

In the responses to the survey 4 pairs of this type occurred:

**Generic term** for ruler: S *pribor za crtanje* (1) vs. Cr *crtaći pribor* (1)

$w_1$	site:.rs	site:.hr		$w_2$	site:.rs	site:.hr	
pribor za crtanje	45 300	67 300		crtaći pribor	102	3760	$k_1 : k_2 = 25,11$
прибор за цртање	1 900			цртаћи прибор	3		
$\Sigma$	47 200	67 300		$\Sigma$	105	3760	
	$k_1 = 0,70$				$k_2 = 0,028$		

**Generic term** for ruler, eraser, pencil: S *pribor za pisanje* (ruler — 1, eraser — 2, pencil — 2) vs. Cr *pisaći pribor* (ruler — 1, eraser — 1, pencil — 2)

$w_3$	site:.rs	site:.hr		$w_4$	site:.rs	site:.hr	
pribor za pisanje	635 000	88 800		pisaći pribor	9160	81 300	$k_3 : k_4 = 63,26$
прибор за писање	2 320			писаћи прибор	64		
$\Sigma$	637 320	88 800		$\Sigma$	9224	81 300	
	$k_3 = 7,18$				$k_4 = 0,11$		

**Generic term** for spoon: S *escajg* (2) vs. Cr *bešteć* (2)

$w_5$	site:.rs	site:.hr		$w_6$	site:.rs	site:.hr	
escajg	180 000	4 350		bešteć	59	11 400	$k_5 : k_6 = 7621,94$
есцајг	319			бештећ	3		
$\Sigma$	180 319	4 350		$\Sigma$	62	11 400	
	$k_5 = 41,45$				$k_6 = 0,0054$		

**Generic term** for pot, spoon, vase, teapot, wineglass: S *posuđe*<sup>5</sup> (pot — 1, spoon — 1, teapot — 4, wineglass — 4) vs. Cr *suđe* (pot — 3, spoon — 1, vase — 1, teapot — 4, wineglass — 4)

$w_7$	site:rs	site:hr		$w_8$	site:rs	site:hr	$k_7 : k_8 = 3,07$
posuđe	458 000	1 340 000		suđe	23 500	241 000	
posudje	5 520	2 790		sudje	9 420	53 600	
посудје	2 220			cyђе	363		
Σ	465 740	1 342 790		Σ	33 283	294 600	
	$k_7 = 0,35$				$k_8 = 0,11$		

The following conclusions can be made:

1) Deverbal adjectives in *-aći* meaning ‘intended for smth.’ are more widespread in Croatian, while in Serbian the construction *za* ‘for’ + deverbal substantive in *-nje* is more frequently used (S *pribor za crtanje* ‘instrument for drawing’ vs. Cr *crtaći pribor* ‘drawing instrument’, S *pribor za pisanje* ‘instrument for writing’ vs. Cr *pisaći pribor* ‘writing instrument’). It is probably not a coincidence that in the entry *pribor* in a Serbo-Croatian-Russian dictionary by I. Tolstoy (1957) which is rather Serbian-oriented one can find exactly the examples *pribor za pisanje* and *pribor za crtanje*.

2) There is a lexical divergence between Serbian and Croatian which to my knowledge has never been properly described before<sup>6</sup>: S *escajg* vs. Cr *bešteak* ‘cutlery’. It is noteworthy that both words are loanwords from German (*Esszeug* resp. *Besteck* ‘cutlery’), but only the second word is in use in modern German. This was also corroborated by the survey: about 60% of German-speaking respondents used the word *Besteck* for categorizing spoon, and none of them wrote *Esszeug*.

3) The differences in frequency of the word for ‘kitchenware’ (*posuđe/suđe*) are not as striking as in the other cases. But some speakers of Serbian and Croatian make notice of the fact that the pair *posuđe* vs. *suđe* distinguishes the two languages<sup>7</sup>. It is interesting that in the Croatian dictionary by V. Anić (2003) the word *posuđe* is given an explication, and the entry *suđe* contains only the reference to *posuđe*. The author of the dictionary was probably influenced by Serbian-oriented dictionaries and was not aware that it is not in accord with the actual Croatian usage.

