

**TERMIZ DAVLAT UNIVERSITETI HUZURIDAGI ILMIY DARAJALAR
BERUVCHI PhD.03/31.03.2021.Fil.78.04
RAQAMLI BIR MARTALIK ILMIY KENGASH**

TERMIZ DAVLAT UNIVERSITETI

ESHMUMINOV ASQAR ALLAMURADOVICH

**O‘ZBEK TILI KORPUSLARI UCHUN TAKSONOMIK VA MERONIMIK
MUNOSABATLI LEKSEMALAR LINGVISTIK BAZASINI
SHAKLLANTIRISH**

10.00.01 – O‘zbek tili

**FILOLOGIYA FANLARI DOKTORI (DSc) DISSERTATSIYASI
AVTOREFERATI**

Termiz – 2022

Filologiya fanlari doktori (DSc) dissertatsiyasi avtoreferati mundarijasi

Оглавление автореферата доктора наук (DSc) по филологическим наукам

Content of Dissertation Abstract of Doctor of Science (DSc) on Philological Sciences

Eshmuminov Asqar Allamuradovich

O‘zbek tili korpuslari uchun taksonomik va meronimik munosabatli
leksemalar lingvistik bazasini shakllantirish..... 3

Eshmuminov Asqar Allamuradovich

Forming linguistic base of taxonomy and meronymically related lexemes
for Uzbek language corpus 33

Eshmuminov Asqar Allamuradovich

Формирование лингвистической базы таксономии и меронимически
связанных лексем для корпуса узбекского языка
..... 65

E’lon qilingan ishlar ro‘yxati

List of published works
Список опубликованных работ
..... 69

**TERMIZ DAVLAT UNIVERSITETI HUZURIDAGI ILMIY DARAJALAR
BERUVCHI PhD.03/31.03.2021.Fil.78.04
RAQAMLI BIR MARTALIK ILMIY KENGASH**

TERMIZ DAVLAT UNIVERSITETI

ESHMUMINOV ASQAR ALLAMURADOVICH

**O‘ZBEK TILI KORPUSLARI UCHUN TAKSONOMIK VA MERONIMIK
MUNOSABATLI LEKSEMALAR LINGVISTIK BAZASINI
SHAKLLANTIRISH**

10.00.01 – O‘zbek tili

**FILOLOGIYA FANLARI DOKTORI (DSc) DISSERTATSIYASI
AVTOREFERATI**

TERMIZ– 2022

Fan doktori (DSc) dissertatsiyasi mavzusi O'zbekiston Respublikasi Vazirlar Mahkamasi huzuridagi Oliy attestatsiya komissiyasida B2021.1.DSc/Fil284 raqam bilan ro'yxatga olingan. Tadqiqot Termiz davlat universitetida bajarilgan.

Dissertatsiya avtoreferati uch tilda (o'zbek, rus, ingliz (rezyume) Termiz davlat universiteti veb-sahifasida (www.tersu.uz) hamda "ZiyoNet" Axborot ta'lim portalida (www.ziynet.uz) joylashtirilgan.

Ilmiy maslahatchi:

Mengliev Baxtiyor Rajabovich
filologiya fanlari doktori, professor

Rasmiy opponentlar:

Pardayev Azamat Baxranovich
filologiya fanlari doktori, professor

Abjalova Manzura Abdurashetovna
filologiya fanlari doktori, dotsent

Ergasheva Guli Ismailovna
filologiya fanlari doktori, dotsent

Yetakchi tashkilot:

Farg'ona davlat universiteti

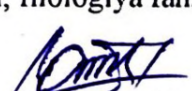
Dissertatsiya himoyasi Termiz davlat universiteti huzuridagi ilmiy darajalar beruvchi PhD.03/31.03.2021.Fil.78.04 raqamli Ilmiy kengashning 2022 yil «22» avgust soat 10⁰⁰ dagi majlisida bo'lib o'tadi (Manzil: 190111, Termiz sh., Barkamol avlod ko'chasi, 43-uy. Tel: (876) 221-74-55; faks: (876) 221-71-17; e-mail: termizdu@umail.uz)

Dissertatsiya bilan Termiz davlat universitetining Axborot-resurs markazida tanishish mumkin (67 raqami bilan ro'yxatga olingan). Manzil: Termiz sh., Barkamol ayloq ko'chasi, 43-uy. Tel: (876) 221-74-55.

Dissertatsiya avtoreferati 2022-yil «11» avgust da tarqatildi.
(2022-yil «11» 08 dagi 12 - raqamli reestr bayonnomasi)




B. Umurqulov
Ilmiy darajalar beruvchi bir martalik ilmiy kengash raisi, filologiya fanlari doktori, professor


N. Omonturdiyev
Ilmiy darajalar beruvchi bir martalik ilmiy kengash kotibi, filologiya fanlari falsafa doktori (PhD)


Sh. Mahmaraimova
Ilmiy darajalar beruvchi bir martalik ilmiy kengash qoshidagi Ilmiy seminar raisi, filologiya fanlari doktori, dotsent

KIRISH (doktorlik (DSc) dissertatsiyasi annotatsiyasi)

Tadqiqot mavzusining dolzarbligi va zarurati. Jahon kompyuter lingvistikasi XX asrning 50-yillaridan rivojlana boshladi, shu vaqtdan boshlab mashina tarjimasi kompyuter lingvistikasida muhim natijalarga erishildi. XX asr oxiriga kelib tabiiy tilga avtomatik ishlov berish, kompyuter dasturlari orqali tarjima masalalarini yechish davri boshlandi. XIX asr oxiri XX asr boshlariga kelib til korpuslari yaratish korpus birliklarini morfologik, semantik teglashning dasturiy vositalarini yaratish, buning uchun lingvistik ta'minot ishlab chiqishga e'tibor oshdi. Til birliklarini korpuslar uchun semantik teglashda, asosan, leksik-semantik munosabatlarni aniqlash, bu munosabatlarni qidiruv tizimida aks ettirish dolzarb muammoga aylandi. Jahondagi peshqadam korpuslarning taksonomik, meronimik munosabatli leksemalar bazasini yaratishning nazariy asoslari ishlab chiqildi.

Dunyo tilshunosligida XXI asr boshlarida lingvistik korpuslarni sifatli izohga ega bo'lgan leksik resurslar bilan ta'minlash, korpusdan foydalanish sifatini yaxshilash, tilni lingvistik modellashtirish kabi masalalar kun tartibiga qo'yildi. Natijada lingvistik korpusga ega bo'lgan tillar korpusida taksonomik va meronimik ma'lumotlar omborini tuzish uchun shu tilda mavjud bo'lgan leksik birliklarni axborot-qidiruv tizimida faol qo'llash ommalashdi. Binobarin, til korpuslarini yaratish, ularni lingvistik annotatsiyalash, taksonomik va meronimik ma'lumotlar omborini tuzish korpus lingvistikasining ilmiy-nazariy muammolariga aylandi. Muammo yechimi sifatida ma'lum bir tabiiy tilning birliklariga ishlov berish, til birliklarini avtomatik morfologik, semantik tahlil qilish vositalari ishlab chiqish tajribasi ommalashdi.

O'zbek korpus lingvistikasida milliy, ta'limiy, mualliflik, parallel korpuslari hamda ularning lingvistik ta'minotini ishlab chiqishning nazariy asoslari sistemali tadqiq etilmoqda. Bular sirasida turli korpuslarni yaratish borasidagi ilmiy va amaliy ishlar diqqatga sazovor. Ushbu ishlarni amalga oshirishda o'zbek tili milliy korpusi uchun taksonomik va meronimik munosabatli leksemalarning ma'lumotlar bazasini tuzish masalasi ilmiy-nazariy planda o'rganilmagan. Zero, "...davlat tilining sofliğini saqlash, uni boyitib borish va aholining nutq madaniyatini oshirish; davlat tilining zamonaviy axborot texnologiyalari va kommunikatsiyalariga faol integratsiyalashuvini ta'minlash"¹ hozirgi kunda o'zbek kompyuter lingvistikasi oldida turgan dolzarb vazifalardan biridir. Shunga ko'ra, o'zbek tili korpuslari uchun taksonomik va meronimik munosabatli leksemalarning ma'lumotlar bazasini tuzish hamda ularning lingvistik ta'minotini yaratish masalalari dolzarblik kasb etmoqda.

O'zbekiston Respublikasi Prezidentining 2016-yil 13-maydagi PF-4997-son "Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universitetini tashkil etish to'g'risida", 2017-yil 7-fevraldagi PF-4947-son "O'zbekiston Respublikasini yanada rivojlantirish bo'yicha Harakatlar strategiyasi to'g'risida", 2019- yil 21- oktabrdagi PF-5850-son "O'zbek tilining davlat tili sifatidagi nufuzi va mavqeyini tubdan oshirish chora-tadbirlari to'g'risida", 2020-yil 20-oktabrdagi PF-6084-son "Mamlakatimizda o'zbek tilini yanada rivojlantirish va til siyosatini

¹Ўзбекистон Республикаси Президенти Шавкат Мирзиёевнинг 2020 йил 20 октябрдаги «Мамлакатимизда ўзбек тилини янада ривожлантириш ва тил сиёсатини такомиллаштириш чора-тадбирлари тўғрисида»ги ПФ-6084-сон фармони // <https://lex.uz/docs/5058351>

takomillashtirish chora-tadbirlari to'g'risida"gi farmonlari, 2017-yil 17-fevraldagi PQ-2789-son "Fanlar akademiyasi faoliyati, ilmiy-tadqiqot ishlarini tashkil etish, boshqarish va moliyalashtirishni yanada takomillashtirish chora-tadbirlari to'g'risida", 2019-yil 4-oktabrdagi PQ-4479-son "O'zbekiston Respublikasining "Davlat tili haqida"gi Qonuni qabul qilinganligining o'ttiz yilligini keng nishonlash to'g'risida"gi qarorlari, 2020-yil 24-yanvardagi Oliy Majlisga Murojaatnomasi hamda boshqa me'yoriy-huquqiy hujjatlarda belgilangan vazifalarni amalga oshirishda ushbu tadqiqot muayyan darajada xizmat qiladi.

Tadqiqotning respublika fan va texnologiyalar taraqqiyoti ustuvor yo'nalishlariga mosligi. Tadqiqot respublika fan va texnologiyalari rivojlanishining I."Axborotlashgan jamiyat va demokratik davlatni ijtimoiy, huquqiy, iqtisodiy, madaniy, ma'naviy-ma'rifiy rivojlantirish, innovatsion iqtisodiyotni rivojlantirish" ustuvor yo'nalishiga muvofiq bajarilgan.

Tadqiqot mavzusi bo'yicha xorijiy ilmiy tadqiqotlar sharhi². Jahon tilshunosligida yaratilayotgan Milliy korpuslar uchun lingvistik ta'minotini o'rganishga yo'naltirilgan ilmiy izlanishlar yetakchi ilmiy markazlar va oliy ta'lim muassasalarda, jumladan, Brown University (AQSh), Oxford University (Buyuk Britaniya), Cambridge University (Buyuk Britaniya), Humboldt Universitat (Germaniya), Boston University, University of New York, University of Pennsylvania, University of California (AQSh); Indiana University (AQSh), The University of Sorbonne (Fransiya), Universitat de Valencia (Ispaniya), University of Zurich (Shveysariya), Upsalla universiteti (Shvetsiya), Kiofo University (Yaponiya), Australian National University (Avstraliya), Istambul texnika universiteti (Turkiya), Moskva davlat universiteti (Rossiya), Moskva davlat lingvistika universiteti (Rossiya), Sankt-Peterburg davlat universiteti (Rossiya), Ural davlat pedagogika universiteti, Xarkov davlat pedagogika universiteti (Ukraina), Baku davlat universiteti (Ozarbayjon), Tatariston Respublikasi Fanlar Akademiyasi amaliy semiotika instituti (Tatariston Respublikasi, Rossiya), Qozon Federal universiteti (Tatariston Respublikasi, Rossiya), V.Vernardskiy nomidagi Tavriya akademiyasi Qrim federal universiteti (Qrim Respublikasi, Rossiya), shuningdek, Mirzo Ulug'bek nomidagi O'zbekiston Milliy universiteti, Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti, Al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti va uning filiallari, Buxoro davlat universiteti, Samarqand davlat universiteti, Termiz davlat universiteti (O'zbekiston)da olib borilmoqda.

Dunyo tilshunosligida lingvistik korpuslar uchun axborot-qidiruv tizimlarining lingvistik ta'minoti, til birliklarini avtomatik teglash, matnni lingvistik va ekstralingvistik annotatsiyalashga oid olib borilgan izlanishlardan quyidagi ilmiy natijalar olingan: tabiiy tilga avtomatik qayta ishlov berish (til birliklarini avtomatik tahlil qilish) xususiyatlari yoritilgan (University of New Jersey, University of California, AQSh; Boston University, New York University, AQSh; Universitat de Valencia, Ispaniya); leksik tahlilning nazariy asoslari, umumiy tavsifi, ishlash tamoyillari, dasturiy ta'minot uchun lingvistik modullar tavsifi ishlab chiqilgan (Kiofo

²Диссертация мавзуси бўйича хорижий илмий тадқиқотлар шарҳи [google.scholar.com.](https://www.google.com/scholar); www.bu.edu; americanenglish.state.gov; micros.uz; www.uv.es; www.zora.uzh.ch; clea@snu.ac.kr; www.u-tokyo.ac.jp/en; www.ras.ru; adwww.ecu.edu.au; unice.fr/en; www.sophia.ac.jp/eng; www.msu.ru; www.vsu.bu; www.education.ua/universities; www.kaznu.kz; www.navoiy-uni.uz ва бошқа манбалар асосида амалга оширилди.

University, Yaponiya; Australian National University, Avstraliya; Istanbul texnika universiteti, Turkiya; Moskva davlat universiteti, Moskva davlat lingvistika universiteti, Sankt-Peterburg davlat universiteti, Rossiya); lingvistik korpusning leksik birliklarga avtomatik ishlov berish moduli va dasturlari (Baku Davlat universiteti, Ozarbayjon); turkiy tillarga avtomatik ishlov berish nazariyasi va amaliyoti ishlab chiqilgan, til korpuslari, mashina tarjimonlari ishlab chiqilgan (Tatariston Respublika Fanlar akademiyasi amaliy semiotika instituti, Qozon federal universiteti, V.Vernardskiy nomidagi Tavriya akademiyasi Qrim federal universiteti, Rossiya); o'zbek tili mualliflik korpusi, til korpusining lingvistik bazasini tuzish tamoyillari va uning lingvistik ta'minoti, o'zbek tili korpuslari uchun taksonomik va meronimik munosabatli leksemalar bazasini yaratish muammolari, o'zbek-ingliz parallel korpusi instrumentariylarini yaratish asoslari ishlab chiqilgan (Mirzo Ulug'bek nomidagi O'zMU, Alisher Navoiy nomidagi ToshDO'TAU, Samarqand davlat universiteti, Buxoro davlat universiteti, Al-Xorazmiy nomidagi TATU, Termiz davlat universiteti O'zbekiston).

Dunyo korpus lingvistikasida tabiiy tillarning ma'lumotlar bazasi asosida lingvistik korpus interfeysi yaratishning quyidagi ustuvor yo'nalishlarida ilmiy tadqiqotlar olib borilmoqda: til korpuslari uchun lingvistik ta'minot ishlab chiqish; ishlab chiqilgan lingvistik ta'minot uchun tegishli teglar biriktirish; grafematik tahrir modullari, ontologiyalar ishlab chiqish, yaratilayotgan lingvistik korpuslar uchun taksonomik, meronimik munosabatli leksemalar bazasini shakllantirish.

Muammoning o'rganilganlik darajasi. Jahon tilshunosligida kompyuter lingvistikasining nazariy, metodologik muammolarini nazariy va amaliy o'rganish o'tgan asrning 50-60-yillarida boshlandi. Jahon kompyuter lingvistikasida tabiiy tilga ishlov berish muammolari A.Keyler, K.V.Linden, N.Vard³, M.A.Mohri⁴ tomonidan yoritildi. Matnga avtomatik ishlov berish texnologiya va vositalari V.Axo Alfred, S.Lam Monika, S.Ravi, D.Ulman Djeffri⁵, morfologik tahlilning nazariy asoslari S.Y.Toldova, AA.Bonch-Osmolovskaya⁶, T.Sadikov⁷ ishlarida tavsiflandi. Rus kompyuter lingvistikasida O.V.Dereza, D.A.Kayutenko, A.S.Fenogenova, M.Hakimovning morfologik analizator, uning umumiy tavsifi, ishlash prinsiplari, tilga oid grammatik kategoriya, dastur uchun lingvistik modullar tavsifi haqidagi ishlarini ta'kidlash o'rinli⁸. Axborot-qidiruv tizimida lingvistik ta'minotning ahamiyati A.B.Antopolskiy, Yu.M.Arskiy, G.G.Belonogov, B.A.Kuznesov, V.N.Beloozerov,

³ Speech and Language Processing. Daniel S. Jurafsky and James H. Martin. Contributing writers: Andrew Kehler, Keith Vander Linden, Nigel Ward 2000y. Prentice Hall, Englewood Cliffs, New Jersey 07632. Pages: 950.

⁴ Mohri M.A. Finite-state transducers in language and speech processing. Computational Linguistics. – С. 269-312.

⁵ Ахо Альфред В., Лам Моника С., Сети Рави, Ульман Джеффри Д. Компиляторы: принципы, технологии и инструментарий, 2-е изд.: Пер. с англ. – М.: ООО “И.Д. Вильямс”, 2008. – 1184 с.

⁶ Толдова С.Ю., Бонч-Осмоловская А.А. Автоматический морфологический анализ // Фонд знаний «Ломоносов». М., 2011. www.lomonosov-fund.ru/enc/ru/encyclopedia:0127430

⁷ Садыков Т., Кочконбаева Б. Об оптимизации алгоритма морфологического анализа // Шестая Международная конференция по компьютерной обработке тюркских языков «Turklang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

⁸ Дереза О.В., Каютенко Д.А., Феногенова А.С. Автоматический морфологический анализ для русского языка сравнительный анализ систем // Шестая Международная конференция по компьютерной обработке тюркских языков «Turklang-2018». (Труды конференции) – Ташкент, 2018.; Хакимов М. Логико-лингвистические модели русского языка // Шестая Международная конференция по компьютерной обработке тюркских языков «Turklang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

R.R.Mdivani, B.A.Glinskiy, A.I.Mixaylov, V.A.Mishin, V.V.Morozov⁹ ishlarida ochib berilgan. Morfoanaliz xususiyatlari D.Rudolf, morfoanalizning amaliy ahamiyati A.Yermakov¹⁰, matnga avtomatik morfologik va sintaktik ishlov berish modeli va dasturlari I.M.Nojoy¹¹ tomonidan o'rganilgan.

Turkiy kompyuter lingvistikasida tabiiy tilning kompyuter ishlovi masalalari atroflicha tadqiq etildi. Bunday ishlar sirasiga G.Eryigit, Ye.Adali¹², A.V.Dibo, A.V.Sheymovich¹³, J.Suleymanov¹⁴, A.R.Gatiatullin, A.M.Bashirov¹⁵, P.V.Jeltov¹⁶, N.A.Israilova, P.S.Bakasova¹⁷, N.A.Leontev¹⁸, V.V.Kukanova, A.Y.Kadjiev¹⁹larning korpus tuzish masalalariga bag'ishlangan tadqiqotlarini kiritish mumkin. Turkiy tillar korpusining tavsifi, tahliliga bag'ishlangan tadqiqotlar²⁰ ko'lamining kengligi ushbu

⁹ Антопольский А.Б. Лингвистическое обеспечение АСНТИ: структура и проблема совместимости // НТИ. Сер.2, 1983. – №7. – С.17-25; Арский Ю.М. Информационные ресурсы для устойчивого развития общества / Ю.М.Арский, А.И.Черный // Междунар. форум по информ. и док. – 2000. – Т. 25, № 1. – С. 3-9.; Белоногов Г.Г. Автоматизированная обработка научно-технической информации. Лингвистические аспекты / Г.Г. Белоногов, Б.А. Кузнецов, А.П. Новоселов // Итоги науки и техники. Сер. Информатика / ВИНТИ. –1984. –Т.8. – 316; Белоозеров В.Н. Ведение и совершенствование Государственного рубрикатора НТИ (ГРНТИ) / В.Н. Белоозеров, В.И. Федосимов, Г.А.Кривенцова // Информ. продукты, процессы и технологии: НТИ-96: Материалы конф. – М., 1996. – С. 60-69.; Мдивани Р.Р. Лингвистическое обеспечение автоматизированной информационной системы по общественным наукам: состав и структура / Р.Р.Мдивани, В.А.Глинский // Вопр. информ. теории и практики. 1985. – №52. – С. 37-55.; Михайлов А.И. Проблемы лингвистического обеспечения АСНТИ / А.И. Михайлов // Вопр. информ. теории и практики. 1984.– № 51. – С. 5-11.; Мишин В.А. Лингвистическое обеспечение библиотечных автоматизированных систем / В.А. Мишин, Л. Санчес-Мехидо // Сов. библиотековедение. 1981. – № 2. – С. 84-93.; Морозов В.В. Лингвистическое обеспечение автоматизированных систем научно-технической информации / В.В. Морозов, Б.Р. Певзнер. –М.: ИПКИР, 1987. – 76 с.

¹⁰ Rodolfe D. Computational Linguistic Text Processing: Lexicon, Grammar, Parsing and Anaphora Resolution. Nova Science Publishers, Inc. – NewYork, 2008. – P. 4-5.; Ермаков А. Морфологический анализатор - основа поисковых систем // <https://www.kv.by/archive/index2004154301.htm>

¹¹ Ножов И.М. Морфологическая и синтаксическая обработка текста (модели и программы): диссертация канд. наук. – Москва, 2003. – 190 с.

¹² Eryigit G., Adali E. An affix stripping morphological analyzer for Turkish \ Proceedings of IASTED International Conference ARTIFICIAL INTELLIGENCE AND APPLICATIONS. February 16-18, 2004, Innsbruck, Austria. – P. 299-304.

¹³ Дыбо А.В., Шеймович А.В. Автоматический морфологический анализ для корпусов хакасского и древнетюркского языков / Научное обозрение саяно-алтая рецензируемый научный журнал Номер 2(08), 2014 (Периодичность – 2 раза в год. серия: Филология.Выпуск 2). С. 9-31.

