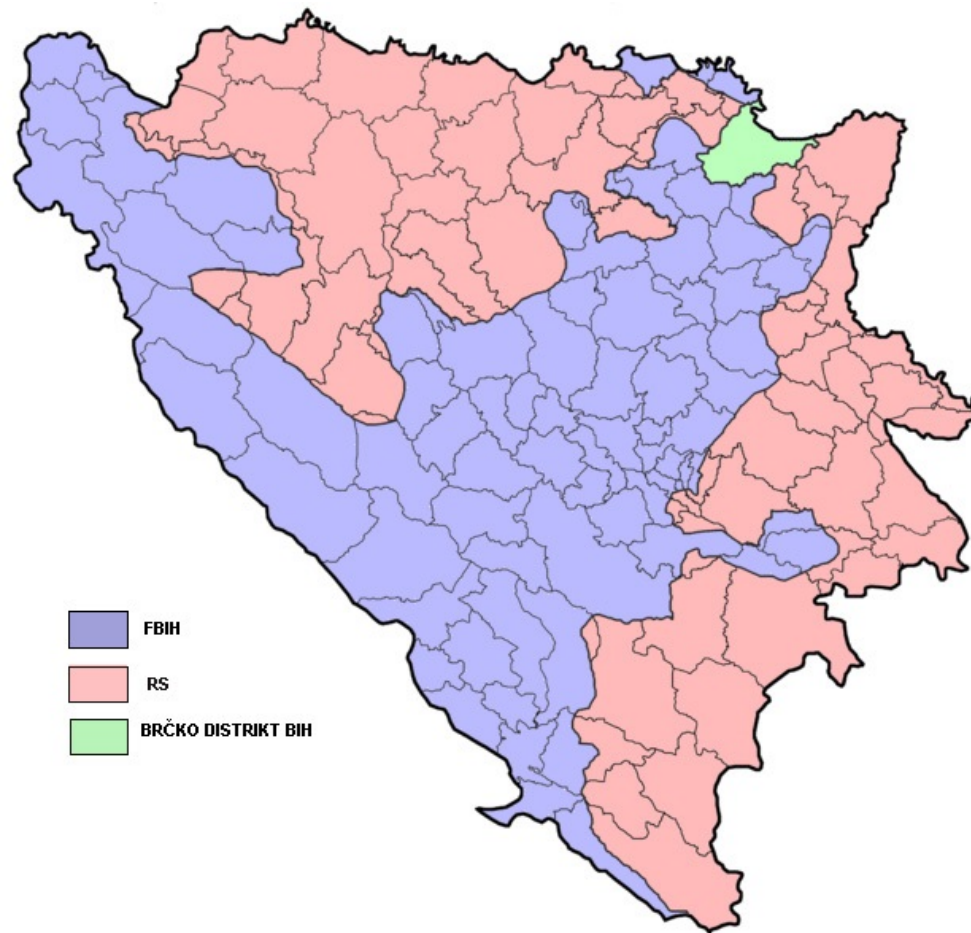


THE CREATION OF SRPKORP RS/СРПКОРП РС

SrpKorp RS/CrпKorp PC

- Corpus of contemporary Serbian language as used in the Republika Srpska (ijekavian/jekavian variants)
- Link: <http://korpus.ffuis.edu.ba>

Map of B&H



Republika Srpska: General Info

- Republic of Srpska Institute of Statistics
(<http://www.rzs.rs.ba>)
- Population: 1.42 million (2015)
- Ethnic makeup of the RS (according to 2013 census):
81.51% the Serbs; 13.99% the Bosniaks, 2.41% the Croats,
3.6% other
- Article 7 of the RS Constitution: “The official languages of the Republika Srpska are: the language of the Serb people, the language of the Bosniak people and the language of the Croat people.”

