

**ҚАРШИ ДАВЛАТ УНИВЕРСИТЕТИ ХУЗУРИДАГИ
ИЛМИЙ ДАРАЖА БЕРУВЧИ
PhD.30.05.2018.Fil.70.01 РАҚАМЛИ ИЛМИЙ КЕНГАШ**

ТЕРМИЗ ДАВЛАТ УНИВЕРСИТЕТИ

ЭШМУМИНОВ АСҚАР АЛЛАМУРАДОВИЧ

**ЎЗБЕК ТИЛИ МИЛЛИЙ КОРПУСИНИНГ СИНОНИМ СЎЗЛАР
БАЗАСИ**

10.00.01. – Ўзбек тили

**ФИЛОЛОГИЯ ФАНЛАРИ БЎЙИЧА ФАЛСАФА ДОКТОРИ (PhD)
ДИССЕРТАЦИЯСИ АВТОРЕФЕРАТИ**

Қарши – 2019

Фалсафа доктори (PhD) диссертацияси автореферати мундарижаси
Contents of Dissertation Abstract of the Doctor of Philosophy (PhD)
Оглавление автореферата диссертации доктора философии (PhD)

Эшмуминов Асқар Алламурадович Ўзбек тили миллий корпусининг
синоним сўзлар базаси..... 5

Eshmuminov Asqar Allamuradovich Synonymous word's basis of the
Uzbek language national corpus 23

Эшмуминов Асқар Алламурадович База синонимичных слов
национального корпуса узбекского языка 41

Эълон қилинган ишлар рўйхати
List of published works
Список опубликованных работ..... 45

**ҚАРШИ ДАВЛАТ УНИВЕРСИТЕТИ ХУЗУРИДАГИ
ИЛМИЙ ДАРАЖА БЕРУВЧИ
PhD.30.05.2018.Fil.70.01 РАҚАМЛИ ИЛМИЙ КЕНГАШ**

ТЕРМИЗ ДАВЛАТ УНИВЕРСИТЕТИ

ЭШМУМИНОВ АСҚАР АЛЛАМУРАДОВИЧ

**ЎЗБЕК ТИЛИ МИЛЛИЙ КОРПУСИНИНГ СИНОНИМ СЎЗЛАР
БАЗАСИ**

10.00.01. – Ўзбек тили

**ФИЛОЛОГИЯ ФАНЛАРИ БЎЙИЧА ФАЛСАФА ДОКТОРИ (PhD)
ДИССЕРТАЦИЯСИ АВТОРЕФЕРАТИ**

Қарши – 2019

Фалсафа доктори (PhD) диссертацияси мавзуси Ўзбекистон Республикаси Вазирлар Маҳкамаси ҳузуридаги Олий аттестация комиссиясида № B2019.1.PhD/Fil 700 рақам билан рўйхатга олинган.

Диссертация Термиз давлат университетида бажарилган.

Диссертация автореферати уч тилда (Ўзбек, инглиз, рус (резюме)) Илмий кенгаш веб-саҳифасида, qarshidu.uz ва «ZiyoNet» Ахборот таълим порталида (www.ziyounet.uz) жойлаштирилган.

Илмий раҳбар:

Менглиев Бахтиёр Ражабович
филология фанлари доктори, профессор

Расмий оппонентлар:

Пардаев Азамат Бахронович
филология фанлари доктори, доцент

Абжалова Манзура Абдурашкетовна
филология фанлари бўйича фалсафа доктори (PhD)

Етакчи ташкилот:

Бухоро давлат университети

Диссертация ҳимояси Қарши давлат университети ҳузуридаги Илмий даражалар берувчи PhD.30.05.2018.Fil.70.01 рақамли Илмий кенгашнинг 2019 йил «5» XII соат 10.00 даги мажлисида бўлиб ўтади (Манзил: 180103, Қарши шаҳри, Кўчабоғ кўчаси 17. Тел.: (0 375) 225-34-13; факс: (0375) 221-00-56; e-mail: qarshidu@umail.uz). Қарши давлат университети, Ўзбек филологияси факультети фаоллар зали.

Диссертация билан Қарши давлат университетининг Ахборот-ресурс марказида танишиш мумкин (19 рақами билан рўйхатга олинган). (Манзил: 180103, Қарши шаҳри, Кўчабоғ кўчаси 17. Тел.: (0 375) 225-34-13; факс: (0375) 221-00-56; e-mail: qarshidu@umail.uz).

Диссертация автореферати 2019 йил «22» XI да тарқатилди.
(2019 йил 22.XI даги 18 рақамли реестр баённомаси).


Н.Н.Шодмонов
Илмий даражалар берувчи илмий
кенгаш раиси, ф.ф.д.
Г.Н.Тожиева
Илмий даражалар берувчи илмий
кенгаш аъзиси, ф.ф.д. (PhD)
Д.Тўраев
Илмий даражалар берувчи илмий
кенгаш қошидаги Илмий семинар
раиси, ф.ф.д., профессор

КИРИШ (Фалсафа доктори (PhD) диссертацияси аннотацияси)

Тадқиқот мавзусининг долзарблиги ва зарурати.

