



UDC 811.16'373.42=111

Original scientific article

Accepted for publication on 24. 11. 2025.

<https://doi.org/10.29162/jez.2025.6>

Anastasia Makarova¹

Tore Nesset²

Laura A. Janda³

¹ Uppsala university, Sweden

² UiT – The Arctic University of Norway

³ UiT – The Arctic University of Norway

Security in Slavic: a linguistic approach

Does the word for ‘security’ mean exactly the same across languages? How can we identify cross-linguistic similarities and differences in word meanings? In this study we carry out a corpus investigation of the words corresponding to English security and safety in the Slavic languages. While our analysis identifies numerous cross-linguistic similarities, we argue that there are differences. Three parameters capture relevant differences: transparent vs. opaque, negative vs. positive, and external vs. internal. With regard to semantic fields and metaphorical uses, we find considerable overlap, but also some differences among the languages under scrutiny. Employing a variety of methods from corpus linguistics, we argue that this “mixed” methodology enables us to avoid the weaknesses of individual methods and to pinpoint both similarities and subtle differences among related languages.

Key words: security, Slavic languages, collocations, semantic vectors, ChatGPT, constructions, corpus

Introduction

In the current tense geopolitical situation, few concepts are more relevant than ‘security’. In our time, when texts are composed and translated by means of artificial intelligence, one must ask whether the words for ‘security’ mean the same across languages, or whether they involve different shades of meaning. Ultimately, can differences in semantic associations hinder communication at the state or sub-state (individual) level? With regard to unrelated or distantly related languages it may seem obvious that the relevant words may display different semantic and pragmatic nuances, but in the present study we compare a group of relatively closely related languages, namely the Slavic branch of the Indo-European language family. While our main focus will be on ‘security’, we will also briefly consider the closely related concept ‘safety’.

As a first approximation, consider the names of the UN Security Council in eleven Slavic languages:

- (1) The UN Security Council in the Slavic languages
 - a. Russian: *Sovet Bezopasnosti OON*
 - b. Belarusian: *Savet Bjaspeki AAN*
 - c. Ukrainian: *Rada Bezpeki OON*
 - d. Polish: *Rada Bezpieczeństwa ONZ*
 - e. Czech: *Rada Bezpečnosti OSN*
 - f. Slovak: *Bezpečnostná rada OSN*
 - g. Slovene: *Varnostni svet ZN*
 - h. Croatian: *Vijeće sigurnosti UN-a*
 - i. Serbian: *Savet bezbednosti UN*
 - j. Macedonian: *Sovetot za bezbednost na ON*
 - k. Bulgarian: *Săvet za sigurnost na OON*

Three observations can be made on the basis of (1). First, different languages use different roots. While Russian has the root *pas*, the remaining East Slavic and West Slavic languages use the root *pek*. In South Slavic, we find *sigur*, *var*, and *bed*. Second, although most roots are Slavic, *var* and *sigur* are borrowings.¹ Third, some languages have the prefix *bez-*, which arguably involves a negative construal (“without”), while other languages have unprefixed words that construe security in a positive way.

These differences suggest that the words for ‘security’ and ‘safety’ may involve different shades of meaning in the Slavic languages. However, how can we investigate this question empirically? The present study is an attempt to provide a principled answer to this research question, applying a variety of methods from cognitive and corpus linguistics:

¹ For the purposes of the present study, we use simple representations of roots which do not do justice to their historical origin and phonological form in all Slavic languages.

- (2) List of methods
 - a. Qualitative analysis of roots
 - b. Translations of security and safety
 - c. Semantic vectors
 - d. Collocations
 - e. Chat-GPT
 - f. Qualitative analysis of constructions

In other words, our research question is whether and how it is possible to pinpoint the differences and similarities between the words for ‘security’ and ‘safety’ in Slavic languages, using the methods listed above. We draw on data from three corpora: Intercorp (a large parallel corpus), the Araneum corpus family (a family of large internet corpora), and the parallel subcorpus of the Russian National Corpus (a curated corpus with parallel texts from a number of languages).²

Our contribution can be summarized as follows. First, our analysis brings out a number of similarities and differences among the Slavic languages. Three pairs of notions capture the different conceptualizations: transparent vs. opaque, negative vs. positive, and external vs. internal threat. Second, we identify a number of semantic fields that are relevant for security; while there is considerable overlap across Slavic, there are also differences. Third, we explore metaphorical conceptualizations of security in Slavic. Fourth, we employ a variety of methods that complement each other. We argue that this methodology enables us to avoid the weaknesses of individual methods and to pinpoint both similarities and subtle differences among related languages.

Our argument is structured as follows. We start with qualitative analysis of roots in section 2, before we consider translations across Slavic (section 3), semantic vectors (section 4), and collocations and metaphors (sections 5-6). A comparison of corpus results with those provided by artificial intelligence is considered in section 7, and grammatical constructions are examined in section 8. We summarize our findings in section 9.

Five roots: qualitative analysis

As illustrated in (1), there are five roots used to represent the meaning ‘security’ across Slavic. In this section, we take a closer look at these roots, considering the notions transparent vs. opaque, negative vs. positive, and external vs. internal threat.

² Intercorp can be accessed at <https://intercorp.korpus.cz/>. The Aranea corpora are available at <http://unesco.uniba.sk/aranea/index.html>. For the Russian National Corpus, see <https://ruscorpora.ru/>.

