

The Challenges of Contrasting Interactional Lexical Layers of Slavic Cultural Identities

DANKO ŠIPKA, *Arizona State University*
 Danko.Sipka@asu.edu

Received: March 18, 2017.
 Accepted: September 8, 2017.

ABSTRACT

The present paper points to the challenges in contrasting interactional lexical layers of Slavic languages using the datasets of Slavic loanword dictionaries. Three main challenges comprise sampling bias, language labeling bias, and field labeling bias. In order to partially mitigate these challenges, the author proposes the deployment of the generality index, metalanguage presence index, metafield presence index, and metafield profiling index. All these indices measure quantities relative to one another at one or more metalevels, which partially eliminates the challenges stemming from the tendency of the dictionary compilers to follow different strategies in selecting materials, and labeling languages and fields. The deployment of the proposed indices was demonstrated in contrasting Klajn and Šipka (2006), a Serbian (de facto Serbo-Croatian) loanword dictionary with Egorova (2014), a Russian loanword dictionary. The analysis has demonstrated the usefulness of the indices in confirming that the general structure of the loanwords is in large part shared between the two languages (despite superficial differences) and that some minor differences stem from different socio-historical circumstances of the two languages. It is important to realize that the solutions proposed in this paper only partially mitigate the challenges of working with the existing datasets. A durable solution would be the development of the guidelines for Slavic dictionaries of this kind. Contrasting two dictionaries compiled using the same set of guidelines would enable much less limited exploration of differences and similarities at all levels.

Keywords: Russian, Serbian, loanwords, lexical interaction, cultural identity.

1. Introduction

The present research is a part of a broader project titled *Lexical Layers of Cultural Identity in Slavic Languages*. A general outline of the model is provided in Šipka (2017a) and results stemming from the model are presented in Šipka (2017b). I will start by outlining the overarching epistemological construct proposed in the project to proceed with discussing the challenges of contrasting the interactional lexical layers of Slavic cultural identities.

Let us take a look at the following three lexical differences in Slavic languages.

- a. Russian²⁹ *глубокий* ‘deep blue’: *серый* ‘grayish blue’ versus Serbo-Croatian *plav* ‘blue’,
- b. Russian *подушка* ‘pillow’ versus Serbo-Croatian *jastuk* ‘pillow’,
- c. Polish *tłok* ‘piston’ versus Serbo-Croatian *klip* ‘piston’.

The first example features a difference of a less or more precise differentiation of a conceptual field. It would be hard to imagine geopolitical, historical generators of this

²⁹ When using a language name without any qualification, I am implying modern standard variety of that language. Other varieties (e.g., regional dialects, sociolects, etc.) may or may not follow the modern standard variety in their lexical systems. Serbo-Croatian (also known as Bosnian/Croatian/Serbian) refers to the features common in all its standard variants.

difference. Similarly, this difference is not a product of linguistic planning. The difference seems to have rose spontaneously and it would be very difficult to trace its underlying mechanisms. Obviously, one can trace the etymology of the words in both languages and link the to the same conceptual field reconstructed for Proto-Slavic, but that still does not explain what caused a more precise division in Russian than in Serbo-Croatian. On a side note, it is important to keep in mind the famous formulation by Jakobson (1959:236): “Languages differ essentially in what they must convey and not in what they may convey.” In this particular case, a speaker of R has to break up the concept of blue into two values while a speaker of Serbo-Croatian does not (although it is possible to do so using compounds).

In the second case the difference is a direct result of historical and geopolitical circumstances. In Russian (just like in Polish and some other Slavic languages) the word for pillow is inherited from the common Slavic lexical pool while Serbo-Croatian features a Turkish loanword. Without intensive and direct cultural contact of the Serbo-Croatian speaking areas with the Turkish language and its culture, this difference would not exist. We can thus easily identify historical and geopolitical underlying mechanisms that have generated this difference. We can also see that this difference is not a result of planned intervention.

