

**Goranka Blagus Bartolec**

## **KOLOKACIJSKA BAZA HRVATSKOGA JEZIKA**

### **1. Uvod**

Unutar rječničkoga članka u *Mrežniku* poveznicom se može pristupiti *Kolokacijskoj bazi hrvatskoga jezika*, mrežnomu izvoru Instituta za hrvatski jezik i jezikoslovlje. *Kolokacijska baza hrvatskoga jezika* javno je dostupna na poveznici [ihjj.hr/kolokacije](http://ihjj.hr/kolokacije). Povezivanjem na *Kolokacijsku bazu* korisnici *Mrežnika* mogu dobiti dodatne informacije o kolokacijskome potencijalu natuknice koja se nalazi u *Mrežniku*.

### **2. O Kolokacijskoj bazi**

*Kolokacijska baza hrvatskoga jezika* oblikovana je kao popis sveza riječi hrvatskoga jezika koje se unose i obrađuju u relacijskoj bazi podataka. Nastala je kao rezultat projekta u okviru osnovne djelatnosti Instituta za hrvatski jezik i jezikoslovlje (<http://ihjj.hr/projekti/projekti-u-okviru-osnovne-djelatnosti/3/>). Baza je dinamična te se može prema potrebi dopunjavati novim svezama i natuknicama, a trenutačno sadržava približno 32 000 jednorječnih natuknica te približno 228 000 sveza. U sklopu projekta izrađena je *Baza frazema hrvatskoga jezika* koja je javno dostupna na mrežnoj adresi [frazemi.ihjj.hr](http://frazemi.ihjj.hr).

Pojedine riječi u hrvatskome jeziku imaju bogat kolokacijski potencijal, dok neke riječi, iako se također učestalo upotrebljavaju u svakodnevnoj komunikaciji, nemaju taj potencijal. Cilj je *Kolokacijske baze* obuhvatiti različite sveze kojima se potvrđuje povezivost pojedinih riječi s drugima s prikazom njihovih glavnih gramatičkih i značenjskih obilježja te pokazati mogućnosti njihove uporabe koje se, u skladu s postojećim teorijskim opisom višerječnih sveza, mogu odrediti kao čvrste ili slobodne sveze. Opširnije o koncepciji *Baze* dostupno je na adresi [ihjj.hr/kolokacije/o-bazi](http://ihjj.hr/kolokacije/o-bazi).