Our survey proves that everyday vocabulary exhibits differences even in closely related languages. The present study has shown five discrepancies of this kind that had not been sufficiently accounted for in the scholarly literature:

<sup>5</sup> One of the Croatian respondents also categorized teapot and wineglass as *posuđe*.

<sup>6</sup> In Brodnjak (1992) *escajg* is listed as a Serbian word, but no one-word Croatian equivalent is given. It is translated (or rather explained) as *pribor za jelo* (*žlica, vilica, nož*) ‘eating utensil (spoon, fork, knife)’.

<sup>7</sup> <http://forum.ffzg.hr/viewtopic.php?p=108806> (retrieved on 31.01.2011)

1) *S štikle* ‘heels’ vs. Cr *štikle* ‘heels; high-heeled shoes’; 2) *S nameštaj* with a broader meaning than its Cr equivalent; 3) *S za -nje* vs. Cr *-aći*; 4) *S escajg* vs. Cr *beštej*; 5) *S posuđe* vs Cr *suđe*.

It is striking that most of these newly observed facts (4 out of 5) concern generic terms and not specific ones. The question arises why speakers and linguists are less aware of differences in superordinates than of differences in subordinates.

A probable explanation is that there is less variation among subordinates even within the same language. In our survey, trunk, gloves, umbrella, slippers, teapot, tablecloth and wineglass were unanimously described as *kofer*, *rukavice*, *kišobran*, *papuče*, *čajnik*, *stolnjak*<sup>8</sup> and *čaša* respectively by all Croatian and Serbian speakers. For many other words responses were pretty similar, e. g. *gumica* or *gumica za brisanje* for eraser (literally: ‘rubber’ or ‘rubber for erasing’). For no generic term there were 11 identical responses nor 5 identical responses among Croatsians were found, and there were only 2 cases where all 6 Serbians agreed with each other (they classified buiding blocks as *igračke* ‘toys’ amd socks as *odeća* ‘clothes’). The uncertainty about generic terms that exists even within one language prevents the speakers from noticing that the speakers of a closely related language use other generic words in their everyday vocabulary.

## References

1. *Anić V.* 2003. Veliki Rječnik Hrvatskoga Jezika.
2. *Berdicevskis A.* Predictors of Pluricentricity: Lexical divergences between Latvian Russian and Russian Russian.
3. *Brodnjak V.* 1992. Rječnik Razlika Između Hrvatskoga i Srpskoga Jezika.
4. *Grčević M.* 2002. Some Remarks on Recent Lexical Changes in the Croatian Language. Lexical Norm and National Language: Lexicography and Language Policy in South Slavic Languages after 1989 :150–163.
5. *Greenberg R. D.* 2004. Language Identity in the Balkans: Serbo-Croatian and Its Disintegration.
6. *Iomdin B. L.* 2009. Everyday life Vocabulary. Search of Standard [Terminologija Byta. Poiski Normy]. Komp’iuternaia Lingvistika i Intellektual’nye Tekhnologii: Trudy Mezhdunarodnoi Konferentsii “Dialog 2009” (Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference “Dialog 2009”), 8 (15) : 127–135.
7. *Iomdin B., Piperski A., Russo M., Somin A.* 2011. How Different Languages Categorize Everyday Items. Komp’iuternaia Lingvistika i Intellektual’nye Tekhnologii: Trudy Mezhdunarodnoi Konferentsii “Dialog 2011” (Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference “Dialog 2011”), 10 (17).
8. *Tolstoj I. I.* 1957. Serbian-Croatian-Russian Dictionary [Serbsko-Khorvatsko-Russkii Slovar’].
9. *Tošović B.* 2009. *Korrelative Grammatik des Bosni(aki)schen, Kroatischen und Serbischen. Teil 1: Phonetik — Phonologie — Prosodie.*

---

<sup>8</sup> One of the Serbian speakers wrote *stolnjak*.