

IVANA MATAS IVANKOVIĆ

# IDENTIFYING NEW MULTI-WORD PREPOSITIONS IN CROATIAN: THE *PREPOSITION + NOUN + PREPOSITION* PATTERN

COBISS: 1.01

[HTTPS://DOI.ORG/10.3986/JZ.31.2.04](https://doi.org/10.3986/JZ.31.2.04)

Multi-word prepositions (MWPs) are complex lexical units, typically formed from prepositions and nouns. Because many are still evolving semantically, morphologically, and syntactically, no definitive list of MWPs exists. This article examines three-word expressions with the structure *preposition + noun + preposition*. Using regular expressions, these patterns were applied to a Croatian language corpus to identify and analyze potential MWPs based on features such as frequency of use, variability of elements, semantics, and other relevant properties.

**Keywords:** multi-word prepositions, Croatian, structure, semantic changes, corpus search

## Prepoznavanje novih večbesednih predlogov v hrvaščini: vzorec *predlog + samostalnik + predlog*

Večbesedni predlogi so kompleksne leksikalne enote, ki so običajno tvorjene iz predlogov in samostalnikov. Ker se mnogi še vedno semantično, morfološko in sintaktično razvijajo, dokončen seznam večbesednih predlogov ne obstaja. Ta prispevek preučuje tribesedne izraze s strukturo *predlog + samostalnik + predlog*. Z uporabo regularnih izrazov so bili ti vzorci uporabljeni na hrvaškem jezikovnem korpusu za identifikacijo in analizo potencialnih večbesednih predlogov na podlagi značilnosti, kot so pogostost uporabe, spremenljivost elementov, semantika in druge pomembne lastnosti.

**Ključne besede:** večbesedni predlogi, hrvaščina, struktura, semantične spremembe, iskanje po korpusu

---

Ivana Matas Ivanković ■ Institut za hrvatski jezik, Zagreb, Hrvatska ■ imatas@ihjj.hr ■  <https://orcid.org/0000-0002-9796-8346>

This work was supported by the Croatian Science Foundation under the project *Croatian Prepositions in Use – Semantic and Syntactic Analysis* (HRPA) (IP-2022-10-6867) and by the European Union – NextGenerationEU, within the framework of which the views and opinions expressed are the sole responsibility of the authors and do not necessarily reflect the official views of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them. The author states that the article is based on research data from existing and publicly available sources that are cited in the list of references.



## 1 INTRODUCTION

Prepositions are relational words that connect full-meaning words both logically and syntactically. Prepositions are generally considered a closed group of words in which there are not too many changes. Such a traditional approach is represented even today in some Croatian grammar books and language manuals (e.g. Barić et al. 1997; dictionaries of Croatian...). Research of secondary prepositions, particularly multi-word ones, is still at an early stage and has not yet been examined in sufficient depth.<sup>1</sup>

In the process of becoming prepositions, multi-word units undergo semantic, syntactic, and pragmatic changes that distinguish them from their individual components. For example, *s obzirom na* ‘in view of’ no longer refers directly to *obzir* ‘consideration’, and *sudeći po* ‘judging by’ does not imply the action of *suditi* ‘to judge, to express an opinion’. Compared to primary prepositions, MWPs are often more precise – while components retain some original meaning, the whole gains a new, specialized sense (e.g., *na temelju* lit. ‘on foundation’ → ‘based on’, *preko puta* lit. ‘across the path’ → ‘across from’). Component words may share meaning (like in *u smjeru* ‘in the direction of’) or differ semantically (like in *s ciljem* ‘with the aim’). MWPs can be two- or three-word expressions (e.g., *u početku* ‘at the beginning of’, *za razliku od* ‘in contrast to’) and may appear as single words or spaced phrases (*nakraj / na kraj* ‘at the end of’). Syntactically, complex prepositions function as mediators between two elements rather than attaching to just one (e.g., *doći ← na kraju → ljeta* ‘to come ← at the end of → summer’). Pragmatically, many MWPs are typical in formal or administrative language (e.g., *od strane* lit. ‘from the side of’, instead of *od* ‘from, by’; *po pitanju* lit. ‘by question’, instead of *o* ‘about’).

Because MWPs are part of an ongoing grammaticalization process – shifting from free word combinations to fixed grammatical units – it is difficult to draw a clear boundary or compile a definitive list of MWPs. This is why the aim of this study is to test various characteristics attributed to MWPs in previous research (e.g., Stosic 2020; Zakharov 2020...), based on a search of the Croatian corpus hrWaC (see more in Ljubešić – Klubička 2014), using the Sketch Engine tool.

## 2 METHODOLOGY

This research is focused on the structure *preposition + noun + preposition*. According to Stosic (2020: 192), in Modern Serbian pattern  $P_1 N P_2$  is the second most productive pattern with 60 items found for this pattern.<sup>2</sup>

1 Some of the more recent works on the topic include Šarić 1999; Silić 2004; Bučar 2009; Glušac – Rišner 2016.

2 The most productive is  $P_1 N$  (e.g. *iz razloga* ‘for reasons’) with 268 of items for this pattern.

Since frequency is an important indicator of language change,<sup>3</sup> the search was based on corpus data from hrWaC. Using the regular expression [tag="S.\*"] [tag="N.\*"] [tag="S.\*"], combinations of *preposition + noun + preposition* were extracted and sorted by frequency (with *word lower case* function on). A list of 200 unique combinations was manually reviewed to identify expressions with prepositional features.<sup>4</sup> From this, a list of 33 potential prepositions was compiled (see Table 1), which were then analyzed according to their semantic and syntactic properties.

Table 1: The list of extracted MWPs (sorted according to frequency)<sup>5, 6</sup>

Expression with the structure <i>preposition + noun + preposition</i>	Number of hits in hrWaC
<i>u odnosu na</i> 'with regard to'	153968
<i>s obzirom na</i> 'with regard to'	116326
<i>za razliku od</i> 'unlike'	109597
<i>u skladu s</i> 'in accordance with'	104707
<i>bez obzira na</i> 'regardless of'	101807
<i>u suradnji s</i> 'in cooperation with'	64643
<i>u vezi s</i> 'in relation to'	34718
<i>u roku od</i> 'within'	29240
<i>u iznosu od</i> 'in the amount of'	26278
<i>u usporedbi s</i> 'in comparison with'	26264
<i>na čelu s</i> 'headed by'	26199
<i>u razdoblju od</i> 'in the period of'*	19627
<i>po cijeni od</i> 'at the price of'	16271
<i>u dobi od</i> 'at the age of'	15166
<i>po uzoru na</i> 'following the example'	13706
<i>u trajanju od</i> 'for the duration of'	12204
<i>u rasponu od</i> 'in the range from'*	7814
<i>u svezi s</i> 'in relation to'	7559

3 “[T]he more frequently a form occurs in texts, the more grammatical it is assumed to be.” (Hopper – Traugott 2003: 106)

4 E. g., constructions such as *na konferenciji za* 'on the conference for', *iz dana u* 'from day to', and similar expressions were discarded.