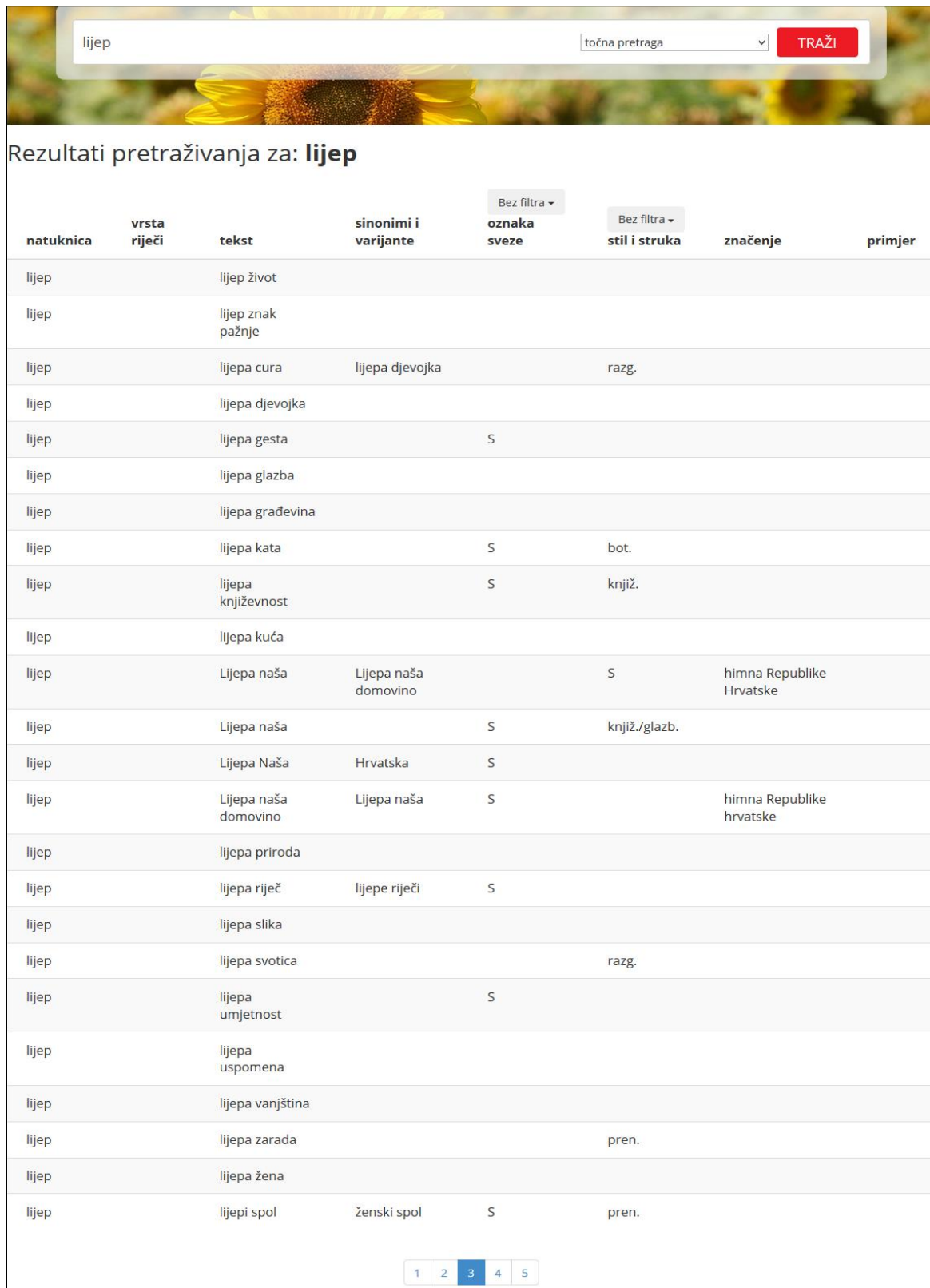
Sveze uvrštene u *Kolokacijsku bazu* temelje se na početnoj građi koju je prikupio vanjski suradnik na projektu dr. Stefan Rittgasser iz Njemačke. Tijekom rada na projektu temeljna građa proširena je i dopunjena novim svezama prikupljenima iz novijih izvora koje pripadaju različitim kontekstima svakodnevne uporabe i svim stilovima suvremenoga hrvatskoga jezika. Popis tiskanih i mrežnih izvora i korpusa naveden je na mrežnoj stranici *Baze* u poglavlju *O bazi* na adresi [ihjj.hr/kolokacije/o-bazi](http://ihjj.hr/kolokacije/o-bazi).

U *Kolokacijskoj bazi* popis sveza ne donosi se samo pod jednom značenjski dominantnom sastavnicom, nego pod svim sastavnicama (samoznačnim i suznačnim) koje tvore pojedinu svezu (npr. kolokacija *cjedilo za čaj* navodi se pod natuknicama *cjedilo*, *čaj* i *za*). Takva se struktura *Baze* temelji na postavci da svaka sastavnica svojim značenjskim potencijalom oblikuje cjelokupni značenjski potencijal sveze (Ivir 1992./1993.: 183) te nije primijenjen hijerarhijski kriterij navođenja sveze samo pod određenom natuknicom kao glavnom ili nosivom.

