

АКТУЕЛНИ ИСТРАЖИВАЧКИ ПРОЈЕКТИ ЗА СТУДЕНТСКУ ПРАКСУ У АКАДЕМСКОЈ 2025/26.

Овај документ садржи преглед истраживачких пројеката и радних задатака у оквиру којих студенти могу обављати стручну праксу у Центру за примењену лингвистику у академској 2025/26. години.

Анотација РАКОПС корпуса

Развојни корпус писаног српског језика (РАКОПС) први је развојни корпус српског писаног језика и састоји се од преко 6000 састава ученика основношколског узраста (од 3. до 8. разреда) из различитих делова Србије. Ученичке радове су прикупљали и дигитализовали студенти, сарадници, докторанди и наставници Катедре за општу лингвистику у периоду од 2011. до 2017. године. Корпус-бази је у целости могуће приступити преко *онлајн* платформе, али га је за његово ефикасно претраживање и оптималну употребу у лингвистичким истраживањима потребно анотирати. Студенти који своју стручну праксу буду обављали у оквиру *Центра за примењену лингвистику* имаће прилику да раде на анотацији овог корпуса на следећим нивоима:

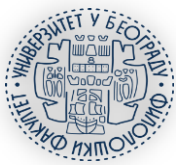
- ✓ нормализацији (одређивању нормализованог облика за речи које нису аутоматски (могле бити) адекватно препознате);
- ✓ лематизацији (одређивању основног облика речи – леме);
- ✓ тагирању (одређивању врста речи);
- ✓ морфосинтаксичкој анотацији (одређивању граматичких категорија речи).

Више о РАКОПС корпусу можете сазнати на следећем линку:

<https://rakops.fil.bg.ac.rs/#>

Издавање и анализа грешака у писаној продукцији ученика основношколског узраста

Уз помоћ РАКОПС корпуса можемо пратити језички развој деце на различитим нивоима, а један од њих је и ниво енкодирања. Енкодирање се односи на записивање говора, односно на преношење усменог језика на писани план. На најнижим узрастима, може се пратити настанак ове способности, док се на старијим може утврдити које грешке опстају, те који су елементи најтежи за усвајање. У складу са тиме, практиканти ће имати прилике да анализирају грешке у писању у оквиру РАКОПС корпуса, као што су мешање ћириличних и латиничних графема, међусобно мешање графема истог алфабета, грешке изостављања или додавања слова, одвојено и спојено писање речи, итд. Анализа ових грешака се у великој мери заснива на познавању фонетских особености говорног језика уопште, као и графичких и фонолошких специфичности српског језика.



Издавање и анализа именичких синтагми у корпусу РАКОПС

Овај задатак се састоји у издвајању и анализи именичких синтагми у писаним радовима ученика у оквиру корпуса РАКОПС. Студенти раде са аутентичним текстовима различитог жанра које су ученици 3-8. разреда, и имају задатак да у додељеним текстовима издвоје и анализирају структуру именичких синтагми.

Главни фокус анализе је на начинима модификације именице у овим текстовима, односно на позицији, типу и разноврсности модификатора којима се проширује структура именичке синтагме. Циљ овог пројекта је да се утврди који тип модификације је најзаступљенији у дејим писаним радовима и да се испита да ли и на који начин долази до усложњавања структуре именичких синтагми са узрастом и да ли жанр текста утиче на структурну сложеност продукованих именичких синтагми.

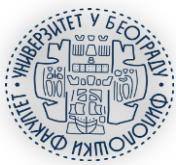
Студенти ће у оквиру овог задатка имати прилику да стекну искуство у раду на развојном корпусу писаног српског језика и припреми језичких података за даљу анализу, да се упознају са начинима анализе синтаксичке сложености специфичних језичких јединица, као и да допринесу бољем разумевању развоја именичке синтагме у дејем писаном језику.

Издавање и анализа језичких средстава за изражавање посесивности у раном дејем говору

У оквиру овог пројекта студенти ће радити са транскриптима аутентичних снимака дејег говора из базе CHILDES (доступна на следећем линку: <https://sla.talkbank.org/TBB/childes>). Ова база представља међународну базу података која садржи транскрипте и аудио снимке дејег говора на различитим језицима и користи се у лингвистичким и психолингвистичким истраживањима за проучавање језичког развоја код деце.

Задатак подразумева анализу дејег говора са циљем да се утврди на које начине деца изражавају категорију посесивности. Сваком студенту ће бити додељен конкретан снимак детета из ове базе и временски сегмент тог снимка који је потребно анализирати. Задатак студената је најпре да у додељеном делу транскрипта издвоје исказе детета у којима се на било који начин изражава припадност. Након тога, студенти треба да идентификују језичка средства којима се та категорија изражава (присвојни придев, присвојна заменица, копулативна конструкција, глагол поседовања и сл.) и да их означе користећи унапред дефинисане ознаке.

Циљ овог задатка је да се студенти обуче како да препознају и анализирају различита језичка средства за изражавање категорије припадности и њихове специфичности (нпр. одсуство конгруенције, редослед елемената) у раном дејем говору, као и да стекну практично искуство у раду са аутентичним језичким подацима.



Анализа употребе придева у склопу евалуација

Теорија оцене (*Appraisal theory*) Мартина и Вајта (2005) у оквиру системско-функционалне лингвистике нуди систематичну класификацију могућности испољавања евалуација у језику, при чему централно место у класификацији има домен *СТАВА* (*ATTITUDE*), у оквиру којег се оцене даље класификују у зависности од мета евалуације. Евалуације у писању и говору остварујемо употребом најразличитијих језичких средстава, при чему разликујемо експлицитне од имплицитних евалуација, те ова област истраживања задаје најразличитије изазове приликом анализе текстова.

У склопу овог пројекта стога се ограничавамо на један сегмент у оквиру могућности испољавања оцена: придеве, пошто они представљају основна, канонична средства за изражавање позитивних и негативних евалуација људских емоција (поддомен *AFFECT* у оквиру теорије), понашања и карактерних црта људи (поддомен *JUDGEMENT*), као и евалуација артефаката (поддомен *APPRECIATION*) (Martin/White 2005: 58).

Ово истраживање студенти практиканти могу спровести над два различита корпуса:

а) Над Развојним корпусом писаног српског језика (РАКОПС),

где је задатак (1) да се у радовима ученика 3-8. разреда основне школе издвоје придеви којима се изражавају евалуације, (2) да се утврди да ли су дате евалуације позитивне или негативне, као и (3) да се одреди ком од наведених поддомена *СТАВА* припадају.

Будући да *РАКОПС* чине текстови различитих жанрова, студенти ће имати прилику да истражују какве евалуативне придеве деца користе у наративним, дескриптивним, експозиторним и аргументативним текстовима, као и то какве мете евалуација у њима бирају, да утврде сличности и разлике према жанру радова, као и то у којој мери се употреба датих придева разликује зависности од узраста деце која су дате радове писала.

б) Над Корпусом књига за децу раног и предшколског узрата (ДециКо),

где је задатак (1) да се у сликовницама, бајкама и причама за децу издвоје придеви којима се изражавају евалуације, (2) да се утврди да ли су дате евалуације позитивне или негативне, као и (3) да се одреди ком од наведених поддомена *СТАВА* припадају.

Будући да је корпус ДециКо још у фази конструисања, рад на овом задатку подразумева и обележавање метаподатака, односно рад на експлицитном обележавању структуре корпуса, за онај подкорпус који буде додељен за анализу.

**Центар за примењену лингвистику жели добродошлицу
свим студентима који желе да се укључе у лингвистичка истраживања
и рад са аутентичним језичким подацима!**