Finally, the third example shows a result of planned intervention. In establishing car mechanic terminology, the Polish language based the name on piston on what that part does in the internal combustion engine. The word *tłok* ‘piston’ is derived from the verb *tłoczyć*, roughly: ‘to compress’. We thus have a function-based word-formation at play. In contrast to that, Serbo-Croatian chose a metaphor based on visual similarity. The word *klip* means ‘corn cob, ear of corn’ and it extended its meaning into ‘piston’ (the process was possibly influenced by the same metaphor in German *Kolben*).

The three aforementioned cases exemplify the three layers of lexical differences and similarities between Slavic languages. These similarities and differences build cultural identity of Slavic languages in a sense that the differences mark the identity of particular Slavic languages and their cultures while the similarities point to a common Slavic cultural identity. The following three layers can be identified.

- a. *The deep layer* – the lexical expression of cultural identity that features the highest degree of stability, which does not seem to be directly influenced by historical developments or political will of the community of speakers,
- б. *The interactional layer* – the lexical expression of cultural identity which is a result of direct or indirect intercultural communication.
- b. *The surface layer* – the lexical expression of cultural identity created by conscious intervention of linguistic and political elites within the community of speakers.

There is a certain degree of interaction between the three layers. For example, a conscious intervention (i.e., a maneuver in the surface layer) may consist of replacing the lexicon from the interactional layer (e.g., loanwords from another language in a terminology of some kind).

Although the three layers can be distinguished in any language or group of languages, Slavic languages constitute a particularly interesting case study. They feature intensive interaction with other languages (which is not the case with many other languages, at least

not to such a degree) and prominent lexical interventionism (again, in other linguistic cultures normativist activities are not that prominent even if they do exist). For these reasons, the epistemological construct of the three layers of identity seems particularly suited for the study of Slavic lexicons.

This particular research concerns the interaction layer. The factor of different historical development is clearly dividing Slavic languages and often the territories within the same language (compare, for example, five centuries of Ottoman Turkish rule in the Balkans versus century and a half of the partition of the Polish language space between Prussia, Russia, and Austria-Hungary). These differences are mirrored in the external linguistic history, most notably in the structure of the respective lexicons of Slavic languages.

Given the rich tradition of lexical interaction in Slavic languages, mostly borrowing from the languages of cultural influence, and strong normativist traditions in the surface layer, all Slavic languages feature numerous dictionaries of lexical borrowings (most commonly called: foreign words dictionaries). To illustrate this, Šipka (2000) attests 68 dictionaries of lexical borrowings for Serbo-Croatian in the period from 1830 to 2000. Numerous dictionaries have multiple editions and numerous new dictionaries of lexical borrowings were published.

It seems that the material from these dictionaries of lexical borrowings can be used to contrast the interaction layers of Slavic lexicons. Assuming that we are contrasting two Slavic languages, interesting questions here are the following. First, how much is the same and how much different in the structure of the borrowed lexicon. In other words, is lexical borrowing something that contributes more to the general Slavic cultural identity or more to the identity of each particular Slavic language? Second, what is the ratio of the sources of borrowing in the two observed Slavic languages? A follow up question can be formulated here as to which socio-cultural and historical circumstances shaped the ratio of the sources of borrowing. Third, what is the distribution of borrowing in the general lexicon relative to those in specialized fields? Finally, for each of the specialized fields, what is the proportion of each language of borrowing? The latter two questions expand on the previous two giving a more detailed picture of the differences and similarities between the interactional lexical layers of the two observed Slavic languages.

In the next section I will outline the methodological challenges of using contrastive data from Slavic loanword dictionaries to answer the aforementioned questions and the solutions intended to mitigate some of these problems.

2. Challenges and (Partial) Solutions

Answering the four aforementioned questions (general differences, the differences in the sources of borrowing, the differences in the structure of the sources, the differences in the participation of the loanwords in the general lexicon versus specialized fields of usage, and the differences in how the sources of borrowing profile specialized fields) assumes a *tertium comparationis* in contrasting the two Slavic languages in question. In an ideal situation, the two compared dictionaries would accurately reflect the body of loanwords in their respective language and the authors of the dictionaries use the same methodology (most notably, lexical selection, language and field labeling). Needless to say, the existing datasets of various Slavic languages deviate significantly from this ideal situation. If we set aside

sheer treatment inconsistencies that can be found in any dictionary, the following challenges can be distinguished.