5 Since *word (lower case)* function was on, forms with *s* (*u vezi s*) and *sa* (*u vezi sa*) were expressed separately in the results, but *s/sa* is one preposition whose form depends on the beginning of the following word, so the number of these forms were manually added together and shown as one result. In the rest of the work *s* is taken as neutral form, which refers to *s* and *sa*.

6 Prepositions with *od* 'from' in final position can appear as part of a more complex structure that includes the antonymous preposition *do* 'to' + genitive as the second part, typically indicating a range – that is, the starting and ending point (e.g., *u razdoblju od 8 do 10 godina* 'in the period from 8 to 10 years'). Formally, these paired structures can be used independently, but their meaning becomes incomplete when one of the parts is missing. The prepositions marked with an asterisk occur more frequently as part of *od – do* construction than in their independent usage without *do*.

Expression with the structure <i>preposition + noun + preposition</i>	Number of hits in hrWaC
<i>u vrijednosti od</i> ‘in the value of’	6962
<i>na putu prema</i> ‘on the way to’	6770
<i>u vremenu od</i> ‘in the period from’*	6602
<i>u periodu od</i> ‘in the period from’*	6432
<i>u suprotnosti s</i> ‘in contrast with’	5676
<i>u visini od</i> ‘at a height of, in the amount of’	5493
<i>u odnosu prema</i> ‘in relation to’	4256
<i>u dužini od</i> ‘in the length of’	3892
<i>na površini od</i> ‘on the area of’	2809
<i>na visini od</i> ‘at a height of’	2749
<i>na rok od</i> ‘for the period of’	2524
<i>u razmaku od</i> ‘in the distance of’	2433
<i>na dubini od</i> ‘at a depth of’	2427
<i>na udaljenosti od</i> ‘at a distance of’	2075
<i>u ovisnosti o</i> ‘depending on’	1918

When looking for prepositional phrases from Table 1 in Croatian language dictionaries, only some can be found, and they are not classified as prepositions. If they are included, they usually appear in the phraseology section, and occasionally the case questions are provided, suggesting that the phrase governs a specific case – i.e., that it has a specific syntactic function, but the nature of this function remains unspecified. In *Hrvatski jezični portal* (HJP) and *Veliki rječnik hrvatskoga standardnog jezika* (VRH), seven of candidates from our list are singled out and described. These are: *s obzirom na, za razliku od (čega), bez obzira na (što), u skladu s, u roku od, po uzoru (na koga, što), u svezi (s čim)*. In VRH they come after the sign ◊, which stands for a *lexicalized locution* or after the sign ♦, which stands for an *idiom*. In HJP they are in the section *Frazeologija (Phraseology)*. This shows that Croatian dictionaries are rather unsystematic when it comes to multi-word units that have the grammatical function of prepositions.

The rest of the paper is structured as follows: Section 3 provides a description of the results, while Section 4 analyzes the results in light of various characteristics of MWPs. Concluding remarks are presented in Section 5.

### 3 DESCRIPTION OF THE RESULTS

In this section, the structure of prepositional candidates will be described. They consist of an initial preposition, a noun, and a final preposition. The prepositions serve a more grammatical role, while the noun carries the semantic content.

### 3.1 Prepositions as parts of multi-word prepositions

The initial preposition in the MWP functions to link the construction with the head word that very often precedes it. As can be seen from Table 2, the most common initial preposition is *u* ‘in’, whereas the preposition *na* ‘on’ ranks second, which corresponds to the general distribution of prepositions in Croatian.<sup>7</sup> The second preposition of the MWP is oriented towards the following word and determines its case. The most frequent preposition in the final position is *od* ‘from’, which, among many other meanings, can express comparison and measure – as seen in many of the multi-word prepositions from our list (e.g., *u iznosu od*, *u roku od*, *za razliku od*). *Od* ‘from’ governs the genitive case, which is often described as the case of connection – reflecting the role that *od* plays as the final preposition in MWPs.

Table 2: Initial and final prepositions

Initial preposition	Number of examples	Final preposition	Number of examples
<i>u</i> ‘in’	21	<i>od</i> ‘from’	19
<i>na</i> ‘on’	7	<i>s</i> ‘with’	7
<i>po</i> ‘over’	2	<i>na</i> ‘on’	4
<i>bez</i> ‘without’	1	<i>prema</i> ‘towards’	2
<i>s</i> ‘with’	1	<i>o</i> ‘about’	1
<i>za</i> ‘for’	1		

The most frequent combinations reflect the previously mentioned frequencies of initial and final prepositions. The most common pair is *u – od* (‘in – from’), which appears 12 times in the following MWPs: *u roku od*, *u razmaku od*, *u dužini od*, *u iznosu od*, *u razdoblju od*, *u dobi od*, *u trajanju od*, *u rasponu od*, *u vrijednosti od*, *u vremenu od*, *u periodu od*, *u visini od*.

The pair *u – s* (‘in – with’) occurs six times in the following MWPs: *u skladu s*, *u suradnji s*, *u usporedbi s*, *u vezi s*, *u svezi s*, *u suprotnosti s*.

The pair *na – od* (‘on – from’) is recorded five times in the MWPs: *na površini od*, *na visini od*, *na dubini od*, *na rok od*, *na udaljenosti od*.

All other pairs appear only once.

### 3.2 Nouns as the part of complex prepositions

All nouns that are part of a multi-word preposition appear in the singular. This is expected, as the plural form gives the noun greater concreteness, which is why plural forms are rare in the formation of complex prepositions (see more in Matas Ivanković 2014: 27). The meanings of these nouns are mainly abstract, and they carry the semantics of the entire complex preposition. They typically denote:

<sup>7</sup> According to hrWaC, *u* ‘in’ comes 2,886,353 times, *na* ‘on’ 1,571,531, *za* ‘for’ 1,134,848, *sa* ‘with’ 820,646, *od* ‘from’ 706,592...