SrpKorp RS/СрпKорп РС:

- **General purpose:** for studies of B&H variety of the Serbian language
- **Contrastive studies:** Serbian and other languages studied at the Faculty of Philosophy in Pale (Departments of Russian, English, Chinese, German)
- **Language Teaching:** Serbian Department

SrpKorp RS: General Properties

SrpKorp RS is :

- A general corpus
- a synchronic corpus
- a corpus of written texts
- a small corpus: fewer than one million words (tokens)

The samples

- Time span: 2001-2016
- Samples collected by: retyping, scanning, obtaining permissions (for novels, stories and scholarly articles), downloading from the Internet
- Size of the samples (uneven!!!): the smallest one-95 words, the biggest one 100054 words
- No translated texts, except translated interviews in journalistic genre

Sample Size

| Genre/Register | Size of the sample: 0-5000 words | Size of the sample: 5000-15.000 words | More than 15.000 words |
|------------------------------------|---|--|-----------------------------------|
| Literary (14 samples) | 9 | 0 | 5 |
| *Journalistic (475 samples) | 474 | 1 | 0 |
| Scholarly (80 samples) | 61 | 19 | 0 |
| Administrative (40 samples) | 36 | 4 | 0 |
| Total: 609 samples | 580 (95%) | 24 (4%) | 5 (1%) |

SrpKorp RS: Structure

| Genre/Register | No. of samples | No. of words | % |
|-----------------------------|-----------------------|---------------------|----------|
| Literary | 14 | 269588 | 28 |
| Journalistic | 475 | 258071 | 27 |
| Scholarly | 80 | 314523 | 33 |
| Administrative/Legal | 40 | 109238 | 12 |
| Total | 609 | 951420 | 100 |

Genres/Styles/Registers

- Literary: novels and short stories (younger-generation writers, who mostly live in the north of the RS); poetry samples excluded
- Journalistic (press reportage on variety of topics): *Buka, Zvornik danas, Glas Srpske, Nezavisne novine, Trebinjelive, Radio Nevesinje*
- Administrative/legal: the Official Gazette of the RS (Službeni glasnik Republike Srpske); judgements of the Supreme Court of Republika Srpska
- Scholarly: articles from various collections of papers published by authors from the RS (mostly humanities, but also natural sciences)

Peculiarities

- Concordances are displayed in both Cyrillic and Latin alphabets, regardless of the one that is used to search the Corpus

SrpKorp RS, sample 1

The screenshot shows a web browser window with the URL `korpus.ffuis.edu.ba/index.php?keyword=knjiga`. The page title is "СрпКорп РС" and there is a link "О Корпусу" in the top right. The search bar contains the word "knjiga" and a "Претрага" button. Below the search bar, it indicates "Фреквенција: > 100". The search results are displayed in a table with 15 rows, each containing a snippet of text from a document. The word "knjiga" is highlighted in bold in the original image.

| Фреквенција: > 100 |
|--|
| e. Spustila sam pogled na stolić ispred njega i vidjela nekoliko pohabanih knjiga . To je bilo sve za prodaju. Moje iznenađenje bilo je na vrhuncu kada sam p |
| ila da ću ikada imati u ruci tako nešto. Prema sjećanju, u pitanju je bila knjiga -dnevnik. Autor je bio čudak i svjetski putnik što je sa svih strana brižlj |
| uguje ljudima, ili je vještac... Možda je usavršio sve vještine iz ovakvih knjiga pa ih isprobava na slučajnim prolaznicima koji se daju zavesti na takve st |
| a na slučajnim prolaznicima koji se daju zavesti na takve stvari? Možda će knjiga imati nada mnomoć? Možda sam, upravo, postala rob mračnog svemoćnog gosp |
| tako ostali. Ovaj susret nije bio slučajan. Oboje postadosmo svjesni: ova knjiga je sama birala put i došla u prave ruke. Ona je odgovor na moje dileme, do |
| ni mi se da sam je nekad davno vidjela u muzeju... Ili je to bila neobična knjiga , koju je jedna djevojka čitala? Ali to na slici nisam mogla biti ja, to je |
| a niječi, iako sam pisac. – Pisac? Drago mi je, i ja sam objavio nekoliko knjiga . Jeste li vi objavljivali, Zorane? – Upravo sam dobio drugu nagradu na Gla |
| da su meteorolozi: У тунелу, усред мрака, сија звијезда петокрака! Од свих књига које сам у животу скупљао, начинили су душмани велику ватру. Гријали су се |
| и, састани се, па једна међу наске читај. Па се ређали, све округ. Било је књига , лијепијех... тада... А сад, највише оне пусте новинетине, ако ишта и чита |
| , локомотива наиђе; уз продоран јаук сирене. Трасак је био библијски... из књига староставних... најављујући „лошљедња времена“. Кô из језгра нâра, крчећи |
| коро осамдесет љета и искуства што га је из свијета донио, уз товар светих књига , Заим још увијек корачаше хитро; чист и замишљен, непорочан у мисли – част |
| а, слушао гусле. У његовој торбици из дјетинства остали су сачувани, поред књига и корице хљеба, звуци гусала и српских народних пјесама, о којима је управ |
| предметну „књижицу“ (боље рећи брошуру) пронашао на гомили бачених издања, књига и часописа: раскодованих „наслова“, након што је ликвидирана подручна Библи |