a. Sampling bias – the dictionaries of loanwords follow different sampling techniques, which produces the differences between them. First, there are differences between between the body of loanwords in the dictionary and in the lexicon of that particular language. Second, any two observed Slavic loanword dictionaries exhibit different patterns of deviating from the general lexicon. To illustrate this bias, let us consider the terminology of philology (linguistics and theory/history of literature). Being philologists, the authors of these dictionaries are prone to including very specific terminology of these fields, which is not the case with the fields with which they may be less familiar (e.g., quantum physics or genetic engineering). This generates the difference toward the situation in the lexicon given that numerous terms of equal specificity in non-philological fields will be omitted. On the same token, the two dictionaries in question may exhibit different levels of depth in covering various fields (e.g., the terminology of computing may be covered exhaustively in one dictionary and only superficially in the other).

b. Language labeling bias – dictionaries of loanwords follow different strategies in labeling the languages of borrowing. The most conspicuous problem in this field is a different way of handling chains of borrowing and hybrids. All Slavic words feature a substantial number of loanwords that are adopted from French and/or German which are ultimately of Greco-Roman origin. In various Slavic loanword dictionaries we can the entire chain of borrowing labeled and even that in different ways (e.g., French from Latin or Latin via French) but we can also find the same loanword labeled using just the final or initial link in the chain (e.g., French or Latin).

c. Field labeling bias – dictionaries of loanwords follow different strategies in labelling specialized fields of usage. The most fundamental difference is that what may be labeled as a word belonging to a specialized field in one dictionary may be left without a label in another dictionary. To exemplify this, numerous computing terms known to the general population (e.g., *keyboard*, *disk*, *port*, *monitor*, etc.) may be labeled as *computing* in one dictionary, *technical* in another, and without a label in yet another dictionary. More broadly, the strategy of one author may be to shy away from using specialized field labels while another author may be prone to using them profusely. A compounding problem is the fact that field labels are inherently inconsistent – they are used to mark the words common in a field of usage and words that are used exclusively in the field, but they are also used to segregate the different senses of the same word (e.g., *port* in computing and in shipping or seafaring).

I will now turn to the solutions intended to mitigate the three aforementioned biases.

There is almost nothing that can be done about the sampling bias if one works with existing dictionary datasets. One natural solution is to acknowledge this research limitation. Furthermore, comparing percentages rather than numbers is an obvious solution to account for a different number of loanwords in the two contrasted dictionaries. A long-term solution would be that the Lexicology and Lexicography Committee at the International Committee of Slavists (see Committee, 2017) come up with guidelines for Slavic monolingual dictionaries compilation with a specific set of guidelines for dictionaries of loanwords. If two dictionaries would stick to such guidelines, e.g., that the collection of loanwords is conducted from representative corpora and monolingual dictionaries of approximately the same size, the selection bias would be largely mitigated (albeit not eliminated).

The language labeling bias can be mitigated in several ways. First, taking into consideration just the last link in the borrowing chain makes the situation somewhat less complicated and mitigates some of the language labeling bias (now the differences between the two languages are reduced to the difference in that language as opposed to the differences in establishing or not establishing the chain, differences in the ordering of the links in the chain, etc.). More importantly, establishing a taxonomy of language groups which enables contrasting at a metalanguage level significantly mitigates language labeling bias. For example, establishing the following in Figure 1 enables us to contrast the languages of borrowing at three different levels (e.g., European vs. Near Eastern, living European vs. classical European, Germanic vs. Romance, etc.). Thus, if the contrast is between European and Near Eastern Languages using just the final link in the chain of borrowing overcomes the problem of the differences of labeling the same entry as French, German, Greek, or Latin (or that in one dictionary Latin is a single category and in another it is divided into classical, medieval, and modern Latin) and, on the same token, labeling the same entry as Turkish in one dictionary and Arabic in the other. Needless to say, the taxonomy is primarily cultural, based on external linguistic history and the sources of cultural influence, rather than on structural and/or genetic criteria.

```

(European cultural influences
(living
(Germanic (German)...)(Romance (French)...
)
(classical
(Latin)(Greek)
)
...
)
...