- relation among things or people (15 of them): *u odnosu na, s obzirom na, za razliku od, bez obzira na, u skladu s, u suradnji s, u vezi s, u usporedbi s, na čelu s, po uzoru na, u svezi s, u odnosu prema, u suprotnosti s, u ovisnosti o, na putu prema;*
- measure or value (eleven of them): *u iznosu od, po cijeni od, u vrijednosti od, u visini od, u dužini od, na površini od, na visini od, u razmaku od, na dubini od, na udaljenosti od, u rasponu od;*
- time<sup>8</sup> (seven of them): *u roku od, u razdoblju od, u dobi od, u trajanju od, u vremenu od, u periodu od, na rok od.*

## 4 ANALYSIS OF THE RESULTS

This section provides an analysis of the results in light of the thesis that, unlike regular expressions, multi-word prepositions are fixed expressions that allow little variation, and that this particular form is dominant compared to other possible variants. For this reason, variation in different elements of each MWP's structure was tested, based on the general assumption that the better candidate is the one with fewer variations in specific positions.

### 4.1 Variations of the initial preposition

Variations of the initial primary preposition in the MWP were tested, based on the assumption that a better candidate for an MWP is the one with no or the fewest variations (see Stosic 2020: 158) – i.e., the one in which the primary preposition in the initial position clearly dominates over other possible prepositions in that slot. The number of hits for the exact form of the MWP and the number of hits with alternative initial prepositions were compared, and the ratio is expressed as a percentage, ranked from highest to lowest in Table 3.<sup>9</sup>

As Table 3 shows, all MWPs exhibit a very low percentage of variation; that is, they predominantly appear in a fixed form, and the first preposition rarely changes. Out of 33 MWPs, 17 (more than half) have the exact same preposition fixed in over 90% of occurrences, with less than 10% variation. Notably, MWPs *u skladu s, na čelu s, u suprotnosti s, u odnosu na, u svezi s, s obzirom na, u dobi od, za*

<sup>8</sup> They denote measure not in a general quantitative sense, but in a temporal one, and are therefore categorized separately.

<sup>9</sup> To count the number of other prepositions in the initial position, e.g., for the preposition *u odnosu na*, the regular expression [tag="S.\*"&lemma!="u"][[lemma="odnos"]][lemma="na"] was used. Three MWPs from the list share the same noun and final preposition components (*s obzirom na – bez obzira na; u roku od – na rok od; u visini od – na visini od*). Since these are potentially new prepositions and occur frequently enough to appear in the list of the 200 most common items, the number of hits of the corresponding pair (shown in Column A) is subtracted from the total number of hits of all other variants of the searched preposition. The result (shown in Column B) is then used when calculating the proportion.

*razliku od, bez obzira na* appear in their fixed forms in more than 99% of instances, with practically no variation in the initial preposition.

**Table 3: Variations of the initial preposition**

Multi-word preposition	A. Number of hits with exact form	B. Number of hits with variations	C. Proportion of exact form in %
			A/(A+B)
<i>u skladu s</i>	104707	64	99.94%
<i>na čelu s</i>	26199	23	99.91%
<i>u suprotnosti s</i>	5676	6	99.89%
<i>u svezi s</i>	7559	13	99.83%
<i>u odnosu na</i>	153968	315	99.80%
<i>s obzirom na</i>	116326	254	99.78%
<i>u dobi od</i>	15166	61	99.60%
<i>za razliku od</i>	109597	446	99.59%
<i>bez obzira na</i>	101807	698	99.32%
<i>u roku od</i>	29240	428	98.56%
<i>u trajanju od</i>	12204	197	98.41%
<i>u rasponu od</i>	7814	246	96.95%
<i>po uzoru na</i>	13706	487	96.57%
<i>u usporedbi s</i>	26264	1009	96.30%
<i>na putu prema</i>	6770	365	94.88%
<i>u vezi s</i>	34718	2569	93.11%
<i>u vrijednosti od</i>	6962	626	91.75%
<i>u dužini od</i>	3892	530	88.01%
<i>na rok od</i>	2524	426	85.56%
<i>u suradnji s</i>	64643	11675	84.70%
<i>u iznosu od</i>	26278	5249	83.35%
<i>u vremenu od</i>	6602	1668	79.83%
<i>po cijeni od</i>	16271	4329	78.99%
<i>na površini od</i>	2809	821	77.38%
<i>u razmaku od</i>	2433	748	76.49%
<i>u periodu od</i>	6432	2043	75.89%
<i>u razdoblju od</i>	19627	6595	74.85%
<i>u odnosu prema</i>	4256	1541	73.42%
<i>na dubini od</i>	2427	1203	66.86%
<i>u visini od</i>	5493	2916	65.32%
<i>na udaljenosti od</i>	2075	1506	57.94%
<i>na visini od</i>	2749	2223	55.29%
<i>u ovisnosti o</i>	1918	1689	53.17%

## 4.2 Variations of the final preposition

Variations of the final primary preposition were analyzed in the same manner as those of the initial preposition, following the same assumption: the better candidate is the one with no variations or the fewest (see Stosic 2020: 158), that is, the one in which the final primary preposition consistently prevails over alternative prepositions.<sup>10</sup> The number of occurrences featuring the exact preposition was compared with the number of occurrences involving other prepositions. This proportion was calculated and expressed as a percentage, ordered from highest to lowest, as presented in Table 4.

This table shows that a large number of MWPs exhibit very limited variation in the final position. As many as 22 examples appear with the given primary preposition in over 90% of occurrences (with less than 10% variation), and all but two occur in over 70% of cases. *Na putu prema* appears at the bottom of the table, with a high proportion of variations (74.13% – such as *na putu do/u/za/od* ‘on the way to/in/for/from’). This indicates that the expression is not yet fully fixed. Interestingly, *u odnosu na* ranks high, appearing in its exact form in 98.18% of cases, with only 1.82% variation, whereas *u odnosu prema* is near the bottom, with just 61.30% of occurrences in the exact form. This suggests that *u odnosu na* is not only more frequently used, but also a stronger candidate for classification as an MWP.

