ISSN: 2181-1741 (print)
ISNN: 2181-1725 (online)
2024 / № 10

FILOLOGIK TADQIQOTLAR:
TIL
ADABIYOT
TA'LIM



Philological research:
LANGUAGE, LITERATURE, EDUCATION

Филологические исследования:
ЯЗЫК, ЛИТЕРАТУРА, ОБРАЗОВАНИЕ

pholologicalresearch.uz

TERMIZ DAVLAT UNIVERSITETI
TerDU

ILMIY-INNOVATSION JURNAL

ТЕРМИЗ ДАВЛАТ УНИВЕРСИТЕТИ

PHILOLOGICAL RESEARCH: LANGUAGE, LITERATURE, EDUCATION

FILOLOGIK TADQIQOTLAR: TIL, ADABIYOT, TA’LIM

ФИЛОЛОГИЧЕСКИЕ ИССЛЕДОВАНИЯ: ЯЗЫК, ЛИТЕРАТУРА, ОБРАЗОВАНИЕ

Илмий-инновацион журнал

Jurnal OAK Rayosatining 2022-yil 30-sentabrdagi 325/6-son qarori bilan filologiya fanlari bo‘yicha doktorlik dissertatsiyalari asosiy ilmiy natijalarini chop etish tavsiya etilgan ilmiy nashrlar ro‘yxatiga kiritilgan

2021-yildan buyon nashr etiladi

2024/ № 10

Termiz – 2024

Muassis:
Termiz davlat universiteti

PHILOLOGICAL RESEARCH: LANGUAGE, LITERATURE, EDUCATION

FILOLOGIK TADQIQOTLAR: TIL, ADABIYOT, TA'LIM

ФИЛОЛОГИЧЕСКИЕ ИССЛЕДОВАНИЯ: ЯЗЫК, ЛИТЕРАТУРА, ОБРАЗОВАНИЕ

ILMIY-INNOVATSION JURNAL

O‘zbekiston Respublikasi Prezidenti Administratsiyasi huzuridagi Axborot va ommaviy kommunikatsiyalar agentligi tomonidan 2020-yil 19-oktyabrda 1122 raqami bilan ro‘yxatga olingan. Matnlarda foydalanilgan misol, ko‘chirma va ma’lumotlar aniqligi uchun mualliflar javobgardir. Jurnaldan ko‘chirib bosilganda manba qayd etilishi shart.

Bosishga ruxsat etildi: 28.10.2024
Qog‘oz bichimi: 70x85
Bosma tabog‘i: 5.94
Ofset bosma. Ofset qog‘ozi
Adadi: 100
Bahosi kelishilgan narxda
Buyurtma: 155-son

«TerDU bosmaxonasi»da chop etildi
Termiz shahri, Barkamol avlod ko‘chasi, 43-uy

Tahririyat manzili:
190111, Termiz shahri,
Barkamol avlod ko‘chasi, 43.
Telefon: +998762231649

E-mail:
philologicalresearch@umail.uz

Rasmiy sayt:
https://journaltersu.uz

BOSH MUHARRIR:
Abduqodir TOSHQULOV,
iqtisod fanlari doktori, professor

TAHRIR HAY’ATI:

Bekpo‘lat Umurqulov, *filologiya fanlari doktori, professor (bosh muharrir o‘rinbosari)*
Shoxista Maxmaraimova, *filologiya fanlari doktori, professor (mas’ul muharrir)*
O‘ral Xoliyorov, f.f.f.d. (PhD), dotsent (O‘zbekiston)
Olmos Ulviy Binnatova, f.f.d., professor (Ozarbayjon)
Karl Rayxl, f.f.d., professor (Germaniya)
Vali Savash yelok, f.f.d., professor (Turkiya)
Dorota Silviya, f.f.f.d. (PhD), dotsent (Polsha)
Muslihiddin Muhiddinov, f.f.d., professor (O‘zbekiston)
Ibodulla Mirzayev, f.f.d., professor (O‘zbekiston)
Abdi Mamatov f.f.d., professor (O‘zbekiston)
Suyun Karimov f.f.d., professor (O‘zbekiston)
Muhammadxon Hakimov, f.f.d., professor (O‘zbekiston)
Nafas Shodmonov f.f.d., professor (O‘zbekiston)
Dilmurod Quronov, f.f.d., professor (O‘zbekiston)
Baxtiyor Mengliyev, f.f.d., professor (O‘zbekiston)
Normat Yo‘ldoshev, f.f.n., dotsent (O‘zbekiston)
Abdulla Xudoyqulov, f.f.f.d. (PhD) (O‘zbekiston)
Bobonazar Murtazayev, f.f.n., dotsent (O‘zbekiston)
Asqar Eshmo‘minov, f.f.d., dotsent (O‘zbekiston)
Obidjon Shofiyev, f.f.f.d. (PhD), dotsent (O‘zbekiston)
Zafar Umurqulov, f.f.f.d., dotsent, (PhD) (O‘zbekiston)
Nurali Amonturdiyev, f.f.f.d. (PhD) (O‘zbekiston)
Hakim Ahmedov, f.f.f.d. (PhD) (O‘zbekiston)
Ahad Shukurov, f.f.f.d. (PhD) (O‘zbekiston)
Eshnazar Jabborov, f.f.f.d. (PhD) (O‘zbekiston)
Abdumurod Arslonov, f.f.f.d. (PhD) (O‘zbekiston)
Sirojiddin Turdimurodov, f.f.f.d. (PhD) (O‘zbekiston)
Umida Raximova, f.f.f.d. (PhD) (O‘zbekiston)
Gulchehra Begmatova, f.f.f.d. (PhD) (O‘zbekiston)
Nargis Qurbonazarova, f.f.f.d. (PhD) (O‘zbekiston)
Muyassar Xolova, f.f.f.d. (PhD) (O‘zbekiston)
Sherzod Turdiqulov, f.f.f.d. (PhD) (O‘zbekiston)
Zaynura Qodirova, f.f.f.d. (PhD) (O‘zbekiston)
Humayun Nadim Muhammad, f.f.f.d. (PhD) (O‘zbekiston)