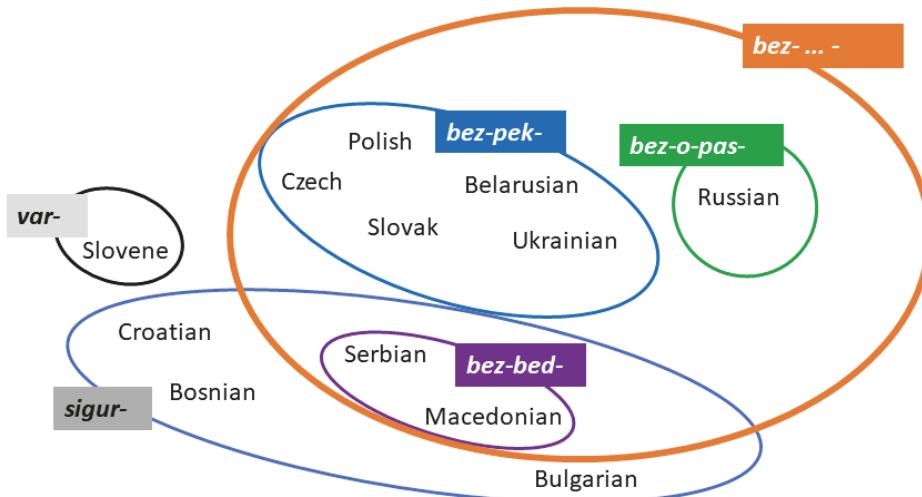


Figure 1. Security in the Slavic languages

Figure 1 shows the geographical distribution of the five roots across Slavic. As shown, the root *pek* (with modifications due to historical changes that have yielded *e* > *i* and *k* > *č* in places) is found in East and West Slavic. This is a root that goes back to Common Slavic and has various connotations in the modern Slavic languages. On the one hand, the root is used for worries and concerns, but on the other hand *pek* may also involve caring about someone. The following examples from Ukrainian illustrate this. We find *pek* in words like *pečal* 'sorrow', but also in *opika* and *pikluvannja*, which can both be glossed 'custody, care':

- (3) Tilky droty telegrafni gudut' sumovyto pid vitrom, roznosjat' **pečal'** po prostorax. 'Only the telegraph wires are sadly buzzing in the wind, spreading the **sorrow** around.' (Ukrainian, Dovženko 1930-1952)
- (4) Zlisne nevykonannja obov'jazkiv po dogljadu za dytynoju abo za osoboju, ščodo jakoji vstanovlena **opika** čy **pikluvannja**. 'Deliberate failure to fulfil the obligations of care of a child, which one has under **custody** or **foster care**'. (Ukrainian, Ukrainian criminal law 2001-2011)

A common denominator for these examples is the fact that they both involve internal psychological states. The root describes what goes on in the mind, without relating it to any external threats or dangers.

In South Slavic, the dominant pattern is *sigur*. This root is a borrowing from Latin *sēcūrus* where *se* (*sin-*) means 'free from' while *cura* means 'care' or 'worry'. However, for speakers of modern Slavic languages, this is not transparent. The root just means 'security', and unlike *pek*, which occurs in a number of other words and

can be related to many words such as the Ukrainian *pečal'* 'sorrow', *opika* 'custody' and *pikluvannja* '(foster) care' in (3) and (4), *sigur* does not come with associations to other concepts like concern, worry or care. We hasten to add that we do not claim that all native speakers make connections between words such as Ukrainian *pečal'*, *opika* and *pikluvannja*. However, there is nevertheless a difference between *pek* and *sigur*, since the latter is only used about security, while the former occurs in words with a variety of meanings across Slavic.

While *sigur* is found in Croatian, Bosnian, and Bulgarian, Serbian and Macedonian instead use the root *bed* in the name of the UN Security Council. This is a Slavic root that is connected to concepts such as misfortune, hardship and sorrow.

(5) Međutim, ni ova zajednička **beda** nije mogla da zbliži konzula i njegovog prvog saradnika. 'However, neither this **misery** that they shared, could bring the consul and his first associate closer.' (Serbian, Andrić 1942)

While this example most likely involves some external problem that both protagonists share, the following Bulgarian example shows that *bed-* is also used about an "internal threat", viz. the lack of knowledge of English:

(6) Vie ne polzvate anglijski, tova e **bedata**. 'You do not speak English, that's the **problem**.' (Bulgarian, Vežinov 1975)

The relationship between *bezbednost* and other words with the root *bed-*, such as Serbian *bedno* 'miserably, poorly', *obezbedenje* 'assurance, security' is straightforward and transparent.

In Slovene, we find the root *var* in the name of the Security Council. This is a borrowing from Germanic, related to German *wahren* 'safeguard, observe, take care of'. The word is morphologically opaque, in the sense that it cannot be decomposed into smaller meaningful units.

Finally, Russian is an outlier among the North Slavic languages insofar as it uses the root *pas-*, which we also find in words like *pasti* 'herd' and *spasti* 'save', suggesting that security is connected with taking care of something. Historically at least, *o-* can be analyzed as a prefix. It has the meaning 'around', so taken together *opas* must have had the meaning of being surrounded by protection or care. Note that *opasnyj* has changed its meaning. In older texts it is attested in the meaning 'protective', as in *opasnaja gramota* 'protective document' (a document that would entitle the holder to help and protection):

(7) **Opasnaja** gramota za velikogo novgoroda pečatmi 'protective document with the Great Novgorod stamps' (Middle Russian, Opasnaja gramota Novgoroda ganzejskim poslam 1480-1500)

Later *opasnyj* became used the meaning 'dangerous', as illustrated in the following example:

(8) Put' učinilsja trudnyj i **opasnyj**. 'The road turned out to be difficult and **dangerous**' (Russian, Verevkin 1782)

The change can be described in terms of metonymy, i.e. a semantic extension within the same domain (see Radden and Kövecses 1999, Peirsman and Geeraerts 2006), since we are dealing with different facets of the same situation involving an external threat: either one is protected from the threat, or one is exposed to it and thus in danger.³

In the same way as *opasnyj*, the related noun *opasnost'* 'danger' refers to an external threat.

In Contemporary Standard Russian, the relationship between *bezopasnost'* and *opasnost'* 'danger' is transparent, so 'security' is construed as the absence of something dangerous, and refers to an external threat.

We have seen that several languages have the prefix *bez-* 'without' in the words for 'security'. We refer to this as a "negative construal", since the prefix implies conceptualizing security as the absence of something.