Table 4: Variations of the final preposition

Multi-word preposition	A. Number of hits with exact form	B. Number of hits with variations	C. Proportion of exact form in %
			A/(A+B)
<i>po uzoru na</i>	13706	16	99.88%
<i>u skladu s</i>	104707	150	99.86%
<i>s obzirom na</i>	116326	347	99.70%
<i>u suradnji s</i>	64643	221	99.66%
<i>za razliku od</i>	109597	592	99.46%
<i>u usporedbi s</i>	26264	170	99.36%
<i>u dužini od</i>	3892	62	98.43%
<i>na čelu s</i>	26199	465	98.26%
<i>u odnosu na</i>	153968	2858	98.18%
<i>u razmaku od</i>	2433	49	98.03%

<sup>10</sup> To count the number of other prepositions in the final position, e.g., for the preposition *u odnosu na*, the regular expression [lemma="u"][[word="odnosu"]][tag="S.\*"&lemma!="na"] was used. *U odnosu na* and *u odnosu prema* share the same noun and the same primary preposition in the initial position. Since both are potentially new prepositions and occur frequently enough to appear in the list of the 200 most common items, the number of hits for *u odnosu prema* (shown in Column A) is subtracted from the total number of other variations of the searched preposition. The resulting number is shown in Column B and used in calculating the proportion.

Multi-word preposition	A. Number of hits with exact form	B. Number of hits with variations	C. Proportion of exact form in %
			A/(A+B)
<i>u suprotnosti s</i>	5676	127	97.81%
<i>u roku od</i>	29240	768	97.44%
<i>u iznosu od</i>	26278	866	96.81%
<i>po cijeni od</i>	16271	638	96.23%
<i>u rasponu od</i>	7814	316	96.11%
<i>u vezi s</i>	34718	1468	95.94%
<i>bez obzira na</i>	101807	4667	95.62%
<i>u svezi s</i>	7559	374	95.29%
<i>na dubini od</i>	2427	161	93.78%
<i>u trajanju od</i>	12204	933	92.90%
<i>u visini od</i>	5493	511	91.49%
<i>u vrijednosti od</i>	6962	676	91.15%
<i>na rok od</i>	2524	288	89.76%
<i>na visini od</i>	2749	327	89.37%
<i>na površini od</i>	2809	373	88.28%
<i>na udaljenosti od</i>	2075	283	88.00%
<i>u dobi od</i>	15166	3058	83.22%
<i>u periodu od</i>	6432	1334	82.82%
<i>u razdoblju od</i>	19627	4151	82.54%
<i>u ovisnosti o</i>	1918	492	79.59%
<i>u vremenu od</i>	6602	2599	71.75%
<i>u odnosu prema</i>	4256	2687	61.30%
<i>na putu prema</i>	6770	19396	25.87%

#### 4.3 Variations of the nominal nucleus in number

As previously mentioned, the singular form is predominantly used in complex prepositions. In all three-word combinations from our list, the nouns appear in the singular. According to Zakharov et al. (2020: 6), “The grammatical number of the noun cannot be changed (it is either singular or plural).” The assumption that the best candidates are those with no or the fewest variations was tested by searching the corpus for instances where the noun appears in the plural instead of the singular for each MWP on the list. These are sorted by the percentage of plural occurrences (with those showing no variation listed first) in Table 5. For some combinations, no plural forms were found at all (*s obzirom na*, *bez obzira na*, *u skladu s*, *na čelu s*, *u dobi od*, *u ovisnosti o*), and in most cases, plural usage accounts for less than 5%. Only *na visinama od* and *na dubinama od* show higher variation, with plural forms occurring in 5.24% and 6.90% of cases, respectively.

Table 5: Variations of the nominal nucleus in number

Multi-word preposition	A. Number of hits with singular	B. Number of hits with plural	C. Proportion of exact form in %
			A/(A+B)
<i>s obzirom na</i>	116326	0	100%
<i>bez obzira na</i>	101807	0	100%
<i>u skladu s</i>	104707	0	100%
<i>na čelu s</i>	26199	0	100%
<i>u dobi od</i>	15166	0	100%
<i>u ovisnosti o</i>	1918	0	100%
<i>u svezi s</i>	7559	1	99.99%
<i>u suradnji s</i>	64643	14	99.98%
<i>u odnosu na</i>	153968	34	99.98%
<i>za razliku od</i>	109597	30	99.97%
<i>u trajanju od</i>	12204	5	99.96%
<i>po uzoru na</i>	13706	7	99.95%
<i>u usporedbi s</i>	26264	15	99.94%
<i>u vremenu od</i>	6602	4	99.94%
<i>u roku od</i>	29240	18	99.94%
<i>u vrijednosti od</i>	6962	5	99.93%
<i>na putu prema</i>	6770	8	99.88%
<i>u visini od</i>	5493	12	99.78%
<i>u razdoblju od</i>	19627	50	99.75%
<i>u rasponu od</i>	7814	29	99.63%
<i>u suprotnosti s</i>	5676	23	99.60%
<i>u periodu od</i>	6432	29	99.55%
<i>u vezi s</i>	34718	166	99.52%
<i>u dužini od</i>	3892	27	99.31%
<i>u razmaku od</i>	2433	23	99.06%
<i>na površini od</i>	2809	29	98.98%
<i>u iznosu od</i>	26278	285	98.93%
<i>na rok od</i>	2524	33	98.71%
<i>po cijeni od</i>	16271	551	96.72%
<i>na udaljenosti od</i>	2075	75	96.51%
<i>u odnosu prema</i>	4256	167	96.22%
<i>na visini od</i>	2749	152	94.76%
<i>na dubini od</i>	2427	180	93.10%

#### 4.4 Modification of the nominal nucleus

In its regular use, a noun can be modified by adjectives or pronouns, as in the example: *Zato bi to trebalo sagledati u relativnom odnosu na BDP* ‘That is why it should be seen **in relative relation to** GDP...’ In such cases, the modification typically clarifies the noun in its original, literal meaning. However, complex prepositions, as fixed expressions, should not permit external elements such as noun modifiers or adverbials. As noted by Zakharov et al. (2020: 6), “[i]n most cases, it does not allow for insertion or separation (as a rule, the noun cannot have a possessive or adjectival determiner).” Based on this, a better prepositional candidate is one that is predominantly used without any modification. This was tested by searching the corpus for examples containing an empty slot before the noun in the MWP.<sup>11</sup> The results are presented in Table 6, with the examples showing the fewest modifications listed at the top.

Twenty-five MWPs have less than 10% of instances with modified nouns, and eight of them show fewer than 1%. These are *na čelu s*, *za razliku od*, *bez obzira na*, *po uzoru na*, *u odnosu na*, *s obzirom na*, *u usporedbi s*, *u skladu s*. At the bottom of the list are *na visini od*, with 25.22% of instances containing inserted modifications (mostly *nadmorskoj*), and *na udaljenosti od*, with 29.61% of modified instances. Common modifiers in the latter include *na sigurnoj udaljenosti od* (‘at a safe distance’), *pristojnoj* (‘decent’), *dovoljnoj* (‘sufficient’), and *istoj* (‘the same’). This marks a significant contrast compared to the other candidates.