```

Figure 1. Metalanguage hierarchy example.

Needless to say, a durable solution for language labeling bias would be that the aforementioned committee (Committee, 2017) establishes guidelines for the treatment of the chains of borrowing.

The principal problem of the field labeling bias is that the two observed dictionaries

may feature a different ratio of general versus specialized vocabulary, which stems from the authors' strategies (more versus less labeling) rather than from objective circumstances in the two observed languages. Nothing can be done about this in the research of existing datasets and this kind of contrasting should not be made. Here too, a set of guidelines by the Committee (2017) would significantly mitigate this problem. Nevertheless, some solutions may be implemented at the level of particular languages and language groups by using language generality index which would measure the level of generality (and, on the same token, specificity) of a language relative to generality in the entire body of loanwords. This index is computed in the following manner: the percentage unlabeled words in that language minus the average percentage of unlabeled words in the dataset. For example, if among the words borrowed from German there are 30 of unlabeled words, and if the average percentage for all languages in the dataset is 35, German in that dataset has a negative generality index of 5 (i.e., -5), which means that it is more specific in that particular field. If we compare this index in two dictionaries, the effect of the authors' strategies would be eliminated. The same index can be tabulated for each of the fields as follows: field presence index is equal to the percentage of the field in the field labeled lexicon minus average percentage for all fields. Thus, if natural sciences comprise 30 of all labeled entries and an average percentage is 25, the field presence index for natural sciences is +5. A similar index can be tabulated for the degree in which a language profiles a particular field. This field profiling index can be computed as: percentage of a language in a given field minus the percentage of the language in the entire dataset. If, for example, the percentage of Italian in the field of music is 20 and its percentage in the entire dataset is 15, Italian field profiling index for the field of music is +5, which means that it contributes to the profiling of that field more than it contributes to the profiling of the entire dataset. This index too is not dependent on the authors' strategies. Another problem with the fields is that the same entry can be labeled differently in two dictionaries. Here too, creating a hierarchy of metafields (e.g., bringing psychology, sociology, anthropology, etc. under the metalabel of social sciences) eliminates labeling bias to some extent.

To summarize, a durable solution for the challenges of contrasting Slavic dictionaries would be a set of guidelines for various types of Slavic dictionaries (the Lexicology and Lexicography Committee at the International Committee of Slavists would be a natural body to come up with such guidelines). Most challenges of contrasting would be eliminated if the two contrasted languages followed such guidelines.

Given the absence of the guidelines, the following strategies of contrasting need to be deployed in the existing loanword dictionaries datasets.

- a. All limitations (e.g., the discrepancy between the language dataset and the dictionary dataset) should be clearly stated,
- b. Contrasting the percentages of language and field metacategories should be used to eliminate some of the existing labeling bias,
- c. Using the indexes of generality, presence, and profiling further mitigates some of the existing challenges.

I will now proceed with exemplifying this by contrasting a S (de facto Serbo-Croatian) and a R loanwords dictionary.

3. An Example of Contrasting

I will now demonstrate the deployment of the four indices in contrasting two Slavic loanwords dictionaries: Egorova (2014), a Russian dictionary, and Klajn and Šipka (2006), Serbian in name but de facto Serbo-Croatian given that the materials from the entire area of this language are included in the dictionary. The two dictionaries cover approximately the same time period with only nine years between them. The two dictionaries differ in size with the Serbian dictionary being more than double in size. Once all cross-referenced entries and the loanwords of uncertain origin and those stemming from the names of people, firms, products, etc. were eliminated, the number of extracted entries (those that featured the label of the language of origin) was 26643 for Klajn and Šipka (2006) and 9426 for Egorova (2014). One should say these numbers are somewhat higher than the numbers of labeled entries given that each meaning with a field label was included as a separate entry. Thus, for example, if a word of Greek origin has one meaning labeled as physics and another labeled as medicine, there would be two entries in the database: Greek – physics and Greek – medicine. The entries that did not feature a field label were counted only once. Given that the same procedure was deployed for both dictionaries, for all involved languages and fields, its limitations are minimal. Only concrete field labels were extracted while all others (such as: humorous, colloquial, outdated, specialized, etc.) were disregarded, i.e., counted as non-labeled entries.