Texnik muharrir: Bahodir To‘raliyev

Учредитель:
Термезский государственный университет

PHILOLOGICAL RESEARCH: LANGUAGE, LITERATURE, EDUCATION

FILOLOGIK TADQIQOTLAR: TIL, ADABIYOT, TA'LIM
ФИЛОЛОГИЧЕСКИЕ ИССЛЕДОВАНИЯ: ЯЗЫК, ЛИТЕРАТУРА, ОБРАЗОВАНИЕ

НАУЧНО-ИННОВАЦИОННЫЙ ЖУРНАЛ

Зарегистрировано Агентством информации и массовых коммуникаций при Администрации Президента Республики Узбекистан от 19 октября 2020 года под номером 1122.

Авторы несут ответственность за точность приведенных в текстах примеров, отрывков и данных. Необходимо указывать источник при копировании из журнала.

Разрешено к печати: 28.10.2024
Размер бумаги; 70x85
Печатный лист: 5.94
Офсетная печать: Офсетная бумага
Тираж: 100. Цена договорная
Заказ: 155

Печатано в «Типографии ТерГУ»
ул. Баркамол авлод 43, г. Термез

Адрес редакции:
190111, г.Термез, ул. Баркамол авлод, 43.
Телефон: +998762231649

E-mail:
philologicalresearch@umail.uz

Официальный сайт:
https://journaltersu.uz

ГЛАВНЫЙ РЕДАКТОР:
Абдукадир ТАШКУЛОВ, доктор экономических наук, профессор

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

Бекпулат Умуркулов, доктор филологических наук, профессор (зам. главного редактора)
Шохиста Махмараимова, доктор филологических наук, профессор (ответственный редактор)
Урал Холияров, ф.ф.ф.д. (PhD), доцент (Узбекистан)
Олмос Улвий Биннатова, д.ф.н., профессор (Азербайджан)
Карл Райхл, д.ф.н., профессор (Германия)
Вали Саваш Елок, д.ф.н., профессор (Турция)
Дорота Сильвия, PhD, доцент (Польша)
Муслихиддин Мухиддинов, д.ф.н., профессор (Узбекистан)
Ибодулла Мирзаев, д.ф.н., профессор (Узбекистан)
Абди Маматов, д.ф.н., профессор (Узбекистан)
Суюн Каримов, д.ф.н., профессор (Узбекистан)
Мухаммадхон Хакимов, д.ф.н., профессор (Узбекистан)
Нафас Шодмонов, д.ф.н., профессор (Узбекистан)
Дилмурод Куронов, д.ф.н., профессор (Узбекистан)
Бахтиёр Менглиев, д.ф.н., профессор (Узбекистан)
Нормат Юлдашев, к.ф.н., доцент (Узбекистан)
Абдулла Худайкулов, д.ф.ф.н. (PhD) (Узбекистан)
Бобоназар Муртазаев, ф.ф.н., доцент (Узбекистан)
Аскар Эшмуминов, д.ф.н., доцент (Узбекистан)
Обиджон Шофиев, д.ф.ф.н. (PhD), доцент (Узбекистан)
Зафар Умуркулов, д.ф.ф.н. (PhD) (Узбекистан)
Нурали Амонтурдиев, д.ф.ф.н. (PhD) (Узбекистан)
Хаким Ахмедов, д.ф.ф.н. (PhD) (Узбекистан)
Ахад Шукуров, д.ф.ф.н. (PhD) (Узбекистан)
Эшназар Джаббаров, д.ф.ф.н. (PhD) (Узбекистан)
Абдумурод Арслонов, д.ф.ф.н. (PhD) (Узбекистан)
Сирожиддин Турдимуродов, д.ф.ф.н. (PhD) (Узбекистан)
Умида Рахимова, д.ф.ф.н. (PhD) (Узбекистан)
Гульчехра Бегматова, д.ф.ф.н. (PhD) (Узбекистан)
Наргис Курбаназарова, д.ф.ф.н. (PhD) (Узбекистан)
Муяссар Холова, д.ф.ф.н. (PhD) (Узбекистан)
Шерзад Турдикулов, д.ф.ф.н. (PhD) (Узбекистан)
Зайнура Кадирова, д.ф.ф.н. (PhD) (Узбекистан)
Хумаюн Надим Мухаммад, д.ф.ф.н. (PhD) (Узбекистан)

