УНИВЕРЗИТЕТ У БЕОГРАДУ
ФИЛОЛОШКИ ФАКУЛТЕТ

Датум 22.2.2016.

ИЗВЕШТАЈ О ПОДОБНОСТИ ТЕМЕ, КАНДИДАТА И МЕНТОРА ЗА ИЗРАДУ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

I ПОДАЦИ О КОМИСИЈИ

1. Датум и орган који је именовао комисију

Одлуком бр. 188 Научно-наставног већа Филолошког факултетаУниверзитета у Београдуод 29. јануара 2016. образована је комисија за оцену подобности докторанта**Саида Сафарија** (**Saeed Safari**) и теме докторске дисертације**"Constructing and Analysing an Error-Tagged Learner Corpus of Persian"** ("Израдаианализаанотираногкорпусаперсијскогјезикакаостраног").

2. Састав комисије са назнаком имена и презимена сваког члана, звања, назив уже научне области за коју је изабран у звање, датум избора у звање и назива факултета, установе у којој је члан комисије запослен:

1. Др Маја Миличевић, доцент (избор 7.12.2010, реизбор 13.10.2015), ужа научна област Општа лингвистика,предмет Примењена лингвистика, Филолошки факултет Универзитета у Београду
2. Др Јулијана Вучо, редовни професор (избор 14.7.2010), ужа научна област Италијански језик, Филолошки факултет Универзитета у Београду
3. Др Јелена Филиповић, редовни професор (избор 30.12.2010), ужа научна област Шпански језик, Филолошки факултет Универзитета у Београду
4. Др Љиљана Марковић, редовни професор (избор 15.7.2009), ужа научна област Јапански језик, књижевност, култура, Филолошки факултет Универзитета у Београду

II БИОГРАФИЈА КАНДИДАТА

Саид Сафари (Saeed Safari)рођен је 1976. године у Ирану. По завршетку средње школе и стручних курсева из области наставе персијског језика за основну школу1994. годинепочиње да ради за Министарство просвете Ирана као наставник персијског језика у државним школама.Године 1998. стиче диплому више школе (Associate Degree) у области наставе енглеског језика као страног, при Центру за обуку наставника Shahid Behshti(Shahid Behshti Teacher Training Center, Mashhad) у Ирану, а затим 2003. стиче диплому основних универзитетских студија(BA) из области превођења на енглески језик.У периоду од 2002. до 2012. године паралелно је стицао радно искуство као наставник и као преводилац за више фирми. Године 2012. завршио је мастер студије(MA) у области наставе персијског језика као страног,наУниверзитету Allameh Tabatabai у Техерану. Током мастер студија био је ангажован на више пројеката везаних за наставу персијског језика као страног, објавио је више научних радова и сарађивао је на изради уџбеника и других материјала за наставу персијског језика. Дипломирао је као најбољи студент у својој генерацији и као добитник универзитетске награде за студенте.

На основу стеченог образовања и искуства у настави позван је од стране Већа за промоцију персијског језика (при Министарству просвете Ирана) да предаје персијски језик у иностранству. Године 2012. почео је да држи наставу персијског на Филолошком факултету Универзитета у Београду, а 2013. је постављен за управника Центра за персијски језик и Иранског културног центра у Београду.Укупно гледано, Саид Сафари има преко 20 година искуства у предавању персијског и енглеског језика, како у државним тако и у приватним школама и при различитим институтима. Током рада на Филолошком факултету у Београду написао је први уџбеник персијског језика намењен матерњим говорницима српског; овај уџбеник заједно су објавили Филолошки факултет и иранска фондација Sa’di. Докторске студије на Филолошком факултету Универзитета у Београду уписао је 2013. године.

III БИБЛИОГРАФИЈА КАНДИДАТА

Библиографија Саида Сафарија садржи осам одредница, од којих је једна самостални уџбеник, док седам чине ауторски и коауторски радови у зборницима. Сви радови посвећени су темама релевантним за области истраживања докторске дисертације – изради корпуса и настави персијског језика као страног.

Safari, S. (2015). *Salam Farsi: Persian Textbook for Serbian Farsi Learners*. Faculty of Philology, University of Belgrade &Sa'di Foundation, Belgrade, Serbia.

Safari, S. (2015). From corpus linguistics to the learner corpora. In: *Proceedings of the 1stNational Conference of Corpus Linguistics*, Iranian Linguistic Association, Tehran, Iran. 131-153.