Following the idea stated in the previous section, only the first stated language of origin was recorded. Again, given that the same procedure was deployed in both dictionaries and for all languages, its limitations are minimal. The initial data from both dictionaries contained two variables with the language of origin as stated in the dictionary in one column and the field label or zero (for unlabeled cases) in the second. This was then recoded into two new variables with metalanguage and metafield values.

	Initial Variables		Recoded Variables	
	Language	Field	Metalanguage	Metafield
<i>Example</i>	<i>Greek</i>	<i>Physics</i>	<i>European – classical</i>	<i>Sciences – natural</i>

Figure 2. Initial and recoded variables.

The two recoded variables were used to contrast the Serbian and the Russian datasets. Metalanguage categories were distinguished based on the principal sources of cultural influences. There were three groups of European languages: Classical (Greek and Latin), Germanic (German, English, etc.), and Romance (French, Italian, etc.), Near Eastern Languages (Turkic, Arabic, Persian), other languages (which included both contact languages such as Hungarian for Serbian and Finish for Russian and “exotic” languages such as Chinese and Japanese), and Slavic languages. The fields were generalized in the categories of Arts (music, fine arts, etc.), Beliefs systems (religion, astrology, etc.), Natural, engineering, and mathematical sciences (physics, geography, computing, etc.), Social sciences and humanities (psychology, linguistics, etc.), Sports, and two categories of vocational terms:

production (such as plumbing or car mechanics nomenclature) and services (e.g., culinary or administrative). Each of these fields represents a different environment for cultural influences, e.g., popular culture in arts and sports, older cultural influences in vocational terminology and belief systems, etc. Just like the metalanguage categories are established based on the sources of cultural influences, the fields are generalized based on their potential to receive cultural influences. The entire research scheme was organized with an eye toward answering the questions about the interaction layer of Slavic cultural identities while minimizing the challenges discussed in the previous section.

I will now discuss the data based on the four indices: the index of generality, metalanguage presence index, metafield presence index, and metafield profiling index. The discussion of each index will be preceded by the description of their tabulation and followed by additional data about concrete language and fields as featured in the two dictionary dataset. S in the tables stands for the Serbian dictionary dataset, R for the Russian.

The index of generality was calculated in the following manner. An average percentage for all six metalanguage groups in the entries that were not field-labeled was first calculated. Then, the index of generality is computed in the following manner: index of generality = metalanguage presence in the general lexicon minus average presence in the general lexicon. The positive value of this index shows that the metalanguage group is used in the general lexicon more than average, i.e. that the level of generality of that metalanguage group is higher than average. The negative value means that the level of generality is lower than average, i.e., that the metalanguage group tends to get more use in the sphere of specialized vocabulary. The following results were found.

Source	S	R
European – Classical	-8.9	-13.1
European – Germanic	-8.5	-6.7
European – Romance	-2.8	-0.3
<i>Average for European Sources</i>	-6.7	-6.7
Near Eastern	6.5	16.2
Other	10.6	5.3
Slavic	3.17	-1.4

Table 1. Generality indices.

From the data in Table 1 we can see that European sources of consistently have negative generality index values and that their average generality index is identical in the two observed dictionaries. This means that the main European sources of influence contribute to the specialized vocabulary much more than Near Eastern, Slavic, and other sources. Another interesting fact is that Near Eastern terms show a higher degree of generality in Russian than in Serbian. The status of other sources is opposite. All this mirrors the cultural history of the two languages, primarily in the fact that Turkish played a more important role in the

development of some terminologies in Serbian.

The metalanguage presence index was tabulated as percentage of the metalanguage in question minus average metalanguage percentage. Its positive value means that the metalanguage in question has a higher-than-average presence and, conversely, its negative value means a lower-than-average presence of the metalanguage. The following values were obtained.

Source	S	R
European – Classical	35.2	17.5
European – Germanic	1.6	15.9
European – Romance	-0.3	12.4
<i>Average for European Sources</i>	<i>12.2</i>	<i>15.3</i>
Near Eastern	-6.0	-15.1
Other	-15.0	-15.6
Slavic	-15.6	-15.1

Table 2. Metalanguage presence indices.