¹⁴ Сулейманов Д.Ш., Гильмуллин Р.А., Гатауллин Р.Р.Морфологический анализатор татарского языка на основе двухуровневой модели морфологии / Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.; Сулейманов Д.Ш., Гатиатуллин А.Р. Модель татарской аффиксальной морфемы и ее реализация // Серия: Интеллект. Язык. Компьютер. – Вып.4. – Казань: Фэн. – 1996. – 113-127.

¹⁵ Гатиатуллин А.Р., Баширов А.М. Морфологический анализатор тюркских словоформ на базе структурно-функциональной модели тюркской морфемы // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академиинаук Республики Татарстан, 2017. – 327 с.

¹⁶ Желтов П.В. Разработка морфологического анализатора чувашского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.

¹⁷ Исраилова Н.А., Бакасова П.С. Морфологический анализатор кыргызского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.

¹⁸ Леонтьев Н.А. Морфологический анализатор якутского языка // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с. – С. 276-279.

¹⁹ Куканова В.В., Каджиев А.Ю. Алгоритм работы морфологического парсера калмыцкого языка // В сборнике: Писменное наследство и информационные технологии. El'Manuscript-2014 Материалы от V международна научной конференции. Отговорни редактори В.А.Баранов, В.Желязкова, А.М.Лаврентьев. 2014. – С. 116-119.

²⁰ Kemal Oflazer. Two-level Description of Turkish Morphology. Literary and Linguistic Computing, – Vol. 9, No 2, – 1994.; Altintas K., Cicekli I. A morphological analyzer for Crimean Tatar // Proceedings of the 10th Turkish Symposium on Artificial Intelligence and Neural Networks (TAINN'2001). – 2001. – P. 180-189.; Çağrı Çöltekin (2014) A Set of

dissertatsiyaning nazariy asoslarini mustahkamladi. Shuningdek, morfoanalizator ishlab chiqish nazariy g'oyalari amaliyotga joriy etilib, Yevropa, rus, turkiy tillarning morfoanalizatorlari ishlab chiqildi²¹. O'zbek tilshunosligida kompyuter lingvistikasi, matnga leksikografik ishlov berish va lingvostatistik tahlil etish borasida birmuncha tadqiqotlar amalga oshirilgan. M.Mahmudov, M.Ayimbetov, S.Karimov, G.Jumanazarova, A.Babanarov, D.O'rinboeva, A.Norov, S.Muhammedov, N.Abdurahmonovalarning tadqiqotlarini ana shunday ishlar sifatida qayd etish mumkin²².

O'zbek tili morfologik analizatorini tuzish maxsus tadqiqot predmeti bo'lmagan, ammo ayrim izlanishlarda masalaning muayyan qirralari haqida fikr bildirilgan. Chunonchi, harakat fe'llari asosida kompyuter dasturlari uchun lingvistik ta'min yaratish²³, o'zbek-ingliz tili mashina tarjimasining lingvistik ta'minoti²⁴, o'zbek tili mualliflik korpusini tuzish tamoyillari²⁵, o'zbek tili birliklarini grafematik tahlil qilish muammolari²⁶, til korpusi lingvistik bazasini tuzish tamoyillari²⁷ monografik planda o'rganilgan. Shuningdek, o'zbek tili atov birliklarini semantik teglashning lingvistik asoslari²⁸, o'zbek-ingliz parallel korpusi tuzish muammolari²⁹,

Open Source Tools for Turkish Natural Language Processing In: Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14) Ed. by N. Calzolari et al. 1079-1086.; G. Kessikbayeva and I. Cicekli, "Rule based morphological analyzer of Kazakh language," in Proceedings of the 2014 Joint Meeting of SIGMORPHON and SIGFSM. Baltimore, Maryland: ACL, June 2014, pp. 46-54.; Tantug C., Adali E., Oflazer K. Computer Analysis of the Turkmen Language Morphology, 5th International Conference on NLP (FinTAL 2006), Turku, Finland, 186-193, 2006.; Orhun M., Tantug C., Adali E. Rule Based Analysis of the Uyghur Nouns, International Journal of Asian Lang. Proc.,19(1), 33-44, 2009.

²¹ Морфологический анализатор текста на русском языке mystem [Электронный ресурс]; // Компания Яндекс [сайт] 2003-2013. URL: <http://company.yandex.ru/technologies/mystem/>; <http://tamgasoft.kg/morfo/>; <http://tadmorphon.pythonanywhere.com/>

²² Махмудов М.А., Пиотровская А.А., Садыков Т. Система машинного анализа и синтеза тюркской словоформы // Переработка текста методами инженерной лингвистики. – Минск, 1982.; Мухаммедов С.А. Статистический анализ лексико-морфологической структуры узбекских газетных текстов: Автореф. дисс... канд. фил. наук. – Ташкент, 1980.; Бабанаров А. Разработка принципов построения словарного обеспечения турецко-русского машинного перевода: Автореф. дисс... канд. фил. наук. – Л., 1981.; Айымбетов М.К. Опыт лингвостатистического анализа лексики и морфологии каракалпакского публицистического текста: Автореф. дисс... канд. фил. наук. – Ташкент, 1987.; Каримов С., Каршиев А., Исроилова Г. Абдулла Қаҳҳор асарлари тилининг луғати. Алфавитли луғат. Частотали луғат. Терс луғат. – Тошкент, 2007.; Ризаев С. Ўзбек тилининг лингвостатистик тадқиқи: Фил. фан. док. дисс...автореф. – Тошкент, 2008.; Пўлатов А. Компьютер лингвистикаси. – Тошкент: Akademnashr, 2011.; Норов А. Компьютер лингвистикаси асослари. – Қарши, 2017. – 136 б. Жуманазарова Г.У. Фозил Йўлдош ўғли дostonлари тилининг лингвопоэтикаси: Фил. фан. док. дисс...автореф. – Тошкент, 2017.; Ўринбоева Д.Б. Ўзбек фольклори матнларининг лингвостатистик тадқиқи. – Тошкент: Фан, 2010.; Ўринбаева Д. Халқ оғзаки ижоди: жанрий-лисоний ва лингвостатистик тадқиқ муаммолари: фил. фан. док. дисс. автореф. – Самарқанд, 2019. – 74 б.

²³ Мухаммедова С. Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. Методик қўлланма. – Тошкент, 2006.;

²⁴ Абдурахмонова Н.З. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти (Содда гаплар мисолида). Филол. фан. бўйича фалсафа доктори (PhD)...дисс. афтореф. – Тошкент, 2018. – 52 б.

²⁵ Хамроева Ш. Ўзбек тили морфологик анализаторининг лингвистик таъминоти. фил. фан. док. дисс. автореф. – Фарғона, 2021. – 74 б.

²⁶ Абжалова М.А. Ўзбек тили онтологиясини яратиш тамойиллари: фил. фан. док. дисс. автореф. – Фарғона, 2022. – 62 б.

²⁷ Эшмўминов А.А. Ўзбек тили миллий корпусининг синоним сўзлар базаси: Филол.фан.бўйича фалсафа доктори (PhD)...дисс. – Қарши, 2019. – 140 б.

²⁸ Akhmedova D.B., Mengliev B.R. Semantic Tag Categories in Corpus Linguistics: Experience and Examination International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-3S, October 2019. – P. 208-212.

²⁹ Karimov R., Mengliev B. Theoretical fundamentals of Uzbek-English parallel corpus / Journal of critical reviews. ISSN- 2394-5125. – VOL 7, ISSUE 17, 2020. – P. 73-76.; Karimov R.A., Mengliev B.R. The Role of the Parallel Corpus

o‘zbek tilidagi otlarni avtomatik tahlil qilish³⁰, so‘z yasashining formal modellari³¹ haqida qator maqolalar nashr etilgan. O‘zbek tilida ot va sifat so‘z turkumi morfologik analizini avtomatlashtirishning dasturiy tomonlari o‘rganilgan³². Bu tadqiqotlar kompyuter lingvistikasi muammolarini tadqiq etganligi bilan dolzarblik kasb etgan, ammo o‘zbek tili avtomatik morfologik tahlil dasturi – morfoanalizator va uning lingvistik ta’minotini yaratish tamoyillari masalasi kun tartibiga maxsus qo‘yilmagan. Dissertatsiyani tayyorlash jarayonida yuqorida sanab o‘tilgan ilmiy izlanishlar atroflicha o‘rganildi, zarur o‘rinlarda ularga munosabat bildirildi va ulardan tadqiqotda foydalanildi.

Tadqiqotning dissertatsiya bajarilgan oliy ta’lim yoki ilmiy-tadqiqot muassasasining ilmiy-tadqiqot ishlari rejalari bilan bog‘liqligi. Dissertatsiya Termiz davlat universitetining ilmiy-tadqiqot ishlari rejasiga muvofiq № 25–1 “Til nazariyasi, amaliy va kompyuter lingvistikasi” mavzusi doirasida bajarilgan.

Tadqiqotning maqsadi o‘zbek tili korpuslari uchun taksonomik va meronimik munosabatli leksemalar lingvistik bazasini yaratishning nazariy-metodologik asoslarini ishlab chiqishdan iborat.

Tadqiqotning vazifalari quyidagilardan iborat:

tilshunoslikda taksonomiya va meronimiya hodisalarining umumiy tavsifi hamda uning til korpuslari yaratishdagi ahamiyatini o‘rganish;

til birliklarining mavzuviy maydon asosida tahlil qilishning nazariy asoslarini o‘rganish;

taksonomik va meronimik munosabatli leksemalarni teglash, ushbu teglarnig qidiruv tizimidagi ahamiyatini aniqlash;

lingvistik korpus interfeysida semantik qidiruvning ko‘rinishi, tarkibi va funksiyalarini tadqiq etish;

til korpusida taksonomik va meronimik munosabatli leksemalar interfeysining mushtarak va farqli jihatlarini tadqiq etish;

o‘zbek tili korpusi uchun semantik ma’lumotlar bazasining tuzilishi va xususiyatlarini aniqlash;

o‘zbek tili korpusi semantik maydon va guruh birliklari qidiruv oynasining interfeysi, unda til birliklarini qidirish, saralash jarayonlarini lingvistik modellashtirish.

Tadqiqotning obyekti sifatida o‘zbek tilidagi taksonomik va meronimik munosabatli leksemalar tanlangan.

in Linguistics, the Importance and the Possibilities of Interpretation International Journal of Engineering and Advanced Technology (IJEAT). ISSN: 2249 – 8958, Volume-8, Issue-5S3 July 2019. – P. 388-391.

³⁰ Орхун М. Computational analysis of Uzbek nouns / Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

³¹ Турсунов А. Вопросы словообразования в формальных моделях тюркских языков (на примере узбекского языка) // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». Труды конференции) – Ташкент, 2018. – 320 с.

³² Болтаев Т.Б., Ибрагимов С.И. О проекте программной системы морфологического анализа узбекского языка // <http://buxdu.uz/index.php/uz/>; Болтаев Т.Б., Кузьминов Т.В., Поттосин И.В. О структурном конструировании программ и инструментах его поддержки // Среда программирования: методы и инструменты. Новосибирск, 1992, С22-37.; The Structured Constructing as a Discipline of Safe Programming and Instruments Supporting It / Aniskov M.I., Boltaev T.B., Kochetov D.V. et al // Instrumental Congress on Computer Systems and Applied Mathematics CSAM’93. St-Petersburg, July 19-23.

Tadqiqotning predmetini o'zbek tili korpuslari uchun taksonomik va meronimik munosabatli leksemalar lingvistik bazasini shakllantirish asoslari tashkil etadi.

Tadqiqotning usullari. Tadqiqot mavzusini yoritishda tasniflash, tavsiflash, qiyoslash, statistik tahlil usullaridan foydalanilgan.

Tadqiqotning ilmiy yangiligi quyidagilardan iborat:

tildagi leksemalararo leksik-semantik munosabatning sinonimik, giperonimik, graduonimik kabi turlari, ularning taksonomik va meronimik munosabat bilan o'zaro bir sistema sifatida bir-biriga ta'siri, leksemalarning so'z ma'nosi asosidagi o'zgarishiga ko'ra umumiy shakliy va mazmuniy jihatdan o'xshash, mavzuviy tuzilishiga ko'ra umumiy va xususiy, maydon va guruhga ajralishiga ko'ra makro va mikro, bir ma'no va ko'p ma'nolik jihatidan farqlari aniqlangan;

til korpuslarining dasturiy ta'minoti uchun lingvistik ma'lumotlar bazasini tuzish jarayonlari hamda taksonomik va meronimik munosabatli leksemalarning o'zbek tili leksik sathidagi miqdori aniqlangan;

o'zbek tilining izohli lug'atlaridagi taksonomik munosabatli leksemalarning meronimik munosabatli leksemalarga nisbatan miqdor jihatdan ko'p, mavzuviy jihatdan teng ekanligi dalillangan;

o'zbek tili korpuslari semantik qidiruv interfeysining monitorli tuzilishi, leksemalarning korpusda birikish strukturasi ochib berilib, o'zbek tili leksik birliklarini mereologik va taksonomik belgisiga asosan teglash jarayoni uchun semantik teglash dasturi loyihalari ishlab chiqilgan;

o'zbek tilidagi leksik-semantik munosabatdosh birliklar qidiruv oynasining interfeysi, mavzuviy va darajalash asosidagi saralash prinsiplari, til korpuslari uchun lingvistik modeli aniqlangan.

Tadqiqotning amaliy natijalari quyidagilardan iborat:

Leksik birliklarning semantik munosabatiga ko'ra taksonomiya va meronimiya hodisalarining umumiy tavsifi yoritilgan; ushbu munosabatlarning til korpuslarini ishlab chiqishdagi ahamiyati dalillangan;

o'zbek tili korpuslari uchun taksonomik va meronimik munosabatli leksemalar lingvistik bazasi shakllantirilgan;

taksonomik va meronimik munosabatli leksemalarni teglash tizimi bo'yicha taklif ishlab chiqilgan;

o'zbek tili korpusi semantik xususiyatli birliklar qidiruvi oynasining interfeysi, unda til birliklarini qidirish, saralash jarayonlari lingvistik modellashdirilgan.

Tadqiqot natijalarining ishonchliligi o'rganilgan materiallarning o'zbek tili xususiyatidan kelib chiqqan holda xulosalar qilishga yordam berganligi, ularning asosli ekanligi, metodologik mukammalligi, o'zbek tili Milliy korpusi tuzish tamoyillarini ishlab chiqishda amalda isbotlangan manbalarga tayanilganligi bilan izohlanadi.

Tadqiqot natijalarining ilmiy va amaliy ahamiyati. Tadqiqot natijalarining ilmiy ahamiyati o'zbek tili Milliy korpusi yaratishning nazariy asoslari, uning lingvistik ta'minotini ishlab chiqish doirasidagi yangi tadqiqotlar uchun ilmiy-nazariy manba bo'la olishi, jumladan, tildagi leksemalarning milliy korpus interfeysiga avtomatik uzatish vositalari bo'yicha nazariy ma'lumotlar berishi bilan belgilanadi.

Tadqiqot natijalarining amaliy ahamiyati amaliy tilshunoslik, kompyuter lingvistikasi, korpus lingvistikasi fanlari o'qitilish jarayonida dastur, reja tuzish, mavzularni bayon etishda manba vazifasini o'tashi, taksonomik va meronimik munosabatli leksemalarning ma'lumotlar bazasini tuzishning umumiy tamoyillari, uning axborot-qidiruv tizimi, til korpuslari, mashina tarjimasini hamda kompyuter leksikografiyasining rivojiga xizmat qilishi bilan izohlanadi.

Tadqiqot natijalarining joriy qilinishi. O'zbek tili korpuslari uchun taksonomik va meronimik munosabatli leksemalar lingvistik bazasini shakllantirishga bag'ishlangan dissertatsion ishda erishilgan ilmiy natijalar asosida:

til korpuslarining dasturiy ta'minoti uchun lingvistik ma'lumotlar bazasini tuzish jarayonlari hamda taksonomik va meronimik munosabatli leksemalarning til leksik sathidagi miqdorini aniqlashga oid natijalardan O'zbekiston Respublikasi Fanlar Akademiyasi Qoraqalpog'iston bo'limi Qoraqalpoq gumanitar fanlar ilmiy-tadqiqot instituti tomonidan olib borilgan FA-F1-G003 «Hozirgi qoraqalpoq tilida funksional so'z yasalishi» mavzusidagi fundamental ilmiy loyihada foydalanilgan (O'zR Fanlar akademiyasi Qoraqalpog'iston bo'limining 2022- yil 28- martdagi 17.01/72-son ma'lumotnomasi). Natijada hozirgi qoraqalpoq tili leksik qatlamining mavzuviy guruhlarini farqlash tamoyillari ishlab chiqilgan;

o'zbek tilining izohli lug'atlarida taksonomiya va meronimiya hodisalarining statistikasini aniqlash va tadqiq etish haqidagi xulosalardan O'zbekiston Respublikasi Fanlar Akademiyasi Qoraqalpog'iston bo'limi Qoraqalpoq gumanitar fanlar ilmiy-tadqiqot instituti tomonidan olib borilgan FA-A1-G007 «Qoraqalpoq naql — maqollari lingvistik tadqiqot ob'ekti sifatida» mavzusidagi fundamental ilmiy loyihada foydalanilgan (O'zR Fanlar akademiyasi Qoraqalpog'iston bo'limining 2022- yil 28- martdagi 17.01/72-son ma'lumotnomasi). Natijada hozirgi qoraqalpoq tilidagi taksonomiya va meronimiya hodisalarining statistik miqdorini aniqlashga erishilgan;

o'zbek tili korpuslari uchun taksonomik va meronimik munosabatli leksemalar lingvistik bazasini shakllantirish asoslari hamda interfeys yaratishga oid natijalaridan O'zbekiston Respublikasi Fanlar Akademiyasi Qoraqalpog'iston bo'limi Qoraqalpoq gumanitar fanlar ilmiy-tadqiqot instituti tomonidan olib borilgan FA-F-1-005 raqamli «Qoraqalpoq folklorshunosligi va adabiyotshunosligi tarixini tadqiq etish» mavzusidagi ilmiy-amaliy loyihada foydalanilgan (O'zR Fanlar akademiyasi Qoraqalpog'iston bo'limining 2022- yil 28- martdagi 17.01/72-son ma'lumotnomasi). Natijada o'zbek va qoraqalpoq tillarida yaratiladigan lingvistik korpuslar uchun algoritmlar ishlab chiqishga erishilgan;

o'zbek tilidagi leksik-semantik munosabatchilik birliklar qidiruv oynasining interfeysi, mavzuviy va darajalash asosidagi saralash prinsiplari haqidagi xulosalardan Surxondaryo viloyati teleradiokompaniyasining "Til-millat ko'zgusi", «Kun mavzusi», «Dialog» teleko'rsatuvida foydalanilgan (Surxondaryo viloyati teleradiokompaniyasining 2022- yil 24- martdagi 01-12/224-son ma'lumotnomasi). Natijada yuqori saviyali barkamol avlodni tarbiyalash, yoshlar orasida til madaniyatini shakllantirish hamda zamonaviy axborot tizimida o'zbek tilining imkoniyatlaridan foydalanish haqidagi targ'ibot ishlari mazmuni mukammallashgan.

Tadqiqot natijalarining aprobatyasi. Mazkur tadqiqot natijalari 3 ta xalqaro, 3 ta respublika anjumanida muhokamadan o'tkazilgan hamda quyidagi profillarda

Tadqiqot natijalarining e'lon qilinishi. Dissertatsiya mavzusi bo'yicha 17 ta ilmiy ish nashr ettirilgan, jumladan, 1 ta monografiya, O'zbekiston Respublikasi Vazirlar Mahkamasi huzuridagi Oliy attestatsiya komissiyasining doktorlik dissertatsiyalari asosiy ilmiy natijalarini chop etish uchun tavsiya etilgan ilmiy nashrlarda 10 ta ilmiy maqola, ulardan 2 tasi xorijiy jurnallarda chop qilingan.

Dissertatsiyaning hajmi va tuzilishi. Dissertatsiya kirish, to'rt bob, xulosa, foydalanilgan adabiyotlar ro'yxati va ilovadan iborat. Hajmi 264 sahifani tashkil etadi.

DISSERTATSIYANING ASOSIY MAZMUNI

Kirish qismida mavzuning dolzarbligi va zarurati asoslangan, tadqiqotning respublika fan va texnologiyalari rivojlanishining ustuvor yo'nalishlariga bog'liqligi ko'rsatilgan, maqsad va vazifalari berilgan, ob'ekti va predmeti tavsiflangan, ilmiy yangiligi va amaliy natijalari bayon qilingan, natijalarning ilmiy va amaliy ahamiyati ochib berilgan, joriylanishi, aprobatyasi, nashr etilgan ishlar va dissertatsiya tuzilishi bo'yicha ma'lumotlar keltirilgan.

Dissertatsiyaning birinchi bobi "Tilshunoslikda taksonomik va meronomik hodisalarni tadqiq qilishning metodologik asoslari" deb nomlangan. Bobning "Taksonomiya va meronomiya hodisalarining tilshunoslikka tatbiqi masalalari" deb atalgan birinchi bo'limida taksonomiya va tasniflash tushunchalari orasidagi farq tavsiflangan. **Taksonomiya** [yun. *taxis* – tartib bilan joylashish + *nomos* + qonun] butun borliqning murakkab, odatda, iyerarxik tuzilmali sohalarini (jonli tabiat, Yerning tuzilishi; geografiya, geologiya, tilshunoslik, etnografiya va sh.k. obyektlarini) tasniflash va tartibga solish nazariyasi, shu haqidagi fan³³. *Taksonomiya* tushunchasi birinchi marta biologiya sohasida paydo bo'ldi. Bu atama 1813-yilda shvetsariyalik botanik, o'simliklarning tasnifini ishlab chiqqan O.Dekandol tomonidan taklif qilingan.