Svaka je sveza u *Bazi* prema svojim leksičkim i semantičkim obilježjima označena posebnom oznakom te korisnik dobiva podatak o vrsti leksičke sveze i, na temelju unesenih primjera, o kontekstu njezine najčešće primjene. Kriteriji za određivanje vrste sveza temelje se na podjeli leksičkih i gramatičkih sveza u postojećoj literaturi (Ivir 1992./1993., Benson i dr. 1997. Meljčuk 1998., Sag i dr. 2002., Baldwin i Kim 2009., Blagus Bartolec 2014.). *Baza* sadržava iscrpan popis frazema (*daleko /komu/ lijepa kuća*), višerječnih naziva (*lijepa kata*, *lijepa književnost*), kolokacijskih sveza (*lijepa riječ*, *lijepi/ljepši spol*), slobodnih sveza (*lijepa djevojka*), pragmema (*hvala lijepa*), poslovice (*Tko će ljepšega kruha od pšenice, željet će i ovsenice.*), prijedložno-padežnih izraza (*niz dlaku*, *bez problema*), gramatičkih sveza kao što su različiti veznički, priložni i prijedložni skupovi (*s obzirom na to da*, *preko puta*, *s pomoću*, *bilo kako*, *budući da*), koligacija, tj. ustaljenih sveza na sintaktičkoj razini (*može se reći da*, *kad je u pitanju*, *nameće se zaključak da*, *ono što najprije upadne u oči jest činjenica da*, *kako sada stvari stoje*) te drugih sveza koje su ustaljene u komunikacijskoj praksi (razgovornoj, službenoj, administrativnoj, publicističkoj, književnoj) i koje su primjeri različitih mogućnosti sintaktičke i stilske upotrebe, pa su zanimljive i na razini prevođenja ili u poučavanju hrvatskoga kao stranoga jezika.

U literaturi se ustaljeno se kao prototipni oblik višerječne sveze najčešće navodi struktura *pridjev + imenica* (*civilno društvo*, *dnevni boravak*, *osnovna škola*), *imenica + prijedložno-padežni izraz* (*lak za nokte*) i *imenica + imenica* (*država članica*). Osim sveza s takvom strukturom u *Bazi* su pobrojani različiti drugi modeli sveza koje tvore i druge vrste riječi, primjerice glagolske, ponajprije perifrazne, sveze (*doći do zaključka*, *održati sjednicu*), prijedložne, vezničke i priložne sveze (*u vezi s*, *osim toga*, *bilo kako*, *malo kasnije*, *kao i*).

Sučelje *Kolokacijske baze* sadržava osam stupaca (natuknica, vrsta riječi, tekst (sadržava oblik sveze), sinonimi i varijante, oznaka sveze, stil i struka, značenje, primjer). Zasad se podatci o svezama navode u prvih šest stupaca, a u planu je *Bazu* dopunjavati podacima i značenju sveze i primjerima. Struktura baze s popisom sveza za natuknicu *lijep* prikazana je na 1. slici.



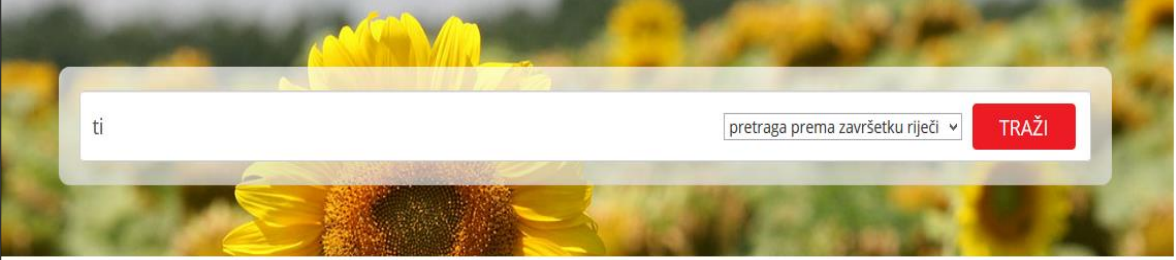
Search results for the word "lijep" in the Hrvatski mrežni rječnik. The interface shows a search bar with the word "lijep" and a search button labeled "TRAŽI". Below the search bar, the results are displayed in a table with columns: natuknica, vrsta riječi, tekst, sinonimi i varijante, oznaka sveze, stil i struka, značenje, and primjer. The table lists various meanings and uses of the word "lijep", including "lijep život", "lijepa cura", "lijepa djevojka", "lijepa gesta", "lijepa glazba", "lijepa građevina", "lijepa kata", "lijepa književnost", "lijepa kuća", "Lijepa naša", "Lijepa naša domovino", "Lijepa naša", "Lijepa Naša", "Lijepa naša domovino", "lijepa priroda", "lijepa riječ", "lijepa slika", "lijepa svotica", "lijepa umjetnost", "lijepa uspomena", "lijepa vanjština", "lijepa zarada", "lijepa žena", and "lijepi spol".