The results from the Table 2 first show that the interactional lexical layer of both languages is shaped primarily by the three main European sources of influence (a high positive average indices for both languages). The structure of the borrowed lexicon is similar in the two languages and the differences between classical and living European languages stem from the tendency of the Serbian authors to label Greek or Latin as the last link in borrowing rather than German or French as is the case in the Russian dictionary. Another interesting fact is a higher presence of Near Eastern metalanguage category in Serbian, which is another consequence of different historical consequences – long Ottoman Turkish presence in the Balkans. If we look in the ratio of the languages in the two groups, the most frequent classical language in Serbian is Latin with 58.2% and in the Russian dictionary it is Greek with 52.8%. This may partially be a consequence of assigning ultimately Latin words German or French etymology in the Russian dictionary as opposed to treating them as Latin in the Serbian. In the Serbian dictionary the most common Germanic language is German with 56.8% followed by English with 39.3%. This ratio is very similar in the Russian dictionary with 57.7% of German and 37.3% for English. The most common Romance language in the Serbian dictionary is French (63.6%) followed by Italian (29%). The dominance of French is even more pronounced in the Russian dictionary (85.5% for French and 10.3% for Italian). This may be a consequence of contact borrowing from Italian in Serbo-Croatian. The structure of the Near Eastern borrowings is another consequence of different historical circumstances. The most common Near Eastern language in the Serbian dictionary is Turkish with whopping 91.1% while in the Russian dictionary the Arabic sources are first (32.2%) followed closely by Turkic languages (29.1%). This again is a clear consequence of different historical circumstances. In the “other” category, the dominant language in the Serbian dictionary

is Hungarian with 43.9% (a clear consequence of contact borrowing) while in the Russian dictionary Sanskrit is the first with 23.4%. In general, there is a higher number of languages if this category in the Russian dictionary, including, for example, visible presence of contact borrowing languages such as Finish and Chinese (7.5% each) or Mongolian (4.7%). Finally, the most dominant Slavic language in the Serbian dictionary is Russian with 84.3% (a clear consequence of cultural influence) and the Russian dictionary has Polish in the top spot with 95.4% of this metalanguage categories. In general, we can see that the big picture of the spheres of cultural influences is very similar in the two observed datasets while at a lower level there are differences stemming from sociohistorical circumstances.

The metafield presence index is tabulated as percentage of the field in question minus average percentage for all fields. Positive value means a higher-than-average presence of the field, negative value a lower-than-average presence of the field. Given that the two observed dictionaries may feature different strategies of labeling, this and the next index should be approached with much more caution than the previous two and the results should be interpreted with a much higher level of limitations. The following values were found.

Field	S	R
Arts	-2.9	-1.5
Belief systems	-3.8	-2.5
Sciences - Natural	14.0	6.8
Sciences - Social	6.0	1.7
Sports	-4.9	-2.4
Vocational - production	-5.1	-1.3
Vocational - services	-3.1	-1.0

Table 3. Metafield presence indices.

Despite all the differences, that most probably stem from different labeling strategies, we can still see that sciences (primarily natural, engineering, and mathematical) are the main field of cultural influences. This is a direct consequence of sociohistorical circumstances where in both languages scientific terminology came from somewhere else and consequently features numerous borrowed words, much more than terminology and nomenclature in other fields. Obviously, this does not change the fact that in both languages there may be domestic equivalents of these borrowed terms. This only reveals relative presence of the science metafields in the group of fields where borrowed words are found.

Finally, metafield profiling index for each metalanguage group was tabulated as: percentage of that particular metalanguage in the metafield minus average percentage for all metalanguage categories in the metafield. Positive value of the index means a higher-than-average presence in the metafield, negative value a lower-than-average presence. From this index we can see how involved any given metalanguage group is involved in profiling any given metafield. Here again, one should be aware of the limitations stemming from different labelling strategies. The following results were obtained.