Технический редактор: Баходир Туралиев

Institutor:
Termez State University

PHILOLOGICAL RESEARCH: LANGUAGE, LITERATURE, EDUCATION

FILOLOGIK TADQIQOTLAR: TIL, ADABIYOT, TA'LIM

ФИЛОЛОГИЧЕСКИЕ ИССЛЕДОВАНИЯ: ЯЗЫК, ЛИТЕРАТУРА, ОБРАЗОВАНИЕ

SCIENTIFIC-INNOVATIONAL JOURNAL

Registered by the Agency for Information and Mass Communications under the President's Administration of the of the Republic of Uzbekistan on October 19, 2020 under number 1122.

The authors are responsible for the accuracy of the examples, passages and data provided in the texts. The source must be specified when copying from the journal.

Approved for printing: 28.10.2024
Size of the paper: 70x85
Printed sheet: 5.94. Offset printing.
Offset paper. Circulation: 100
Price is negotiable
Order: 155. Printed at the «Printing House of the TerSU»
43, Barkamol avlod str. Termez city.

Address of Editorial Office:
190111, Barkamol avlod str., 43,
Termez city.
Phone: +998762231649

E-mail:
philologicalresearch@umail.uz

Official website:
https://journaltersu.uz

EDITOR IN-CHIEF:
Abdukodir TASHKULOV, doctor of economic sciences, professor

EDITORIAL BOARD:

Bekpulat Umurkulov, doctor of philological sciences, professor (Uzbekistan) (deputy chief editor)
Shohista Makhmaraimova, doctor of philological sciences, professor (Uzbekistan) (editor-in-chief)
Ural Kholiyorov, PhD, associate professor (Uzbekistan)
Olmos Ulviy Binnatova, doct.of philol. sciences, professor (Azerbaijan)
Karl Rayxl, doct.of philol. sciences, professor (Germany)
Vali Salvash Elok, doct. of philol. sciences, professor (Turkey)
Dorota Silvia, PhD, associate professor (Poland)
Muslihiddin Mukhiddinov, doct.of philol. sciences, professor (Uzbekistan)
Ibodulla Mirzaev, doct.of philol. sciences, professor (Uzbekistan)
Abdi Mmatov, doct.of philol. sciences, professor (Uzbekistan)
Suyun Karimov, doct.of philol. sciences, professor (Uzbekistan)
Muhammadkhon Khakimov, doct.of philol. sciences, professor (Uzbekistan)
Nafas Shodmonov, doct.of philol. sciences, professor (Uzbekistan)
Dilmurod Kuronov, doct. of philol. sciences, professor (Uzbekistan)
Bakhtiyor Mengliev, doct. of philol. sciences, professor (Uzbekistan)
Normat Yuldashev, cand.of philol. sciences, associate professor (Uzbekistan)
Abdulla Khudaykulov, PhD (Uzbekistan)
Bobonazar Murtazaev, cand.of philol. sciences, associate professor (Uzbekistan)
Askar Eshmuminov, doct.of philol. sciences, associate professor (Uzbekistan)
Obidjon Shofiev, PhD, associate professor (Uzbekistan)
Zafar Umurkulov, PhD (Uzbekistan)
Nurali Amonturdiev, PhD (Uzbekistan)
Hakim Akhmedov, PhD (Uzbekistan)
Ahad Shukurov, PhD (Uzbekistan)
Eshnazar Jabbarov, PhD (Uzbekistan)
Abdumurod Arslonov, PhD (Uzbekistan)
Sirojiddin Turdimurodov, PhD (Uzbekistan)
Umida Rakhimova, PhD (Uzbekistan)
Gulchekhra Begmatova, PhD (Uzbekistan)
Nargis Kurbanazarova, PhD (Uzbekistan)
Muyassar Kholova, PhD (Uzbekistan)
Sherzod Turdikulov, PhD (Uzbekistan)
Zaynura Kadirova, PhD (Uzbekistan)
Humayun Nadim Muhammad, PhD (Uzbekistan)