To summarize, we have seen five different roots that represent 'security' in Slavic. Some languages have transparent word-formation patterns, while other languages use opaque morphology from the point of view of modern language users. Our data show that security may be conceptualized as the absence of something, and while some of the words refer to external threats which can appear at state-level or at sub-state level, other words for 'security' involve individual internal psychological states.

Intercorp: translations of *security* and *safety*

So far we have analyzed the names of the UN Security Council. We now zoom out and consider translations of the two English words *security* and *safety* into eleven Slavic languages.⁴ Our data come from Intercorp, an internet corpus with parallel data from a number of languages.⁵ English has two words with closely related mean-

³ The presence of antonymous etymological descendants from a single root is not unusual. Note, for example, that both English *black* and the Slavic root *bel* 'white' (as in Russian *belyj*) derive from the same Indo-European root containing *b(V)l*, cf. Norwegian *bål* 'fire'; a fire produces both black coals and white ashes.

⁴ An anonymous reviewer points out that our "findings do not reflect usage in present-day language", as long as we are not studying the actual modern languages, but only focus on translations. While we agree that translations can have their specific features, we believe that translations, made by the users of the respective languages, do reflect the actual usage in the languages.

⁵ Searches were carried out in spring 2023. The version of InterCorp used for the searches is "Release 15", which has 1 588 mil. word forms in foreign language texts. For more information on the composition of this version of the corpus, see <https://wiki.korpus.cz/doku.php/en:cnk:intercorp:verze15>. The "Translations equivalents database" (Treq, <https://treq.korpus.cz/index.php>) was used for separate searches in

ings, *security* and *safety*. We took the first one hundred examples of each English word as returned by the corpus in the query and analyzed the translations of them into the Slavic languages in question.⁶ Table 1 and Figure 2 summarize the situation.

Table 1. The distribution of roots in translations of *security* and *safety* into eleven Slavic languages. The total in each row equals 100 examples.

		pas-	pek-	bed-	sigur-	var-	xran-	other
Russian	security	95	2				3	
	safety	89					11	
Belarusian	safety		100					
Ukrainian	security		100					
	safety		100					
Polish	security		80				5	15
	safety		100					
Czech	security		80				3	17
	safety		96				1	3
Slovak	security		91					9
	safety		99				1	
Slovene	security					100		
	safety					100		
Croatian	security				100			
	safety				100			
Serbian	security			32	43			25
	safety			41	59			
Macedonian	security			85	15			
	safety			68	32			
Bulgarian	security	2	3		80		1	14
	safety	90			7			3

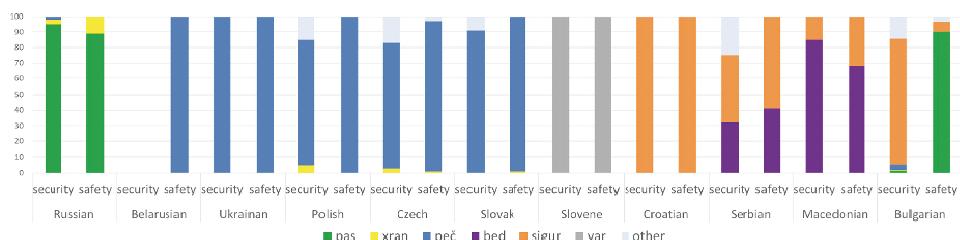


Figure 2. The distribution of roots in translations of *security* and *safety* into eleven Slavic languages

the relevant Slavic languages in order to get information about the relative frequencies of the various translations. Search results were then exported to an Excel file using the available function in the corpus.

⁶ For Belarusian, the corpus did not have enough data for *security*, so for this language we only consider translations of *safety*.

Figure 2 confirms the analysis from the previous section, but adds some nuances to the picture. It is striking that there is considerable variation, and it stands to reason that a larger dataset would involve even more variation. A number of observations can be made. First, we see that a sixth root, *xran*, is attested in Russian, Polish and Czech. This root is associated with (security) guards. Consider this Polish example:

(9) Olivetti was emerging from the **security** center. 'Olivetti wychodził właśnie z centrali sił **ochrony**'. (Polish, InterCorp v15: Brown 2000)

Second, we find a miscellaneous category of "other roots" in Polish, Czech, Slovak, Serbian and Bulgarian. This category includes examples of the following type:

(10) „Kde je?“ zeptal se a obrátil se k dělníkům, kteří se pro **jistotu** stáhli do hlučku. 'Where is she?' he demanded, turning on them grouped together for **safety**'. (Czech, InterCorp v15: Bainbridge 1974)

Third, in addition to variation between languages, we also see that there is variation within individual languages. A case in point is the competition between *sigur* and *bed* in Serbian and Macedonian.

Fourth, the figure facilitates comparison of translations of *security* and *safety*. For some languages, e.g., Ukrainian, Slovene, Croatian, there are no differences between security and safety. In other languages, notably Russian, Polish, Czech, Slovak, Serbian, and Macedonian, there are quantitative differences. In Serbian, for instance, we have a competition between *bed*, *sigur* and other roots for both *security* and *safety*, but the distribution is different for the two English nouns. In Bulgarian, *safety* and *security* show completely different distributions, whereby *sigur* dominates for *security*, while *pas* is widely used for *safety*.

To summarize, from the analysis of translations, we see that some languages display more than one root, and some, but not all Slavic languages treat *security* and *safety* differently. A larger database is likely to involve more variation, but our analysis is sufficient to show that there is more variation than suggested by the names of the UN Security Council.

Araneum: semantic vectors

Semantic vectors are a method for measuring the similarity between word meanings that has gained popularity in recent years. The basis is the so-called Distributional Hypothesis, whereby words that occur in similar contexts are assumed to be semantically close to each other (see Jurafsky and Martin 2025 for discussion and Nesset and Makarova 2023 for a recent application of the method on Russian data). Simply put, the method involves computing a vector for each word based on

all the contexts where it is attested. Each vector can be represented as a point in a multidimensional space, which makes it possible to measure the distance between each vector. The shorter the distance, the more similar the words.