Table 6: Modification of the nominal nucleus

Multi-word preposition	A. Number of exact hits	B. Number of hits with modified noun	C. Proportion of exact form in %
			A/(A+B)
<i>na čelu s</i>	26199	4	99.98%
<i>za razliku od</i>	109597	43	99.96%
<i>bez obzira na</i>	101807	59	99.94%
<i>po uzoru na</i>	13706	8	99.94%
<i>u odnosu na</i>	153968	190	99.88%
<i>s obzirom na</i>	116326	235	99.80%
<i>u usporedbi s</i>	26264	55	99.79%
<i>u skladu s</i>	104707	473	99.55%
<i>u visini od</i>	5493	59	98.94%
<i>u suradnji s</i>	64643	981	98.51%
<i>u dobi od</i>	15166	308	98.01%
<i>u svezi s</i>	7559	154	98.00%
<i>na dubini od</i>	2427	64	97.43%

<sup>11</sup> For this search, e.g., for the preposition *u odnosu na*, the regular expression [lemma="u"][] [word="odnosu"] [lemma="na"] was used.

Multi-word preposition	A. Number of exact hits	B. Number of hits with modified noun	C. Proportion of exact form in %
			A/(A+B)
<i>u roku od</i>	29240	930	96.92%
<i>u trajanju od</i>	12204	452	96.43%
<i>u razmaku od</i>	2433	93	96.32%
<i>u vremenu od</i>	6602	324	95.32%
<i>u razdoblju od</i>	19627	1201	94.23%
<i>u rasponu od</i>	7814	576	93.13%
<i>na rok od</i>	2524	187	93.10%
<i>u dužini od</i>	3892	333	92.12%
<i>u vrijednosti od</i>	6962	626	91.75%
<i>u ovisnosti o</i>	1918	183	91.29%
<i>u periodu od</i>	6432	615	91.27%
<i>u vezi s</i>	34718	3358	91.18%
<i>u iznosu od</i>	26278	2924	89.99%
<i>na površini od</i>	2809	359	88.67%
<i>na putu prema</i>	6770	902	88.24%
<i>u suprotnosti s</i>	5676	943	85.75%
<i>u odnosu prema</i>	4256	801	84.16%
<i>po cijeni od</i>	16271	3365	82.86%
<i>na visini od</i>	2749	927	74.78%
<i>na udaljenosti od</i>	2075	873	70.39%

#### 4.5 Insertion of an adverbial modifier

The property tested in this section is modification by an adverbial modifier, which appears between the noun and the final preposition (*U odnosu samo na tržište BiH mogu dosegnuti oko 100 milijuna eura. ‘In relation to the BiH market alone, they can reach around 100 million euros.’*). Such modifications can even be separated by commas (*Dakle, riječ je o igri koju ljudska psiha igra sa samom sobom u odnosu, naravno, na kompletni svijet. ‘So, it is a game that the human psyche plays with itself in relation, of course, to the entire world.’*).

The possibility of adverbial modification was tested by searching the corpus for examples with an empty slot before the final primary preposition.<sup>12</sup> All candidates show only a few instances of this type of insertion and are listed in Table 7, with the examples showing the fewest modifications listed at the top. Based on this feature, they are all strong candidates, as such modifications are very rare.

<sup>12</sup> For this search, e.g., for the preposition *u odnosu na*, the regular expression [lemma="u"] [word="odnosu"] [word=","] ? [tag="R.\*"] [word=","] ? [lemma="na"] was used.

Table 7: Adverbial modification

Multi-word preposition	A. Number of exact hits	B. Number of hits with adverbial modification	C. Proportion of exact form in % A/(A+B)
<i>s obzirom na</i>	116326	12	99.99%
<i>u skladu s</i>	104707	19	99.98%
<i>u odnosu na</i>	153968	32	99.98%
<i>u suradnji s</i>	64643	16	99.98%
<i>u usporedbi s</i>	26264	9	99.97%
<i>bez obzira na</i>	101807	40	99.96%
<i>u periodu od</i>	6432	3	99.95%
<i>po uzoru na</i>	13706	8	99.94%
<i>u razdoblju od</i>	19627	12	99.94%
<i>za razliku od</i>	109597	73	99.93%
<i>u roku od</i>	29240	20	99.93%
<i>na površini od</i>	2809	2	99.93%
<i>u dobi od</i>	15166	11	99.93%
<i>u vremenu od</i>	6602	5	99.92%
<i>u vezi s</i>	34718	28	99.92%
<i>na čelu s</i>	26199	22	99.92%
<i>u suprotnosti s</i>	5676	5	99.91%
<i>u rasponu od</i>	7814	7	99.91%
<i>na putu prema</i>	6770	7	99.90%
<i>u iznosu od</i>	26278	31	99.88%
<i>na dubini od</i>	2427	3	99.88%
<i>u svezi s</i>	7559	9	99.88%
<i>na rok od</i>	2524	4	99.84%
<i>u visini od</i>	5493	9	99.84%
<i>u odnosu prema</i>	4256	7	99.84%
<i>na visini od</i>	2749	5	99.82%
<i>u dužini od</i>	3892	9	99.77%
<i>u razmaku od</i>	2433	6	99.75%
<i>u trajanju od</i>	12204	44	99.64%
<i>u ovisnosti o</i>	1918	7	99.64%
<i>na udaljenosti od</i>	2075	10	99.52%
<i>u vrijednosti od</i>	6962	63	99.10%
<i>po cijeni od</i>	16271	212	98.71%

### Interim overview of variations

The average values of all percentages in the tables were calculated and are presented in Table 8, ordered from highest to lowest. For easier comparison between the individual tables and the average values, and to highlight the best and worst candidates in the different tables, percentages ranked first to fifth in each table are shaded dark gray, sixth to tenth are shaded light gray, and the last five are bolded and underlined.

The table shows that *u skladu s*, *s obzirom na*, *za razliku od*, *na čelu s*, *u odnosu na*, *po uzoru na*, *u usporedbi s*, *bez obzira na* are both frequent and highly consistent in use, ranking near the top across most categories.

The weakest results appear in initial preposition variation, with the lowest scores being 53.17% for *u ovisnosti o* and 55.29% for *na visini od*. This also affects *u suradnji s*, which is ranked 14th overall despite placing within the top five or ten in frequency and four other categories, making it a strong example of an MWP.