Technical editor: Bakhodir Turaliyev

МУНДАРИЖА

TILSHUNOSLIK

JO‘RAYEVA BIBISH MUXSINOVNA,
TO‘XTAYEVA DILAFRUZ ASHUROVNA
Intensivlik va darajalanish munosabati 7

ТОЖИЕВ ОЛИМЖОН МУЙДИНОВИЧ
Ўзбек тилида равиш/сифат типидаги лексемаларда мазмуний синкретизм 12

UMURQULOV ZAFAR BEKPO‘LATOVICH
Til – hissiyotlar ifodalash vositasi 16

MARDIYEV BEGALI XUDOYQUL O‘G‘LI
Shaxs nomlarining lug‘aviy qatlamdagi o‘rni va tavsifiy belgilari 21

QODIROVA MUKADDAS TOG‘AYEVNA
Ingliz tili biznes diskursida dunyoning paremiologik manzarasi.....24

JO‘RAYEVA NASIBAXON GANIYEVNA
Tarixiy romanlar diskursida so‘zlarning
o‘zaro ma’no munosabatlari 29

RO‘ZIMATOVA DILNOZA MURODOVNA
Ijtimoiy maqomning sotsiolingvistik talqini 32

ZAYNUTDINOVA GULNOZA ANVARJON QIZI
Presedent birliklar va turg‘un birikmalarning
badiiy matndagi poetik ahamiyati (Shuhrat Orif ijodi misolida).....37

VASLIDDINOVA KAMOLA QODIRJON QIZI
Korpus platformalarining til tadqiqotidagi amaliy ahamiyati.....42

НАРМУРОДОВА ДИЛФУЗА АБДУХОЛИК КИЗИ
Типологические и генетические характеристики
терминосистемы civil engineering/строительство 46

YADIGAROVA SITORA BAXRAMOVNA
Phraseological units with clothing components in english language.....49

МУНДАРИЖА

ADABIYOTSHUNOSLIK

BOLTAYEVA GULCHEHRA SHOKIROVNA
Muxammaslarda ijtimoiy muammolar tasviri.....51

СМАМУТОВА ШИРИН АЗАТБАЕВНА
Қазіргі қазақ-қарақалпақ прозасындағы әдеби байланыстар.....54

АБДУЛЛАЕВ МУҲАММАД ҲАМЗАЕВИЧ
Убайдий рубоийнавис шоир.....57

ГАСЫМОВА СЕВИНДЖ АШУР КЫЗЫ
Традиции праздника новруз в Азербайджане.....60

NADIM MUHAMMAD HUMAYUN
Alisher Navoiying «Hayratul-abror» dostoniga ishlangan «Hammomdagi uchrashuv» miniatyurasiga sharh va Navoiy qurdirgan ba'zi inshootlar.....66

JUMAYEV ILHOM HAMROYEVICH
O'zg'urmish – komil zohid timsoli.....71

КЕНЖЕБАЕВА РОЗА САЙЛАУОВНА
Фариза оңғарсынова және қарақалпақ Ақын-жазушылары арасындағы достық әдеби байланыс.....78

ALIYEVA ORZIGUL
Abdulla Oripov va Shavkat Rahmonning olam hamda insonga munosabatidagi o'ziga xoslik.....81

Filologik tadqiqotlar: til, adabiyot, ta'lim. 2024; 10
ISNN: 2181-1741 (Print)
ISNN: 2181-1725

KORPUS PLATFORMALARINING TIL TADQIQOTIDAGI AMALIY AHAMIYATI

Vasliddinova Kamola Qodirjon qizi,
O'zbekiston davlat jahon tillari universiteti tayanch doktoranti

Tilshunoslikka korpus konsepsiyasining kirib kelishi natijasida tilda turli nazariya va amaliy platformalar tendensiyasi oshib bormoqda. Dunyo miqyosida tilshunoslik nuqtayi nazaridan korpus til nazariyasida tilga doir amaliy loyihalar revolyutsiyasiga sabab bo'ldi desak mubolag'a bo'lmaydi. Ijtimoiy hayotda turli qatlamlar leksikasini ifodalovchi korpus platformalari paydo bo'ldi. Ayniqsa Yevropa mamlakatlarida inson ijtimoiy-ruhiy holatini ifodalash yoki insonlar psixologik holatini baholash tahlillari uchun maxsus korpus yordamida turli sentiment analizlar olib borish amaliy tilshunoslikda tobora ommalashib bormoqda. Bugungi kunda o'zbek tilshunosligi oldida turgan muhim vazifalardan biri til tadqiqiga doir turli amaliy platformalar nazariyalarini tadqiq qilish va o'zbek milliy korpusini yanada mukammallashtirish orqali kelgusi tilga doir amaliy tadqiqotlar uchun tahlillar olib borishdir.