The Araneum corpora include an automatic function for computing vectors, and the corpus returns lists of the closest “neighbors” for each word. While it was possible to study vectors for all eleven languages, we prefered to limit the analysis to only four languages for practical reasons.⁷ We focus on the words for security in four languages, Russian, Ukrainian, Polish and Czech, and analyze the twenty most closely related words in each language. The lists are given in Table 2. As shown, the lists are quite heterogeneous, ranging from proper names and acronyms of government agencies to common nouns, some of which are near synonyms to the target word in the languages under scrutiny. Nevertheless, some trends can be identified.

Table 2. The most closely related words to ‘security’ in Russian, Ukrainian, Polish and Czech. Data from the Araneum corpora

Russian	Ukrainian	Polish	Czech
bezopasnost’ ‘safety’	bezpeka ‘safety’	bezpieczeństwo ‘safety’	bezpečnost ‘safety’
oxrana ‘protection, guards’	bezpečnist’ ‘safety’	obronność ‘defense’	bezpečnostní ‘safety’
požarnyj ‘fire’	kiberbezpeka ‘cybersecurity’	stabilność ‘stability’	zabezpečení ‘security’
kiberbezopasnosti ‘cybersecurity’	bezpekovýj ‘safety’	bezpieczeństwo_ pożarowy ‘fire safety’	ochrana ‘protection, ward’
obespečenie ‘providing’	energobezpeka ‘energetic safety’	niezawodność ‘reliability’	zabezpečování ‘security’
radiacionnoj ‘radiation’	nacbezpeka ‘national security’	Bezpieczeństwa ‘safety’	bezpečný ‘safe’
zaščitennost’ ‘protection’	požežnyj ‘fire’	komfort ‘comfort’	zajištování ‘securing’
protivopožarnyj ‘against fire’	kiberzaxyst ‘cybersecurity’	pięciopunktowy_szelka ‘five-point belt’	zajištění ‘insurance’
vzryvobezopasnosti ‘explosion safety’	zaxyščenist’ ‘safety’	ochrona_przeciwożarowy ‘fire safety’	zdraví ‘health’
pravoporjadok ‘law and order’	garantuvannja ‘guaranteed’	higiena ‘hygiene’	dodržování ‘compliance’
Ènergobezopasnosti ‘energetic safety’	èkobezpeka ‘ecosafety’	poufność ‘confidentiality’	hospodárnost ‘economy’
sanitarii ‘sanitary’	biobezpeka ‘biosafety’	FRSC (Federal Road Safety Corps)	provozuschopnost ‘operability’
antiterorističeskij_ zaščitennost’ ‘antiterrorist_protection’	zabezpečennja ‘software’	NHTSA (National Highway Traffic Safety Administration)	kybernetický ‘cyber’

⁷ In order to reproduce our analysis, the reader is invited to visit <https://www.juls.savba.sk/sem%C3%A44/?lang=ru&kio=lemma&visualsel=gnuplot&topn=24&wpos=&wneg=#>, select a language, choose “lemma”, then choose “gnuplot” and search by pressing the “go” button. The corpus then returns a list of the related words and a plot, which can both be downloaded.

Russian	Ukrainian	Polish	Czech
protivodejstvie_terrorizmu 'counteracting_terrorism'	radiacijnyj 'radiation'	wygoda 'convenience'	BOZP (Bezpečnost a ochrana zdraví při práci) 'working safety'
élektrobezopasnost' 'electrical safety'	oborozdatnist' 'defense capability'	interoperacyjność 'interoperability'	nezávadnost 'harmlessness'
soxrannost' 'preservation'	pryvatnist' 'privacy'	efektywność_kosztowy 'efficacy_cost'	ekonomičnost 'economy'
pb 'fire safety'	oborona 'defense'	bezpieczny 'safe'	kyberbezpečnost 'cybersecurity'
požarobezopasnosti 'fire safety'	bezopasnist' 'safety'	ruch_drogowy 'traffic safety'	stabilita 'stability'
besperebojnosti 'continuity'	bezpečnyj 'safe'	przeciwporażeniowej 'anti-shock'	bezporuchovost 'flawlessness'
biobezopasnost' 'biosafety'	žyttjedijal'nist' 'life activity'	trzypunktowy_pas 'three-point belt'	efektivita 'efficacy'

In Russian and Ukrainian, the majority of words relate to what we may call “collective security”. Such words are typically associated with external threats, such as fire, radiation, natural disasters and terrorism, as well as infrastructure (e.g., electricity and cybersecurity). This tendency seems less pronounced in the two West-Slavic languages. In both Polish and Czech, we find words referring to ‘stability’, ‘comfort’, ‘health’, and ‘hygiene’. Such words are arguably more relevant for an individual’s well-being than for the collective security of a society or state, although we hasten to add that it is impossible to draw clear-cut boundaries between “collective” and “individual” words on the lists in Table 2.

Although our data do not permit strong conclusions, our findings suggest that there may be a stronger focus on collective security in East Slavic than in West Slavic. Furthermore, our data support the idea that the concept of security may invoke different connotations in different languages.

Araneum: collocations – nouns and adjectives

The Araneum Corpora also identify collocations, i.e., words that occur together more frequently than one should expect based on the frequency of each word.⁸ Using collocations to explore word meanings is a well-established method in corpus linguistics (see, e.g., Divjak and Gries 2009; Hanks 2013, 2020; Pivovarova et al. 2018 for discussion and Nesset and Makarova 2023 for a recent application of the method on Russian data). We entered the words for ‘security’ in five languages and the cor-

⁸ The anonymous reviewer suggests that the “method that has been proven solid for identifying word meanings is using Word sketches” and advocates for the use of www.sketchengine.eu. While we are familiar with the method and agree that sketchengine is a well-tested tool, we prefer to stick to the functions available in the Araneum Corpora.

pus yielded ranked lists of collocations for each word. In addition to the languages analyzed in the previous section, we also received analyzable data for Bulgarian.