natuknica	vrsta riječi	tekst	sinonimi i varijante	Bez filtra ▾ oznaka sveze	Bez filtra ▾ stil i struka	značenje	primjer
lijep		lijep život					
lijep		lijep znak pažnje					
lijep		lijepa cura	lijepa djevojka			razg.	
lijep		lijepa djevojka					
lijep		lijepa gesta		S			
lijep		lijepa glazba					
lijep		lijepa građevina					
lijep		lijepa kata		S		bot.	
lijep		lijepa književnost		S		knjiž.	
lijep		lijepa kuća					
lijep		Lijepa naša	Lijepa naša domovino		S		himna Republike Hrvatske
lijep		Lijepa naša		S		knjiž./glazb.	
lijep		Lijepa Naša	Hrvatska	S			
lijep		Lijepa naša domovino	Lijepa naša	S			himna Republike hrvatske
lijep		lijepa priroda					
lijep		lijepa riječ	lijepa riječi	S			
lijep		lijepa slika					
lijep		lijepa svotica				razg.	
lijep		lijepa umjetnost		S			
lijep		lijepa uspomena					
lijep		lijepa vanjština					
lijep		lijepa zarada				pren.	
lijep		lijepa žena					
lijep		lijepi spol	ženski spol	S		pren.	

1. slika: Prikaz natuknice *lijep* u *Kolokacijskoj bazi*

Osim osnovne pretrage prema jednorječnoj natuknici u *Bazi* je moguće pretraživati natuknice prema početku riječi, prema završetku riječi te prema pojedinačnoj stručnoj ili stilskoj

odrednici. Na 2. slici prikazan je rezultat pretrage prema završetku riječi s pomoću koje su izdvojene natuknice koje završavaju na *-ti*, kojom se može dobiti popis glagola infinitivu.



Rezultati pretraživanja za: **ti**

natuknica	vrsta riječi	tekst	sinonimi i varijante	Bez filtra ▾	Bez filtra ▾	značenje	primjer
				oznaka sveze	stil i struka		
jesti		anđeli bi jeli		F			
raditi		aparat radi na baterije					
ponašati	povr.	apatično se ponašati					
apelirati		apelirati na /čiju/ savjest					
apelirati		apelirati na građane					
aplaudirati		aplaudirati izvođačima					
apsolvirati		apsolvirati pravo					
apstinirati		apstinirati od pušenja					
aranžirati		aranžirati cvijeće					
argumentirati		argumentirati tezu					
argumentirati		argumentirati tvrdnju					
arhivirati		arhivirati povijesni dokument					
arhivirati		arhivirati spis					
artikulirati		artikulirati svoje zahtjeve					

2. slika: Prikaz rezultata pretrage prema završetku riječi u *Kolokacijskoj bazi*