XX asrning ikkinchi yarmida taksonomiya muammolari nafaqat biologiyada, balki iyerarxik ravishda tashkil qilingan alohida diskret obyektlar to'plamlari bilan shug'ullanadigan boshqa bir qator fanlarda ham muhim rol o'ynay boshladi. Natijada, bu atama boshqa bilim sohalarida: tilshunoslik, geografiya, geologiyada ham murakkab tizimlarni tasniflash, tizimlashtirishning umumiy nazariyasini ifodalash uchun qo'llanila boshlandi. Shu bilan birga, taksonomik toifalarining aniq toifalaridagi tabiiy farqlardan tashqari, bilimlarning turli sohalarida taksonomiya haqidagi boshlang'ich tushunchalar va uning vazifalari har xil talqin etiladi. Masalan, tilshunoslikda taksonomiya matndagi lisoniy birliklarni ajratib olish, ularning tartibi, tarqalishini tahlil qilish orqali ularning xususiyatlarini o'rganishga tayanadi. Shunga ko'ra, lingvistik nazariya elementlar sinfi toifalari, elementlar va sinflar o'rtasidagi munosabatlar turi bilan ishlaydi. Ba'zan tilshunoslikda taksonomiya turli tillardagi o'xshash grammatik kategoriyalarni bir tizimlashtirilgan toifaga birlashtirish nuqtayi nazaridan talqin qilinmoqda. Shunday qilib, taksonomiyaning nazariy fan sifatidagi umumiy tamoyillari ularning boshlang'ich bosqichida bo'lib, bunda semiotika,

³³ Ўзбек тилининг изоҳли луғати. – Тошкент: Ўзбекистон миллий энциклопедияси, III том. – Б. 646.

lingvistika va avtomatik tarjimaning dolzarb masalalari, lingvistik faktlarni tasniflashga (tizimlashtirishga) qaratilgan til tadqiqotlarining ahamiyati o'smoqda.

Ingliz botanik olimi Charlz Djeffriy taksonomiya hodisasini quyidagicha tavsiflaydi: "Fanda taksonomik yondashuv generativ usulga qarshi qo'yiladi. Matematik jihatdan taksonomiya – bu ma'lum obyektlar to'plamining daraxtga o'xshash tuzilishi. Ushbu tuzilmaning yuqori qismida birlashtiruvchi yagona tasnif – ildiz taksoni mavjud bo'lib, u ma'lum taksonomiyadagi barcha subyektlarga tegishli. Ildiz ostidagi taksonlar – tasniflangan obyektlar umumiy to'plamining kichik guruhlariga tegishli aniq tasniflardir. Zamonaviy biologik tasnif iyerarxik tizim bo'lib, uning asosini individual organizmlar (individlari) tashkil qiladi, yuqorisida esa bitta takson joylashadi; poydevor va yuqori pog'ona orasidagi iyerarxiyaning turli darajalarida taksonlar mavjud bo'lib, ularning har biri yuqori darajadagi bitta va faqat bitta taksonga bo'ysunadi"³⁴.

Taksonomiyaning eng asosiy muammolaridan biri bu taksonning usulini aniqlashdir.³⁵ Reyting, taksonomik toifa – bu biologik taksonomiya tushunchasi bo'lib, u iyerarxik tashkil etilgan tirik organizmlar tizimidagi darajani tavsiflaydi (masalan, sinf, tartib, oila). Maqom(rus. *ранг*)larning ishlatishiga doir bir qator muammolarga qaramay, ular qo'llanilishda davom etmoqda, chunki nomenklatura kodlari (organizmlarga ilmiy nom berishning qoidalar to'plami)da belgilangan tirik organizmlar guruh(takson)larining nomlarini shakllantirish qoidalari aynan daraja (*ранг*) prinsipiga asoslanadi³⁶.

"Kelurnoma" asari tilshunosligimiz tarixida birinchilardan bo'lib taksonomik jihatdan tahlil qilingan. Asarning oxirgi XV bobida ismlar taksonomik qatorlarga bo'linib, 9 ta guruhga ajratilganligini ko'rish mumkin. Ular sirasida (1) osmon va u bilan bog'liq tushunchalar; (2) yer va u bilan bog'liq tushunchalar; (3) ot va egarlanadigan hayvonlar; (4) burgutlar va yovvoyi qushlar; (5) odam jismi a'zolari; (6) qarindoshlik nomlari; (7) qurol-aslaha nomlari; (8) sanoq nomlari; (9) olmosh, ravish, bog'lovchi va qo'shimchalar³⁷ farqlanadi.

Til – yaxlit bir butunlikni tashkil qiluvchi elementlar sistemasi. Uni tashkil etuvchi har bir elementning ma'nosi bir vaqtning o'zida sistemani tashkil qilayotgan boshqa elementlarning mavjud bo'lishidan kelib chiqadi³⁸. Mohiyatini haligacha hech kim anglamagan, umri davomida imkoniyatlarining milliard-milliarddan bir qismidan ham hech kim foydalanishga ulgurolmaydigan, cheksiz sanalgan til sanoqli ma'no g'ishtchalari (semalar), shakl birliklari va munosabat turlarining o'zaro rang-barang birikib, ko'z ko'rolmaydigan, quloq eshitib nihoyasiga yetolmaydigan turli-tumanlikni tashkil etadi, deydi- H.Ne'matov.

Tilshunoslikda semantik maydon nazariyasi sohasi asosida sohalararo alohida tadqiqot ishlari olib borish XX asrning 70-80-yillaridan boshlanadi. Shu munosabat bilan o'zbek tili leksikologiyasi qiyosiy-tavsifiy bosqichdan yangi nazariy bosqichga ko'tarildi. Bu bosqich davomida leksikani ma'lum bir semantik guruhlardan, muayyan

³⁴ Джеффри Ч.. Биологическая номенклатура Charles Jeffrey. Biological Nomenclature. Second edition / Пер. с англ. — М.: Мир, 1980. — С. 14. — 120 с.

³⁵ Зуев В. В., Розова С. С. Проблема способа бытия таксона в биологической таксономии // Философия науки : журнал. — Издательство Сибирского отделения РАН, 2001. — № 2 (10). — ISSN 1560-7488.

³⁶ O'sha yerda.

³⁷ Искандарова Ш. Тил системасига майдон асосида ёндашув. — Ташкент: Фан, 1982. — В. 135.

³⁸ Соссюр Ф. Де. Труды по языкознанию. — М.: Прогресс, 1977. — С.147.

ma'no elementlarining munosabatlaridan tashkil topgan yaxlit bir tizim-sistema sifatida o'rganish bilan xarakterlanishi muhim ahamiyat kasb etdi. Professor A.Sobirovning fikricha, o'zbek tilshunosligiga maydon nazariyasining kirib kelishi til birliklariga nisbatan sistem-struktur metodlarning qo'llanila boshlashi bilan bog'liq. Sh.Iskandarova shaxs mikromaydoni misolida o'zbek tili leksikasini semantik maydon sifatida o'rganish muammolarini ko'tarib chiqdi. Sh.Iskandarova olamning guruhlariga bo'linishi va uning tilda aks etishi haqida batafsil fikr yuritib, shaxs mikromaydonida faoliyat ko'rsatuvchi 100 ga yaqin leksik-semantik guruh haqida ma'lumot beradi. Masalan, *savodxon*, *donishmand*, *zakovatli*, *farosatli*, *idrokli*, *qoloq* kabi leksemalar saviya asosida farq qiluvchi nomlar guruhiga kiritiladi, ularning umumiy va farqli tomonlari izohlanadi³⁹. Demak, taksonomiya, tasniflash, sistema atamalarining bir-biriga bo'lgan munosabatining ilmiy tadqiqi o'zbek tilshunosligida hali o'z yechimini kutayotgan muammolardan biri ekanligi shubhasiz.

Birinchi bobning ikkinchi bo'limi "*Fanda leksik-semantik maydon (LSM) va leksik-semantik guruh (LSG) atamaları talqini*" deb nomlangan bo'limda borliqdagi barcha narsa-hodisalar o'zaro sistema sifatida birlashgan. XX asrning 40-yillarida shveytsariyalik tilshunos Sharl Balli "assotsiativ maydon" atamasini fanga kiritdi. Uning fikriga ko'ra, "assotsiativ maydon" tushunchasi qayishqoq (elastik) tushuncha, bu assotsiatsiyalar hajmining har xil odamlarda bir xil emasligi bilan dalillanadi.

Dunyo tilshunosligida tilga strukturalizm yo'nalishini olib kirgan Ferdinand de Sossyur tilshunoslar diqqatini til birliklari orasidagi munosabatni ochishga qaratib, munosabatlarning paradigmatic va syntagmatic turlari mavjudligini aniqlagan edi. Rus tili lingvistik ensiklopedik lug'atda "Lingvistik birliklarning muayyan birlashtiruvchi ma'no asosida ma'lum paradigmalariga birlashuvi keyinchalik tilshunoslikda maydon nazariyasini vujudga keltirdi. Lingvistikada esa "ma'no" (mazmun) umumiyliги bilan birlashgan va belgilanayotgan hodisalarning tushunchaviy, predmetlik yoki vazifaviy (funktional) o'xshashligini aks ettiruvchi til birliklari yig'indisi maydon sifatida belgilanadi"⁴⁰ deb ta'kidlanadi.

Tilshunoslik tarixiga murojaat qilsak, til birliklarning ma'lum bir ma'no asosida birlashuvchanligi, ma'lum bir tildagi lug'aviy birliklarni ana shunday mazmuniy uyalarga birlashtirish qadimgi Sharq tilshunosligida rivojlanganligini ko'rish mumkin. Keyinchalik tilshunoslikda strukturalizm yo'nalishining kirib kelishi natijasida XIX asrning o'rtalarida til birliklarni mazmuniy guruhlariga birlashtirish haqidagi fikrlar G'arb tilshunosligida avj oldi. Shunga ko'ra, bu nazariya G'arb tilshunosligi nomi bilan bog'landi. XIX asrda til birliklarini umumiylik asosida o'rganishga mashhur tilshunos olim M.Pokrovskiy tomonidan biri e'tibor qaratildi. Tilshunos G.S.Shchur maydon nazariyasi haqida quyidagicha fikr bildirgan: "Lingvistikada maydon nazariyasining keyingi yillardagi rivoji shu darajaga yetdiki, uning g'oyalari va metodlari tilning leksik sathidan boshqa sathlarga nisbatan ham tatbiq qilina boshlandi"⁴¹. Ushbu sohaga oid tadqiqotlarda mazmuniy maydon doirasidan kelib

³⁹ Искандарова Ш. Ўзбек тили лексикасини мазмуний майдон сифатида ўрганиш (шахс микромайдони): Филол. фанлари док. . дисс. автореферат. – Тошкент, 1999.

⁴⁰ Лингвистический энциклопедический словарь. – М., СЭ, 1990. – С. 380.

⁴¹ Щур Г.С. О морфосемантических полях, функционально-инвариантных группах и категориях // Вопросы лингвистики. – Томск., 1970. – С. 3-10; Филичева Н.И. Синтаксические поля. – М., – ВЯ, 1977.

chiqsak, turli mavzuiy lug‘atlar (ko‘proq tezauruslar) tuzish uchun umumiy qonuniyatlarni ochishda nazariy asos bo‘lib xizmat qilishi aniqlangan. Keyinroq bu yo‘nalishda tilning barcha sathlari, jumladan, morfosemantika, grammatika singarilar doirasida ham jiddiy tadqiqotlar yuzaga kelganligini ham ta’kidlash lozim. Maydon tushunchasining tavsifi hamda unga xos xususiyatlarga tayansak, ayrim o‘rinlarda yondashuv maydon nazariyasini tadqiq etishning turlicha muammolarini vujudga keltiradi. Bunda maydon chegaralarining aniqligi, maydonning alohidaligi, butunligi, cheksizligi jihatidan til birligi sifatida qaralgani holda yoki maydon nazariyasining nisbiy mustaqilligini, maydonlar o‘zaro munosabatining keng xarakterga ega ekanligi bo‘yicha lingvistik birliklaridan farqlanuvchi tomonlari asosidagi muammolar; maydonlar tuzilishi muammolari; ko‘p ma’noli so‘zlarning faqat bir maydonga mansubligi yoki bir necha maydonning elementlari bo‘lishi mumkinligi asosida maydon va ko‘p ma’noilik kabi munosabatlari muammolari o‘rtaga tashlanmoqda. Maydon nazariyasining turli tomonlama xususiyatlariga tayanib, ushbu muammolar tadqiqotchilar tomonidan turli aspektda o‘z yechimini topib bormoqda. O‘zbek tilshunosligida ham maydon talqini bilan bog‘liq bir qator qarashlar yuzaga chiqmoqda⁴². O‘zbek tilshunosligida til birliklarini maydon asosida o‘rganish keng yo‘lga qo‘yilmoqda. A.Nurmonov “Bu usul leksikani o‘rganishda katta ahamiyatga ega. O‘zbek tili leksikasini makromaydon sifatida qaralgani holda, uni mikromaydonlarga ajratish tezaurus, ideografik lug‘atlar tuzishda katta samara beradi. Shuning uchun lingvistik maydon va til lug‘at tarkibini bunday maydonlarga ajratishning nazariy asoslarini chuqur, atroflicha o‘rganish hozirgi o‘zbek tilshunosligi oldidagi eng dolzarb muammolardan sanaladi”⁴³ deb izohlaydi.

Bobning uchinchi bo‘limi o‘zbek tilida leksik-semantik munosabat turlari va ularning o‘rganilishiga bag‘ishlangan. A.Sobirovning fikricha, semantik maydondagi paradigma a‘zolari o‘rtasidagi munosabatni sinonimik, giper-giponimik (jism-tur), xolo-meronimik (butun-bo‘lak), graduonimik (leksik darajalanish), funksionimik (vazifadoshlik), zidlik munosabati⁴⁴ kabi turlarga ajratadi. Leksik-semantik munosabat turlari ikki xil asosda: leksik-shakliy va leksik-ma’noviy nuqtayi nazardan o‘rganilmoqda. *Meronim* boshqasining ajralmas qismi bo‘lgan tushunchadir. *Meronim partonim* deb ham ataladi. *Meronimiya* (partonimiya) atamasi o‘zbek tilshunosligida ham, Yevropa tilshunosligida ham yangi yo‘nalish termini sanalib, ilmiy adabiyotlarda ilk marta “Sistem leksikologiya tezislar”⁴⁵da uchraydi. O‘zbek tilshunosligida *meronimiya* atamasi *partonimiya* nomi ostida XX asrning 90-yillarida Buxoro tilshunoslik maktabi vakillari tomonidan fanga olib kirilgan⁴⁶. *Meronimiya* termini

⁴² Мирзакулов Т. Ўзбек тили морфем парадигматикаси ва синтагматикаси масалалари: Филол. фанл. докт....дисс. автореф. – Т., 1994; Абдувалиев М. Тўсиқсизлик майдони ва уни ташкил этувчи синтактик бириклар // Ўзбек тили ва адабиёти. -1988. - № 4. – Б. 62-66.

⁴³Нурмонов А. Танланган асарлар III жилд. Тошкент. 2012. – Б. 55.

⁴⁴ Собиров А. Ўзбек тилининг лексик сатҳини системалар системаси тамойили асосида тадқиқ этиш. – Тошкент: Маънавият, 2004. –140 б.

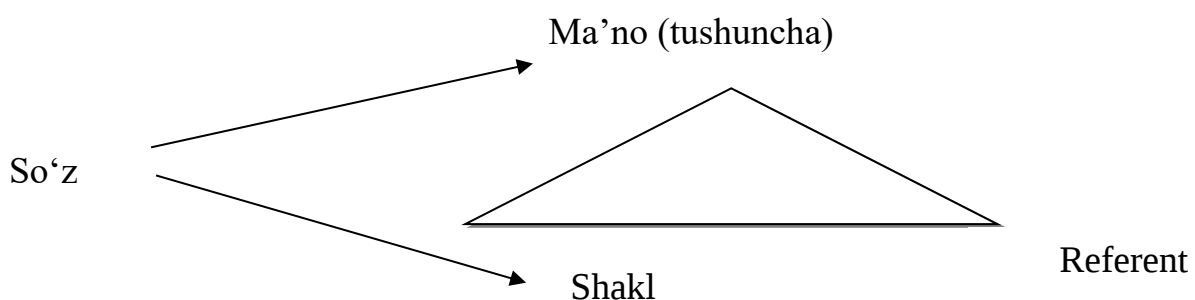
⁴⁵ Бегматов Э., Нематов Х., Расулов Р. Лексик микросистема ва унинг тадқиқ методикаси. //Ўзбек тили ва адабиёти. 1989. 6-сон. – Б. 35-40.

⁴⁶ Неъматов Х., Расулов Р. Ўзбек тили систем лексикологияси асослари. – Т.:Ўқитувчи. 1995. – 127 б.; Қиличев Б.Э. Партонимик қатор ва кўп маънолилик. Ўзбек тили доимий анжумани учинчи йиғинининг тезислари. –Т.: Ўқитувчи, 1995. -138 б.

M.V.Nikitin tomonidan bo‘lak ma’nosini ifodalash uchun fanga kiritilgan⁴⁷. O‘zbek tilshunosligida *meronimya* terminini B.Qilichev⁴⁸, A.Sobirovlar⁴⁹ kiritishsa, H.Ne‘matov, R.Rasulovlar⁵⁰ unga munosabat bildirishgan. Professor A.Sobirov meronimiyasi hodisasini: “Uy” butun nomi *poydevor, devor, pol, tom, eshik, deraza* bo‘laknomlaridan iborat. Ular ichidan “tom”ni ajratib olib, unga aloqador so‘zlar asosida maydoncha hosil qiladigan bo‘lsak, *to‘sin, shifer, shift, tarnov, cherdakka* nisbatan “tom” butun nom bo‘lib qoladi” deb izohlaydi⁵¹. Tilshunosligimizda meronimiyasi hodisasi partonimiyasi nomi ostida o‘rganilmoqda.

Ikkinchi bob “O‘zbek tilshunosligida leksikani mazmuniy maydon asosida o‘rganishning lingvistik asoslari” deb nomlangan. Bobning “Korpus materiallarini teglashda mazmuniy maydonning ahamiyati” deb ataladigan birinchi bo‘limida leksik sistema borasida bahs yuritiladi. Leksika ichki bo‘linuvchanlik xususiyatiga ega. O‘zbek tili leksikasi yuzlab semantik maydonlardan iborat. Maydon leksik-semantik guruhdan, guruh esa o‘z navbatida, tarkibiga kirgan elementlardan iborat. Leksik qatlamning bu holati uning diskretligidir. Masalan, **vaqt** (zamon) semantik maydoni tarkibiy tuzulishiga ko‘ra *era, asr, yil, fasl* mikromaydonlarini o‘z ichiga oladi. **Fasl** o‘z navbatida *qish, bahor, yoz, kuz* super mikromaydonlariga bo‘linadi. *Dekabr, yanvar, fevral, qor, sovuq, muz, sumalak, sovuq olmoq, chilla, qahraton, sovuq olmoq, pechka, o‘tin, ko‘mir, chana, sirpanmoq, qorbo‘ron, yangi yil, archa, qorbobo, qorqiz, qo‘lqop* kabi semalar (semantik maydondagi eng kichik elementlar) umumiy yagona chiziqli ostida **qish** mikromaydoniga to‘planadi va bir-biri bilan assotsiativ bog‘lanishni hosil qiladi⁵². Shu bilan birga leksik sistemadagi elementlar o‘zaro bir-biri bilan iyerarxik munosabatda bo‘ladi. Buni pog‘onali, ya’ni o‘zaro shartlangan munosabat deyishadi. Professor E.Begmatov tematik guruh bir umumiy mavzu, g‘oya atrofiga birlashgan (uyg‘unlashgan) so‘zlar majmuyini, yig‘indisini tushunadi. Olim o‘zbek tilidagi leksemalarni 28 ta mavzuviy guruhga ajratadi.

1-jadval



⁴⁷ Собиров А. Ўзбек тилининг лексик сатҳини системалар системаси тамойили асосида тадқиқ этиш. – Тошкент: Маънавият, 2004. –6-142.

⁴⁸ Қиличев Б. Ўзбек тилида партонимия. Филол. фан. номз... дисс. автореферати. – Тошкент, 1997.

⁴⁹ Собиров А. Ўзбек тилининг лексик сатҳини системалар системаси тамойили асосида тадқиқ этиш. – Тошкент: Маънавият, 2004. –154 б.

⁵⁰ Ҳ.Нематов, Р.Расулов. Ўзбек тили систем лексикологияси асослари. – Тошкент: Ўқитувчи, 1195. –128 б.

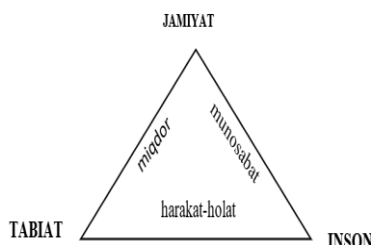
⁵¹ Собиров А. Ўзбек тилининг лексик сатҳини системалар системаси тамойили асосида тадқиқ этиш. – Тошкент: Маънавият, 2004. –154 б.

⁵² Собиров А. Ўзбек тилининг лексик сатҳини системалар системаси тамойили асосида тадқиқ этиш. – Тошкент: Маънавият, 2004. –140 б.

Soʻz – mazmun va shakldan iborat belgi. Mazmun, shakl va referent oʻrtasidagi aloqani quyidagicha ifodalanadi⁵³:

Yuqoridagi fikrlarga tayanib, biz falsafiy nuqtayi nazardan Tabiat-Inson-Jamiyat munosabatlarining bitta maydon asosidagi tahlilini quyidagicha jadvalga joylashtirib uchburchak asosida tahlil qildik.

2-jadval



Ikkinchi bobning “Oʻzbek tilining izohli lugʻatlarida taksonomik va meronimik birliklarning berilishi” deb nomlangan ikkinchi boʻlimida oʻzbek tilidagi taksonomik va meronimik leksemalarning oʻzbek tilining izohli lugʻatlaridagi statistik tadqiqi haqida fikrlar bildiradi. **Shaxs** taksonomik mavzuiy guruhi «Oʻzbek tili izohli lugʻati»da quyidagicha beriladi:

SHAXS [a. شخص kimsa, kishi, zot] Jamiyatdagi alohida bir kishi, odam. *Notanish shaxs. Tarixiy shaxs Shaxsga sigʻinish*⁵⁴.

Oʻzbek tilining izohli lugʻatlarida (1981, 2006, 2020) “shaxs” taksonomik mavzuiy guruhiga kiruvchi shaxs otlari quyidagicha koʻrsatgichga ega:

- 1) 1981-yilda chop qilingan “Oʻzbek tili izohli lugʻati”da 1927 ta;
- 2) 2006-2020-yillarda chop qilingan “Oʻzbek tili izohli lugʻatlari”da 2760 ta shaxs otlari borligi aniqlandi (Qarang: 2.1-jadval).