Safari, S. (2015). Constructing Persian learner corpora: An essential requirement for teaching Persian to foreigners. In: *Proceedings of the 3rdNational Conference of Linguistics*. Ferdowsi University, Mashhad, Iran.

Safari. S. (2013). The development of corpus linguistics and types of corpora. In: *Proceedings of the 4thSummer Academy of Iranian Linguistic Association*, QayemShahr Azad University, Iran.

Sahrai, R, & Safari. S. (2012). Teaching Persian using Internet: Challenges and solutions. In: *Proceedings of the International Conference of Teaching Persian and Spanish*. Complutense University of Madrid, Spain. 33-53.

Sahrai, R, & Safari. S. (2011). Teaching the Persian language using web applications a new way to expand teaching the Persian language. In: *Proceedings of the The 7thInternational Conference of The Persian Language Professors*. University of Tehran. Iran.

Safari, S. (2010). Collocations in compound words. In: *Proceedings of the 1stNational Conference of Linguistics*. Ferdowsi University, Mashhad, Iran.

Safari, S. &Shahabi. Sh. (2010). The essential need tobuild a Persian learner corpus. In: *Proceedings of the 1stNational Conference of Linguistics*. Ferdowsi University, Mashhad, Iran.

IV ОЦЕНА ДА ЈЕ КАНДИДАТ ПОДОБАН ДА РАДИ ДИСЕРТАЦИЈУ

Увидом у биографију и библиографију кандидата, комисија сматра да кандидатSaeed Safariиспуњава све услове за израду докторске дисертације.

V ОЦЕНА ПОДОБНОСТИ ПРЕДЛОЖЕНОГ МЕНТОРА

За ментора је предложена др Маја Миличевић, доцент Филолошког факултета у Београду, Катедра за општу лингвистику.

Библиографија ментора са најмање 5 јединица релевантних за област из које се ради докторска дисертација.

(1) Miličević, M. i T. Kraš (u štampi) Translation between L2 Acquisition and L1 Attrition: Anaphora Resolution in Italian by English-Italian Trainee Translators. *Applied Linguistics*. doi: 10.1093/applin/amu070

(2) Miličević, M. (uštampi)Korpus srpskog kao stranog jezika (KSKS): opis građe i plan izrade. U V. Krajišnik (prir.) *Srpski kao strani jezik u teoriji i praksi III*. Beograd: Filološki fakultet.

(3) Bernardini, S., A. Ferraresi i M. Miličević (u štampi)From EPIC to EPTIC – Exploring simplification in interpreting and translation from an intermodal perspective. *Target. International Journal of Translation Studies.* [https://m.benjamins.com/#catalog/journals/target.28.1/toc]

(4) Miličević, M., S. Bernardinii A. Ferraresi (2014) Costruzione semi-automatica di corpora orientati al genere in linguemorfologicamentericche: Un paragonefral'italiano e ilserbo. *ItalicaBelgradensia* 1/2014. 99-114.

(5) Samardžić, T. i M. Miličević(2013) Constructing a learner-friendly corpus-based dictionary of Serbian verbal aspect.*Primenjenalingvistika*14. 77-89.

(6) Miličević, M. i T. Samardžić (2010) Korpusikaosredstvozapostizanjeautonomije u učenjustranogjezika.U J. Vučo i B. Milatović (prir.), *Autonomijaučenika i nastavnika u nastavijezika i književnosti.*Nikšić: FilozofskifakultetCrne Gore. 387-399.

(7) Miličević, M. (2007) Modularity in SLA: Transfer, innate mechanisms and input in the L2 acquisition of Italianreflexive and reciprocalconstructions. U L. Roberts, A. Gürel, S. Tatar i L. Martı (prir.), *EUROSLA Yearbook* 7. Amsterdam: John Benjamins. 89-120.

Комисија сматра дадоц.др Маја Миличевић испуњава све услове за ментора ове докторске дисертације за област примењена лингвистика.