In Table 3 we have singled out seven semantic fields that are dominant in our data when it comes to nouns and adjectives that collocate with the words for security. Each semantic field is represented as a row in the table. Although other classifications are possible, the seven semantic fields we propose offer a good overview of collocations characteristic for security. The table furthermore enables us to compare the languages under scrutiny. Each field is represented in several languages.

Table 3. Dominating semantic fields associated with 'security' in Russian, Ukrainian, Polish, Czech and Bulgarian. Data from the Araneum corpora

	Semantic fields	Ru	Uk	Pl	Cz	Bu
a	cyber, information	informacionnyj 'information'	informacijnyj 'information'		kybernetický 'cyber'	informacionen 'information'
b	traffic	dorožnyj 'road', remen 'belt', dviženie 'traffic', poduška 'airbag'	dorožnij 'road', rux 'traffic'	drogowy 'road', ruch 'traffic', pas 'belt'	silniční 'road', provoz 'traffic'	
c	providers, organizations, procedures	trebovanie 'request', texnika 'technique', služba 'service', sistema 'system', mera 'measure', pravilo 'rule', sobljudenie 'compliance'	texnika 'technique', pravilo 'rule', služba 'service', dotrymannja 'compliance'	služba 'service', system 'system'		služba 'service', sistema 'system'
d	nuclear, biology, ecology	ékologičeskij 'ecological'	ekologičnyj 'ecological', jadernyj 'nuclear'		jaderný 'nuclear'	
e	energy		energetičnyj 'energetic'	energetyczny 'energetic'		energien 'energy'
f	fire	požarnyj 'fire'	požežnyj 'fire'		požární 'fire'	
g	food		prodovol'čyj 'food'		potravina 'food'	

While no field is found in all languages, three fields are represented in four out of five languages. This applies to (a) cyber and information security, (b) traffic safety, and (c) the field we refer to as "providers, organizations, procedures". This latter field encompasses collocates such as words for 'system', 'technique' and 'service'. Russian and Ukrainian have collocates in all three fields (a)-(c), while the South and West Slavic languages in the table lack one of these fields.

The fields referred to as (d) "nuclear, biology, ecology", (e) "energy" and (f) "fire" are found in three out of five languages. Ukrainian is the only language that has collocates from all three fields.

The field labeled “(g) food” in the table is represented in two languages: Ukrainian and Czech. The remaining languages do not display collocates in this category.

Two conclusions can be drawn, although the material is limited and therefore calls for caution. First, we see considerable overlap when we compare languages: the same fields are found in several languages. Second, there are differences among the languages. Ukrainian is the only language that has collocates in all semantic fields in the table, while Polish and Bulgarian at the other end of the spectrum have collocates in only three out of seven fields. Taken together, these findings indicate that the concept of security is similar across the Slavic languages, while at the same time there may be differences between languages. If ‘security’ meant exactly the same in all (Slavic) languages, we would expect exactly the same or, at least, very similar sets of collocates across languages.

Araneum: collocations – verbs

The previous section only addresses nominal collocates, but the words for ‘security’ also occur together with verbs (and deverbal nouns with the same meaning as the corresponding verbs). What can you do with security? Certain verbs are found as collocations in more than one language in our dataset. These verbs represent a variety of conceptualizations of ‘security’. These conceptualizations are metaphorical in the sense of Lakoff and Johnson (1980) and Lakoff (1993).

Verbs for ‘provide, ensure, maintain’ often co-occur with ‘security’:

(11) Ru: *obespečit* ‘provide, ensure’
Pl: *zapewnić* ‘ensure’, *dbać* ‘provide, ensure’
Cz: *zajistit/zajistřovat* ‘provide, ensure’

Such verbs suggest that ‘security’ can be conceptualized as a commodity that someone can give to someone else.

Closely related are verbs meaning ‘guarantee’:

(12) Pl: *gwarantować* ‘guarantee’
Bu: *garantiram* ‘guarantee’

Verbs of this type indicate that someone can promise to do whatever is needed for someone to have the ‘commodity’ of ‘security’.

The words for ‘security’ furthermore collocate with verbs meaning ‘increase’:

(13) Pl: *zwiększyć* ‘increase’
Cz: *zvýšit/zvyšovat* ‘increase’

Here, security is conceptualized as a scalar property that can be higher or lower on a *quantitative* scale. However, *qualitative* scales are also relevant, as suggested by verbs meaning ‘improve’ (or deverbal nouns with the meaning ‘improvement’):

(14) Pl: *poprawa* ‘improvement’

Finally, we have collocates of the following type:

(15) Ru: *soblijudenie* ‘observance, abidance’

Such collocates testify to a construal of ‘security’ as a set of rules one must obey.

Taken together, the verbs and deverbal nouns that collocate with the words for ‘security’ suggest that security is conceptualized as a transferable commodity, a scalar property, and a set of rules.

Chat GPT

We additionally carried out a small experiment with artificial intelligence (Chat GPT) to explore the associations it offered for security in various languages.⁹ This experiment was purely exploratory, and as such is not suggested as a research method on a par with corpus study. AI does not function as a full-fledged methodological tool for data collection in the same sense as the corpus-based methods. Our aim was to probe whether an AI model that has access to vast data on cooccurrences of words produces results that align with findings obtained through established linguistic methods. We gave the model prompts in the ten languages listed in Table 4. For Bulgarian, we asked the model for information about both *sigurnost* and *bezopasnost*, as shown in the table.

For each language, we gave the bot the following prompt in English, asking it to respond in the relevant Slavic language. By way of example, consider the following prompts for Slovene:

(16) “You have access to a large amount of language data in Slovene. I would like to ask you to take a look at that data, and in particular please look at how the word *varnost* is used in the language data in Slovene that you have access to. Based on your data, please answer for me the following questions about how the word *varnost* is used in Slovene:

- 1) What kinds of semantic associations are there with the word *varnost* in Slovene?
- 2) What kinds of collocations do you find with the word *varnost* in Slovene?