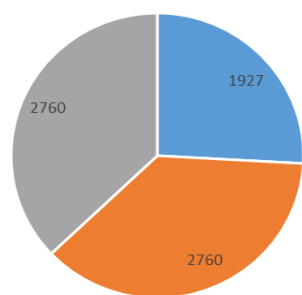
“Oʻzbek tili izohli lugʻati”ning 2006-2020-yillarda nashrida koʻrganimizdek, shaxs taksonomik guruhiga kiruvchi shaxs otlarining 69.81 % 1981-yilda nashr qilingan nashrida mavjuddir. Demak, izohli lugʻatimizning keyingi nashrida keltirilgan shaxs taksonomik guruhining 69.81 % 1981-yilgi nashrdan olinganligini koʻrish mumkin. Keyingi nashrlarda miqdorning keskin oshib ketishiga sabab shuki, keyingi 30 yilda oʻzbek tilshunosligiga sistem-struktur tilshunoslik tamoyillarining kirib kelishi natijasida oʻzbek tili leksikasida juda katta miqdoriy oʻzgarishlar yuzaga keldi.

3-4-jadval

⁵³ Эшмуминов А. Ўзбек тили миллий корпусининг синоним сўзлар базаси : Филол. фан. бўйича фалсафа докт. (PhD) диссертацияси. – Термиз, 2019. – Б. 52.

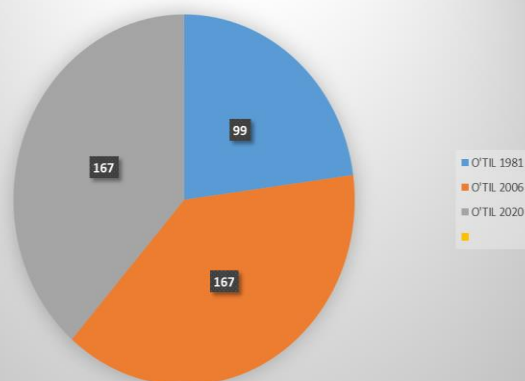
⁵⁴ Ўзбек тилининг изоҳли луғати. Тўртинчи жилд. – Тошкент: «Ўзбекистон миллий энциклопедияси» Давлат илмий нашриёти, 2006. – Б.557.

Shaxs taksonomik guruhi



■ O'TIL 1981 ■ O'TIL 2006 ■ O'TIL 2020 ■

Etnonimlar taksonomik mavzuiy guruhi



■ O'TIL 1981
■ O'TIL 2006
■ O'TIL 2020

“O‘zbek tili izohli lug‘ati”ning 2006-yilda chop qilingan nashri asosida tuzatishlar bilan chop qilingan 2020-yilgi nashrlari bir-birining aynan takrori bo‘lgani holda 2760 ta shaxs taksonomik guruhiga kiruvchi birliklar aniqlandi. **Etnonimlar** taksonomik mavzuiy guruhining «O‘zbek tili izohli lug‘ati»da quyidagicha izohlanadi: **Etnonim** [yun. ethnos - qabila, xalq + onyma — nom] Xalq, qabila, urug‘ va b. turli xil etnik uyushmalar nomi⁵⁵.

O‘zbek tili izohli lug‘atlarida taksonomik mavzuiy guruhlarining statistik miqdori quyidagicha ekanligi aniqlandi:

T/r	Taksonomik ma’noviy guruh nomi	1981-yil nashri	2006-2008 -yil nashri	2020-yil nashri
1	Shaxs	1927	2760	2760
2	Etnonim	99	167	167
3	Qarindoshlik otlari	49	87	87
4	Afsonaviy mavjudotlar	17	25	25
5	Jonivorlar	295	505	505
6	O‘simliklar	365	543	543
7	Jism va materiallar	84	135	135
8	Makon va fazo	32	67	67
9	Qurilish va inshootlar	255	324	324
10	Qurol va uskunalar	72	123	123
11	Anjomlar	71	116	116
12	Asboblar va mexanizmlar	97	225	225
13	Transport vositalari	38	84	84
14	Qurol	37	62	62
15	Musiqiy anjomlar	40	56	56
16	Mebel	19	30	30
17	Idish	67	147	147

⁵⁵ Ўзбек тилининг изоҳли луғати. Бешинчи жилд. – Тошкент: «Ўзбекистон миллий энциклопедияси» Давлат илмий нашриёти, 2006. – Б.62.

18	Kiyim va kechak	102	156	156
19	Ovqat va ichimliklar	128	183	183
20	Matnlar	16	54	54
Jami		3810	5849	5849

O‘zbek tili izohli lug‘atlarida meronimik mavzuviy guruhlarining statistik miqdori quyidagicha ekanligi aniqlandi:

Tadqiqotning III bobi **“O‘zbek tili milliy korpusida taksononiya va meronimiya hodisalarini razmetkalashning nazariy masalalari”** deb ataladi. Birinchi bo‘lim *“Jahon va o‘zbek tilshunosligida korpus lingvistikasining tadrijiy taraqqiyoti”* deb nomlanadi.

Murakkab lisoniy butunlik sifatida korpus o‘zida nutqiy/lisoniy manbaning tarkibiga doir keng hajmli axborotni saqlabgina qolmay, shu axborotni taqdim etishning formal usullari (so‘zlarni indeksatsiyalash, morfologik ma’lumot va b. singlarlar) ni ham jamlaydi. Shunga mos tarzda, korpusni maxsus qurilgan semiotik tizim qabilida sharhlash ham mumkin⁵⁶.

1980-yillar o‘rtalarida turli til/hajmdagi korpus loyihalari yaratildi. Bu davrda korpus lingvistikasida erishilgan natijalar til o‘qitish jarayoniga keng tatbiq etila boshlandi, jahonning yetakchi oliy ta’lim muassasalarida korpus lingvistikasi bo‘yicha yaratilgan loyihalar amaliyot va kun mavzusiga aylana bordi. Bu orqali korpus yondashuvning tarixiy, geografik, ijtimoiy variatsiyasi va uning lisoniy tizimdagi

T/r	Meronimik ma’noviy guruh nomi	1981-yil nashri	2006-2008 -yil nashri	2020-yil nashri
1	Odam tanasi a’zolari	136	181	181
2	Kiyim va kechak qismlari	34	48	48
3	Mexanizm va asboblarning qismi	123	150	150
4	O‘simlik qismlari	29	24	24
5	Ob’ektlar jamlanmasi qismlari	498	511	511
6	Idishlarning qismlari	88	149	149
7	Inshoot qismlari	30	46	46
8	Transport vositalari qismlari	27	35	35
Jami		965	1144	1144

o‘zgaruvchan shakllarini namoyon etuvchi optimal tabiati ochila bordi. Bu, o‘z navbatida, lingvistik tahlilning korpusga asoslangan usullarga doir asosiy tamoyillarni o‘zlashtirish imkonini yaratdi⁵⁷. Korpus lingvistikasining rivoji, korpuslarning

⁵⁶Рыков В.В. Корпус текстов как новый тип словесного единства // Труды Междунар. семинара «Диалог-2003». – М.: Наука, 2003. – С. 21.

⁵⁷Information and Communications Technology for Language Teachers. Введение в прикладное значение корпуса. Режим доступа: http://www.ict4lt.org/en/en_mod2-4.htm

yaratilish masalasi o'z kelajak taqdiriga qayg'uradigan millatlar uchun dolzarb masalalardan biri. XX asrning birinchi yarmida milliy korpusni faqat qo'lbola usulda tuzish imkoni mavjud edi. Bu esa uzoq muddat va ko'p xarajat talab qilar edi.

Butun dunyoning oliy ta'lim muassasalarida talabalarga nazorat savollari va ma'ruza matnlari tayyorlashda lingvistik korpuslardan foydalaniladi. Ko'plab talabalar mustaqil ish va loyihalarni tayyorlashda mustaqil tarzda korpusga murojaat qiladilar. Ushbu jarayonda korpusdan foydalanuvchi hamda undan foydalanmaydigan talabalarni o'zaro qiyoslab, shunday xulosa qilish mumkin: korpusdan unumli foydalanuvchi talaba til qonuniyatlari, o'rganayotgan xorijiy til xususiyatlarini ikkinchi toifadagi talabaga nisbatan oson va tez o'zlashtiradi.

An'anaviy va korpus lingvistikasining o'zaro farqlarga ega ular quyidagi o'rinlarda ko'rinadi (1.4-jadval).

I. 4-jadval. An'anaviy va korpus lingvistikasining farqli jihatlari

Xususiyatlari	An'anaviy lingvistikada	Korpus lingvistikasida
Yo'naltirilganligi	Til/lison tadqiqi	Nutq tadqiqi
Tadqiq obyekti	Nazariya orqali nutq hodisalarini tushuntirish	Matn korpusi
Usullari	Kvalitativ usullar	Kvantativ usullar
Matnga munosabati	Mavhumlik	Aniqlik
Tahlilga e'tibor	Asosiy e'tibor mazmunga	Asosiy e'tibor shaklga
Tahlil usullari	Cheklanganlik	Cheklanmaganlik
Tahlil turlari	Nutq materialini sezgi asosida tahlil qiladi	Nutqiy faoliyatni matn holida ifodalaydi
Ma'lumot yetkazish usullari	Mantiqiy fikrlashni afzal ko'radi	Nutqiy materialini birlamchi qayta ishlash uchun ehtimollik nazariyasi va statistik ma'lumotlardan foydalaniladi
Tahlil natijadorligi	Matndagi so'zlar ajratib olib tahlil qilinadi	Matndagi lingvistik ma'lumotlar (so'z qo'llanilishi) bilan ish olib boradi
Yondashuv metodi	Deduktiv	Induktiv
Vaqtga nisbatan	Diaxron	Sinxron

Korpusning mohiyati va uni shakllantirishdan mavjud ma'lumotlarni matn holida taqdim eta olish; korpusning hajmiga ko'ra imkon darajada ko'proq axborot berishning imkoniyati; bir marta yaratilgan korpus ma'lumotlaridan takroran, turli masalalarni hal qilishda foydalanish imkoniyati kabi maqsadlar ko'zlanadi. Buning uchun esa korpus oldida turgan korpus asosini tashkil etuvchi tamoyillarni aniqlash,

turli lingvistik o'lichamlar bo'yicha standart bo'lgan korpus razmetkasi qanday bo'lishi kerakligini aniqlash (bunda semantik, janriy, uslubiy, semantik, morfologik razmetka ko'zda tutilmoqda) kabi vazifalarni hal etish lozim bo'ladi.

Bobning *“Taksonomik va meronimik munosabatli leksemalarni korpusda teglash masalalari”* degan bo'limida taksonomik va meronimik munosabatli leksemalarni korpusda teglashning ahamiyati to'g'risida bahs yuritiladi. Korpusga matnlar muammoga tegishlilik nuqtayi nazaridan, ya'ni tilshunos tadqiqotchilar uchun muhim bo'lgan tematik aloqadorlik asosida tanlanadi. Muammolar sohasi – tematik ko'lam ikki aspektdan iborat, o'zida til va nutqqa doir matnlarni jamlaydi. Til aspekti eng tadqiqotlab soha bo'lsa, nutq aspekti til aspektiga doir holatlarni aks ettiruvchi kontekstlardan tarkib topadi. Matnlarni hajmiga ko'ra saralash muammosi esa korpus yaratishning muhim masalalaridan. Mazkur jihat korpusning *mukammallik* jihatini ta'minlab beradi. Biroq korpusga tegishli aynan shu qirra unga taalluqli bo'lgan reprezentativlik jihatiga zid. Negaki matn hajmi korpus ustida olib boriladigan amaliyotni murakkablashtiradi, bu korpusga doir ma'lumotlar tadqiqi uning hajmiga bog'liq tarzda, to'liq realizatsiyani ta'minlab bera olmaydi degani. Demak, korpus oldiga qo'yiladigan maqsad uning yechimiga yetaklovchi tor yondashuvni – tematik sohani tadqiq doirasining ko'lamiga qarab belgilab olishni taqozo etadi. Bunday tanlov tadqiq etilayotgan tilning u yoki bu mezonlarini nutqdagi hodisalarga muvofiq tarzda qamrab olishi lozim. Matnlarni korpusga saralash yuzasidan korpuslar quyidagi turlarga bo'linadi:

- muvozanatga solingan korpuslar;
- monitorli korpuslar.

Taksonomik va meronimik munosabatli leksemalarni korpusda teglash jarayonini amalga oshirish uchun o'zbek tili leksikasida ana shu mavzuviy guruhlarda qancha so'z borligi aniqlanishi lozim. Buning uchun *“O'zbek tilining izohli lug'at”*larida taksonomik va meronimik munosabatli leksemalarning miqdorini aniqlash lozim. Tadqiqotimizda tanlanma sifatida etnonimlar taksonomik mavzuviy guruhining izohli lug'atlarda berilishini tahlil qildik va statistikasini aniqladik Qarang: (1-jadval).

Ko'rganimizdek, *“O'zbek tilining izohli lug'at”*ining 2006, 2020-yilgi nashrlarida etnonimlar mavzuviy 72 tani tashkil qilishi aniqlandi. Aniqlangan birliklarni yuqoridagi jadvaldan to'liqroq ustunni JavaScript kodlari orqali tegga solamiz.

t:hum:etn: (abxaz, avar, adigey, ajam, aymoq, alban, aleut, arab, arman, afg'on, balkar, bantu, barlos, baxtiyoriylar, belarus, beluj, bengal, bijanak, bilkut, bobi, bolgar, bolqor, boshqird, britt, bulog'ochun, bulg'or, burbon, buryat, barqut, vandal, varvar, venger, gabr, gagauz, german, golland, grek, gruzin, gurji, darg'ali, darg'in, dungan, durraniy, do'g'lat, yevrey, yobu, yoqut, jaloyir, jung'or, juhud, jo'gi, jo'jiylar, zangi, zangi, indeys, indonez, ispan, italyan, kabardin, kaltatoy, komi, koreys, kreol, kurd, lazgin, latin, latish, laqay, mavr, major, malik, malika)

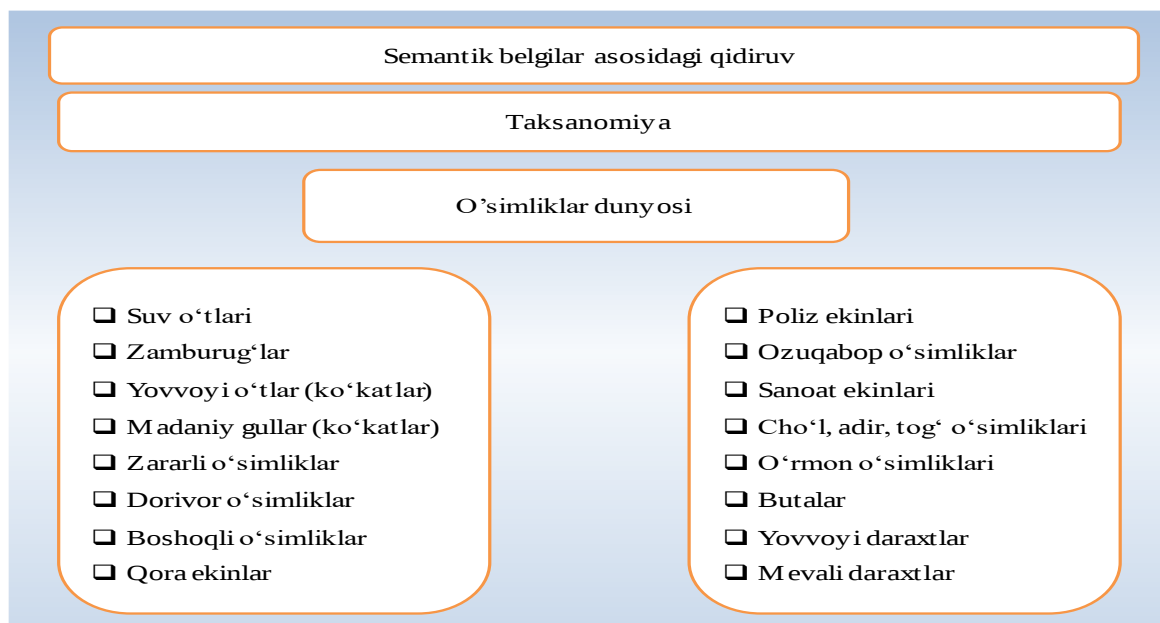
Korpus qidiruv oynasida kerakli belgilarni qo'yib qidiruv amalga oshirilganda shu so'zlar qatnashgan matnlarni topadi. Buning uchun albatta, kerakli tegni tanlash va tanlangan tegni belgilash talab qilinadi (Qarang: 1-rasm). Shundan so'ng korpus interfeysining so'z qidiruvi oynasi orqali tanlanma yoki belgilangan matn, jumla yoki so'zlarni qidiruvga berishimiz lozim hisoblanadi. Natijada bizga tizim xotirasidagi

barcha etnonimlar mavzuviy guruhining barcha soʻzlari va ularning qoʻllanish statistikasi va matnlarda berilish davriyligi boʻyicha barcha maʼlumotlarni olish imkoniga ega boʻlamiz (Qarang: 2-rasm).

Dissertatsiyaning IV bobi **“Oʻzbek tili korpuslari uchun taksonomik va meronimik munosabatli leksemalar lingvistik bazasini shakllantirish asoslari”** deb nomlangan. Bobning *“Oʻzbek tili korpuslari uchun taksonomik munosabatli leksemalar lingvistik bazasi interfeysi mazmuniga qoʻyiladigan talablar”* deb ataluvchi boʻlimida taksonomik birliklarni lingvistik korpusda interfeysga joylashtirish muammosi haqida fikr yuritiladi. Jahon korpus lingvistikasining tadrijiy taraqqiyotidan shu narsa maʼlumki, shu vaqtgacha yaratilgan Milliy korpuslar deyarli hammasining qidiruv menyusida taksonomik munosabatli leksemalar lingvistik bazasi interfeysiga ega boʻladi. Ular maʼlum bir ketma-ketlikda qidiruv tizimiga boʻysindirilgan.

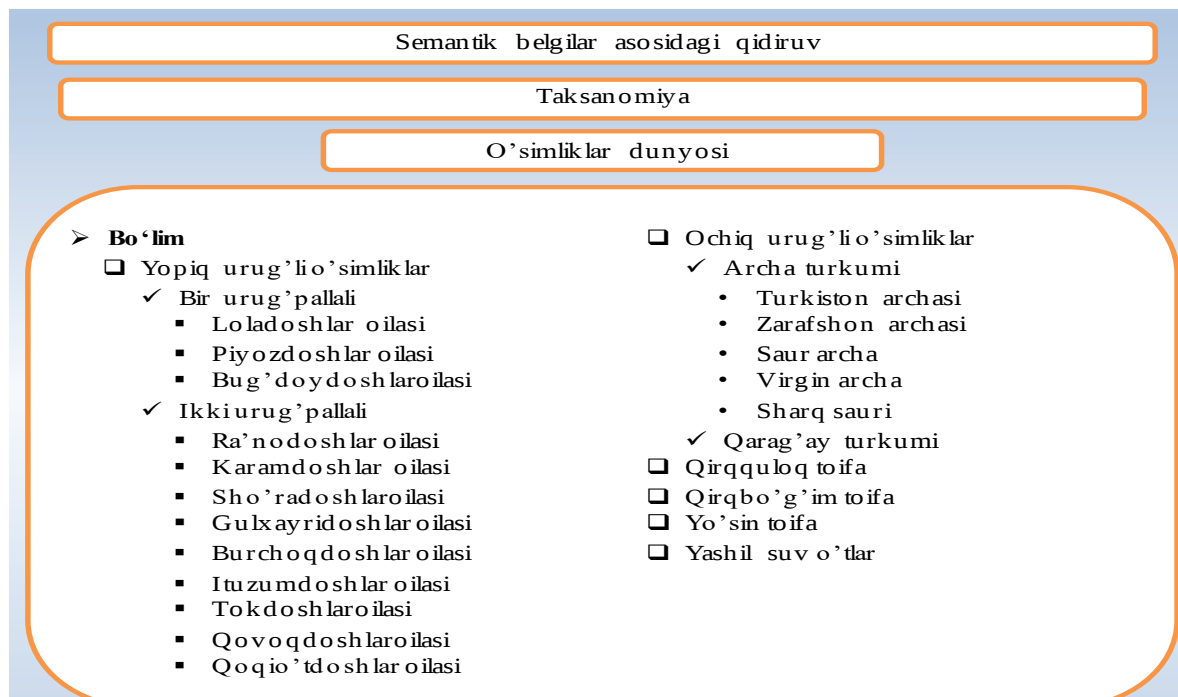
Rus tili Milliy korpusida taksonomik va meronimik munosabatli leksemalar lingvistik bazasi interfeysini tahlil qilish shuni koʻsatdiki, uyadoshlik (taksonomik) va butun-boʻlak (meronimik) munosabat ifodalovchi birliklar ketma-ketlikda muayyan tasnif asosida joylashtirilgan. Shuningdek, tadqiqotchi Oʻ.Xoliyovrov *“Oʻzbek tili taʼlimiy korpusini tuzishning lingvistik asoslari”* nomli dissertatsiyasida oʻzbek tili Milliy korpusining bosh menyusida interfeysi joylashuvi haqida shunday deydi: *“Oʻzbek tili korpusini ishlab chiqishda umumiy talablar bilan birga oʻzbek tili xususiyatlari, oʻquvchi yoshiga mos jihatlariga alohida eʼtibor qaratish lozim”*⁵⁸.

A.Sobirov oʻsimliklarni taksonomik birliklarga boʻlib tahlil qilish haqida fikr bildiradi. Oʻsimliklar hozirgi zamon sistematiikasida quyidagi birliklarga (taksonlarga) boʻlib yuboriladi: boʻlim, sinf, tartib, oila, avlod, tur⁵⁹. A.Sobirov tasnifiga asosan taksonomik qidiruv interfeysini quyidagicha loyihalash mumkin:



⁵⁸ Холиёров Ў. Ўзбек тили таълими корпусини тузишнинг лингвистик асослари: Филол. фан. бўйича фалсафа доктори (PhD) дисс. – Термиз, 2021. – Б 38.