Certain candidates consistently perform lower. The last five are *u ovisnosti o*, *u odnosu prema*, *na visini od*, *na udaljenosti od*, *na putu prema*. Comparing the candidates that appear most often in the bottom five positions, *na udaljenosti od* occurs four times, while *u ovisnosti o*, *po cijeni od*, *na visini od*, and *u odnosu prema* each occur three times. These frequent low rankings suggest that such expressions may be less grammaticalized or more syntactically flexible, making them weaker candidates for classification as fixed complex prepositions.

Table 8: All values and average value

	Frequency	Variations of initial preposition	Variations of final preposition	Variations of nominal nucleus in number	Modification of nominal nucleus	Insertion of adverbial modifier	Average value
<i>u skladu s</i>	104707	99.94%	99.86%	100%	99.55%	99.98%	99,87%
<i>s obzirom na</i>	116326	99.78%	99.70%	100%	99.80%	99.99%	99,85%
<i>za razliku od</i>	109597	99.59%	99.46%	99.97%	99.96%	99.93%	99,78%
<i>na čelu s</i>	26199	99.91%	98.26%	100%	99.98%	99.92%	99,61%
<i>u odnosu na</i>	153968	99.80%	98.18%	99.98%	99.88%	99.98%	99,56%
<i>po uzoru na</i>	13706	96.57%	99.88%	99.95%	99.94%	99.94%	99,26%
<i>u usporedbi s</i>	26264	96.30%	99.36%	99.94%	99.79%	99.97%	99,07%
<i>bez obzira na</i>	101807	99.32%	95.62%	100%	99.94%	99.96%	98,97%
<i>u svezi s</i>	7559	99.83%	95.29%	99.99%	98.00%	99.88%	98,60%
<i>u roku od</i>	29240	98.56%	97.44%	99.94%	96.92%	99.93%	98,56%
<i>u trajanju od</i>	12204	98.41%	92.90%	99.96%	96.43%	<b><u>99,64%</u></b>	97,47%
<i>u rasponu od</i>	7814	96.95%	96.11%	99.63%	93.13%	99.91%	97,15%
<i>u suprotnosti s</i>	5676	<b><u>99.89%</u></b>	97.81%	99.60%	<b><u>85,75%</u></b>	99.91%	96,59%
<i>u suradnji s</i>	64643	84.70%	99.66%	99.98%	98.51%	99.98%	96,57%

	Frequency	Variations of initial preposition	Variations of final preposition	Variations of nominal nucleus in number	Modification of nominal nucleus	Insertion of adverbial modifier	Average value
<i>u dobi od</i>	15166	99.60%	83.22%	100%	98.01%	99.93%	96,15%
<i>u vezi s</i>	34718	93.11%	95.94%	99.52%	91.18%	99.92%	95,93%
<i>u dužini od</i>	3892	88.01%	98.43%	99.31%	92.12%	99.77%	95,53%
<i>u vrijednosti od</i>	6962	91.75%	91.15%	99.93%	91.75%	<b>99.10%</b>	94,74%
<i>u razmaku od</i>	2433	76.49%	98.03%	99.06%	96.32%	99.75%	93,93%
<i>u iznosu od</i>	26278	83.35%	96.81%	98.93%	89.99%	99.88%	93,79%
<i>na rok od</i>	2524	85.56%	89.76%	98.71%	93.10%	99.84%	93,39%
<i>u visini od</i>	5493	<b>65.32%</b>	91.49%	99.78%	98.94%	99.84%	91,07%
<i>po cijeni od</i>	16271	78.99%	96.23%	<b>96.72%</b>	<b>82.86%</b>	<b>98.71%</b>	90,70%
<i>na površini od</i>	2809	77.38%	88.28%	98.98%	88.67%	99.93%	90,65%
<i>u razdoblju od</i>	19627	74.85%	<b>82.54%</b>	99.75%	94.23%	99.94%	90,26%
<i>na dubini od</i>	2427	<b>66.86%</b>	93.78%	<b>93.10%</b>	97.43%	99.88%	90,21%
<i>u periodu od</i>	6432	75.89%	82.82%	99.55%	91.27%	99.95%	89,90%
<i>u vremenu od</i>	6602	79.83%	<b>71.75%</b>	99.94%	95.32%	99.92%	89,35%
<i>u ovisnosti o</i>	1918	<b>53.17%</b>	<b>79.59%</b>	100%	91.29%	<b>99.64%</b>	84,74%
<i>u odnosu prema</i>	4256	73.42%	<b>61.30%</b>	<b>96.22%</b>	<b>84.16%</b>	99.84%	82,99%
<i>na visini od</i>	2749	<b>55.29%</b>	89.37%	<b>94.76%</b>	<b>74.78%</b>	99.82%	82,80%
<i>na udaljenosti od</i>	2075	<b>57.94%</b>	88.00%	<b>96.51%</b>	<b>70.39%</b>	<b>99.52%</b>	82,47%
<i>na putu prema</i>	6770	94.88%	<b>25.87%</b>	99.88%	88.24%	99.90%	81,75%

#### 4.6 Semantics

Among the features of MWPs, Zakharov et al. (2020:6) note that an MWP inherits the semantics of the notion word (noun or verb). Its nominal components tend to have abstract semantics. Moreover, it is idiomatized – that is, its nominal component loses its lexical meaning to some extent, which is why MWPs are sometimes referred to as “prepositional idioms.”

“From very early times researchers on issues related to grammaticalization have observed that it involves loss of semantic content. This has been described by the metaphor of ‘fading’ or ‘bleaching’” (Hopper – Traugott 2003: 94). In the process of grammaticalization, MWPs begin to lose the categorical meaning typical of their original word class – for example, nouns tend to lose their **objectivity (objecthood)** and take on more abstract, relational functions. The lexical meaning of the noun is not entirely lost but becomes less prominent. As a result, these newly formed prepositions tend to have more specific meanings than primary prepositions, which typically cover a wide range of meanings. This may explain the emergence of more semantically precise constructions functioning as prepositions. Grammaticalization often affects words with abstract semantics, especially those expressing relational meaning. This is confirmed by the semantic profiles of

the MWPs in the list: fourteen candidates from the list express abstract relations between people or things, while the remaining ones indicate measures or values (e.g., *u iznosu od*, *po cijeni od*, *u vrijednosti od*), often in a temporal context (e.g., *u roku od*, *u razdoblju od*, *u dobi od*), which also reflects abstract meaning.