### **3. Mrežnik i Kolokacijska baza hrvatskoga jezika**

Temeljna razlika u donošenju sveza u *Kolokacijskoj bazi* i *Mrežniku* jest koncepcijska. U *Kolokacijskoj bazi* donosi se popis sveza hrvatskoga jezika te se uz jednorječne natuknice pod kojima su pojedine sveze unesene ne navode značenja i definicije, što je obilježje općejezičnoga rječnika, nego samo primjeri sveza za koje se, ako je potrebno, navodi odrednica za vrstu sveze te stručna ili stilska odrednica kojom se upućuje na komunikacijski kontekst u kojemu se neka sveza najčešće pojavljuje. *Mrežnik* kao općejezični rječnik sadržava iscrpan opis značenja, tj. definicije natuknice, a pod značenjima se navode rečenični primjeri i korpusno izdvojene kolokacije, kojima se potvrđuje značenje opisano u definiciji. Kolokacije u *Mrežniku* unose se prema skicama riječi iz korpusa, te se sustavno nabrajaju prema strukturiranim pitanjima. Takva rječnička koncepcija nije zastupljena u *Kolokacijskoj bazi*, čija je struktura znatno jednostavnija, pri čemu se naglasak stavlja na leksikološka obilježja pojedinih sveza u skladu s više-manje ustaljenom podjelom višerječnih jedinica u postojećoj literaturi te na uporabni kontekst u kojemu se višerječna sveza najčešće pojavljuje. U skladu s takvom koncepcijom, poveznica na *Kolokacijsku bazu* u *Mrežniku* ne donosi se pod značenjem natuknice ili, ako je natuknica višeznačna, pod pojedinim značenjima, nego se navodi na kraju rječničkoga članka, dakle uvodi se iz natuknice. Višerječne sveze u *Bazi* navode se pod svim sastavnicama koje tvore svezu, što u *Mrežniku* nije automatski provedeno za sva kolokacijska pitanja (iz razloga što npr. sveza određenoga pridjeva i imenice može imati visoku pojavnost kod imenice, a ona ne mora nužno biti takva kod pridjeva jer on, primjerice, može biti potvrđeniji u svezi s kojom drugom imenicom: nije, naime, riječ o frekvencijski simetričnome odnosu, a frekvencija pojedine kolokacije iščitana iz skica riječi ključni je element odabira kolokacija koje se u *Mrežniku* pojavljuju u kolokacijskome bloku). U *Mrežniku* se antonimi i sinonimi povezuju na razini pojedinih značenja natuknice, a u *Kolokacijskoj bazi* sinonimi se, ako ih ima, navode uz svaku svezu pojedinačno (npr. uz svezu *lijep dar* navodi se sinonim *lijep poklon* i obratno). Na 3. slici prikazana je obrada pridjeva *lijep* u *Mrežniku* s poveznicom na *Kolokacijsku bazu* na kraju rječničkoga članka.

lijep **lijep** prid. (G lijepa; odr. lijepi, G lijepōg(a); ž. lijepa, s. lijepo; komp. ljepši, G ljepšēg(a); sup. najljepši, G najljepšēg(a))

**1 Lijep je koji je ugodan oku, koji je skladna izgleda.**  
 - Nakon radnog djela goste smo proveli kroz naš lijepi grad: do fontane, bolničke kapelice, Zavičajnog muzeja, Glazbene škole, a susret smo završili ukusnim ručkom u Hotelu Park.  
 - Na otoku Hvaru više je lijepih mjesta i gradića koje morate posjetiti.  
 - Milla Jovovich, poznata glumica i jedna od najljepših žena svijeta, gostovala je u televizijskom showu na njemačkoj televiziji.  
**Što je lijepo?** čovjek, djevojka, glumica, žena, lice, kosa, zubi, grad, mjesto, plaža, zemlja; izgled, pogled  
**Koliko je što lijepo?** iznimno, izuzetno, jako, toliko, veoma, vrlo  
**Koordinacija:** lijep i bogat, lijep i dobar, lijep i pametan, lijep i topao, lijep i zdrav, lijep i zgodan, mlad i lijep, velik i lijep; lijep ili ružan  
**ANTONIM:** ružan :1

**2 Lijep je koji izaziva ugodu.**  
 - Iz Turske smo se vratili puni dojmova, lijepih uspomena i s još većom željom da i nagodinu sudjelujemo na jednom takvom velikom festivalu.  
 - Nadam se i vjerujem da ću u životu postići još mnogo toga i da me čeka još puno lijepih trenutaka.  
 - Ponekad je dovoljna samo lijepa riječ, običan razgovor, komunikacija koja beskućniku daje do znanja kako nije sasvim otpisan.  
**Što je lijepo?** iskustvo, osjećaj, ponašanje, priča, riječ, sjećanje, trenutak, uspomena, vijest, želja  
**Koliko je što lijepo?** iznimno, izuzetno, jako, toliko, veoma, vrlo  
**Koordinacija:** lijep i topao, lijep i ugodan, velik i lijep; lijep ili ružan  
**ANTONIM:** ružan :2