⁵⁹ Собиров А. Ўзбек тилининг лексик сатҳини системалар системаси тамойили асосида тадқиқ этиш. – Тошкент: Маънавият, 2004. – Б 93.



Bobning “*O‘zbek tili korpuslari uchun meronimik munosabatli leksemalar lingvistik bazasi interfeysi tuzulish*” deb ataluvchi bo‘limida taksonomik birliklarni lingvistik korpusda interfeysga joylashtirish muammosi haqida fikr yuritiladi. O‘zbek tili korpuslari uchun meronimik munosabatli leksemalar lingvistik bazasi o‘zbek tilshunosligida shu vaqtga qadar mana shu sohada amalga oshirilgan tadqiqotlarga tayanadi. Xususan, tilshunos olim B.Qilichev meronimik hodisani quyidagicha semantik maydonlarga bo‘lib tahlil qiladi: “Tana” partonimi, “Bosh” partonimi, “Oyoq” va “qo‘l” bo‘laknomlari, “Bosh” butunnom, “O‘simta” semali partonimlar, “Ko‘z”, “quloq”, “og‘iz” “burun” partonimlari, “Tekislik” semali partonimlar, “Hayvon” partonimik LMGdagi “bosh” bo‘laknomi⁶⁰. Ushbu tasnifni biz o‘zbek tili milliy korpusi interfeysida o‘ziga xos tarzda ifodalash mumkin. Meronimik munosabatlar qatoridan birgina tana meronimini tahlil qilsak, tilshunos olim B.Qilichev quyidagicha tasnif qiladi. Tana – butunnom sifatida ham “tana” o‘ziga xos xususiyatlarga ega. Chunonchi, “odam” tanasining bo‘laknomlari quyidagilar: Tana I: yelka, ko‘krak, qorin, bel.

“Mol” partonimik LMGsidagi “tana” butunnomini quyidagilar yuzaga keltirgan: Tana II: to‘sh, qorin, orqa, bel.

“Daraxt” partonimik LMGsidagi “tana” butunnomi quyidagi bo‘laknomlardan iborat: Tana III: po‘stloq, parda, yog‘och, mag‘z kabi.

Ko‘rinib turibdiki, “tana I” va “tana II” o‘zaro bo‘laklar tasnifiga ko‘ra ham o‘xshash va “tana III”dan keskin farq qiladi. Bulardan xulosa qilib, quyidagilarni bayon qilishimiz mumkin:

Birinchidan, lug‘aviy qatorlar hosil qilish jihatdan “tana I” va “tana II” bir-biriga yaqin hodisa, “tana III”da esa bunday xususiyat sezilmaydi.

Ikkinchidan, qurshov jihatdan ham “tana I” va “tana II” butunliklarining bo‘laknomlari bir xil xususiyatga ega. Bu o‘rinda ham “tana III” o‘z bo‘laklari tabiatidan kelib chiqib, “tana I”, “tana II” qurshovidan chetda turadi. Faqatgina ma’no

⁶⁰Qilichev B. Leksikada so‘zlararo butun-bo‘lak munosabati. Monografiya. Toshkent. 2019. – 78.B.

zanjiriga munosabati masalasida bu uchala butunnom bir guruhda tura olishi mumkin. Chunonchi, har uchala butunnomning ifodalagan ma'nosi bitta: "Ma'lum butunliklarning qolgan bo'laklarini birlashtirib turuvchi vosita"⁶¹. O'zbek tili milliy korpusida bitta "tana" meronimi uch xil tasnifda ifodalanishini ko'rish mumkin. Inson, mol va daraxt partonimik leksik-ma'noviy guruhida uchala holatda bir-biri bilan bog'liq holda ifodalash lozim.

O'zbek tili milliy korpusida muayyan qismlardan tuzilgan butunlik va shu butunlikni tashkil qiluvchi qism(meronim)larga o'tishini ko'rish mumkin. Masalan, tana xolonimi qo'l, oyoq, bosh meronimlardan tashkil topgan:

1. Qo'l – tirnoq, barmoq, kaft, bilak, tirsak, yelka.
2. Oyoq – tirnoq, kaft, tovon, to'piq, ilik, tizza, boldir, son.
3. Bosh – soch, qosh, qovoq, kiprik, ko'z, quloq, burun, og'iz, peshona, ensa, iyak⁶².

XULOSA

1. Taksonomiya butun borliqning murakkab iyerarxik tuzilmali sohalarini jonli tabiat, yerning tuzilishi; geografiya, geologiya, tilshunoslik, etnografiya va obyektlarini tasniflash va tartibga solish nazariyasi, shu haqdagi fan. Tilshunoslikda taksonomiya matndagi lisoniy birliklarni ajratib olish va ularning tartibini va tarqalishini tahlil qilish orqali ularning xususiyatlarini o'rganish vositasi.

2. Taksonomiya hodisasi falsafa, biologiya, matematika, pedagogika, tilshunoslik fanlarida tasniflash va sistema sifatida jarayonlarni tizim sifatida o'rganishga asoslangan metod.

3. Real borliqdagi butun-bo'lak munosabatining leksikada aks etishi masalasi azaldan tilshunos olimlarning e'tiborini tortib kelganligi sir emas. *butun* tushunchasini *xolonim*, *qismni meronim* yoki *partonim* deb atashni ushbu hodisani chuqurroq tahlil qilish imkonini beradi. Xolo-meronimiya yoki partonimiyaning giper-giponim uyadan farqi shundaki, giponimiyada maydonning bir uzvi (a'zosi) boshqalariga nisbatan qarama-qarshi qo'yilib, ular o'rtasidagi ma'no munosabatlari aniqlansa, partonimiyada alohida olingan uzv (a'zo)ning uni tashkil qilayotgan ichki a'zolari bilan munosabati tushuniladi. *Partonim* atamasi faqat bo'lakka nisbatan olingan. Shu bois butun va bo'lakni ifodalash uchun taklif qilingan *xolonim* va *meronim* atamalaridan foydalanish lozim.

4. O'zbek tili izohli lug'atlarining turli nashrlarida taksonomik va meronimik munosabatli leksemalarning miqdorini tanqidiy ko'rib chiqish shuni ko'rsatdiki, taksonomik mavzuviy guruhlardagi leksemalar miqdori o'sib borish ko'rsatkichiga ega ekanligi meronimik mavzuviy guruhlar esa ba'zi o'rinlarda kamayish ko'rsatkichiga ega ekanligi aniqlandi.

5. Yuqoridagi tasnif natijasi shuni ko'rsatadiki, o'zbek tili izohli lug'atidagi taksonomik va meronimik munosabatli leksemalar leksik-semantik razmetkalashda to'laqonli leksikografik baza sifatida xizmat qilolmaydi. Bu esa razmetkalash jarayonining aniqligiga ta'sir etadi. Shunga qaramay, ulardagi taksonomik va

⁶¹ Qilichev B. Leksikada so'zlararo butun-bo'lak munosabati. Monografiya. Toshkent. 2019. – B 42.

⁶² Собиров А. Ўзбек тилининг лексик сатҳини системалар системаси тамойили асосида тадқиқ этиш. – Тошкент: Маънавият, 2004. – Б. 43.

meronimik munosabatli birliklardan sharh va tavsif uchun asos manba sifatida foydalanish lozim

6. Korpusga taksonomik va meronimik munosabatli leksemalar bazasini kiritish mavjud ma'lumotlarni matn holida taqdim eta olish; korpusning hajmiga ko'ra imkon darajada ko'proq axborot berish; bir marta yaratilgan korpus ma'lumotlaridan takroran, turli masalalarni hal qilishda foydalanish imkonini beradi. Buning uchun taksonomik va meronimik munosabatli leksemalar bazasini tuzishning tamoyillarini ishlab chiqish, bu boradagi muammolarni bartaraf etish lozim. Korpusni tashkil etuvchi birliklarni aniqlash, turli lisoniy parametrlar bo'yicha standart razmetkasi ishlab chiqish asosiy masalalardan biri.

7. Lingvistik korpuslarda taksonomik va meronimik munosabatli leksemalarni tizim interfeysi orqali yuzaga chiqarish jarayoni ma'lum bir algoritm asosida amalga oshadi. Tizim interfeysida ma'lumotlar paydo bo'lishi uchun birinchi navbatda taksonomik va meronimik munosabatli leksemalarning ma'lumotlar ombori yaratilishi kerak undan so'ng esa, JavaScript kodi orqali tizim menyusiga maxsus biriktirilgan teglar vositasida kerakli bo'lgan taksonomik birlik qidiruv oynasiga yozilib qidiruv beriladi ana shundan so'ng biz izlayotgan taksonomik birlik qidiruvga beriladi va biz tomonimizdan kiritilgan barcha leksik birliklar tizim interfeysida yaqqol foydalanish uchun tayyor holga keltiriladi.

8. Sun'iy intellekt tizimi doirasidagi tabiiy tillarning korpus interfeys, ekspert tizimlari, neyron tarmoqlar, lingvoanalizatorlar, gapiruvchi avtomatlar – barchasi kompyuter modellashtirish natijasi hisoblanadi. Milliy korpus interfeysi ham qaysidir ma'noda modellashtirish natijasidir.

**ONE-TIME SCIENTIFIC COUNCIL AT SCIENTIFIC COUNCIL № PhD.
03/31.03.2021.Fil.78.04 ON AWARD OF SCIENTIFIC DEGREES AT
TERMEZ STATE UNIVERSITY**

TERMEZ STATE UNIVERSITY

ESHMUMINOV ASQAR ALLAMURADOVICH

**FORMATION OF LINGUISTIC BASE OF TAXONOMIC AND
MERONYMIC LEXEMES FOR UZBEK LANGUAGE CORPUS**

**DISSERTATION OF DOCTOR OF PHILOLOGY (DSc)
ABSTRACT**

10.00.01 – Uzbek language

TERMEZ– 2022

The theme of the dissertation for the doctoral dissertation (DSc) was registered at the Supreme Attestation Commission of the Cabinet of Ministers of the Republic of Uzbekistan under B2021.1.DSc / Fil284.

The dissertation has been prepared at the Termez State University.

The abstract of the dissertation is posted in three languages (Uzbek, English, Russian (resume)), is placed on the website of Scientific Council (www.navoiy-uni.uz) and on the website of 'Ziyonet' information and educational portal www.ziyonet.uz

Scientific supervisor:

Mengliyev Bakhtiyor Rajabovich
Doctor of Philological Sciences, Professor

Official opponents:

Pardayev Azamat Baxranovich
Doctor of Philological Sciences, Professor

Abjalova Manzura Abdurashetovna
Doctor of Philological Sciences, Docent

Ergasheva Guli Ismailovna
Doctor of Philological Sciences, Docent
Fergana State University

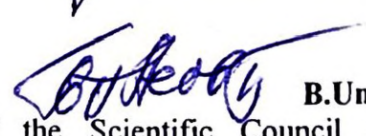
Leading organization:

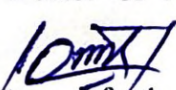
The defence of the dissertation will be held on august «22», 2022 at 10⁰⁰ at the meeting of the Scientific Council № PhD. 03/31.03.2021.Fil.78.04 at the Termez State University. Address: 43 Barkamol Avlod Street, Termez, Surkhandarya Region, 190111 Uzbekistan. Tel: (876) 221-74-55; fax: (876) 221-71-17; e-mail: termizdu@umail.uz).

The dissertation can be found in the Information Resource Center of Termez State University (registered with the number 67). Address: 43 Barkamol Avlod Street, Termez, Surkhandarya Region, 190111 Uzbekistan. Website: www.tersu.uz; E-mail: termizdu@umail.uz

The abstract of dissertation was distributed on “11” august 2022.
(Registry record №. 12 dated “11” august 2022.)




B. Umurkulov
Chairman of the Scientific Council awarding scientific degrees, Doctor of Philological Sciences, Professor


N. Omonturdiyev
Scientific Secretary of the Scientific Council awarding scientific degrees, PhD on philological sciences


Sh. Mahmaraimova
Chairman of the Scientific Seminar under the Scientific Council awarding scientific degrees, Doctor of Philological Sciences, docent

INTRODUCTION (doctoral (DSc) dissertation annotation)

Relevance and necessity of the research topic. In the middle of the twentieth century in the world corpus linguistics began the period of intensive processing of natural language, the active application of translation problems in computer programs through machine translation. So far, all of the world's leading languages: English, German, French and Russian have created a database of taxonomic and meronymically related lexemes, which are conveniently located directly on the system interface of the linguistic corps. This underscores the practical importance of the ability of system users to obtain sufficient information from their native language or from linguistic corporations created in advanced languages.

In world linguistics, at the beginning of the XXIst century, issues such as providing linguistic corps with quality lexical resources, improving the quality of corpus use, and linguistic modeling of language were on the agenda. As a result, the active use of lexical units in the language in the information retrieval system to create a taxonomic and meronymic database in the corpus of languages with a linguistic corpus has become popular. Consequently, the creation of language corpus, their linguistic annotation, and the creation of taxonomic and meronymic databases have become the scientific and theoretical problems of corpus linguistics.

In the field of Uzbek corpus linguistics, attention is paid to scientific research on the creation of a national corps, an educational corps. In the course of this work, the issue of creating a database of taxonomically and meronymically related lexemes for the Uzbek National Corps has not been studied in a scientific and theoretical way. After all, "... to preserve the purity of the state language, to enrich it and to increase the speech culture of the population; Ensuring⁶³ the active integration of the state language into modern information technologies and communications is one of the most important tasks facing the Uzbek corpus linguistics today. Accordingly, issues such as the creation of a database of taxonomically and meronymically related lexemes and the creation of their linguistic support are important.

The given research work serves to a certain extent to perform the tasks specified in the Resolutions of the President of the Republic of Uzbekistan PF-4997 «On the establishment of the Tashkent State University of Uzbek Language and Literature named after Alisher Navoi» dated May 13, 2016, PF-4947 "On the Strategy of Actions for the Further Development of the Republic of Uzbekistan" dated February 7, 2017, PF-5850 "On measures to radically increase the prestige and status of the Uzbek language as the state language" dated October 21, 2019, PF-6084 "On measures to further development of the Uzbek language and improve language policy in the country" dated October 20, 2020, PQ-2789 "On measures about measures of further improvement of the activities, organization, management and funding of research works of the Academy of Sciences February dated February 17, 2017, PQ-4479 " On the wide celebration of the thirtieth anniversary of the "Law of the Republic of

⁶³Ўзбекистон Республикаси Президенти Шавкат Мирзиёевнинг 2020 йил 202 октябрдаги «Мамлакатимизда ўзбек тилини янада ривожлантириш ва тил сиёсатини такомиллаштириш чора-тадбирлари тўғрисида»ги ПФ-6084-сон фармони // <https://lex.uz/docs/5058351>

Uzbekistan on State Language” dated October 4, 2019, and the The Prezident`s Address to the Oliy Majlis in January 24, 2020 and other normative legal acts.

The relevance of the research to the priorities of the development of science and technology in the country. The study was conducted in accordance with the priority of the development of science and technology of the republic I. “Social, legal, economic, cultural, spiritual and educational development of an informed society and a democratic state, the development of an innovative economy.”

Review of foreign research on the topic of research.⁶⁴ Scientific research aimed at studying the linguistic supply for the National Corps created in world linguistics is carried out in leading research centers and higher education institutions, including Brown University (USA), Oxford University (UK), Cambridge University (UK), Humboldt University (Germany), Boston University, University of New York, University of Pennsylvania, University of California (USA); Indiana University (USA), The University of Sorbonne (France), Universitat de Valencia (Spain), University of Zurich (Switzerland), Upsalla University (Sweden), Kiofo University (Japan), Australian National University (Australia), Istanbul Technical University (Turkey), Moscow State University (Russia), Moscow State Linguistic University (Russia), St. Petersburg State University (Russia), Ural State Pedagogical University, Kharkiv State Pedagogical University (Ukraine), Baku State University (Azerbaijan), Republic of Tatarstan Institute of Applied Linguistics and Semiotics of the Academy of Sciences (Republic of Tatarstan, Russia), Kazan Federal University (Republic of Tatarstan, Russia), V. Vernardsky Tavriya Academy, Crimean Federal University (Republic of Crimea, Russia), as well as Uzbekistan named after Mirzo Ulugbek National University, Tashkent State University of Uzbek Language and Literature named after Alisher Navoi, Al-Khwarizmi i Tashkent University of Information Technologies and its branches, Bukhara State University, Samarkand State University, Termez State University (Uzbekistan).

In world linguistics, the following scientific results have been obtained from the research on linguistic support, automatic tagging of lexemes, tagging and marking of texts in the information retrieval system for linguistic corpora: the features of automatic processing of natural language (analysis of language units) (University of New Jersey, University of California, the USA; Boston University, New York University, the USA; Universitat de Valencia, Spain); Theoretical bases of lexical analysis, general description, principles of operation, description of linguistic modules for software (Kiofo University, Japan; Australian National University, Australia; Istanbul Technical University, Turkey; Moscow State University, Moscow State Linguistic University, St. Petersburg State University, Russia); models and programs for processing lexical units collected for linguistic corps (Baku State University, Azerbaijan); development of natural language processing, language corpus, translators in computer linguistics of Turkic languages (Institute of Applied Linguistics and Semiotics of the Academy of Sciences of the Republic of Tatarstan, Kazan Federal University, V. Vernardsky Tavriya Academy, Crimean Federal University, Russia);

⁶⁴ Диссертация мавзуси бўйича хорижий илмий тадқиқотлар шарҳи [google.scholar.com](https://www.google.com/scholar); www.bu.edu; americanenglish.state.gov; micros.uz; www.uv.es; www.zora.uzh.ch; cleee@snu.ac.kr; www.u-tokyo.ac.jp/en; www.ras.ru; adwww.ecu.edu.au; unice.fr/en; www.sophia.ac.jp/eng; www.msu.ru; www.vsu.bu; www.education.ua/universities; www.kaznu.kz; www.navoiy-uni.uz ва бошқа манбалар асосида амалга оширилди.

The principles of formation of the Uzbek language author's corps, the linguistic base of the language corps and its linguistic support, the problems of creating a database of taxonomic and meronymic related lexemes for the Uzbek language corps, the basis for creating tools of the Uzbek-English parallel corps (National University of Uzbekistan named after Mirzo Ulugbek, Tashkent State University of Uzbek Language and Literature named after Alisher Navoi, Samarkand State University, Bukhara State University, Tashkent State Information Technologies university named after Al-Khorazmi, Termez State University, Uzbekistan).

In world corpus linguistics, research is being conducted in the following priority areas of creating a linguistic corpus interface based on a database of natural languages: development of linguistic support for language corpus; attach tags in the appropriate order for the developed linguistic support; forming a base of taxonomic and meronymically related lexemes for the linguistic corporations being created.

Degree studiedness of the problem. The study of theoretical and methodological problems of linguistic support in world linguistics began in the 60s of the last century. It is recognized that without the language of information retrieval, the computer will remain a machine, from the «mechanical» stage to the «logical-linguistic» stage, the information retrieval system will be developed⁶⁵. Problems of natural language processing in world computer linguistics A.Keyler, K.V.Linden, N.Vard⁶⁶. Commented and disclosed by M.A.Mohri⁶⁷. Technology and means of automatic word processing V.Axo Alfred, S.Lam Monica, S.Ravi, D.Ulman Jeffrey⁶⁸. Theoretical bases of morphological analysis S.Y.Toldova, AA.Bonch-Osmolovskaya.⁶⁹ T.Sadikov⁷⁰. described in his work. In Russian computer linguistics it is worth noting the works of O.V.Dereza, D.A.Kayutenko, A.S.Fenogenova, M.Hakimov on morphological analyzer, its general description, principles of operation, grammatical category of language, description of linguistic modules for the program⁷¹. The importance of linguistic support in the information retrieval system is revealed in the works of A.B.Antopolsky, Y.M.Arsky, G.G.Belonogov, B.A.Kuznesov, V.N.Beloozerov, R.R.Mdivani, B.A.Glinsky, A.I.Mikhailov, V.A.Mishin, V.V. Morozov.⁷² Features of morphoanalysis were studied by D.Rudolf, the

⁶⁵ Соколов А.В. Автоматизация библиографического поиска / А.В.Соколов. – М.: Книга, 1981. – 167 с. – С.91.

⁶⁶ Speech and Language Processing. Daniel S. Jurafsky and James H. Martin. Contributing writers: Andrew Kehler, Keith Vander Linden, Nigel Ward 2000y. Prentice Hall, Englewood Cliffs, New Jersey 07632. Pages: 950.

⁶⁷ Mohri M.A. Finite-state transducers in language and speech processing. Computational Linguistics. – С. 269-312.

⁶⁸ Ахо Альфред В., Лам Моника С, Сети Рави, Ульман Джеффри Д. Компиляторы: принципы, технологии и инструментарий, 2-е изд.: Пер. с англ. – М.: ООО “И.Д. Вильямс”, 2008. – 1184 с.

⁶⁹ Толдова С.Ю., Бонч-Осмоловская А.А. Автоматический морфологический анализ // Фонд знаний «Ломоносов». М., 2011. www.lomonosov-fund.ru/enc/ru/encyclopedia:0127430

⁷⁰ Садыков Т., Кочконбаева Б. Об оптимизации алгоритма морфологического анализа // Шестая Международная конференция по компьютерной обработке тюркских языков «Turklang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

⁷¹ Дереза О.В., Каютенко Д.А., Феногенова А.С. Автоматический морфологический анализ для русского языка сравнительный анализ систем // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018.; Хакимов М. Логико-лингвистические модели русского языка // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

⁷² Антопольский А.Б. Лингвистическое обеспечение АСНТИ: структура и проблема совместимости // НТИ. Сер.2, 1983. – №7. – С.17-25; Арский Ю.М. Информационные ресурсы для устойчивого развития общества / Ю.М.Арский, А.И.Черный // Междунар. форум по информ. и док. – 2000. – Т. 25, № 1. – С. 3-9.; Белоногов Г.Г. Автоматизированная обработка научно-технической информации. Лингвистические аспекты / Г.Г. Белоногов, Б.А. Кузнецов, А.П. Новоселов // Итоги науки и техники. Сер. Информатика / ВИНТИ. –1984. –Т.8. – 316;

general significance of morphoanalysis by A.Yermakov⁷³, models and programs of automatic morphological and syntactic processing of text by I.M.Nojov.⁷⁴ In Turkish computer linguistics, the issues of computer processing of natural language, including the construction of morphoanalysts, have been studied in detail. Among such works are G.Eryigit, Ye.Adali⁷⁵, A.V.Dibo, A.V.Sheymovich⁷⁶, J.Suleymanov⁷⁷, A.R.Gatiatullin, A.M.Bashirov⁷⁸, P.V.Jeltov⁷⁹, N.A.Israilova, P.S.Bakasova⁸⁰, N.A.Leontev⁸¹, V.V.Kukanova, A.Y.Kadjiev⁸² can be included in the study of morphological analyzers. The system of morphological analysis of Turkic languages and the description of their models, the breadth of research devoted to the analysis strengthened the theoretical⁸³ foundations of this dissertation.