Parallel korpus sohasidagi eng mashhur loyihalardan biri Lancaster universiteti tadqiqotchilari Paul Baker, Andrew Hardie, Tony McEnery, Richard Xiao va Shiffield universiteti tadqiqotchilari Kalina Bontcheva Hamish Cunningham, Robert Gaizauskas, Oana Hamza, Diana Maynard, Valentin Tablan, and Cristian Ursu tomonidan yaratilgan ***"Emille (Enabling Minority Language Engineering) loyihasi"*** hisoblanib, 14ta Janubiy Osiyo mamlakatlari tillari doirasida 96 million so'zdan iborat monolingval va parallel korpusdan iborat katta hajmli platforma hisoblanadi. Uning funksional imkoniyatlari ham murakkab bo'lib, 14ta til doirasida izohli komponent, nomlash va nutq qismlari bo'yicha teglash kabi imkoniyatlarga ega. Keyinchalik bu platforma Janubiy Osiyo (Bengal, Gujarot, Hind, Panjab, Sinhalis, Tamil, Urdu) mamlakatlari tillari uchun umumiy til injiniringini qo'llash xususiyatlarini tadqiq qilish va til injiniringi uchun funksional qurilmalar yaratish uchun asos bo'ldi.

Pos teglash ingliz tilidan PoS tagging (Parts of Speech tagging) so'zining aynan tarjimasi bo'lib, matn tarkibida kelgan har bir so'zning grammatik kategoriyasi bo'yicha avtomatik ravishda aniqlab berishdir.

Loyihaning ilmiy jihatdan yuqori salohiyatda bajarilganligi ingliz va urdu tillari parallel korpusida teglash jarayonlari "DG XIII Linguistic research and Engineering" dasturi asosida tuzilgan Yevropa komissiyasining ***"The Expert Advisory Group on Language Engineering Standards (EAGLES)"***[1] yo'riqnomasi asosida shakllantirganligi bilan izohlashimiz mumkin. Yevropa komissiyasining ***"EAGLES"*** yo'riqnomasi til injiniringi kabi amaliy tilshunoslik masalalarida tadqiqot olib borish uchun algoritmlar izchilligini shakllantirib bergan va Yevropa tilshunoslari jamoasi hamda yirik tadqiqot markazlar tomonidan muhim funksional tamoyil sifatida baholanadi. Yevropa mintaqasida yaratilgan korpuslarning til injiniringi funksional imkoniyatlari ***"EAGLES"*** yo'riqnomasiga ko'ra ish ko'rishini izlanishlarimiz davomida guvohi bo'ldik. Masalan, websaytlarni qurishda foydalaniladigan *HTML markup* (rasmlar uchun alternative matn nomlari, mobil qurilmalar uchun ovozli boshqaruv), *screen reader compatibility* (ekranga matnni qaratish orqali avtomatik tarjima qilish funksiyasi), *Pdf*, *Adobe Reader* (matn bilan ishlovchi dasturlar) kabi dunyo miqyosidagi eng ommaboq platformalarning barchasi til injeneringi qoidalari va algoritmlarini ***"EAGLES"*** yo'riqnomasiga binoan ishlab chiqadi.

Korpus platformalarini tahlil qilar ekanmiz asosan analitik, analiz, sintez, qiyosiy, metodlarga bevosita murojat qildik. Til platformalarining amaliy ahamiyati natijasida tillarda sodir bo'lgan yangi terminlar va progreslarni solishtirdik. Ayniqsa, platformalar yordamida tilda monolingval, bilingval va multilingval lug'atlarning tuzilishi natijasida leksikografik tadqiqotlarga ham manba vazifasini bajarganini ta'kidlashimiz joiz.

Paralel korpus sohasidagi eng mashhur platforma bu – **"EUROPARL"** bo'lib, olimlar tomonida dastlab, statistik mashina tarjimasi uchun manba sifatida yaratilgan bo'lsada keyinchalik, NLP texnologiyalari uchun muhim lingvistik masalalarni hal etishda yordam berdi. Masalan, matn tarkibidagida ko'p ma'noli so'zlarning mos variantini saralab olish, *anaphoric resolution* ya'ni birinchi gap tarkibida kelgan so'zlarning ikkinchi gapda grammatik vositalar yoki diskurs sifatida mantiqiy bog'lanishi, tektsdan ma'noni ajratib olish kabi NLP texnologiyalarining rivojlanishi sabab

[1] Eagles Guidelines | Semantic Scholar. 19.06.2024.

bo‘ldi.