⁹ We used the Chat GPT 3 version, <https://chatgpt.com>. The experiment was carried out in the spring of 2023.

- 3) When people use the word *varnost* in Slovene, what kinds of things are they thinking about?
- 4) What kinds of domains (politics, society, health, economics, etc.) are relevant when people use the word *varnost* in Slovene? Please write in Slovene when answering these questions. Thank you."

Here are the examples we received for Slovene:

(17) "1. S besedo *varnost* v slovenščini so povezane različne semantične konotacije, kot so zaščita pred nevarnostjo, ohranjanje reda in miru, zanesljivost, zaupanje, stabilnost, varovanje človekovih pravic in svoboščin, fizična in finančna varnost ter varnost na delovnem mestu."

'The word *varnost* has a number of various semantic connotations in Slovene, such as protection from danger, maintaining peace and order, reliability, trust, stability, protection of human rights and freedoms, physical and financial safety, as well as workplace safety.'

"2. V slovenščini se beseda *varnost* pogosto pojavlja v različnih kolokacijah, kot so nacionalna varnost, varnostni sistem, varnostna politika, varnostna služba, varnostni ukrepi, varnostni pas, varnostni list, varnostna oprema, varnostni standardi itd."

'In Slovene, the word *varnost* often occurs in various collocations, such as national security, safety system, security policy, security service, safety measures, safety belt, safety list, safety equipment, safety standards etc.'

"3. Ko ljudje v slovenščini uporabljajo besedo *varnost*, lahko razmišljajo o različnih stvareh, kot so osebna varnost, varnost v družbi in skupnosti, varnost pri delu, finančna varnost, varnost v prometu, varnost v primeru naravnih nesreč, nacionalna varnost in varnost države itd."

'When people use the Slovene word *varnost*, they may be thinking of various things, such as personal safety, social and community safety, work safety, financial security, traffic safety, safety in the event of natural disasters, national security and state security, etc.'

"4. Beseda *varnost* se nanaša na različna področja in domene, kot so politika, družba, zdravje, delo, finance, transport, energetika, okolje in naravne nesreče, kibernetska varnost, obramba in vojska ter mednarodna varnost in stabilnost."

'The word *varnost* refers to various domains, such as politics, society, health, work, finances, transport, energy, environment and natural disasters, cybersecurity, defense and military security, international security and stability.'

Based on the answers we received, we classified the results and summarized the findings as shown in Table 4. We identified sixteen semantic fields that recur across

the languages in question. In the table, we registered whether a semantic field was mentioned by Chat GPT for the relevant languages. A plus sign indicates that the relevant field was represented in the answers from the bot for the language in question. In the table, the most frequently attested fields are placed at the top. The right-most column shows the total number of plus signs for a semantic field, while the bottom row gives the total number of plus signs for each language.

Table 4. Semantic fields in the experiment with ChatGPT

	Rus	Ukr	Pol	Cz	Slk	Sln	Cr	Srb	Mac	Blg s	Blg b	Total
politics and society	+	+	+	+	+	+	+	+	+	+	+	11
health	+	+	+	+	+	+	+	+	+	+	+	11
IT and technology	+	+	+	+	+	+	+	+		+	+	10
transportation	+	+	+	+	+	+		+	+	+	+	10
state	+	+	+		+	+	+	+	+		+	9
person and property	+	+			+	+	+	+	+	+	+	9
economy and industry	+	+	+		+	+	+		+		+	8
workplace	+			+	+	+		+		+	+	8
energy		+	+		+	+	+			+		6
crime					+	+			+	+	+	6
military		+	+		+	+						4
environment	+			+		+						4
accidents and natural disasters	+				+							3
terrorism	+							+				2
global, international	+					+						2
education and science		+	+									2
Total	9	13	11	7	11	13	7	9	7	8	9	

The results resemble the findings from the previous section. While there are numerous similarities across the Slavic languages, there are also some differences, suggesting that ‘security’ may not involve exactly the same associations in all the languages under scrutiny. The following points merit discussion. First, some fields are represented in more languages than others. In particular, we see that “politics and society” and “health” are attested in all ten languages. Second, some languages have associations to more fields than other languages. Ukrainian and Slovene, for instance, have more plus signs in the table, suggesting that the concept of security has a broader range of associations in these languages. Third, we note that the two Bulgarian words *sigurnost* and *bezopasnost* have very similar associations. Finally, it is worth mentioning that the bottom right corner in the table has few plus signs. This suggests that the South Slavic languages Croatian, Serbian, Macedonian and Bulgarian do not associate security with the fields “military”, “environment”, “accidents and natural disasters”, “terrorism”, “global, international”, and “education and

science". Possibly, therefore, the concept of security is somewhat more narrow in these languages compared to the East and West Slavic languages.

Summing up, our experiment with artificial intelligence suggests that some fields might be better attested across some Slavic languages than others, and conversely, that some languages might have more associations to more fields – a more even distribution across fields. While the material is limited and there are many uncertainties with the use of artificial intelligence, our findings tie in with findings from previous sections, suggesting that the words for security may not have identical meanings across the Slavic languages.

On a methodological level, while we do observe convergence between the results in this and previous sections, this does not mean that AI is a viable option for the kind of analysis we pursue in the present article. More evidence is needed before definite conclusions can be drawn.

Constructions and conceptualizations

Constructions with prepositions tell us how security is conceptualized across Slavic.¹⁰ In many languages, the words for 'security' combine with prepositions meaning 'in':

(18) Pilat otkryl glaza, znaja, čto on teper' **v bezopasnosti** – osuždennyx on videt' uže ne mog. 'Pilate opened his eyes knowing that he was **safe** now - he could not see the sentenced ones any longer.' [Russian, Bulgakov 1929-1940]

In this example, security is conceptualized as a container where one can be located; the fact that Pilate feels safe is described with the preposition *v* 'in' with 'bezopasnost' in the locative, in this way we are dealing with a linguistic manifestation of the CONTAINER image schema (Johnson 1987, Lakoff and Johnson 1980, Kövecses 2010). This is the default construction for describing location within a three-dimensional space in Russian, and 'security' is therefore conceptualized as such a space.