Белоозеров В.Н. Ведение и совершенствование Государственного рубрикатора НТИ (ГРНТИ) / В.Н. Белоозеров, В.И. Федосимов, Г.А.Кривенцова // Информ. продукты, процессы и технологии: НТИ-96: Материалы конф. – М., 1996. – С. 60-69.; Мдивани Р.Р. Лингвистическое обеспечение автоматизированной информационной системы по общественным наукам: состав и структура / Р.Р.Мдивани, В.А.Глинский // Вопр. информ. теории и практики. 1985. – №52. – С. 37-55.; Михайлов А.И. Проблемы лингвистического обеспечения АСНТИ / А.И. Михайлов // Вопр. информ. теории и практики. 1984. – № 51. – С. 5-11.; Мишин В.А. Лингвистическое обеспечение библиотечных автоматизированных систем / В.А. Мишин, Л. Санчес-Мехидо // Сов. библиотековедение. 1981. – № 2. – С. 84-93.; Морозов В.В. Лингвистическое обеспечение автоматизированных систем научно-технической информации / В.В. Морозов, Б.Р. Певзнер. – М.: ИПКИР, 1987. – 76 с.

⁷³ Rodolfe D. Computational Linguistic Text Processing: Lexicon, Grammar, Parsing and Anaphora Resolution. Nova Science Publishers, Inc. – NewYork, 2008. – P. 4-5.; Ермаков А. Морфологический анализатор - основа поисковых систем // <https://www.kv.by/archive/index2004154301.htm>

⁷⁴ Ножов И.М. Морфологическая и синтаксическая обработка текста (модели и программы): диссертация канд. наук. – Москва, 2003. – 190 с.

⁷⁵ Eryigit G., Adali E. An affix stripping morphological analyzer for Turkish \\\ Proceedings of IASTED International Conference ARTIFICIAL INTELLIGENCE AND APPLICATIONS. February 16-18, 2004, Innsbruck, Austria. – P. 299-304.

⁷⁶ Дыбо А.В., Шеймович А.В. Автоматический морфологический анализ для корпусов хакасского и древнетюркского языков / Научное обозрение саяно-алтая рецензируемый научный журнал Номер 2(08), 2014 (Периодичность – 2 раза в год. серия: Филология.Выпуск 2). С. 9-31.

⁷⁷ Сулейманов Д.Ш., Гильмуллин Р.А., Гатауллин Р.Р. Морфологический анализатор татарского языка на основе двухуровневой модели морфологии / Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.; Сулейманов Д.Ш., Гатиатуллин А.Р. Модель татарской аффиксальной морфемы и ее реализация // Серия: Интеллект. Язык. Компьютер. – Вып.4. – Казань: ФЭН. – 1996. – 113-127.

⁷⁸ Гатиатуллин А.Р., Баширов А.М. Морфологический анализатор тюркских словоформ на базе структурно-функциональной модели тюркской морфемы // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.

⁷⁹ Желтов П.В. Разработка морфологического анализатора чувашского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.

⁸⁰ Исраилова Н.А., Бакасова П.С. Морфологический анализатор кыргызского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.

⁸¹ Леонтьев Н.А. Морфологический анализатор якутского языка // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с. – С. 276-279.

⁸² Куканова В.В., Каджиев А.Ю. Алгоритм работы морфологического парсера калмыцкого языка // В сборнике: Писменото наследство и информационните технологии. El'Manuscript-2014 Материалы от V международна научной конференции. Отговорни редактори В.А.Баранов, В.Желязкова, А.М.Лаврентьев. 2014. – С. 116-119.

⁸³ Kemal Oflazer. Two-level Description of Turkish Morphology. Literary and Linguistic Computing, – Vol. 9, No 2, – 1994.; Altintas K., Cicekli I. A morphological analyzer for Crimean Tatar // Proceedings of the 10th Turkish Symposium on Artificial Intelligence and Neural Networks (TAINN'2001). – 2001. – P. 180-189.; Çağrı Çöltekin (2014) A Set of Open Source Tools for Turkish Natural Language Processing In: Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14) Ed. by N. Calzolari et al. 1079-1086.; G. Kessikbayeva and I. Cicekli, "Rule based morphological analyzer of Kazakh language," in Proceedings of the 2014 Joint Meeting of SIGMORPHON and SIGFSM. Baltimore, Maryland: ACL, June 2014, pp. 46-54.; Tantug C., Adali E., Oflazer K. Computer Analysis of the Turkmen Language Morphology, 5th International Conference on NLP (FinTAL 2006), Turku, Finland, 186-193,

Theoretical ideas for the development of morphoanalyzers were also put into practice, and morphoanalyzers for European, Russian and Turkish languages were developed⁸⁴. In Uzbek linguistics, a lot of research has been done on computer linguistics, lexicographic processing of text and linguostatistical analysis. The researches of M.Mahmudov, M.Ayimbetov, S.Karimov, G.Jumanazarova, A.Babanarov, D.Urinbaeva, A.Norov, S.Muhammedov, N.Abdurahmanova can be mentioned as such works⁸⁵. The development of a morphological analyzer of the Uzbek language has not been the subject of special research, but some studies have suggested specific aspects of the problem. For example, the creation of linguistic software⁸⁶ for computer programs based on action verbs, Linguistic support of Uzbek-English machine translation, principles of formation of Uzbek author's corpus, problems of graphic⁸⁷ analysis of Uzbek language units, principles of formation⁸⁸ of linguistic base of language corpus are studied in the monographic plan⁸⁹. Also, the linguistic basis of semantic tagging of Uzbek nouns⁹⁰. Problems of creating Uzbek-English parallel corps⁹¹, Automatic analysis of Uzbek nouns,⁹² published a number of articles on formal word formation models⁹³, Software aspects of automating the

2006.; Orhun M., Tantug C., Adali E. Rule Based Analysis of the Uyghur Nouns, International Journal of Asian Lang. Proc.,19(1), 33-44, 2009.

⁸⁴ Морфологический анализатор текста на русском языке mystem [Электронный ресурс]; // Компания Яндекс [сайт] 2003-2013. URL: <http://company.yandex.ru/technologies/mystem/>; <http://tamgasoft.kg/morfo/>; <http://tatmorphan.pythonanywhere.com/>

⁸⁵ Махмудов М.А., Пиотровская А.А., Садыков Т. Система машинного анализа и синтеза тюркской словоформы // Переработка текста методами инженерной лингвистики. – Минск, 1982.; Мухаммедов С.А. Статистический анализ лексико-морфологической структуры узбекских газетных текстов: Автореф. дисс... канд. фил. наук. – Ташкент, 1980.; Бабанаров А. Разработка принципов построения словарного обеспечения турецко-русского машинного перевода: Автореф. дисс... канд. фил. наук. – Л.,1981.; Айымбетов М.К. Опыт лингвостатистического анализа лексики и морфологии каракалпакского публицистического текста: Автореф. дисс... канд. фил. наук. – Ташкент, 1987.; Каримов С., Каршиев А., Исроилова Г. Абдулла Қаҳҳор асарлари тилининг луғати. Алфавитли луғат. Частотали луғат. Терс луғат. – Тошкент, 2007.; Ризаев С. Ўзбек тилининг лингвостатистик тадқиқи: Фил. фан. док. дисс...автореф. – Тошкент, 2008.; Пўлатов А. Компьютер лингвистикаси. – Тошкент: Akademnashr, 2011.; Норов А. Компьютер лингвистикаси асослари. – Қарши, 2017. – 136 б. Жуманазарова Г.У. Фозил Йўлдош ўғли дostonлари тилининг лингвопoэтикаси: Фил. фан. док. дисс...автореф. – Тошкент, 2017.; Ўринбоева Д.Б. Ўзбек фольклори матнларининг лингвостатистик тадқиқи. – Тошкент: Фан, 2010.; Ўринбаева Д. Халқ оғзаки ижоди: жанрий-лисоний ва лингвостатистик тадқиқ муаммолари: фил. фан. док. дисс. автореф. – Самарканд, 2019. – 74 б.

⁸⁶ Мухаммедова С. Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. Методик қўлланма. – Тошкент, 2006.;

⁸⁷ Абжалова М.А. Ўзбек тили онтологиясини яратиш тамойиллари: Филол.фан.бўйича фалсафа доктори (PhD)...дисс. автореф. – Фарғона, 2022. – 62 б.

⁸⁸ Хамроева Ш. Ўзбек тили морфологик анализаторининг лингвистик таъминоти» фил. фан. док. дисс. автореф. – Фарғона, 2021. – 74 б.

⁸⁹ Эшмўминов А.А. Ўзбек тили миллий корпусининг синоним сўзлар базаси: Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Қарши, 2019. – 140 б.;

⁹⁰ Akhmedova D.B., Mengliev B.R. Semantic Tag Categories in Corpus Linguistics: Experience and Examination International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-3S, October 2019. – P. 208-212.

⁹¹ Karimov R., Mengliev B. Theoretical fundamentals of uzbek-english parallel corpus / Journal of critical reviews. ISSN- 2394-5125. – VOL 7, ISSUE 17, 2020. – P. 73-76.; Karimov R.A., Mengliev B.R. The Role of the Parallel Corpus in Linguistics, the Importance and the Possibilities of InterpretationInternational Journal of engineering and Advanced Technology (IJEAT). ISSN: 2249 – 8958, Volume-8, Issue-5S3 July 2019. – P. 388-391.

⁹² Орхун М. Computational analysis of Uzbek nouns / Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

⁹³ Турсунов А.Вопросы словообразования в формальных моделях тюркских языков (на примере узбекского языка) // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». Труды конференции) – Ташкент, 2018. – 320 с.

morphological analysis of nouns and adjectives in Uzbek are studied⁹⁴. These studies are relevant because they study the problems of computer linguistics, but the issue of automatic morphological analysis of the Uzbek language - morphoanalyzer and the principles of creating its linguistic support is not on the agenda. In the process of preparing the dissertation, the above-mentioned scientific researches were studied in detail, treated where necessary and used in research.

The connection of the research with the research plans of the higher education or scientific-research institution where the dissertation was completed.

The dissertation was completed within the framework of the practical project of the Tashkent State University of Uzbek Language and Literature AM-FZ-201908172 “Creation of the educational corps of the Uzbek language.”

The aim of the research is to create a linguistic base of taxonomically and meronymically related lexemes for Uzbek language corpora.

The tasks of the study are:

to study the general description of the phenomena of taxonomy and meronymy in linguistics and its role in the formation of language corpora;

study the theoretical foundations of thematic field analysis of language units;

study the structure, composition and functions of the linguistic corpus interface;

determine the importance of the process of tagging taxonomic and meronymically related lexemes;

to study the commonalities and differences of the interface of taxonomic and meronymically related lexemes in the language corpus;

determine the structure and content of the linguistic database for the Uzbek language corpora;

The interface of the search box of the Uzbek language corpus and the search for language units in it, linguistic modeling of the selection process.

The object of research. Taxonomic and meronymic lexemes in Uzbek language were selected as the object of research.

The subject of the research is the basis for the formation of the linguistic base of taxonomic and meronymically related lexemes for the Uzbek language corpus.

Research methods. Methods of classification, description, comparison, statistical analysis, modeling were used to cover the research theme.

The scientific novelty of the research is:

types of inter-lexeme lexical-semantic relations in the language, such as synonymic, hyperonymic, graduonymic, their interaction with taxonomic and meronymic relations as a mutual system, general formal and similar in terms of content, general and specific according to the thematic structure, macro and micro according to division into field and group, differences in terms of single meaning and multiple meaning are identified;

⁹⁴ Болтаев Т.Б., Ибрагимов С.И. О проекте программной системы морфологического анализа узбекского языка // <http://buxdu.uz/index.php/uz/>; Болтаев Т.Б., Кузьминов Т.В., Поттосин И.В. О структурном конструировании программ и инструментах его поддержки // Среда программирования: методы и инструменты. Новосибирск, 1992, С22-37.; The Structured Constructing as a Discipline of Safe Programming and Instruments Supporting It / Aniskov M.I., Boltaev T.B., Kochetov D.V. et al // Instrumental Congress on Computer Systems and Applied Mathematics CSAM'93. St-Petersburg, July 19-23.

the processes of creating a linguistic database for the software of language corpora and the amount of lexemes related to taxonomy and meronymy at the lexical level of the Uzbek language is identified;

It is substantiated that taxonomically related lexemes in the explanatory dictionaries of the Uzbek language are quantitatively more than meronymically related lexemes and thematically equal;

The monitor structure of the semantic search interface of the Uzbek language corpora, the structure of the connection of lexemes in the corpus, and the semantic tagging program projects were developed for the process of tagging the lexical units of the Uzbek language based on their mereological and taxonomic signs;

The interface of the search window for lexical-semantic related units in the Uzbek language, the principles of classification based on thematic and ranking, and the linguistic model for language corpora have been defined.

The practical results of the study are as follows:

From the point of view of lexical-semantic relations the general description of phenomena of taxonomy and meronymy is described and its importance in creation of language corpus is proved;

a linguistic base of taxonomically and meronymically related lexemes has been formed for the Uzbek language corpus;

suggested the process of tagging taxonomic and meronymically related lexemes;

The interface of the search box of the Uzbek language corpus and the process of searching and sorting language units are linguistically modeled.

The reliability of the results of the study is explained by the fact that the studied materials helped to draw conclusions based on the nature of the Uzbek language, their validity, methodological excellence, reliance on proven sources in the development of the principles of the Uzbek National Corps.

Scientific and practical significance of research results.

The scientific significance of the research results Theoretical basis for the creation of the National Corpus of the Uzbek language, its ability to be a scientific and theoretical source for new research in the field of linguistic support, including theoretical knowledge on the means of automatic transmission of language lexemes to the interface of the national corpus data.

Practical significance of the research results: the role of the program in the process of teaching linguistics, computer linguistics, corpus linguistics, programming, planning, description of topics, general principles of creating a database of taxonomic and meronymic related lexemes, its information retrieval system, language corporations, machine translation, and the development of computer lexicography.

Implementation of research results.

Based on the scientific results achieved in the dissertation work dedicated to the formation of the linguistic base of taxonomic and meronymic related lexemes for the corpora of the Uzbek language:

The process of creating a linguistic database for the software of language corpora and the results of determining the amount of taxonomic and meronymic related lexemes at the language lexical level was used in the fundamental scientific project on the topic «Formation of functional words in the modern Karakalpak language» conducted by the Karakalpak Research Institute of Humanities, Karakalpak

Department of the Academy of Sciences of the Republic of Uzbekistan –FA-F1-G003 (Certificate No. 17.01/72 dated March 28, 2022 of the Karakalpakstan branch of the Academy of Sciences of the Republic of Uzbekistan). As a result, the principles of distinguishing the thematic groups of the lexical layer of the current Karakalpak language were developed;

from the conclusions about the identification and research of the statistics of taxonomy and meronymy phenomena in the explanatory dictionaries of the Uzbek language was used in the fundamental scientific project FA-A1-G007 on the topic «Karakalpak proverbs and sayings as an object of linguistic study» carried out by the Karakalpak Department of the Karakalpak Institute of the Academy of Sciences of the Republic of Uzbekistan» (Certificate No. 17.01/72 dated March 28, 2022 of the Karakalpakstan branch of the Academy of Sciences of the Republic of Uzbekistan). As a result, it was possible to determine the statistical amount of taxonomy and meronymy phenomena in the current Karakalpak language;

From the basis of the formation of the linguistic base of taxonomic and meronymic related lexemes for the corpora of the Uzbek language, as well as the results of the interface creation was used in the fundamental scientific project FA-F-1-005 on the topic «Researching the history of Karakalpak folklore and literary studies» conducted by the Karakalpak Humanities Research Institute of the Karakalpak Department of the Academy of Sciences of the Republic of Uzbekistan (Certificate No. 17.01/72 of March 28, 2022 of the Karakalpak branch of the Academy of Sciences of the Republic of Uzbekistan). As a result, the development of algorithms for linguistic corpora created in Uzbek and Karakalpak languages was achieved;

Conclusions about the interface of the search window for lexical-semantic related units in the Uzbek language, and the principles of sorting based on thematic and ranking were used in the television programs «Til-millat ko`zgusi», «Topic of the day», «Dialogue» of the Surkhondarya region television and radio company (Certificate number 01-12//224 dated March 24, 2022 of the Surkhondarya region television and radio company). As a result, the content of propaganda work on raising a high-level competent generation, forming language culture among young people and using the opportunities of the Uzbek language in the modern information system has been perfected.

Approbation of research results. The results of this study were discussed at 3 international and 3 national conferences and are available in the following profiles:

<https://orcid.org/0000-0003-4821-5919>, <https://scholar.google.com/citations>

Publication of research results.

Published 17 scientific papers on the topic of the dissertation, including one monograph, 10 scientific articles in scientific publications recommended for publication of the main scientific results of doctoral dissertations of the Higher Attestation Commission under the Cabinet of Ministers of the Republic of Uzbekistan, including 2 in foreign journals.

The scope and structure of the dissertation. The dissertation consists of an introduction, four chapters, a conclusion, a list of references and an appendix. The size is 264 pages.

MAIN CONTENT OF THE DISSERTATION

The introduction is based on the relevance and necessity of the topic, the relevance of research to the priorities of science and technology of the republic, the goals and objectives, the object and subject, scientific novelty and practical results, scientific and practical significance of the results. information on disclosure, implementation, approbation, published works and the structure of the dissertation.

The first chapter of the dissertation is entitled “Methodological basis for the study of taxonomic and meronomic phenomena in linguistics.” The first section of the chapter, “Problems of the Application of Taxonomy and Meronymy to Linguistics”, classifies the differences between the concepts of taxonomy and classification. Taxonomy [yun. taxis - the theory of classification and regulation of complex, usually hierarchically structured areas of all beings (living nature, the structure of the Earth; objects of geography, geology, linguistics, ethnography, etc.). The concept of taxonomy first appeared in science in the field⁹⁵ of biology (the term was proposed in 1813 by the Swiss botanist O. Decandol, who developed the classification of plants.

In the second half of the twentieth century, taxonomic problems began to play an important role not only in biology, but also in a number of other disciplines dealing with hierarchically organized sets of discrete objects. As a result, the term began to be used in other fields of knowledge: linguistics, geography, and geology to describe the general theory of classification and systematization of complex systems.

However, in addition to the natural differences in specific categories of taxonomy, the basic concepts of taxonomy and its functions are interpreted differently in different areas of knowledge. For example, in linguistics, taxonomy relies on the study of linguistic units in a text by distinguishing them and analyzing their order and distribution; accordingly, linguistic theory deals with the categories of element classes and the type of relationships between elements and classes. In linguistics, taxonomy is sometimes interpreted in terms of combining similar grammatical categories in different languages into one structured category. Thus, the general principles of taxonomy as a theoretical science are at their initial stage, with the growing importance of semiotics, linguistics and automatic translation, and the importance of linguistic research to classify (systematize) linguistic facts.

The English botanist Charles Jeffrey describes the phenomenon of taxonomy as follows: “In science, the taxonomic approach is opposed to the generative method. Mathematically, a taxonomy is a tree-like structure of a set of known objects. At the top of this structure there is a single unifying classification - the root taxon, which applies to all subjects in a particular taxonomy. Rooted taxa are specific classifications that belong to subgroups of the general set of classified objects. Modern biological classification is a hierarchical system based on individual organisms (individuals), with a single taxon at the top; there are taxa at different levels of the hierarchy between the base and the upper level, each of which is subordinate to one and only one higher level taxon”⁹⁶.

⁹⁵ Ўзбек тилининг изоҳли луғати. – Тошкент: Ўзбекистон миллий энциклопедияси, III том. – Б. 646.

⁹⁶ Джеффри Ч.. Биологическая номенклатура Charles Jeffrey. Biological Nomenclature. Second edition / Пер. с англ. — М.: Мир, 1980. – С. 14. – 120 с.

One of the main problems of taxonomy is to determine the method of this taxon. Rating,⁹⁷ a taxonomic category, is a concept of biological taxonomy that describes a level in a hierarchically organized system of living organisms (e.g., class, order, family). Despite a number of problems with the use of status (Russian: color), they continue to apply, because the rules for forming the names of groups (taxa) of living organisms, defined in the nomenclature codes (set of rules for scientific naming of organisms) are exactly the degree (color). based on the principle.⁹⁸

Kelurnoma was one of the first taxonomic analyzes in the history of our linguistics. In the last chapter of the work, the names are divided into 9 groups, divided into taxonomic rows. They are:

- 1) heaven and its concepts;
- 2) land and related concepts;
- 3) horses and saddled animals;
- 4) eagles and wild birds;
- 5) organs of the human body;
- 6) names of relatives;
- 7) names of weapons;
- 8) number names;
- 9) diamonds, rhymes, conjunctions and suffixes.⁹⁹

Language is a system of elements that make up a whole. The meaning of each element that makes it up stems from the fact that there are other elements that make up the system at the same time¹⁰⁰. “No one still understands its essence, no one will be able to use even a billionth of their potential in their lifetime,” an infinite number of languages, semaphores, units of form, and types of relations, all of which combine to form a variety that cannot be seen or heard.