Dastlab **“EUROPARL”** Yevropa parlamenti websaytida e’lon qilingan 11ta mamlakatning tillari asnosida tuzilgan bo‘lib, bugungi kunda 21ta tillarni (fransuz, italyan, ispan, portugal, roman, nemis, golland, daniya, shved, bolgar, chexiya, polsha, slovakiya, sloveniya, fin, vengar, latviya, litva, yunon) ingliz tilidagi muqobili bilan parallel korpuslari yaratilgan.

Language	Days	Chapters	Speaker Turns	Sentences	Words
Danish (da)	492	4,120	90,017	1,032,764	27,153,424
German (de)	492	4,119	90,135	1,023,115	27,302,541
Greek (el)	398	3,712	66,928	746,834	27,772,533
English (en)	488	4,055	88,908	1,011,476	28,521,967
Spanish (es)	492	4,125	90,305	1,029,155	30,007,569
French (fr)	492	4,125	90,335	1,023,523	32,550,260
Finnish (fi)	442	3,627	81,370	941,890	18,841,346
Italian (it)	492	4,117	90,030	979,543	28,786,724
Dutch (nl)	492	4,122	90,112	1,042,482	28,763,729
Portuguese (pt)	492	4,125	90,329	1,014,128	29,213,348
Swedish (sv)	492	3,627	81,246	947,493	23,535,265

2.2-rasm. **“EUROPARL”ning dastlabki varianti. Tokenlash va lemmalash jarayonidan so‘ng gaplar miqdori**[1]

Bizning fikrimizcha, bugungi kunda har bir til kesimida 60 milliondan ziyod so‘zni qamrab olgan yuqorida keltirib o‘tilgan platforma statistik mashina tarjimasi, neyron mashina tarjimasi, NLP texnologiyalari, bilingval lug‘atlar tuzish, tilning funksional imkoniyatlarini solishtirish uchun yuqori samaradorlikni ta’minlab beruvchi manba hisoblanadi.

Parallel korpusning vazifaviy xususiyatlari jihatidan har bir tilda parallel korpuslar yaratish dunyo olimlari tomonidan jadallik bilan boshlab yuborilgan. Jumladan Boltiqbo‘yi mamlakatlarida ham korpusshunoslik sohasida jadal ilmiy tadqiqot ishlari olib borildi. Masalan adabiyot janrlari asosida ***Multext-East corpus, SEE-ERA.net Literary Corpus***, parallel korpuslar asnosida ***RuN-Euro Corpus, ParaSol, Regensburg Parallel Corpus***, ***Bulgarian-Polish-Lithuanian Corpus,*** siyosiy matnlardan tarkib topgan ***SEE-ERA.NET Administrative Corpus,*** yangiliklardan tarkib topgan matnlardan iborat ***SETimes news corpus***, tibbiyotga oid matnlardan iborat ***EMEA*** kabi korpuslar tuzilgan[2]. Bu korpuslarning har biri funksional jihatdan tokenlash, POS teglash, lemmalash xususiyatlariga ega. Lekin bu korpuslardagi matnlar asosan badiiy va rasmiy uslubga xos materyallardan iborat bo‘lib, keyinchalik uslub va miqdor jihatidan yuqori salohiyatdagi ***the Bulgarian National Corpus (BulNC)*** tarkibida ***“Bulgarian X-language Parallel Corpus (Bul-X-Cor)”*** subkorpusi web resursidagi matnlarni Internet browsing, web crawling (umumiy, xususiy) orqali avtomatik shakllantirish yondashuvlaridan foydalanib tuzildi[3]. Bizning fikrimizcha, ***“Bulgarian X-language Parallel Corpus (Bul-X-Cor)”*** subkorpusi boshqa parallel korpuslardan gapni segmentlarga ajratish va gapni muqobili bilan juftlash uchun annotatsiyalashning bir necha bosqichlarga ajratib olgan, vaholanki boshqa parallel korpuslarda bu kabi funksiyalar maxsus avtomatik funksional qurilmalar (morfoanalizator, morfo-sintaktik teglash, lemmalash, sintaktik tahlil, so’z ma’nosini aniqlashtirish) mavjud ekanligiga bevosita bog‘liq bo‘ladi.