Stosic (2020: 164) also mentions context extension, noting that “it corresponds to the possibility of combining a given expression with NPs that could not be selected in their referential usages.” Since the nouns forming MWPs in our list express relational meanings or measures – both of which are abstract – there are no obvious semantic shifts like those found in the examples discussed by Stosic: *obilazak mnogih atraktivnih lokaliteta u srcu kanjona Tare* ‘a visit of many attractive locations in the heart of Tara canyon’. The only preposition in which such semantic shift can be observed is *na čelu s*. The noun *čelo* ‘forehead’ refers to a part of the human body, and, in its literal sense, it does not usually appear in sentences like *Odatle su ustanci na čelu s Kvaternikom 9. listopada krenuli u Plaški*. ‘From there, the rebels, led by Kvaternik, set off for Plaški on October 9th.’ This suggests a degree of semantic bleaching or idiomatization.

#### 4.7 Prepositional synonymy

The possibility of replacing an MWP with another primary or secondary preposition is often taken as evidence of its prepositional function (cf. Bučar 2009: 199; Mardale 2011: 67; Stosic 2020: 162; Zakharov et al. 2020). The list below shows that almost all of the MWPs can be replaced by one or more prepositions, supporting the interpretation of these expressions as prepositions.

<i>u odnosu na &gt; prema</i>	<i>u vrijednosti od &gt; od, za</i>
<i>s obzirom na &gt; prema</i>	<i>na putu prema &gt; prema</i>
<i>za razliku od &gt; nasuprot</i>	<i>u vremenu od &gt; za, tijekom, od – do</i>
<i>bez obzira na &gt; neovisno o</i>	<i>u periodu od &gt; za, tijekom, od – do</i>
<i>u skladu s &gt; prema</i>	<i>u svezi s &gt; povodom, o, glede</i>
<i>u suradnji s &gt; s</i>	<i>u visini od &gt; od</i>
<i>u roku od &gt; za</i>	<i>u odnosu prema &gt; prema</i>
<i>u vezi s &gt; povodom, o, glede</i>	<i>u dužini od &gt; od</i>
<i>u iznosu od &gt; od</i>	<i>u suprotnosti s &gt; protiv, nasuprot</i>
<i>u razdoblju od &gt; od</i>	<i>na površini od &gt; na</i>
<i>u usporedbi s &gt; prema</i>	<i>na visini od &gt; na</i>
<i>na čelu s &gt; s?</i>	<i>na rok od &gt; na</i>
<i>po cijeni od &gt; po</i>	<i>u razmaku od &gt; za, u?</i>
<i>u dobi od &gt; od, s</i>	<i>na dubini od &gt; na</i>
<i>po uzoru na &gt; prema</i>	<i>na udaljenosti od &gt; na</i>
<i>u trajanju od &gt; od</i>	<i>u ovisnosti o &gt; prema</i>
<i>u rasponu od &gt; od – do</i>	

Some of these constructions are even fully or partially synonymous with each other, such as *u odnosu na*, *s obzirom na*, *u skladu s*, *u usporedbi s*, *u odnosu prema*, *u ovisnosti o*, *u vezi s*, and *u svezi s*, which express ‘in relation to’; *u roku od*, *u razdoblju od*, *u trajanju od*, *u vremenu od*, and *u periodu od*, which express time; *u iznosu od*, *u vrijednosti od*, and *u visini od*, which indicate amount or value; and *za razliku od* and *u suprotnosti s*, which express contrast. *Na površini od*, *na visini od*, *na dubini od*, and *na udaljenosti od* are not fully interchangeable, but they share the common idea of location measured from a point, with each specifying a different spatial dimension.

#### 4.8 Coordination

Coordination is possible only between free constituents that are functionally identical,<sup>13</sup> which makes the coordination of MWPs and simple prepositions a useful test for assessing their categorical status. In this case, we looked for examples such as *U slučaju spornih pitanja koja proizađu iz ili u svezi s Ugovorom...* ‘In the event of any disputes arising out of or in connection with the Agreement...’, with MWP before or after the coordinator.

The corpus was searched for examples of coordination,<sup>14</sup> and for 14 out of the 33 MWPs from the list, instances of coordination were found. The most frequent examples involved *u vezi s* (21 hits) and *u svezi s* (19 hits), followed by *u skladu s* (nine hits), *u suradnji s* (eight hits), *u odnosu na* and *s obzirom na* (four hits each), *na putu prema* (three hits), and *bez obzira na* and *u odnosu prema* (two hits each). *U roku od*, *u dobi od*, *po uzoru na*, *na dubini od* and *na udaljenosti od* had only one instance of coordination.

Since prepositions are relational and grammatical in nature, coordination of prepositions is not very common. When it does occur, it typically involves the omission of the nominal constituent. This explains the relatively low number of coordinated examples. Nonetheless, coordination can serve as an additional indicator of prepositional status.

### 5 CONCLUSION

The aim of this study was to test multi-word expressions consisting of the structure *preposition + noun + preposition*, retrieved from hrWaC, to determine

<sup>13</sup> “A construction [A B] is considered coordinate if the two parts A and B have the same status (in some sense that needs to be specified further), whereas it is not coordinate if it is asymmetrical and one of the parts is clearly more salient or important, while the other part is in some sense subordinate.” (Haspelmath 2004 :3).

<sup>14</sup> For this search, e.g. for the preposition *u odnosu na*, the regular expression [lemma="u"] [word="odnosu"][[lemma="na"][[lemma="ili"]][[tag="S.\*"]][[tag="S.\*"]][[lemma="ili"]][[lemma="u"]][word="odnosu"]][lemma="na"] was used.

whether they exhibit features of grammaticalization and can be treated as multi-word prepositions, especially given that such expressions are often underrepresented or not officially recognized in Croatian dictionaries. The underlying assumption was that grammaticalized expressions, in addition to certain semantic, syntactic, and grammatical specificities, would exhibit minimal variation in their internal structure and would predominantly occur in fixed form. This assumption was tested through several corpus-based analyses, which revealed the following: (i) High-frequency phrases (e.g., *u skladu s, s obzirom na, za razliku od, na čelu s, u odnosu na, po uzoru na, u usporedbi s, bez obzira na* including *u suradnji s*) appear at the top of the table, with high average values across all parameters. They show near-total stability across most parameters. These expressions are well-established and largely fixed in usage, operating in Croatian as multi-word prepositions. (ii) Other prepositions from the list still exhibit very high values but are slightly more variable in some categories. This supports the idea of gradual grammaticalization: complex prepositions emerge from frequent multi-word expressions that, over time, lose internal flexibility. (iii) The lowest-ranked prepositions on the list (e.g., *u ovisnosti o, u odnosu prema, na visini od, na udaljenosti od, na putu prema* including *po cijeni od*) display lower average values, sometimes dropping below 85%. They exhibit greater syntactic flexibility, particularly in initial or final preposition variation and in modification of the nominal nucleus. MWPs typically inherit meaning from their notional noun, but in grammaticalization this noun undergoes semantic bleaching and the expression takes on a relational or idiomatic role. Their prepositional status is supported by substitutability with primary/secondary prepositions and, in some cases, by coordination.