VI ОЦЕНА ПОДОБНОСТИ ТЕМЕ:

ОЦЕНА:

1. Формулације назива тезе (наслова)

Комисија закључује да је радни назив тезе "Constructing and Analysing an Error-Tagged Learner Corpus of Persian" ("Израдаианализаанотираногкорпусаперсијскогјезикакаостраног")**прикладан и да добро репрезентује суштину предложене теме истраживања.**

2. Предмета (проблема) истраживања

Опис теме: Предложена тема обједињује две области значајне за проучавање процеса усвајања страног (другог, нематерњег) језика: израду корпуса ученичког језика (где се под ученицима подразумевају све особе које уче одређени језика као страни, без обзира на узраст, контекст учења и сл.)и корпусно засновану анализу грешака. Кандидат предложеним истраживањем обухвата усвајање персијског језика од стране ученика којима је матерњи језик српски, аграђу за корпус (под називом *Salam Farsi Learner Corpus* – SFLC) чиниће састави прикупљени од студената који су током претходних година слушали персијски језик као изборни предмет на Филолошком факултету у Београду.Текстови, који су уз сагласност студената већ сакупљени, биће пребачени у дигитални формат, а грешке ће, како би њихова анализа била олакшана, бити посебно означене на основу претходно разрађене теоријски засноване класификације.

Ова тема је изузетно актуелна у контексту брзог развоја корпусне лингвистике и у светлу све чешће употребе ученичких корпуса у истраживањима процеса усвајања другог језика, али и настави (Nesselhauf 2004, Granger2009, 2015).Значај ученичких корпуса пре свега лежи у томе што они чине доступнима веће количине аутентичне ученичке продукције, која доприноси ближем опису ученичког језика, односно *међујезика* – система чије су одлике делом потекле из матерњег, делом из другог језика, а делом су резултат креативности ученика и његове/њене способности за продуктивну употребу језика, уз евентуални паралелни допринос нејезичких фактора.Ученички корпуси већ постоје за језике попут енглеског, италијанског, чешког, арапског, и бројних других (као што показује листа доступна на страници https://www.uclouvain.be/en-cecl-lcworld.html). Посебно су битни корпуси у којима су означене (анотиране) ученичке грешке. Иако грешке нису једино што чини ученички језик, који одликују и бројне тачне форме и правилности, грешке одувек привлаче посебну пажњу како истраживача који желе да утврде факторе одговорне за њих, тако и наставника који желе да користе сазнања о типичним грешкама у припреми наставног материјала и развоју стратегија за истрављање грешака, у сврху олакшавања напредовања ученика. Анотирати се могу грешке на различитим нивоима језичке структуре, од фонетских и/или ортографских до семантичких и прагматичких (типично са посебним нагласком на морфосинтакси), што омогућује обухватан опис ученичког језика. Уз то су честе и ознаке грешака према површинском типу (додавање елемената, изостављање елемената, и сл.), што такође доприноси стварању потпуније слике ученичке продукције.Другим речима, класификација грешака представља још једно значајно и развијено поље истраживања у оквиру пручавања процеса усвајања страног језика, и то поље које је лингвистички гледано изузетно значајно будући да захтева како општија сазнања о међујезику (посебно у случајевима корпуса који обухватају продукцију ученика са различитим матерњим језицима), тако и поређење матерњег и нематерњег језика ученика и детаљну анализу циљног језика(в. нпр. Nicholls 2003, S□tindlová isar. 2003, Izumiisar. 2005). Анализа грешака као процедура анализе ученичког језика стога је последњих деценија поново постала актуелнакаометода истраживања(и делимична теорија усвајања другог језика, иако у мањој меринего приликом њеног настанка ’70-их година 20. века).

Тема о којој је реч у овом предлогуиновативна је у погледу комбинације укључених језика, као и у погледу самог персијског као страног језика, будући да ће израђени корпус бити први ученички корпус за овај језик. Са стране лингвистичке теорије и теорије усвајања другог језика треба истаћи да су и персијски и српски језик чланови породице индоевропских језика, али их одликују бројне разлике везане како за писмо (персијски језик користи делимичноприлагођено арапско писмо)тако и за морфосинтаксу, што их чини занимљивим за контрастивну анализу и анализу грешака. Предложено анотирање ученичких грешака стога је вишеструко значајно – са једне стране оно представља допринос развоју корпусне лингвистике уопште и уклапа се у врло актуелне трендове израде анотираних ученичких корпуса за друге језике, док са друге стране представља допринос теорији усвајања другог језика захваљујући томе што се

таксономија грешака утврђује унапред, на основу постојеће литературе из области анализе грешака и контрастивне анализе персијског и српског језика. Битно је истаћи и да ће корпус бити слободно доступан, захваљујући чему ће моћи да га користе сви заинтересовани истраживачи, предавачи и ученици персијског језика као страног.