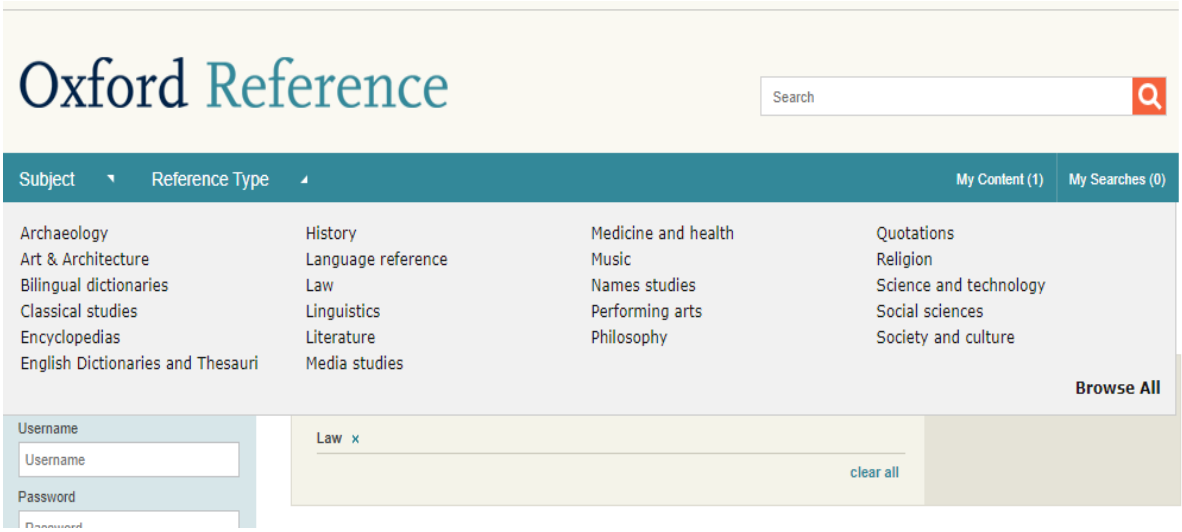
Korpusga asoslangan deskriptiv leksikografik yondashuv orqali korpusga yeg‘ilgan til birliklarining dialektal, sotsiologik, idiomatik, pragmatik kontekstual analizlar orqali lug‘at tuzish ishlari jadallashdi. Ayniqsa so‘nggi yillarda korpus yordamida sotsiolingvistik kontekstual analizlar yordamida jins, yosh, muayyan

[1] Koehn, P 2005, Europarl: A Parallel Corpus for Statistical Machine Translation. in J Hutchins (ed.), The Tenth Machine Translation Summit Proceedings of Conference. International Association for Machine Translation, pp. 79-86.

[2] Svetla Koeva, Ivelina Stoyanova, Rositsa Dekova, Borislav Rizov, Angel Genov. Bulgarian X-language Parallel Corpus. International Conference on Language Resources and Evaluation.2012. Pp 2480-2486. https://api.semanticscholar.org/CorpusID:12984431

[3] Koeva S, Stoyanova I, Dekova R, Rizov B, Genov A. Bulgarian X-language Parallel Corpus. International Conference on Language Resources and Evaluation.2012. Pp 2480-2486. CorpusID:12984431 https://api.semanticscholar.org/

ijtimoiy guruhlarning ehtiyojiga moslashtirilgan lug‘atlar yaratildi. Masalan elektronlashtirilgan 25 turdagi sohaga doir so‘zlarni guruhlarga tizimlashtirgan ***“OXFORD REFERENCE PLATFORMASI”*** ning asosi ***“Oxford English Dictionary”*** ga borib taqaladi. Bu platforma sotsiologik guruhlar uchun maxsus lug‘atlarni shakllantirgan.



3.1-rasm. **“OXFORD REFERENCE PLATFORMASI”da so‘zlarning guruhlashtirish kategoriyasi**[1]

Parlce *(http:// malheildir.arnastofnun yoki http:// www. malfong)* yana bir muhim, yirik loyiha hisoblanadi. Til texnologiyalarining rivoji, lug‘at va ontologiya tuzish, lingvistik matnlarni klassifikatsiyalash, turli tillaro annotatsiyalash muammolariga yechim topish maqsadida island va ingliz tillari doirasida parallel korpus tuzilgan. Islandiyada amalga oshirilgan bu yirik platformada 3.5 million gap juftliklari mavjud bo‘lib, ***BiLSTM tagger*** funksiyasi orqali teglanadi, ***Nefnir*** orqali lemmalarga ajratiladi, ***GIZA++*** orqali so‘zlar juftlanadi[2]. Lekin ***Parlce***ga boshqa mashhur parallel korpus platformalaridagi tayyor segmentlarga ajratilgan gaplari yuklab olingan, masalan, asosan The **OPUS** til platformasidan, **The TILDE MODEL CORPUS**idan, **ELRC** va **The Bible** yirik hajmli multilingval til korpusidagi tayyor matnlardan ham island tiliga oid misollar olingan.