	Arts		Belief Systems		Sciences - Natural		Sciences - Social		Sports		Vocational - Production		Vocational - Services	
	S	R	S	R	S	R	S	R	S	R	S	R	S	R
European – Classical	-1.9	0.3	0.7	-0.1	15.0	8.5	2.3	1.6	-1.1	-0.7	0.3	-0.1	-2.3	-0.6
European - Germanic	0.1	0.6	-0.8	-0.3	6.5	2.5	-3.3	0.8	2.0	0.7	1.2	1.6	1.1	2.7
European - Romance	4.7	1.1	-1.0	-0.3	-2.3	1.1	-4.6	0.1	0.4	-0.6	0.1	0.7	3.0	0.7
<i>European - Average</i>	<i>1.0</i>	<i>0.7</i>	<i>-0.3</i>	<i>-0.3</i>	<i>6.4</i>	<i>4.0</i>	<i>-1.9</i>	<i>0.8</i>	<i>0.5</i>	<i>-0.2</i>	<i>0.5</i>	<i>0.8</i>	<i>0.6</i>	<i>1.0</i>
Near Eastern	-2.9	-0.9	-0.7	1.3	-8.0	-3.4	-1.1	-1.3	-1.2	-0.8	-0.8	-0.3	-1.6	-1.2
Other	0.3	-0.9	-0.4	0.2	-6.0	-6.0	0.9	-3.9	0.0	2.1	-0.2	-1.0	0.1	-1.2
Slavic	-0.3	-0.2	2.1	-0.7	-5.3	-2.7	5.8	2.7	-0.2	-0.8	-0.6	-1.0	-0.3	-0.5

Table 4. Metafield profiling indices.

From Table 4 we can see that European languages play a strong role in profiling natural, engineering, and mathematical sciences, that Slavic languages are important in profiling the terminology of Social sciences, and that Romance languages are important in arts (more in Serbian than in Russian) and that Romance and Germanic languages are important in vocational terminology. With some minor exceptions (e.g., other languages in sports or Near Eastern languages in belief systems in the Russian dictionary), in all other cases metalanguage groups make a lower-than-average contribution to profiling the fields. These main trends are again consequences of socio-historical circumstances. The data from this table shows that the quantitatively most dominant field of natural, engineering, and mathematical terminology came in great degree from the West.

4. Conclusion

The present paper pointed to the challenges of using loanword dictionaries to contrast interactional lexical layers of cultural identity in Slavic languages. These challenges stem primarily from different dictionary compilation strategies. There are only limited ways in providing the answers to these challenges if we work with existing data. Four indices were proposed to provide partial answers to the challenges of working with the existing datasets of Slavic loanword dictionaries: the generality index, the metalanguage presence index, the metafield presence index, and the metafield profiling index. Their deployment was demonstrated in contrasting a Serbian (de facto Serbo-Croatian) and a Russian dictionary. The analysis has revealed similarities in the main trends of development of the interactional lexical layer and some minor differences stemming from different socio-cultural circumstances. Needless to say, a durable solution for the challenges of comparisons of this kind (and many others) would be the development of dictionary compilation guidelines

for Slavic dictionaries, which would eventually overcome the shortcomings of this kind of contrasts.

REFERENCES

- Committee. (2017). *Комиссия по лексикологии и лексикографии при Международном комитете славистов*. Retrieved from: <http://www.ruslang.ru/agens.php?id=lexcomission>
- Egorova, T.V. (2014). *Словарь иностранных слов современного русского языка*. М.: «Аделант».
- Jakobson, R. (1959). On Linguistic Aspects of Translation. In: Reuben A. Brower (Ed.), *On Translation* (pp. 232–39). New York: Oxford University Press, 1966 [1959].
- Klajn, I. & Šipka, M. (2006). *Veliki rečnik stranih reči i izraza*. Novi Sad: Prometej.
- Šipka, D. (2000). *A Bibliography of Serbo-Croatian Dictionaries*. Springfield: Dunwoody Press.
- Šipka, D. (2017a). Leksički slojevi slovenskog kulturnog identiteta. In: Rajna Dragičević (Ed.), *Putevima reči* (pp. 131-143). Beograd: Čigoja.
- Šipka, D. (2017b). Semantic Change in Slavic Inherited Lexicon: An Initial Analysis. *Wiener Slawistischer Almanach*, 79, forthcoming.