Комисија закључује да је предложени предмет истраживања својом сложеношћу, актуелношћу и значајем у областима дисертације.

3. Библиографија прелиминарних истраживања

Кандидат је приложио библиографију од **80** библиографскихјединица релевантних за истраживање. Овде наводимо јединице најзначајније за истраживање:

Alfaifi, A. and Atwell, E. (2013). Arabic Learner Corpus v1: A New Resource for Arabic Language Research. In *Proceedings of the Second Workshop on Arabic Corpus Linguistics (WACL 2)*. Lancaster University.

Alfaifi, A., Atwell, E. and Abuhakema, G. (2013) Error Annotation of the Arabic Learner Corpus: A New Error Tagset. In I. Gurevych, C. Biemann and T. Zesch (Eds),*Language Processing and Knowledge in the Web* (GSCL 2013). Berlin and Heidelberg: Springer Verlag. 14-22.

Buttery, P., & Caines, A. (2012). Normalising frequency counts to account for ‘opportunity of use’ in learner corpora. In Y. Tono, Y. Kawaguchi and M. Minegishi (Eds.), *Developmental and Cross-Linguistic Perspectives in Learner Corpus Research*. Amsterdam: John Benjamins. 187-204.

Corder, S. P. (1976). The Study of Interlanguage.In *Proceedings of the Fourth International Conferenceof Applied Linguistics*. Munich: Hochschulverlag.

Dulay, H.C. Burt, M.K. and Krashen, S. (1982). *Language Two*. New York: Oxford University Press.

Farrokh, P. (2011). Analyzing of EFL Learners' Linguistic Errors: Evidence from Iranian Translation Trainees. *Theory and Practice in Language Studies* 1(6): 676-680.

Gilquin, G., Granger, S. and Paquot, M. (2007). Learner Corpora: The Missing Link in EAPPedagogy. *Journal of English for Academic Purposes* 6(4): 319-335.

Granger, S. (2015). The contribution of learner corpora to reference and instructional materials design. InS. Granger, G. Gilquin andF. Meunier (Eds),*The Cambridge Handbook of Learner Corpus Research*. Cambridge: Cambridge University Press. 486-510.

Granger, S. (2009). The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation. In K. Aijmer (Ed.),*Corpora and Language Teaching*. Amsterdam and Philadelphia: John Benjamins. 13-32.

Granger, S. (2003). Error-tagged Learner Corpora and CALL: A promising Synergy. *CALICO Journal* 20(3): 465-480.

Granger, S. (2002).A bird's-eye view of computer learner corpus research. In S. Granger, J. Hung and S. Petch-Tyson (Eds), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam/Philadelphia: Benjamins. 3-33.

Hana, J., Rosen, A., S□kodová, S. and B. S□tindlová (2010). Error-tagged learner corpus of Czech. In *Proceedings of the Fourth Linguistic Annotation Workshop (LAW-10).* Uppsala: Association forComputational Linguistics.11-19.

Izumi, E., Uchimoto, K.and Isahara, H. (2005). Error annotation for corpus of Japanese learner English. In*Proceedings of the Sixth International Workshop on Linguistically Interpreted Corpora (LINC 2005)*. Jeju Island, Korea. 71-80.

Lozano, C., & Mendikoetxea, A. (2013). Learner corpora and second language acquisition: The design and collection of CEDEL2. In A. Díaz-Negrillo, N. Ballier and P. Thompson (Eds), *Automatic Treatment and Analysis of Learner Corpus Data*. Amsterdam: John Benjamins. 65-100.

Nesselhauf, Nadja. (2004). Learner corpora and their potential in language teaching. In J. Sinclair (Ed.), *How to Use Corpora in Language Teaching*. Amsterdam & Philadelphia: Benjamins. 125-152.

Nicholls, D. (2003). The Cambridge Learner Corpus - Error coding and analysis for lexicography and ELT. In D. Archer, P. Rayson, A. Wilson and T. McEnery (Eds),*Proceedings of the Corpus Linguistics 2003 Conference*. Lancaster: Lancaster University. 572-581.

Ro□mer, U. (2008). Corpora and Language Teaching. In A. Lüdeling and M. Kytö (Eds),*Corpus Linguistics. An International Handbook (Vol. 1)*. Berlin: Mouton de Gruyter.112-130.