## REFERENCES

**Barić et al. 1997** = Eugenija Barić – Mijo Lončarić – Dragica Malić – Slavko Pavešić – Mirko Peti – Vesna Zečević – Marija Znika, *Hrvatska gramatika*, Zagreb: Školska knjiga, 1997.

**Bučar 2009** = Mirjana Bučar, Gramatikalizacija i sekundarni prijedlozi glagolskoga i imeničkoga porijekla, *Jezikoslovje* 10.2 (2009), 183–216.

**Glušac – Rišner 2016** = Maja Glušac – Vlasta Rišner, O uporabi i značenju sekundarnih prijedloga nastalih gramatikalizacijom imenica, *Rasprave: časopis Instituta za hrvatski jezik i jezikoslovje* 42.2 (2016), 409–442.

**Haspelmath 2007** = Martin Haspelmath, Coordination, in: *Language Typology and Syntactic Description*, ed. Timothy Shopen, Cambridge: Cambridge University Press, 2007, 1–51, DOI: <https://doi.org/10.1017/CBO9780511619434.001>

**HJP** = *Hrvatski jezični portal*, <https://hjp.znanje.hr/>.

**Hopper – Traugott 2003** = Paul J. Hopper – Elizabeth Closs Traugott, *Grammaticalization*, Cambridge: Cambridge University Press, 2003, DOI: <https://doi.org/10.1017/CBO9781139165525>.

**Ljubešić – Klubička 2014** = Nikola Ljubešić – Filip Klubička, {bs,hr,sr}WaC – Web corpora of Bosnian, Croatian and Serbian, in: *Proceedings of the 9th Web as Corpus Workshop (WaC-9)*, eds. Felix Bildhauer – Roland Schäfer, Gothenburg: Association for Computational Linguistics, 2014, 29–35, DOI: <https://doi.org/10.3115/v1/W14-0405>.

**Mardale 2011** = Alexandru Mardale, Prepositions as a semilexical category, *Bucharest Working Papers in Linguistics* 13.2 (2011), 57–72, <https://shs.hal.science/halshs-00663128v1>.

**Matas Ivanković 2014** = Ivana Matas Ivanković, *Izražavanje prostora i vremena prijedlozima s genitivom u hrvatskom i ruskom jeziku*, Zagreb: Institut za hrvatski jezik i jezikoslovje, 2014.

**Matas Ivanković 2018** = Ivana Matas Ivanković, Prijedlozi kao rječničke natuknica, in: *Od dvojbe do razdvojbe. Zbornik radova u čast profesorici Branki Tafri*, eds. Petra Košutar – Mislav Kovačić, Zagreb: Ibis grafika, 2018, 113–127.

**Silić 2004** = Josip Silić, Pretvorba imenskih vrsta riječi u neimenske, in: *Zbornik Zagrebačke slavističke škole 2003.*, ed. Stipe Botica, Zagreb: FF press, 2004, 43–48.

**Stosic 2020** = Dejan Stosic, Complex adpositions in Slavic languages: the case of Serbian, in: *Complex Adpositions in European Languages*, eds. Benjamin Fagard – Jose Pinto de Lima – Dejan Stosic – Elena Smirnova, Berlin – Boston: Walter de Gruyter GmbH, 2020, 139–194, DOI: <https://doi.org/10.1515/9783110686647-004>.

**Šarić 1999** = Ljiljana Šarić, Gramatikalizacija i sekundarni prijedlozi (*pomoću, s pomoću, uz pomoć*), *Riječ* 5.1 (1999), 81–95.

**VRH** = Ljiljana Jojić – Anuška Nakić – Nada Vajs Vinja – Vesna Zečević, *Veliki rječnik hrvatskoga standardnog jezika*, Zagreb: Školska knjiga, 2015.

**Zakharov et al. 2020** = Victor Zakharov – Anastasia Golovina – Elena Alexeeva – Vadim Gudkov, Russian Secondary Prepositions: Methodology of Analysis, in: *Proceedings of the Computational Models in Language and Speech Workshop (CMLS 2020) co-located with 16th International Conference on Computational and Cognitive Linguistics (TEL 2020)*, eds. Alexander Elizarov – Natalia Loukachevitch, Saint Petersburg: St. Petersburg University, 2020, 187–201, <https://ceur-ws.org/Vol-2780/>.

## POVZETEK

### Prepoznavanje novih večbesednih predlogov v hrvaščini: vzorec *predlog + samostalnik + predlog*

V prispevku so bili preizkušeni izrazi v strukturi *predlog + samostalnik + predlog*, pridobljeni iz korpusnih podatkov, z namenom ugotoviti, ali izkazujejo značilnosti gramatikalizacije in jih je mogoče obravnavati kot večbesedne predloge. Izhodiščna predpostavka je bila, da se gramatikalizirani izrazi pojavljajo v fiksnih oblikah z minimalnimi notranjimi variacijami ter izkazujejo specifične semantične in sintaktične lastnosti. Za preverjanje teh predpostavk je bilo izvedenih več analiz na podlagi korpusnih podatkov. Analiza potrebuje, da večina naštetih večbesednih predlogov v hrvaščini izkazuje značilnosti gramatikalizacije, zlasti zaradi svoje strukturne fiksnosti, omejene variacije in sintaktičnega vedenja. Večina izrazov kaže minimalne ali nikakršne variacije v začetnih in končnih predlogih, daje prednost edninski samostalniški obliki ter se upira modifikaciji s pridevnškim ali prislovnim vrvanjem. Dodatne lastnosti, kot sta semantično beljenje in zamenljivost s primarnimi predlogi, dodatno potrjujejo njihovo predložno funkcijo, medtem ko koordinacijski vzorci nudijo dopolnilne skladenjske dokaze. Čeprav nekateri izrazi po določenih merilih, uporabljenih v tej raziskavi, kažejo šibkeje rezultate, večina enot z začetnega seznama izkazuje jasne značilnosti gramatikalizacije in jih je mogoče šteti za večbesedne predloge.