In linguistics, separate interdisciplinary research based on the field of semantic field theory begins mainly in the 70s and 80s of the twentieth century. In this regard, the lexicology of the Uzbek language has risen from a comparative-descriptive stage to a new theoretical stage. During this stage, it was important to characterize the study of lexicon as a whole system-system consisting of certain semantic groups, the relationships of certain elements of meaning. According to Professor A. Sobirov, the introduction of field theory into Uzbek linguistics is associated with the introduction of system-structural methods in relation to language units. A.Nurmonov, H.Nematov, R.Rasulov, I.Kuchkartoev, E.Begmatov, T.Mirzakulov, Sh.Iskandarova, O.Bozorov, R.Safarova, B The scientific observations and conclusions of Kilichev, M.Narziyeva, H.Hojiyeva are important in the study of the lexicon of the Uzbek language on the basis of field theory. In particular, Sh.Iskandarova raised the problem of studying the lexicon of the Uzbek language as a semantic field on the example of a personal microfield. Iskandarova gives a detailed account of the division of the universe into groups and its reflection in the language, and provides information about about 100 lexical-semantic groups operating in the micro-field of the individual. For example,

⁹⁷ Зуев В. В., Розова С. С. [Проблема способа бытия таксона в биологической таксономии](#) // [Философия науки](#) : журнал. — Издательство Сибирского отделения РАН, 2001. — № 2 (10). — ISSN 1560-7488.

⁹⁸ Ibid.

⁹⁹ Искандарова Ш. Тил системасига майдон асосида ёндашув. — Ташкент: Фан, 1982. — В. 135.

¹⁰⁰ Соссюр Ф. Де. Труды по языкознанию. — М.: Прогресс, 1977. — С.147.

lexemes such as literate, wise, intelligent, prudent, perceptive, backward are included in the group of nouns that differ on the basis of level, their common and different aspects are explained.¹⁰¹

Thus, the scientific study of the relationship between the terms taxonomy, classification, system is undoubtedly one of the problems in Uzbek linguistics.

The second part of the first chapter is called «The Lexical-Semantic Field (LSM) and the Lexical-Semantic Group (LSG) Interpretation of Terms in Science», in which all things in Being are interconnected as a system. In the 1940s, the Swiss linguist Charles Balli coined the term “associative field”. As he writes, the concept of “associative field” is a flexible (elastic) concept, as evidenced by the fact that the size of these associations is not the same in different people: it should be borne in mind that the field can include near and far associations.

Ferdinand de Saussure, who introduced the direction of structuralism to language in world linguistics, drew the attention of linguists to the discovery of the relationship between language units and identified the existence of paradigmatic and syntagmatic types of relations. In the Linguistic Encyclopedic Dictionary of Russian Linguistics, “The unification of linguistic units into certain paradigms based on a certain unifying meaning later gave rise to field theory in linguistics. In linguistics, a field is defined as a set of linguistic units (mainly lexical units) that are combined with a common «meaning» and reflect the conceptual, subjective, or functional similarity of the events being defined¹⁰².”

If we look at the history of linguistics, we can see that the unity of linguistic units on the basis of a certain meaning, the combination of lexical units of a particular language into such semantic cells, developed in ancient Eastern linguistics. Later, with the advent of structuralism in linguistics, the idea of grouping linguistic units into semantic groups flourished in Western linguistics in the mid-nineteenth century. Accordingly, this theory is associated with the name of Western linguistics. In the 19th century, the study of linguistic units on the basis of generalization was focused on by the famous linguist M. Pokrovsky.

Linguist GS Schur commented on field theory: «In linguistics, the development of field theory in recent years has reached such a level that its ideas and methods have begun¹⁰³ to be applied to levels other than the lexical level of language». Research in this area has shown that, given the scope of the content field, it serves as a theoretical basis for the discovery of general laws for the creation of various thematic dictionaries (more thesauruses). Later, we can say that serious research has been done in this area at all levels of language, including morphosemantics and grammar. Based on the description of the concept of field and its specific features, in some places the approach raises various problems in the study of field theory.

In this case, the problems based on the accuracy of the boundaries of the field, the separation of the field as a linguistic unit in terms of individuality, integrity,

¹⁰¹ Искандарова Ш. Ўзбек тили лексикасини мазмуний майдон сифатида ўрганиш (шахс микромайдони): Филол. фанлари док. . . дисс. автореферат. – Тошкент, 1999.

¹⁰² Лингвистический энциклопедический словарь. – М., СЭ, 1990. – С. 380.

¹⁰³ Щур Г.С. О морфосемантических полях, функционально-инвариантных группах и категориях // Вопросы лингвистики. – Томск., 1970. – С. 3-10; Филичева Н.И. Синтаксические поля. – М., – ВЯ, 1977.

infinity, or the relative independence of the field theory, differing from linguistic units in the broad nature of field interactions; field structure problems; problems of relations such as field and polysemy are raised based on the fact that polysemous words belong to only one field or can be elements of several fields. Based on the multifaceted nature of field theory, these problems are being addressed by researchers in various aspects. In Uzbek linguistics, there are a number of views on the interpretation of the field.¹⁰⁴

Thus, in Uzbek linguistics, the study of language units on a field basis is widely practiced. Well-known linguist professor A.Nurmonov commented on this issue as follows. «It simply came to our notice then. Considering the Uzbek lexicon as a macro-field, dividing it into micro-fields is very effective in creating thesauruses and ideographic dictionaries. That is why an in-depth study of the theoretical basis for the division of the linguistic field and the structure of the language dictionary into such fields is one of the most pressing problems facing modern Uzbek linguistics¹⁰⁵». In our linguistics today, the study of language on the basis of field theory has a great influence on the development of applied linguistics.

The third part of the first chapter is called «Types of lexical-semantic relations in the Uzbek language and their study». The textbook “Modern Uzbek Literary Language” co-authored by Professor R.Sayfullayeva comments on lexical-semantic relations as follows: Synonymous, antonymic, graduonymic, partonymic, hyponymic relations are such types of linguistic relations”. One describes it as.¹⁰⁶

According to A. Sobirov, the types of relationships between members of the paradigm in the semantic field can be divided into the following groups: Synonymous relationship; Hyper-hyponymic (body-type) relationship; Holo-meronymic (whole-part) relationship; Graduonymic (lexical hierarchy) approach; Functional (functional) approach and Contradictory attitude¹⁰⁷.

Types of lexical-semantic relations are studied on two different bases: lexical-formal and lexical-semantic. A meronym is a concept that is an integral part of another. Meronim is also called partonim. The term meronymy (partonymy) is a new term in both Uzbek and European linguistics, and appears for the first time in the scientific literature in the Theses of Systemic Lexicology¹⁰⁸.

In Uzbek linguistics, the term meronymy was introduced under the name of partonymy in the 1990s of the XX century by representatives of the Bukhara Linguistics School.¹⁰⁹ In linguistics, the term meronymy was first introduced to science by the Russian scholar M.V.Nikitin to describe the meaning of a part.¹¹⁰

¹⁰⁴ Мирзакулов Т. Ўзбек тили морфем парадигматикаси ва синтагматикаси масалалари: Филол. фанл. докт....дисс. автореф. – Т., 1994; Абдувалиев М. Тўсиқсизлик майдони ва уни ташкил этувчи синтактик бирликлар // Ўзбек тили ва адабиёти. -1988. - № 4. – Б. 62-66.

¹⁰⁵ Нурмонов А. Танланган асарлар III жилд. Тошкент. 2012. – Б. 55.

¹⁰⁶ Sayfullayeva R. va b. Hozirgi o`zbek adabiy tili. O`quv qo`llanma. – Т.,:Fan va texnologiya, 2009. – В.112.

¹⁰⁷ Собиров А. Ўзбек тилининг лексик сатҳини системалар системаси тамойили асосида тадқиқ этиш. – Тошкент: Маънавият, 2004. –140 б.

¹⁰⁸ Бегматов Э., Нематов Х., Расулов Р. Лексик микросистема ва унинг тадқиқ методикаси. //Ўзбек тили ва адабиёти. 1989. 6-сон. – Б. 35-40.

¹⁰⁹ Неъматов Х., Расулов Р. Ўзбек тили систем лексикологияси асослари. – Т.:Ўқитувчи. 1995. -127 б.; Қиличев Б.Э. Партонимик қатор ва кўп маънолилик. Ўзбек тили доимий анжумани учинчи йиғинининг тезислари. –Т.: Ўқитувчи, 1995. – 138 б.

¹¹⁰ Собиров А. Ўзбек тилининг лексик сатҳини системалар системаси тамойили асосида тадқиқ этиш. – Тошкент: Маънавият, 2004. – 6-142.

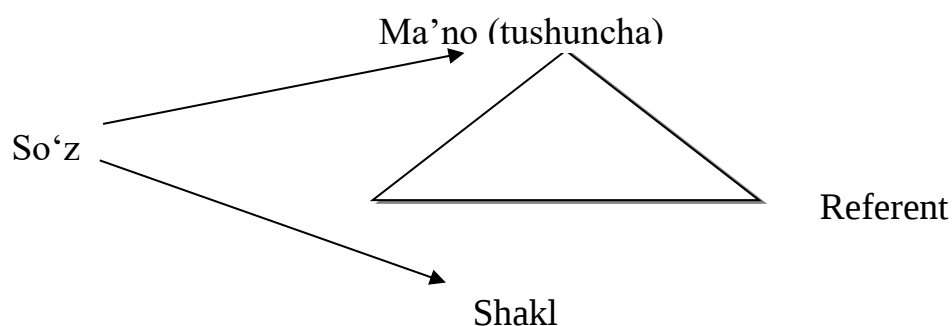
If in Uzbek linguistics the term *meronymy* firstly was introduced by linguists B.Qilichev¹¹¹ and A.Sobirov¹¹², H.Nematov and R.Rasulov¹¹³ also commented on this phenomenon. Professor A.Sobirov explains the phenomenon of meronymy as follows: The whole «house» consists of the *foundation, walls, floor, roof, doors, windows* parts. If we separate the «roof» from them and create a square based on the words related to it, the «roof» will be the whole name for the *beam, slate, ceiling, gutter (chute), garret*¹¹⁴.

In Uzbek linguistics, the phenomenon of meronymy is studied under the name of partonymy.

The second chapter is entitled “**Linguistic basis of the study of lexicon in Uzbek linguistics on the basis of semantic field**». The first paragraph of the chapter, named “The role of semantic field in tagging corpus materials» discusses the lexical system. Vocabulary is internally divisible. The Uzbek lexicon consists of hundreds of semantic fields. The field consists of a lexical-semantic group, and the group, in its turn, consists of the elements which it contains. This state of the lexical layer is its discreteness. For example, the semantic field of *time* includes the micro-fields of *era, century, year, and season* according to its structure. The season is divided into winter, spring, summer and autumn superseconds. *December, January, February, snow, cold, ice, sumalak, chill, rage, stove, wood, coal, sled, slip, snowstorm, new year, spruce, christmas, snowman, glove-like emas* (the smallest elements in the semantic field) gather in a *winter* microfield under a common single line and form an associative bond with each other¹¹⁵. At the same time, the elements of the lexical system are hierarchically related to each other. This is called a hierarchical relationship. Each century has 100 anniversaries, each year has four seasons, twelve months, and each month has 30-31 days. The day is divided into hours and hours. If an element in the hierarchy seems to fall, the balance is disturbed and the system breaks down.

According to professor E.Begmatov a thematic group (or series) as a general theme, a set of words united around an idea. Commenting on the theoretical ideas, the scientist divides the Uzbek lexemes into 28 thematic groups.

Table 1.



¹¹¹ Қиличев Б. Ўзбек тилида партонимия. Филол.фан.номз...дисс.автореферати. – Тошкент, 1997.

¹¹² Собиров А. Ўзбек тилининг лексик сатҳини системалар системаси тамойили асосида тадқиқ этиш. – Тошкент: Маънавият, 2004. –154 б.

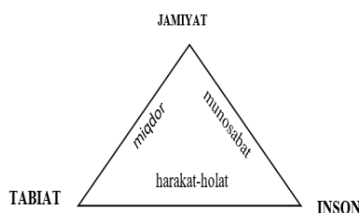
¹¹³ Ҳ.Нематов, Р.Расулов. Ўзбек тили систем лексикологияси асослари. – Тошкент: Ўқитувчи, 1195. –128 б.

¹¹⁴ Собиров А. Ўзбек тилининг лексик сатҳини системалар системаси тамойили асосида тадқиқ этиш. – Тошкент: Маънавият, 2004. –154 б.

¹¹⁵ Собиров А. Ўзбек тилининг лексик сатҳини системалар системаси тамойили асосида тадқиқ этиш. – Тошкент: Маънавият, 2004. –140 б.

A word is a symbol consisting of content and form. The relationship between content, form and referent is expressed as follows: Based on the above points, we have analyzed the nature-man-society relations based on one field from a philosophical point of view and analyzed them based on a triangle.

Table 2.



The second section of the second chapter, entitled «Assignment of taxonomic and meronymic units in the explanatory dictionaries of the Uzbek language» provides an overview of the statistical study of taxonomic and meronymic lexemes in the Uzbek language in the explanatory dictionaries of the Uzbek language.

The taxonomic thematic group of the **person** is given in the «Explanatory dictionary of the Uzbek language» as follows:

PERSON [a. شخص kimsa, kishi, zot] A special person in society, a man. Stranger. Worship of a historical figure¹¹⁶

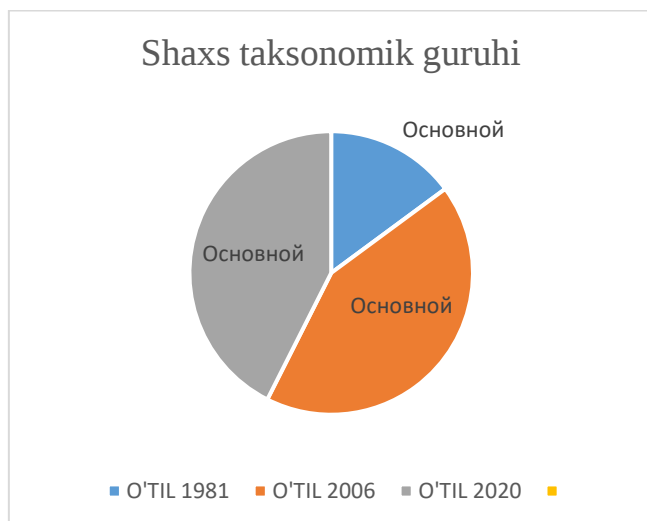
In the explanatory dictionaries of the Uzbek language (1981, 2006, 2020) the names of the persons belonging to the taxonomic thematic group «person» have the following index:

1. In the «Annotated Dictionary of the Uzbek language» published in 1981, 458;
2. The Uzbek Explanatory Dictionaries, published in 2006-2020, found 1,309 personal names (see Table 2.1).

As we saw in the 2006-2020 edition of the Uzbek Explanatory Dictionary, 34.9% of the names of individuals belonging to the taxonomic group of individuals are contained in the 1981 edition. This means that 34.9% of the taxonomic group of individuals mentioned in the next edition of our annotated dictionary is derived from the 1981 edition. The reason for the sharp increase in the number of subsequent editions is that over the past 30 years, the introduction of system-structural linguistic principles into Uzbek linguistics has led to significant quantitative changes in the Uzbek lexicon.

Table 3-4.

¹¹⁶ Ўзбек тилининг изоҳли луғати. Тўртинчи жилд. – Тошкент: «Ўзбекистон миллий энциклопедияси» Давлат илмий нашриёти, 2006. – Б.557.



Based on the 2006 edition of the Annotated Dictionary of the Uzbek Language, the 2020 editions with corrections were identical, identifying units belonging to the taxonomic group of 1,309 individuals. **Ethnonyms** in the «Explanatory Dictionary of the Uzbek language» is explained as follows: **Ethnonym** [yun. ethnos - tribe, people + onyma - name] People, tribe, clan, etc. the name of various ethnic associations¹¹⁷.

The statistical number of taxonomic thematic groups in the Uzbek dictionaries is as follows:

T/R	The name of the spiritual group	1981 edition	2006-2008 edition	2020 edition
1	Person	1927	2760	2760
2	Ethnonym	99	167	167
3	Kinship names	49	87	87
4	Legendary creatures	17	25	25
5	Animals	295	505	505
6	Plants	365	543	543
7	Body and materials	84	135	135
8	Space and space	32	67	67
9	Construction and structures	255	324	324
10	Weapons and equipment	72	123	123
11	Equipment	71	116	116
12	Tools and mechanisms	97	225	225
13	Means of transport	38	84	84
14	Weapons	37	62	62
15	Musical instruments	40	56	56
16	Furniture	19	30	30
17	Container	67	147	147

¹¹⁷ Ўзбек тилининг изоҳли луғати. Тўртинчи жилд. – Тошкент: «Ўзбекистон миллий энциклопедияси» Давлат илмий нашриёти, 2006. – Б.557.

18	Clothes and clothing	102	156	156
19	Food and beverages	128	183	183
20	Texts	16	54	54
Total		3810	5849	5849

Chapter III of the study is entitled “**Theoretical problems marking the taxonomy and meronymics in the Uzbek National Corps**». The first section is entitled «The evolutionary development of corpus linguistics in world and Uzbek linguistics».

As a complex linguistic entity, the corpus not only stores a large amount of information about the content of the speech/linguistic source, but also includes formal methods of presenting this information (word indexing, morphological information, etc.). Correspondingly, the corpus can be interpreted as a specially constructed semiotic system¹¹⁸.

In the mid-1980s, various language/size corpus projects were created. During this period, the results achieved in corpus linguistics began to be widely applied to the language teaching process, and the projects created on corpus linguistics in the world's leading higher education institutions became a topic of practice and day. Through this, the historical, geographical, social variation of the corpus approach and its optimal nature showing its changing forms in the linguistic system were revealed. This, in turn, made it possible to master the basic principles of corpus-based methods of linguistic analysis¹¹⁹. The development of corpus linguistics, the creation of corpora, is one of the urgent issues for nations concerned about their future. In the first half of the XXth century, it was possible to create the national corps only manually. This required a long time and a lot of money.

T/r	Name of the meronomic semantic group	Edition in 1981	2006-2008 – years edition	2020-year edition
1	Parts of human body	136	181	181
2	Parts of clothes	34	48	48
3	Parts of mechanism and instruments	123	150	150
4	Parts of plants	29	24	24
5	Parts of object complex	498	511	511
6	Parts of dishes	88	149	149
7	Parts of constructions	30	46	46
8	Parts of transport means	27	35	35
Jami		965	1144	1144

¹¹⁸ Рыков В.В. Корпус текстов как новый тип словесного единства // Труды Междунар. семинара «Диалог-2003». – М.: Наука, 2003. – С. 21.

¹¹⁹ Information and Communications Technology for Language Teachers. Введение в прикладное значение корпуса. Режим доступа: http://www.ict4lt.org/en/en_mod2-4.htm

In higher education institutions around the world, linguistic corpora are used to prepare test questions and lecture texts for students. Many students independently refer to the corpus when preparing independent works and projects. In this process, by comparing the students who use the corpus and those who do not use it, it is possible to conclude that: a student who uses the corpus effectively learns the laws of the language, features of the foreign language he is learning more easily and quickly than the student of the second category. The differences between traditional and corpus linguistics can be seen in the following places (Table 1.4).

I.4 Table. Different aspects of traditional and corpus linguistics

Features	In traditional linguistics	In Corpus Linguistics
Orientation	Language / language research	Speech research
Object of research	Explain speech phenomena through theory	Text corpus
Methods	Qualitative methods	Quantitative methods
Attitude to the text	Abstraction	Accuracy
Focus on analysis	Focus on content	The focus is on shape
Methods of analysis	Limitations	Infinity
Types of analysis	Analyzes speech material based on intuition	Represents speech activity in the form of text
Methods of conveying information	Prefers logical thinking	Probability theory and statistics are used for the primary processing of speech material
The effectiveness of the analysis	The words in the text are analyzed separately	Works with linguistic information (word usage) in the text
Approach method	Deductive	Inductive
In relation to time	Diachron	Synchronous

The essence of the corpus and the ability to present the available information from its formation in text form; the possibility of providing as much information as possible according to the size of the corpus; purposes such as the ability to repeatedly use the once-created corpus data to solve various problems. For this, it is necessary to determine the principles that form the basis of the corpus in front of the corpus, to determine what the standard corpus format should be according to different linguistic dimensions (semantic, genre, stylistic, semantic, morphological format is envisaged). it will be necessary to solve such tasks.

The chapter entitled “*Issues of tagging taxonomic and meronymic lexemes in the corpus*” discusses the importance of taxonomic and meronymic lexemes in the corpus. Computer linguistics is a field that literally describes the operation of linguistic

machines and the set of laws that govern them. The strength of computer linguistics lies in its ability to manage linguistic cues that are «saturated» with a particular set of meanings.¹²⁰ However, there is a direct «interaction» between computer and language; it also appears in the concept of computer-program-knowledge-language¹²¹.

The basic concept of corpus linguistics is the linguistic corpus. Corpus texts are selected from the point of view of relevance to the problem, that is, from the thematic relevance that is important for linguists. The problem area is the thematic scale, which consists of two aspects, including texts on language and speech. While the language aspect is the most researched area, the speech aspect consists of contexts that reflect the language aspects. The problem of sorting texts by size is one of the most important issues in creating a corpus. This aspect ensures the perfection of the case. However, this aspect of the corpus contradicts its representativeness. Because the size of the text complicates the operation on the corpus, it means that the study of data on the corpus, depending on its size, can not provide a complete implementation. Thus, the goal of the corpus requires a narrow approach to its solution - to define the thematic area according to the scope of research. Such a choice should cover a particular criterion of the language under study in accordance with the events in the speech.

Corporations are divided into the following types for sorting texts:

- balanced housings;
- Monitor housings.

In order to carry out the process of tagging taxonomically and meronymically related lexemes in the corpus, it is first necessary to determine how many words are in these lexical groups in the Uzbek lexicon. To do this, it is necessary to determine the number of taxonomic and meronymically related lexemes in the «Explanatory dictionaries of the Uzbek language» and determined the statistics. See Table 1.