SrpKorp RS, sample 2

СрпКорп РС

Претрага

knjiga

Претрага

Префреквенција: > 100

е. Spustila sam pogled na stolicu ispred nje
ila da ću ikada imati u ruci tako nešto. Pre
uguje ljudima, ili je vještac... Možda je usa
a na slučajnim prolaznicima koji se daju za
tako ostali. Ovaj susret nije bio slučajan. O
ni mi se da sam je nekad davno vidjela u p
a riječi, iako sam pisac. – Pisac? Drago mi
da su meteorolozi: У тунелу, усред мрак
и, састани се, па једна међу наске читај
, локомотива наиђе; уз продоран јавк си
коро осамдесет љета и искуства што га је из свијета донио, уз товар светих **књига**, Заим још увијек корачаше хитро; чист и замишљен, непорочан у мисли – част
а, слушао гусле. У његовој торбици из детињства остали су сачувани, поред **књига** и корице хљеба, звуци гусала и српских народних пјесамa, о којима је управ
предметну „књижицу“ (боље рећи брошуру) пронашао на гомили бачених издања, **књига** и часописа: расхатованих „наслова“, након што је ликвидирана подручна

О Корпусу

Подаци о извору

| | |
|----------|---|
| Аутор(и) | Тања Ступар Трифуновић |
| Наслов | Сатови у мајчиној соби |
| Издавач | ЈП Завод за уџбенике и наставна средства а.д. |
| Година | 2014 |

Текст

ом, тај туђи поглед у очима. Некад. И можда. Добро је, жива сам. Обухватила је рукама сопствена рамена и стисла јако. Жива сам. Само то. Довољно је. Понекад није била сигурна чак ни у то, када би јој мисли наједном склизнуле у онај процијеп између живота и смрти у којем се задржала пређуго. Скида комад по комад своје одјеће и гледа тијело у огледалу. Скоро да је исто. Избјегава сусрет с очима. Лијеже у каду. Само да не мисли. Не опет. Вода смирује. Поподне. Ручак. Спавање. Шетња. Телевизија или **књига**. Понекад разговарају њих двоје, она и муж. Политичка ситуација. Економска криза. Ништа лично. Избјежавати лично. То боли. Дијете се увуче крај ње ујутро у кревет. Опушта је мирис дјечје косе и мекоћа под прстима. Свјетло упада у собу рано ујутро, неочекивано. Тако љети, сунце улази мимо завјеса, увијек нађе неки мали отвор. И свјетлост се разлије по лицу. И пробуди је. Гледа његово тијело крај себе. Правилност свега тога је плаши; Облика, црта лица, руку које леже опуштено. Обузме је језа од пом

Изађи

CONCLUSION

- The future of SrpKorp RS is uncertain