**3 Lijep je koji je vedar, sunčan, u kojemu nema kiše.**  
 - Dolazak lijepog vremena u zooškom vrtu ćemo iskoristiti za puno građevinskih radova na nastambama i poboljšanju uvjeta života naših stanovnika.  
 - Zatvorili smo se u vreće i zaspali nakon napornog, ali lijepog dana.  
**Što je lijepo?** dan, vrijeme  
**Koordinacija:** lijep i sunčan, lijep i topao, lijep i ugodan  
**ANTONIM:** ružan :3

**4 Lijep je koji je ugodan uhu, koji skladno zvuči.**  
 - Premda u tom periodu Gibonni stvara svoje najljepše pjesme (Dvije duše, Ozdravi mi ti, Lipa moja...) tek nakon raspada grupe sazrijeva vrijeme za prva službena priznanja.  
 - U jednoj od najljepših ljubavnih pjesama u hrvatskom pjesništvu i u svjetskoj literaturi Ti koja imaš nevinije ruke sadržana je sva ljubav jedne žene.  
**Što je lijepo?** pjesma  
**ANTONIM:** ružan :1

**5 Lijep je koji je skladno napravljen, izveden.**  
 - Osobito su lijepo slike tih svetih žena što ih naslikaše velemajstori slikarske umjetnosti Albrecht Dürer i Beato Angelico.  
 - Pored svih tih lijepih stvari koje krase vaš i naš muzej pojavila se i velika briga koja je na početku bila slatka, a lagano prerasta u ozbiljan problem – financije.  
**Što je lijepo?** slika, stvar  
**Koordinacija:** lijep ili ružan  
**ANTONIM:** ružan :1  
**TVORENICE:** lijepo, liepota  
**Kolokacijska baza hrvatskoga jezika:** [http://ihjj.hr/kolokacije/search/?q=lijep&search\\_type=basic](http://ihjj.hr/kolokacije/search/?q=lijep&search_type=basic)

3. slika: Obrada natuknice *lijep* u Mrežniku s navođenjem poveznice na *Kolokacijsku bazu*

## Izvori i literatura

- Baldwin, Timothy; Kim, Su Nam. 2009. Multiword expressions. *Handbook of Natural Language Processing* (2nd edition). Ur. Indurkha, Nitin; Damerau, Fred J. CRC Press. Boca Raton. 1–39.
- Benson, Morton i dr. 1997. = Benson, Morton; Benson, Evelyn; Ilson, Robert. 1997. *The BBI Dictionary of English Word Combinations*. Hohn Benjamins Publishing Co. Amsterdam – Philadelphia.
- Blagus Bartolec, Goranka. 2012. *Riječi i njihovi susjedi: Kolokacijske sveze u hrvatskom jeziku*. Institut za hrvatski jezik i jezikoslovlje. Zagreb.
- Ivir, Vladimir. 1992./1993. Kolokacije i leksičko značenje. *Filologija* 20–21. 181–189.
- Meljčuk 1998. = Mel'čuk, Igor. 1998. Collocations and Lexical Functions. *Phraseology: Theory, Analysis and Applications*. Ur. Cowie, Anthony Paul. Oxford University Press. Oxford – New York. 23–53.
- O bazi. *Kolokacijska baza hrvatskoga jezika*. <http://ihjj.hr/kolokacije/o-bazi/> (pristupljeno 14. studenoga 2020.).
- Sag i dr. 2002. = Sag, Ivan A.; Baldwin, Timothy; Bond, Francis Copestake; Ann i Dan Flickinger. 2002. Multiword expressions: A pain in the neck for NLP. *Lecture notes in computer science* 2276. 1–15.