Metaphorical uses of prepositions and prepositional phrases are well-described for Slavic languages (Janda and Clancy 2002, 2006; Šarić 2008; Parizoska 2009; Kiš Žuvela and Parizoska 2023), it is therefore not surprising that the *v* + locative construction is very widespread across Slavic with words for 'security'. Parallel examples are well attested in several languages:

¹⁰ The relationship between metaphor and grammar is an important field of cognitive studies in general (see Deignan 2005, 2012; Stefanowitsch 2006; Sullivan 2013) and Slavic cognitive linguistics in particular (among others, Janda & Solov'yev 2009; Stanojević et al. 2014; Stanojević 2019; Šarić & Stanojević 2019).

- (19) Belarusian: u bjасpeci
- (20) Ukrainian: u bezspeci
- (21) Czech: v bezpečí
- (22) Bulgarian: v bezopasnost

A different conceptualization is found in languages where the relevant words are governed by a preposition meaning 'on':

- (23) Serbian, Croatian: na sigurnom
- (24) Slovene: Grob je daleč, sem si rekел, tam bo **na varnem**. 'The grave is far away, I told myself, there he will be **safe**' [Kosmač 1950]

Notice that Serbian and Slovene use adjectives in this construction. *Na* is used when something is located on a (horizontal or vertical) plane, so here security is conceptualized as a metaphorical plane, which one can be located on. This conceptualization appears to be frequent in Serbian, Croatian and Slovene.

Although both conceptualizations with 'in' and 'on' are widespread, these constructions are not used in all languages. In Polish, the preferred construction appears to be with an adjective:

- (25) Czy są **bezpieczni**? 'Whether they are **safe**?' [Sapkowski 2006]

Rather than a metaphorical container or plane, examples of this type involve conceptualizing security as a property one can have or lack.

To summarize, constructions with prepositions indicate that 'security' can be conceptualized as a metaphorical container or plane. However, while such conceptualizations are widespread, some languages appear to prefer conceptualizing 'security' as a property.

Conclusions

In this article, we have explored the notion of 'security' across the Slavic languages, employing a number of different methodologies in corpus linguistics. Although our dataset is not very large, it is possible to draw some conclusions. Our contribution can be summarized as follows. First, we have seen that the relevant words involve five different roots in Slavic. In order to describe differences among the languages, we have advanced three pairs of concepts:

- Transparent ~ opaque
- Negative ~ positive
- External threat ~ internal psychological state

Examples of transparent patterns are words such as Serbian *bezbednost* where it is clear to modern language users that the word consists of *bez* ‘without’ and *bednost* ‘misery’. Croatian *sigurnost* and Slovene *varnost*, on the other hand, are *opaque*, since a layman cannot divide this into smaller meaningful units (except the abstract suffix *-nost*, which does not contribute to the understanding of the word as a whole). Words like Serbian *bezbednost* also illustrate that some words involve (indirect) negation, while others like *sigurnost* and *varnost* do not from the perspective of the modern Slavic languages. External threats, which can appear at the sub-state level, state level, and even at the level of the entire planet, are relevant for words like Russian *bezopasnost*, which implies the absence of danger (an external threat). Ukrainian *bezpeka*, on the other hand, contains a root that is more about internal psychological states, as is evident from words such as *pečal* ‘sorrow’ with the same root.

Second, we have identified a small set of semantic fields relevant for security across Slavic, such as fire, radiation, natural disasters and terrorism, as well as infrastructure (e.g., electricity and cybersecurity). These fields involve what we may call “collective” security, since they are about factors relevant for groups of people (communities, states). However, especially in West Slavic, we also find fields more relevant for individuals, such as ‘stability’, ‘comfort’, ‘health’ and ‘hygiene’.

Third, we have identified a number of metaphorical conceptualizations of ‘security’ in Slavic. ‘Security’ may be construed as a transferable commodity, a scalar property, as well as a set of rules one may obey. Security is furthermore conceptualized as a metaphorical container in many Slavic languages, while some languages, e.g. Slovene, rather see security as a plane, as shown by expressions such as *na varnem* with the preposition *na* ‘on’.

Fourth, our investigation offers a methodological contribution, insofar as we have tested out a number of different methodologies, including semantic vectors, collocations, and compared these results with findings from artificial intelligence, albeit on a limited dataset. We conclude that such methodologies offer a valuable supplement to more traditional analytical tools in corpus linguistics. Our analysis demonstrates that the tools we have employed complement each other; taken together they enable us to both identify cross-linguistic patterns and pinpoint subtle differences among the Slavic languages.

While our primary focus was on a linguistic investigation, these results can be of value also for political scholars as well as politicians, who must take into account possible mismatches in crucial concepts across supposed translation equivalents. As shown in this article, there are more similarities than differences, but subtle divergences may have consequences for communication even among speakers of closely related languages.

References

Deignan, Alice. 2005. *Metaphor and corpus linguistics*. Amsterdam, Philadelphia: John Benjamins.

Deignan, Alice. 2012. Corpus linguistics and metaphor. In Gibbs, Raymond W. (ed.), *The Cambridge handbook of metaphor and thought*, 280–294. Cambridge: Cambridge University Press.

Divjak, Dagmar & Gries, Stefan. 2009. Behavioral profiles. A corpus-based approach to cognitive semantic analysis. In Evans, Vyvyan & Pourcel, Stéphanie (eds.), *New directions in Cognitive Linguistics* (Human Cognitive Processing, vol. 24), 57–75. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/hcp.24.07gri>

Janda, Laura A. & Clancy, Steven J. 2002. *The case book for Russian*. Slavica Publishers.