S□tindlová, B., S□kodová, S., Rosen, A. and Hana, J. (2010). Annotating foreign learners' Czech. In M.Ziková and M. Doc□ekal (Eds),*Slavic Languages in Formal Grammar. Proceedings of FDSL*.Frankfurt am Main: Peter Lang. 205-219.

Tenfjord, K., Meurer, P. and Hofland, K. (2006). The ASK corpus: A language learner corpus of Norwegian as a second language. In: *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*. Genoa, Italy: European Language Resources Association. 1821-1824.

4. Циљеви истраживања

Предложена дисертација има два главна циља:

1) Израду слободно доступног анотираног корпуса персијског језика као страног, која укључује и израду система класификације и ознака за грешке, у сврху употребе аутентичних ученичких података у истраживањима и настави персијског језика као страног. Корпус ће поред података о грешкама садржати и метаподатке о ученицима и њиховим текстовима, који ће омогућити претрагу корпуса према критеријумима какви су пол ученика или њихов ниво знања персијског језика, тема текста и сл.

2) Утврђивање типова иучесталости грешака, према површинској структури, нивоу језичке структуре и врсти речи, које праве ученици персијског језика којима је матерњи језик српски. На основу оватри нивоа таксномије грешака (детаљније описаних у одељку ”Методологија истраживања”) биће утврђено које се грешке најчешће јављају, а које су мање заступљене. Типови грешака затим ће бити доведени у везу са одликама двају језичких система (система персијскоги српског језика).

Хипотезе:Централни део дисертације посвећен је изради ресурса – ученичкогкорпуса и ознака за грешке које ће у њему бити употребљене.У погледу анализе грешака, општа хипотеза је да ће се услед негативног трансфера највећи број грешака – на ортографском, морфолошком и синтаксичком нивоу – тицати појава у којима постоје разлике између персијског језика и српског језика, који је матерњи језик ученика чији састави улазе у корпус.

5. Очекивани резултати

Очекивани резултати предложене дисертације јесу: (1) теоријски заснован модел анотације ученичких грешака за персијски језик; (2) скуп ознака (*tagset*) за грешке; (3) алат за рачунарски потпомогнуто означавање грешака; (4) слободно доступан анотиран ученички корпус персијског језика, који нуди могућност приступа путем графичког корисничког интерфејса на наменској страници на Интернету; (5) анализа грешака у персијском као страном језику код ученика којима је матерњи језик српски, заснована на израђеном корпусу и вишеслојној анотацији грешака која у њему буде примењена.

6. План рада

Предложени план рада, према активностима и њиховом трајању, приказан је у табели испод. Предвиђено је укупно пет фаза израде дисертације, почевши од довршавања прегледа литературе до самог писања рада. Централни део плана чине активности везане за израду и анализу корпуса. Предложени план рада у потпуности одговара захтевима појединачних фаза израде дисертације.

Фаза	**Активност**	**Потребно време**
1	Завршетак прегледа литературе и сродних пројеката; финално дефинисање критеријума израде корпуса	2 месеца
2	Транскрипција и нормализација текстова; утврђивање скупа ознака за грешке и метаподатке	3 месеца
3	Развој алата за означавање грешака; означавање текстова	3 месеца
4	Употреба алата за претрагу и анализу корпуса	2 месеца
5	Постављање Интернет странице на којој ће корпус бити доступан; анализа података; завршетак писања рада	3 месеца

7. Методе истраживања

Методологија израде корпуса персијског језика као страног састојаће се из три главне фазе, (1) израда корпуса, (2) означавање грешака, и (3) израда и/или примена алата за обраду и претрагу корпуса,од којих је свака подељена на више подзадатака.

(1) Израда корпуса

а) Дефинисање критеријума израде корпуса

У процесу израде ученичких корпуса битно је одлучити какав тип језика ће у корпусу бити заступљен (у погледу регистра, жанрова, тема), какви задаци ће бити употребљени за елицитацију текстова (спонтана продукција, писање уз припрему, и сл.), као и то чија продукција ће бити заступљена (где битне варијабле везане за ученике чинематерњијезик, узраст, пол, ниво знања другог језика). Будући да корпус који се израђује у оквиру предложене дисертације настаје у институционалним условима, критеријуми су у извесној мериодређени тиме. Текстови су сакупљени од студената којима је матерњи језик српски, а на Филолошком факултету у Београду (у Иранском културном центру) су слушали персијски језик као страни. Узорком су обухваћени ученици који су предмет пратилитоком академских година 2012-2013. и 2013-2014. Текстове који су прикупљени чине домаћи задаци есејског типа које су писали ученици на средњем и напредном нивоу персијског језика.