As we have seen, in the 2006 and 2020 editions of the Annotated Dictionary of the Uzbek Language, 72 thematic ethnonyms were identified. We tag the identified units in a more complete column from the table above using JavaScript code.

t:hum:etn: (*Abkhaz, Avar, Adyghe, Ajam, Aymak, Albanian, Aleutian, Arab, Armenian, Afghan, Balkar, Bantu, Barlos, Bakhtiari, Belarusian, Baluch, Bengal, Bijanak, Bilkut, Bobi, Bulgarian, Bulgarian, Bashkir, British, Bulogochun , bulgur, bourbon, buryat, velvet, vandal, barbarian, Hungarian, gabr, gagauz, german, holland, Greek, georgian, georgian , jo'gi, jojiylar, zangi, zangi, indeys, indonesian, ispan, italian, kabardin, kaltatoy, komi, koreys, kreol, kurd, lazgin, latin, latish, laqay, mavr, major, malik, malika*)

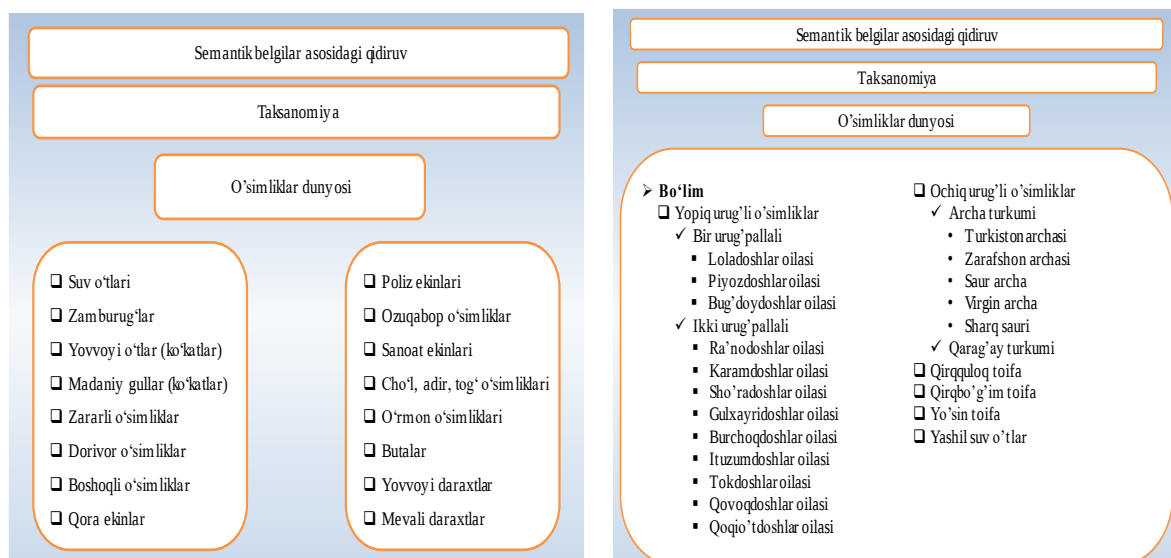
The body finds the text in which these words are present when you search by placing the required characters in the search box. This, of course, requires you to select the desired tag and select the selected tag. After that, we need to search for a selected or selected text, sentence or word through the word search window of the corpus interface. As a result, we will be able to get all the information about all the words of the thematic group of all ethnonyms in the system memory and their usage statistics and the periodicity of their presentation in the texts.

¹²⁰ Шемякин Ю.И. Начала компьютерной лингвистики: Учеб. пособие. – М.: Изд-во МГОУ, А/О "Росвузнаука", 1992.

¹²¹ Ibid .

Chapter IV of the dissertation is entitled «**Fundamentals of the formation of the linguistic base of taxonomic and meronymically related lexemes for the corpus of the Uzbek language**». The first paragraph of the chapter «Requirements for the content of the interface of the linguistic base of taxonomically related lexemes for Uzbek language corpus» discusses the problem of placing taxonomic units in the interface of the linguistic corpus. It is known from the gradual development of World Corpus Linguistics that almost all of the National Corpus that have been created so far will have the interface of the taxonomically related lexemes linguistic base in the search menu. They are submitted to the search engine in a certain sequence. The analysis of the interface of the linguistic base of lexemes with taxonomic and meronymic relation in the National Corpus of the Russian language showed that units representing nested (taxonomic) and whole-piece (meronymic) relations are arranged in a sequence based on a certain classification. Also, the researcher U. Kholiyorov in his dissertation «Linguistic basis of creating the educational corpus of the Uzbek language» says about the location of the main menu interface of the National corpus of the Uzbek language: It is necessary to pay special attention to features of the Uzbek language, aspects suitable for the age of the student¹²².

A. Sobirov comments on the analysis of plants into taxonomic units. In modern systematics, plants are divided into the following units (taxa): section, class, order, family, genus, species¹²³. Based on A. Sobirov's classification, the taxonomic search interface can be designed as follows:



In the section of the chapter called «*Development of the interface of the linguistic base of meronymic lexemes for the corpora of the Uzbek language*» the problem of placing taxonomic units in the interface in the linguistic corpus is discussed. The linguistic base of meronymic related lexemes for Uzbek language corpora is based on the researches carried out in this field in Uzbek linguistics. In

¹²² Холиёров Ў. Ўзбек тили таълимий корпусини тузишнинг лингвистик асослари: Филол. фан. бўйича фалсафа доктори (PhD) дисс. – Термиз, 2021. – Б 38.

¹²³ Собиров А. Ўзбек тилининг лексик сатҳини системалар системаси тамойили асосида тадқиқ этиш. – Тошкент: Маънавият, 2004. – Б 93.

particular, the linguist scientist B. Qilichev analyzes the meronymic phenomenon into the following semantic fields: the partonym «Body», the partonym «Head», the participles «Leg» and «hand», the full noun «Head», the term «Tumor» partonyms, partonyms «Eye», «ear», «mouth», «nose», partonyms with the symbol «Plain», partonym «Animal» partonym «head» in Lexical Semantic Group (LSG)¹²⁴. This classification can be expressed in a unique way in the interface of the national corpus of the Uzbek language. If we analyze only one body meronym from the range of meronymic relations, the linguist scientist B. Qilichev classifies it as follows. Body - even as a noun, «body» has its own characteristics. For example, the parts of the «human» body are as follows: Body I: shoulder, chest, abdomen, waist.

The full name «body» in the partonymic LSG «Animal» was created by: Body II: chest, abdomen, back, waist.

The full noun «body» in the partonymic LSG «tree» consists of the following participles: Body III: such as bark, veil, wood, core.

It can be seen that «body I» and «body II» are similar to each other according to the classification of fragments, and sharply different from «body III». In conclusion, we can state the following:

First of all, «body I» and «body II» are close to each other in terms of forming lexical lines, while «body III» does not have such a feature.

Secondly, the parts of «body I» and «body II» wholes have the same characteristics. In this case, «body III» is separated from «body I» and «body II» due to the nature of its parts. Only in terms of their relationship to the chain of meaning, these three nouns can be in one group. For example, the meaning of all three whole names is the same: «A tool that unites the remaining parts of known wholes». In the national corpus of the Uzbek language, it can be seen that one meronym «body» is expressed in three different classifications. In the partonymic lexical-semantic group of man, cattle and tree, it is necessary to express them in connection with each other in all three cases.

In the national corpus of the Uzbek language, it can be seen that the whole made up of certain parts and the parts (meronyms) that make up this whole. For example, the body holonym consists of arms, legs, and head meronyms:

1. Hand - nail, finger, palm, wrist, elbow, shoulder.
2. Foot - nail, palm, heel, ankle, marrow, knee, calf, thigh.
3. Head - hair, eyebrows, eyelids, eyelashes, eyes, ears, nose, mouth, forehead, nape, chin¹²⁵.

¹²⁴ Qilichev B. Leksikada so'zlararo butun-bo'lak munosabati. Monografiya. Toshkent. 2019. – 78.B.

¹²⁵ Собиров А. Ўзбек тилининг лексик сатҳини системалар системаси тамойили асосида тадқиқ этиш. – Тошкент: Маънавият, 2004. – Б. 43.

CONCLUSIONS

1. Taxonomy includes complex hierarchical structures of the whole being, living nature, the structure of the earth; geography, geology, linguistics, ethnography and the theory of classification and arrangement of objects, science about it. In linguistics, taxonomy is a tool for studying the properties of linguistic units in a text by analyzing their order and distribution.

2. The phenomenon of taxonomy is a method based on classification and systematic study of processes in the sciences of philosophy, biology, mathematics, pedagogy, linguistics.

3. It is no secret that the issue of the reflection of the whole-part relationship in the real world in the lexicon has attracted the attention of linguists since time immemorial. Calling the whole concept a holonym, a part a meronym or a partonym allows a deeper analysis of this phenomenon. The difference of holo-meronymy or partonymy from hyper-hyponym nest is that in hyponymy one part (member) of the field is contrasted with others and the meaning relationship between them is determined, while in partonymy the part (member) taken separately the relationship with the internal members that make it up is understood. The term partonym is used only for a fragment. Therefore, it is necessary to use the proposed terms holonym and meronym to express whole and part.

4. A critical review of the amount of taxonomic and meronymic related lexemes in various editions of explanatory dictionaries of the Uzbek language showed that the number of lexemes in taxonomic thematic groups has an increasing index, while meronymic thematic groups have some. It was found that it has a decreasing index.

5. The result of the above classification shows that lexemes with taxonomic and meronymic relations in the explanatory dictionary of the Uzbek language cannot serve as a full-fledged lexicographic base for lexical-semantic classification. This affects the accuracy of the sorting process. Nevertheless, the taxonomic and meronymic related units in them should be used as a basic source for interpretation and description.

6. To introduce the base of taxonomic and meronymic related lexemes into the corpus, to be able to present the existing information in the form of text; provide as much information as possible according to the size of the corpus; it allows to use the once created corpus data repeatedly in solving different problems. For this, it is necessary to develop the principles of creating a database of taxonomic and meronymic related lexemes, to eliminate problems in this regard. One of the main issues is to determine the units that make up the corpus, to develop a standard format for various linguistic parameters.

7. The process of revealing taxonomic and meronymic lexemes in linguistic corpora through the system interface is carried out on the basis of a certain algorithm. In order for information to appear in the system interface, first of all, a database of taxonomic and meronymic related lexemes must be created, and then, by means of tags specially attached to the system menu through JavaScript code, the necessary taxonomic unit is searched by typing in the search box. After that, the taxonomic unit we are looking for will be searched, and all the lexical units entered by us will be ready for obvious use in the system interface.

8. Corpus interface, expert systems, neural networks, linguanalyzers, speaking automata of natural languages within the framework of the artificial intelligence system are all the result of computer modeling. The national corpus interface is also a result of modeling in some sense.

**РАЗОВЫЙ НАУЧНЫЙ СОВЕТ НА ОСНОВЕ НАУЧНОГО СОВЕТА
PhD.03/31.03.2021.FIL.78.04 ПО ПРИСУЖДЕНИЮ УЧЕНЫХ СТЕПЕНЕЙ
ПРИ ТЕРМЕЗСКОМ ГОСУДАРСТВЕННОМ УНИВЕРСИТЕТЕ**

ТЕРМЕЗСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ЭШМУМИНОВ АСКАР АЛЛАМУРАДОВИЧ

**ФОРМИРОВАНИЕ ЯЗЫКОВОЙ БАЗЫ ТАКСОНОМИЧЕСКИХ И
МЕРОНИМИЧЕСКИХ ЛЕКСЕМ ДЛЯ КОРПУСОВ УЗБЕКСКОГО
ЯЗЫКА**

10.00.01 – Узбекский язык

АВТОРЕФЕРАТ ДИССЕРТАЦИИ ДОКТОРА ФИЛОЛОГИЧЕСКИХ НАУК (DSc)

ТЕРМЕЗ– 2022

Тема диссертации доктора философии по филологическим наукам (PhD) в Высшей аттестационной комиссии при Кабинете Министров Республики Узбекистан B2021.1.DSc/Fil зарегистрирована под номером 284.

Диссертация доктора филология (DSc) выполнена в Термезском государственном университете.

Автореферат диссертации выполнен на трех языках (узбекский, английский, русский) размещён на веб-странице по адресу www.terasu.uz и на информационно-образовательном портале «ZiyoNET (www.ziyo.net.uz)».

Научный руководитель:

Менглиев Бахтиёр Ражабович
доктор филологических наук, профессор

Официальные оппоненты:

Пардаев Азамат Бахронович
доктор филологических наук, профессор

Абжалова Манзура Абдурашитовна
доктор филологических наук, доцент
Эргашева Гули Исмаиловна
доктор филологических наук, доцент

Ведущая организация:

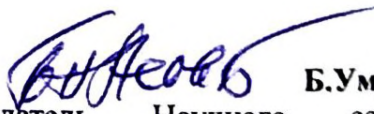
Ферганский Государственный Университет

Предоставление ученых степеней в рамках защиты диссертации состоится в 10⁰⁰ часов 22 08 2022 года на заседании научного совета PhD. 03/31. 03.2021.Fil.78.04 при Термезском государственном университете (Адрес: 190111, город Термиз., улица Баркамол авлод, 43. Тел.: (376) 221-71-17; e-mail: termizdu@umail.uz).

С диссертацией можно ознакомиться в информационно-ресурсном центре Термезского государственного университета (зарегистрирован под номером 67). Адрес: 190111, город Термез, улица Баркамол авлод, 43. Тел.: (376) 221-74-55.

Автореферат диссертации разослан « 11 » 08 2022 года (Реестр протокола рассылки № 12 от « 11 » 08 2022 года).



 Б.Умуркулов
Председатель Научного совета по
присуждению ученой степени доктор
филологических наук, профессор

 Н.Амонтурдиев
Учёный секретарь научного совета по
присуждению учёных степеней, PhD.

 Ш. Махмараймова
Председатель разового научного семинара
при Научном Совете по присуждению
учёных степеней, доктор филологических
наук, доцент

ВВЕДЕНИЕ (аннотация диссертации доктора филологических наук (DSc))

Актуальность и востребованность темы диссертации. В 50-х годах XX века в мировой корпусной лингвистике начался период интенсивной обработки естественного языка, активного применения задач перевода в компьютерных программах посредством машинного перевода. На сегодняшний день в крупнейших языках мира: английский, немецкий, французский и русский созданы базы таксономических и меронимически связанных лексем, которые удобно расположены непосредственно в системном интерфейсе лингвистического корпуса. Это подчеркивает практическую важность возможности для пользователей системы получать достаточную информацию из своего родного языка или из лингвистических корпусов, созданных на передовых языках.

В мировой лингвистике в начале XXI века на повестке дня стоят такие вопросы, как обеспечение лингвистического корпуса качественными лексическими ресурсами, повышение качества использования корпусов, лингвистическое моделирование языка. В результате стало популярным активное использование лексических единиц языка в информационно-поисковой системе для создания таксономической и меронимической базы данных в корпусе языков с лингвистическим корпусом. Следовательно, создание языковых корпусов, их лингвистическая аннотация, создание таксономических и меронимических баз данных стали научно-теоретическими проблемами корпусной лингвистики.

В области узбекского корпусного языкознания уделяется особое внимание научным исследованиям по созданию национального корпуса, учебного корпуса. Вопрос создания базы данных таксономически и меронимически родственных лексем для Узбекского национального корпуса не изучался в научно-теоретическом плане. Ведь «...сохранить чистоту государственного языка, обогатить его и повысить речевую культуру населения, обеспечение активной интеграции государственного языка в современные информационные технологии и коммуникации»¹²⁶ является одной из важнейших задач, стоящих сегодня перед узбекским корпусным языкознанием. Соответственно, такие вопросы, как создание базы данных таксономически и меронимически связанных лексем и создание их лингвистического обеспечения являются важными в узбекском языкознании.

Целью исследования является создание лингвистической базы таксономически и меронимически родственных лексем для корпуса узбекского языка.

Задачами исследования являются:

изучить общее описание явлений таксономии и меронимии в языкознании и их значение при создании языковых корпусов;

¹²⁶ Ўзбекистон Республикаси Президенти Шавкат Мирзиёевнинг 2020 йил 202 октябрдаги «Мамлакатимизда ўзбек тилини янада ривожлантириш ва тил сиёсатини такомиллаштириш чора-тадбирлари тўғрисида»ги ПФ-6084-сон фармони // <https://lex.uz/docs/5058351>

изучить теоретические основы анализа языковых единиц исходя из тематической области;

тегирование таксономических и меронимически связанных лексем, определение важности этих тегов в поисковой системе;

исследование внешнего вида, содержания и функций семантического поиска в интерфейсе лингвистического корпуса;

исследование общих и различных аспектов интерфейса таксономически и меронимически связанных лексем в языковом корпусе;

определение структуры и характеристик семантической базы данных корпуса узбекского языка;

лингвистическое моделирование процессов поиска и сортировки языковых единиц и окна поиска интерфейса семантического поля и единиц группы корпуса узбекского языка.

В качестве **объекта исследования** были выбраны таксономические и меронимические лексемы узбекского языка.

Предметом исследования является основа для формирования лингвистической базы таксономически и меронимически родственных лексем корпуса узбекского языка.

Методы исследования. Для освещения темы исследования использовались методы классификации, описания, сравнения, статистического анализа, моделирования.

E'LON QILINGAN ISHLAR RO'YXATI

LIST OF PUBLISHED WORKS

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

I бўлим (I часть; I part)

1. Эшмуминов А. Ўзбек тили миллий корпусининг синоним сўзлар базаси Monografiya. – Termiz davlat universiteti nashriyoti, 2021. – 116 б. (ISBN 978-9943-7244-5-7)
2. Eshmuminov A. Semantik razmetka: tushuncha, vosita va usullar.// Til va adabiyot ta'limi. – Toshkent, 2022. №1. – Б. 34-35. (10.00.00 № 9)
3. Eshmuminov A. Implementation of maydon's theory in linguistics // International Scientific Journal Theoretical & Applied Science. Volume:102, Issue:05 - Philadelphia, USA, 2022. – P. 468-470 (Impact factor –8,716).
4. Eshmuminov A. O'zbek tili korpuslari uchun taksonomik munosabatli leksemalar lingvistik bazasi interfeysi mazmuniga qo'yiladigan talablar.// Til va adabiyot ta'limi. – Toshkent, 2022. №4. – Б. 38-41. (10.00.00 № 9)
5. Eshmuminov A. Tilshunoslikda maydon nazariyasining tadqiqi// ЎЗМУ хабарлари. Илмий журнал. – Тошкент, 2022. № 1/4. – Б. 292-295. (10.00.00 № 15)
6. Eshmuminov A. Tilshunoslikda meronimiya hodisasi tadqiqi// Ta'lim va innovatsion tadqiqotlar. – Buxoro, 2022. № 7. – Б. 64-68. (10.00.00 № 36)
7. Eshmuminov A. The study of the meronymic phenomena in linguistics//Journal of Advanced Research and Stability. Volume: 02 Issue: 04 | 2022 265-266.
8. Эшмуминов А. Ўзбек миллий корпусининг лексик-семантик муносабатдош бирликлар базасини яратиш масалалари//Science and Education. Volume 3 Issue 04 | 2022 2025-2032 (Impact factor –3,567).
9. Eshmuminov A. O'zbek tili korpuslari uchun taksonomik munosabatli leksemalar lingvistik bazasi interfeysi mazmuniga qo'yiladigan talablar.// Til va adabiyot ta'limi. – Toshkent, 2021. №10. – Б. 28-30. (10.00.00 № 9)
10. Eshmuminov Asqar. Problems of creating national corpus of the Uzbek language (level of synonyms) // International Scientific Journal Theoretical & Applied Science. Volume:73, Issue:05 - Philadelphia, USA, 2019. – P. 47-50 (Impact factor – 8,716).
11. Eshmuminov A. Requirements for the content of the interface of the linguistic base of taxonomical lexemas for Uzbek language corporations//Science and Education. Volume 3 Issue 05 | 2022 222-225 (Impact factor –3,567).
12. Eshmuminov A. Semantic system: concepts, means and methods//Science and Education. Volume 3 Issue 05 | 2022 323-326 (Impact factor –3,567).
13. Eshmuminov A. Issues of creating a database of lexically and semantically related units of the Uzbek national corpus // Zamonaviy fan, ta'lim va tarbiyaning dolzarb muammolari (elektron ilmiy to'plam) Urganch. iyun, 2020-III B.1-8. (10.00.00 № 22).

14. Эшмуминов А. Лингвистик корпусда разметкаланишнинг аҳамияти// Filologik tadqiqotlar: Til, adabiyot, ta'lim, – Termiz. 2021. №3-4. – Б.16-19.

II бўлим (II часть; II part)

15. Эшмуминов А. Корпусда синоним сўзларни теглашнинг дастурий таъминоти масалалари // Uluslararası Mersin Simpozyumu, 22-24 Ekim, 2020 Mersin. Turkiya , 2020. B.210- 219.

16. Эшмуминов А. Корпусда разметка ва унинг хусусиятлари//“Ўзбек миллий ва таълимий корпусларини яратишнинг назарий ҳамда амалий масалалари” мавзусидаги халқаро илмий-амалий конференция материаллари. Тошкент, 2021. Б.98-100.

17. Eshmuminov A. Jahon va o‘zbek tilshunosligida korpus lingvistikasining tadrijiy taraqqiyoti// O‘zbek tilini keng targ‘ib qilish bo‘yicha hamkorlik istiqbollari” mavzusidagi xalqaro ilmiy-amaliy konferensiya materiallari. Toshkent, 2020. B.98–100.

18.Эшмуминов А. Ўзбек тилшунослигида тил корпусларининг тузилиши зарурат ва эҳтиёж ҳақида // «Давр, Замон, Шахс» Республика илмий-амалий анжуман материаллари. – Термиз, 2017. – Б. 713-714.

19. Эшмуминов А. Ўзбек тилининг изохли луғатларида синонимлар тавсифи ва талқини // “O‘zbek tilshunosligining dolzarb masalalari” onlayn ilmiy anjuman materiallari. Termiz, 2021. – B. 21-24

20. Eshmuminov A. Taksonomiya hodisasining tilshunoslikka kirib kelishi haqida// “O‘zbek tili imkoniyatlaridan foydalanish masalalari” Respublika ilmiy-amaliy konferensiyasi materiallari. – Qarshi, 2022. –B. 179-181.

Автореферат “Сурхондарё илм ва фан” журнали таҳририятида таҳрирдан
ўтказилди (11.08.2022).

Босишга рухсат этилди 11.08.2022 й
Офсет босма қоғози. Қоғоз бичими 60 1/16.
«Times New Roman» гарнитураси.
Офсет босма усули. Ҳажми 3,5 босма табоқ
Адади 100. Буюртма №39

Термиз давлат университети нашр-матбаа босмахонасида чоп этилди.
Манзил: Термиз шаҳри, “Баркамол авлод” кўчаси, 43-уй.