Janda, Laura A. & Clancy, Steven J. 2006. *The case book for Czech*. Slavica Publishers.

Janda, Laura A. & Solovyev, Valery D. 2009. What constructional profiles reveal about synonymy: A case study of Russian words for sadness and happiness. *Cognitive Linguistics* 20 (2). 367–393.

Johnson, Mark. 1987. *The body in the mind: the bodily basis of meaning, imagination, and reason*. Chicago, London: The University of Chicago Press.

Jurafsky, Daniel & Martin, James H. 2025. *Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition with language models*, 3rd edition. Online manuscript released January 12, 2025. <https://web.stanford.edu/~jurafsky/slp3>.

Hanks, Patrick. 2013. *Lexical analysis: norms and exploitations*. MIT Press.

Hanks, Patrick. 2020. How context determines meaning. In Corpas Pastor, Gloria & Colson, Jean-Pierre (eds.), *Computational Phraseology*, 297–309. Amsterdam, Philadelphia: John Benjamins.

Kiš Žuvela, Sanja & Parizoska, Jelena. 2023. Constructions with lice ('face') in Croatian: lexicoc-grammar and cultural models. In Pattillo, Kelsie & Waśniewska, Małgorzata (eds.), *Embodiment in cross-linguistic studies: the 'face'*, 170–190. Leiden: Brill.

Kövecses, Zoltán. 2010. *Metaphor: a practical introduction* (2nd ed). Oxford: Oxford University Press.

Lakoff, George & Johnson, Mark. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.

Lakoff, George. 1993. The contemporary theory of metaphor. In Ortony, Andrew (ed.), *Metaphor and thought* (2nd ed.), 202–251. Cambridge University Press.

Nesset, Tore & Makarova, Anastasia. 2023. Threatening in Russian with or without *sja*: Grozit' vs. grozit'sja. *Journal of Slavic Linguistics* 31(1-2). 215–244.

Parizoska, Jelena. 2009. The idiom variability in Croatian: the case of the container schema. *Cognitive Studies* 9, 171–180.

Peirsman, Yves & Geeraerts, Dirk. 2006. Metonymy as a prototypical category. *Cognitive Linguistics* 17(3). 269–316.

Pivovarova Lidia & Kormacheva, Daria & Kopotov, Mikhail. 2018. Evaluation of collocation extraction methods for the Russian language. In Kopotov, Mikhail & Lyashevskaya, Olga & Mustajoki Arto (eds.), *Quantitative approaches to the Russian language*, 137–157. Routledge.

Radden, Günther & Zoltán Kövecses. 1999. Towards a theory of metonymy. In Klaus-Uwe Panthe, & Radden, Günther (eds.), *Metonymy in language and thought*, 17–59. John Benjamins. <https://doi.org/10.1075/hcp.4.03rad>

Stanojević, Mateusz-Milan & Tralić, Ivo & Ljubičić, Mateja. 2014. Grammatical information and conceptual metaphors: the case of anger. In Peti-Stantić, Anita & Stanojević, Mateusz-Milan (eds.), *Language as information: proceedings from the CALS conference 2012*, 131–154. Peter Lang.

Stanojević, Mateusz-Milan. 2019. Metaphorical and non-metaphorical dimensions of the term nacija in Croatian online discourse. In Šarić, Liljana & Stanojević, Mateusz-Milan (eds.), *Metaphor, nation and discourse*, 259–286. Amsterdam, Philadelphia: John Benjamins.

Stefanowitsch, Anatol. 2006. Words and their metaphors: A corpus-based approach. In Stefanowitsch, Anatol & Gries, Stefan Th. (eds.), *Corpus-based approaches to metaphor and metonymy*, 63–105. Mouton de Gruyter.

Sullivan, Karen. 2013. *Frames and constructions in metaphoric language*. Amsterdam, Philadelphia: John Benjamins.

Šarić, Ljiljana. 2008. *Spatial concepts in Slavic. A Cognitive Linguistic study of prepositions and cases*. Harrasowitz Verlag.

Šarić, Ljiljana & Stanojević, Mateusz-Milan (eds.). 2019. *Metaphor, nation and discourse*. Amsterdam, Philadelphia: John Benjamins

SIGURNOST U SLAVENSKIM JEZICIMA: LINGVISTIČKI PRISTUP

Znači li riječ za 'sigurnost' isto u svim jezicima? Kako možemo prepoznati međujezične sličnosti i razlike u značenju riječi? U ovom istraživanju provodimo korpusnu analizu riječi koje odgovaraju engleskim pojmovima *security* i *safety* u slavenskim jezicima. Iako naša analiza otkriva brojne međujezične sličnosti, tvrdimo da postoje i razlike. Tri parametra obuhvaćaju relevantne razlike: transparentno nasuprot ne-prozirnom, negativno nasuprot pozitivnom te vanjsko nasuprot unutarnjem. U pogledu semantičkih polja i metaforičke uporabe pronalazimo znatna preklapanja, ali i određene razlike među jezicima koji su predmet istraživanja. Primjenom različitih metoda korpusne lingvistike tvrdimo da nam ovaj „mješoviti“ metodološki pristup omogućuje izbjegavanje slabosti pojedinih metoda i precizno uočavanje i sličnosti i suptilnih razlika među srodnim jezicima.

Ključne riječi: *sigurnost*, slavenski jezici, kolokacije, semantički vektori, Chat-GPT, konstrukcije, korpus

Authors' addresses:

Anastasia Makarova

Uppsala university
Department of Modern languages
Box 636
751 26 Uppsala
SWEDEN
anastasia.makarova@moderna.uu.se

Tore Nesset

UiT – The Arctic University of Norway
Department of Language and Culture
PO Box 6050 Stakkevollan
NO-9037 Tromsø
NORWAY
tore.nesset@uit.no

Laura A. Janda

UiT – The Arctic University of Norway
Department of Language and Culture
PO Box 6050 Stakkevollan
NO-9037 Tromsø
NORWAY
laura.janda@uit.no