б) Прикупљање података

Прикупљање ученичких текстова већ је спроведено. Ученици су приликом пријаве за слушање предмета дали пристанак за употребу радова у сврху израде корпуса, а задатке су писали у наменски припремљеним свескама. Укупно је прикупљено приближно 700 текстова. Уз текстовепостоје забележени подаци о полу и узрасту ученика, као и о њиховом нивоу знања и теми рада; ови подаци биће употребљени у корпусу као метаподаци путем којих ће претрага корпуса моћи да се ограничи на подскупе ученика и/или тема.

в) Дефинисање стандарда за транскрипцију података

Персијски језик у одређеним случајевима дозвољава различите начине писања, што у појединим ситуацијама доводи до постојања више варијанти записа. Будући да избор у оваквим случајевима донекле спада у домен личног стила писања, како би се олакшало означавање грешака на ортографском нивоу и како би се спречило означавање допуштених варијанти као грешака, биће развијена детаљна правила за конзистентну транскрипцију текстова. Ово је фаза израде корпуса која је тренутно у току.

г) Транскрипција података

Сви текстови који ће бити укључени у корпус писани су руком. Током процеса транскрипције биће пребачени у електронски текстуални формат (*plain text*). Текстови ће прво бити скенирани како би био сачуван дигитални запис оргинала, а затим ће се транскрипцији приступити ручно, у складу са дефинисаним страндардима.

д) Додатна обрада корпуса

Једна од могућности складиштења корпуса јесте у виду база у програму Microsoft Access, а по потреби ће бити размотрене и друге опције.

(2)Означавање грешака

а) Развијање таксономије грешака

НИВО А **Површинска структура**	**НИВО Б** **Ниво језичке структуре**	**НИВО В** **Врста речи**
Погрешан облик	Морфологија	Именица
Изостављање	Синтакса	Заменица
Додавање	Семантика	Глагол
Замена	Стил	Придев
Пермутација	Ортографија	Прилог
		Предлог
		Везник

Таксономијакоја ће бити примењена на корпус биће заснована на нивоима и могућим врстама грешака приказаним у горњој табели, уз евентуална мања прилагођавања. Класификација грешака заснована је на

моделу који предлажу Dulay i sar. (1982), који укршта поделу преманивоима језичке структуре(морфологија, синтакса, итд.) и поделу према површинској структури грешака (изостављање, додавање, замена, пермутација). Овај модел ће бити прилагођен потребама SFLC, при чему ће узор бити и систем означавања грешака примењен у Кембричком ученичком корпусу (*Cambridge Learner Corpus*;Nicholls 2003), по принципу <#ОЗНАКА>погрешна реч|исправна реч</#ОЗНАКА> (<#CODE>wrong word|corrected word</#CODE>, Nicholls 2003: 573), где ознаке комбинују различите нивое означавања, нпр.CN = *Countability of Noun error* – грешкавезана за бројивост именице); тачан формат ознака биће по потреби прилагођен техничким захтевима софтвера који буде кориштен за обраду и претрагу корпуса (в. ниже).

б) Дефинисање скупа ознака

На основу горње класификације грешака биће развијен скуп ознака (*Salam Farsi Error Tagset* – SFET) које ће комбиновати различите нивое (уп. нпр.OMN заизостављањеморфеманаименици– "Omission-Morphology-Noun", Nicholls 2003).

(3)Израда и/или примена алата за обраду и претрагу корпуса

а) Развијање алата за означавање грешака

У оквиру дисертације биће израђен алат за означавање грешака који ће имати за циљ олакшавање процеса означавања у овом пројекту, као и могућу будућу употребу у другим пројектима. Овај алат ће укључивати менија у којима ће бити наведене ознаке, што ће спречити евентуално ручно уношење непостојећих ознака.Биће размотрена и могућност делимичне аутоматизације означавања грешака.