Corpus	Sentence pairs	Bad align-ments (%)
The Bible[4]	31,085	0.5
EMA[8]	420,297	3.3
Gnome[9]	5,431	n/a
KDE4[9]	87,575	45.0
OpenSubtitles[9]	1,368,170	8.3
Statistics Iceland[10]	2,360	8.0

[1] https://www.oxfordreference.com/browse?t0=ORO:LAW00010.24.072024

[2] Starkadur B. Compiling and Filtering Parlce: An English-Icelandic Parallel Corpus //Proceedings of the 22nd Nordic Conference on Computational Linguistics, - “Turku, Finland”, 2019 pp 140--145, https://www.aclweb.org/anthology/W19-6115

Tatoeba[9]	8,139	0.0
Ubuntu[9]	2,127	2.5
TOTAL	**1,923,060**	

2.3-jadval. **Island-Ingliz parallel korpusida boshqa til platformalaridan olingan (alignment) islan-ingliz gap muqobillari juftliklari miqdori**[1]

Yuqorida keltirib o‘tilgan platformalarda OPUS va BIBLE platformalari haqida fikrlar bayon etildi. Parallel korpus yaratish nazariyasi borasida olib borilgan bu ikki yirik hajmli platformalar har qanday tadqiqot uchun katta hajmli ma’lumotlar bazasiga egadir. Ulardan OPUS 744ta tillar parallel korpusidan iborat bo‘lib, 45 milliarddan oshiq gap fragmentlari o‘z ichiga oladi[2]. BIBLE bilan OPUSning imkoniyatlarini solishtiradigan bo‘lsak, har ikkovi ham multilingval parallel korpus bo‘lib, BIBLE asosan xristianlik dining muqaddas kitobi Bibliyaning 102 tildagi tarjimalari asosida Christos Christodoulopoulos va Mark Steedman boshchiligida tuzilgan bo‘lib, 5,148 ta bitext, 107 ta fayl, 56 mingdan oshiq token va 2,84 ming gap fragmentlaridan iboratdir[3].

Yuqorida keltirib o‘tilgan qiyoslashdan shunday xulosaga kelishimiz mumkinki: OPUS orqali turli tillar kesimida gap muqobillarini bilingval va multilingval holatida saralash, turli tillaro matnlarni qayta ishlash, stilistik jihatdan tarjimada qo‘llanish uslubiyatini solishtirish imkonini beradi, BIBLEda joylangan parallel matnlar esa bir asarning turli tillardagi tarjima variantlarini solishtirish orqali tarjima usullari, foydalanuvchiga ma’noni yetkazib berish imkoniyatlarini o‘zlashtirish orqali metodologik yondashuv sifatida yangi nazariyalarga ham asos bo‘ladi. Yuqorida keltirib o‘tilgan tilshunoslikka oid platformalarning har biri nafaqat tillarning saqlanib qolishi, uning ustida lingvistik, sotsiologik tadqiqot ishlari olib borish uchun ishonchli manba vazifasini bajarib kelmoqda.

Rezyume: Ushbu maqolada korpus platformalarining til nazariyasi va amaliy tilshunoslik sohasidagi ahamiyati borasida fikrlar misollar orqali batafsil yoritib berilgan. Ma’lumki, korpus konsepsiyasining tilshunoslikda rivoji yangi terminlarga, kompyuter lingvistikasida esa yangi funksional lingvistik instrumentlarga sabab bo‘ldi. Maqolada bugungi kunda xorijiy tillardagi eng ommabop til platformalari hamda platformalar asosida tadqiq qilingan lingvistik yutuqlar haqida ma’lumotlar berilgan.

Resume: This article illustrates the importance of corpus platforms in the fields of linguistic theory and applied linguistics through examples. It is known that the development of the concept of a corpus in linguistics has led to the emergence of new terms, and in computational linguistics, to new functional linguistic tools. The article provides information about the most popular language platforms today in foreign languages and the linguistic achievements studied based on them.

Резюме: В этой статье на примерах раскрывается важность корпусных платформ в области лингвистической теории и прикладной лингвистики. Известно, что развитие понятия корпуса в лингвистике привело к появлению новых терминов, а в компьютерной лингвистике – к новым функциональным лингвистическим инструментам. В статье представлена информация о наиболее популярных сегодня языковых платформах на иностранных языках и лингвистических достижениях, исследованных на их основе.

[1] Barkarson S. Compiling and Filtering ParIce: An English-Icelandic Parallel Corpus //Proceedings of the 22nd Nordic Conference on Computational Linguistics: - “Turku, Finland”, 2019 pp 140--145, https://www.aclweb.org/anthology/W19-6115

[2] https://opus.nlpl.eu/. 28.06.2024

[3] A massively parallel corpus: the Bible in 100 languages, Christos Christodoulopoulos and Mark Steedman, *Language Resources and Evaluation*, 49 (2)