б) Примена система за претрагу корпуса

Како би се омогућили напредни видови претраге, корпус ће употребити једну од постојећих слободно доступних платформи за управљање корпусима, највероватније*IMS Open Corpus Workbench* (CWB). Ова платформа спада међу најчешће кориштене у корпусној лингвистици уопште и укључује опцију индексирања корпуса, као и могућност дефинисања више нивоа анотације, што је у случају корпуса SFLC посебно значајно јер ће омогућити унос посебног слоја анотације са ознакама за грешке.Употреба платформе CWB дозволиће, употребом језика за претрагу корпуса(*Corpus Query Language* – CQL), и напредне видове претраге, на пример претрагу према ознакама грешака. Платформа укључује и могућност уноса метаподатака, који ће моћи да буду употребљени као критеријуми за сужавање претраге на одређени тип текстова или одређену подгрупу ученика.

в) Постављање Интернет странице корпуса

Како би корпус био доступан широј заједници истраживача, наставника и ученика персијског језика као страног, биће израђен (или употребљен неки од постојећих) интерфејс за приступ, који ће моћи да се користи путем наменске Интернет странице.

Друга методолошки битна димензија предложеног докторског рада јесте лингвистичка анализа корпуса, односно анализа ученичких грешака.Анализа грешака спроведена у оквиру предложене дисертације следиће приступ који се у литератури типично означава као **рачунарски потпомогнута анализа грешака**(*Computer-aided Error Analysis* – CEA, в. нпр. Granger 2002).

Анализа грешака развијена током '70-их година 20. века подразумевала је пет основних корака: (1) прикупљање узорка ученичког језика, (2) идентификацију грешака, (3) опис грешака (у који спада и њихова класификација), (4) објашњење грешака, и (5) евалуацију грешака (са педагошким импликацијама). Будући да је целокупан процес био спровођен ручно, анализа је нужно трајала дуго и било је немогуће анализирати веће количине материјала. Рачунарски потпомогнута анализа грешака у одређеној мери наслеђује исте кораке, али су ови кораци у високомстепену аутоматизовани. Кораци (1)-(3) тесно су повезани са израдом корпуса и означавањем грешака (чија методологија у оквиру предложене дисертације је описана у горњем тексту). Као поткорак у оквиру корака (3) може се у случају рачунарски обрађених корпуса додати употреба ознака унетих приликом израде корпуса за аутоматско добијање података о учесталости различитих типова грешака – лакше (уједно и егзактније) спровођење овог поткорака једна је од главних предности савремених ученичких корпуса. У оквиру предложене дисертације посебна пажња биће посвећена грешкама на нивоима ортографије, морфологије и синтаксе – на основу претходно унетих ознака биће одређена учесталост различитих типова грешака на овим језичким нивоима. Потом ће добијени подаци бити употребљени на начин сличан корацима (4) и (5) из стандардне анализе грешака – биће тражена објашњења забележених грешака (уз претпоставку да ће велики део бити узрокован одликама матерњег језика ученика, а затим ће на основу добијених података

бити изведене импликације за наставу персијског као страног језика.

VII ЗАКЉУЧАК СА ОБРАЗЛОЖЕНОМ ОЦЕНОМ О ПОДОБНОСТИ ТЕМЕ КАНДИДАТА: Саида Сафарија (Saeed Safari)

На основу наведених података о кандидату, Комисија констатује да је Saeed Safari подобан за израду докторске дисертације.

На основу постављених циљева и очекиваних резултата, Комисија констатује да је предложена тема подобна за израду докторске дисертације у области примењене лингвистике.

На основу наведених података о предложеном ментору, Комисија констатује да је др Маја Миличевић подобан/на за ментора предложене докторске дисертације.

На основу изнетог, Комисија предлаже Наставно-научном већу Филолошког факултета Универзитета у Београду да донесе позитивне оцене о подобности кандидата**Саида Сафарија(Saeed Safari)** и предложене теме докторске дисертације **"Constructing and Analysing an Error-Tagged Learner Corpus of Persian"** ("Израдаианализаанотираногкорпусаперсијскогјезикакаостраног") а да се за ментора именује **др Маја Миличевић**, доцент Филолошког факултета у Београду.

ПОТПИСИ ЧЛАНОВА КОМИСИЈЕ

1.Др Маја Миличевић, доцент (предложени ментор)
Филолошки факултет, Београд

2. Др Јулијана Вучо, редовни професор
Филолошки факултет, Београд

3. Др Јелена Филиповић, редовни професор
Филолошки факултет, Београд

4. Др Љиљана Марковић, редовни професор
Филолошки факултет, Београд