Жаҳон тилшунослигида синонимларга илмий муносабат ўз тараққиёт босқичларига эга, асосан, XIX асрдан жадал такомиллашиб бормоқда. Фарқли ва кўпинча қарама-қарши талқинларга эга синонимларнинг электрон базасини яратиш ва уни миллий тил корпусларига киритиш масаласига эса жаҳондаги энг йирик ҳажмга эга бўлган тил корпусларидагина мурожаат қилинган. Корпус лингвистикасининг ривожда гуманитар ва техника фанларига оид кўплаб ижобий жиҳатлар ўзаро бирлашади. Ҳозирги кунда ушбу соҳада амалга оширилган назарий ишлар шаклланаётган миллий тил корпуси тузувчилари учун асос вазифасини бажаради.

Дунё тилшунослигида XXI аср ва унгача бўлган даврда миллий корпусларнинг лексикография асосидаги шакллари нисбатан мукаммал ишланган, хусусан, синонимлар луғати ҳам. Бугунги кунда, корпус лингвистикаси илмий-назарий жиҳатдан мукаммал босқичга кўтарилаётган бир даврда, корпуснинг семантик разметкасида мазкур луғатлар беназир аҳамият касб этмоқда. Автоматик таржима сифатини яхшилаш, тилни лингвистик моделлаштириш, сўзларни леммалаш назарияси, алгоритмини яратиш ҳамда тилнинг кўп асрлик миллий-маданий меросдан фойдаланиш имкониятини ошириш мақсадида уларни электронлаштириш жаҳон тилшунослигида долзарб масалага айланди. Турли хил корпуслар яратиш, мавжуд корпуслар ҳажмини кенгайтириш, матнни автоматик қайта ишлайдиган дастур(таггер, парсер)ларни ишлаб чиқиш глобаллашаётган даврда тилларнинг яшовчанлик имконияти ошишига хизмат қилади.

Мустақиллик йилларида ўзбек тилини ахборот технологиялари тилига айлантириш жараёни жадаллашди. Ўзбек тилининг интернет тили даражасига кўтарилмаганлиги ва бу йўналишда ишлар жуда суст амалга оширилаётганлиги ўзбек тилшунослиги олдида турган долзарб муаммолардан бири бўлиб, бугунги куннинг кечиктириб бўлмас вазифаларидандир. Ўзбек тилининг миллий корпусини тузиш учун лингвистик маълумотлар омборини яратиш ўзбек тилшунослари олдидаги муҳим муаммо эканлигини қайд қилган ҳолда “Давлат тилининг ахборот ва коммуникация технологиялари, хусусан, Интернет жаҳон ахборот тармоғида муносиб ўрин эгаллашини таъминлаш, ўзбек тилининг компьютер дастурларини яратиш”¹ ўзбек тилшунослиги олдида турган долзарб вазифалардан бири эканлигини таъкидлаш лозим. Шу маънода корпус лингвистикаси истикболли йўналиш сифатида амалий тилшуносликнинг янги соҳаси бўлиб, бугунги кунда ўзбек тилшунослигида амалга ошириладиган амалий тадқиқотлар лингвистик корпусга асосланиши зарур. Ўзбек тили миллий корпуси учун синонимлар базасини яратиш, синонимларни теглашнинг лингвистик асосларини ишлаб чиқиш сингари

¹Ўзбекистон Республикаси Президентининг “Ўзбек тилининг давлат тили сифатидаги нуфузи ва мавқеини тубдан ошириш чора-тадбирлари тўғрисида”ги ПФ-5850-сон Фармони.// Халқ сўзи, 2019, 22 октябрь. №218(7448)

масалаларни замонавий тилшунослик тамойиллари асосида тадқиқ этиш долзарблик касб этмоқда.

Ўзбекистон Республикаси Президентининг 2016 йил 13 майдаги “Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университетини ташкил этиш тўғрисида”ги ПФ-4997-сон, 2017 йил 7 февралдаги “Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси тўғрисида”ги ПФ-4947-сон, 2019 йил 21 октябрдаги “Ўзбек тилининг давлат тили сифатидаги нуфузи ва мавқеини тубдан ошириш чора-тадбирлари тўғрисида”ги ПФ-5850-сон Фармонлари, 2017 йил 17 февралдаги “Фанлар академияси фаолияти, илмий тадқиқот ишларини ташкил этиш, бошқариш ва молиялаштиришни янада такомиллаштириш чора-тадбирлари тўғрисида”ги ПҚ-2789-сон ва 2019 йил 4 октябрдаги “Ўзбекистон Республикасининг “Давлат тили ҳақида”ги Қонуни қабул қилинганлигининг ўттиз йиллигини кенг нишонлаш тўғрисида”ги ПҚ-4479-сон Қарорлари ва мазкур фаолиятга тегишли бошқа меъёрий-ҳуқуқий ҳужжатларда белгиланган вазифаларни амалга оширишга ушбу диссертация иши муайян даражада хизмат қилади.

Тадқиқотнинг Ўзбекистон Республикаси фан ва технологиялар тараққиётининг устувор йўналишларига мослиги. Тадқиқот республика фан ва технологиялар ривожланишининг I. «Ахборотлашган жамият ва демократик давлатни ижтимоий, иқтисодий, ҳуқуқий, маданий ва маънавий-маърифий ривожлантиришда инновацион ғоялар тизимини шакллантириш ва уларни амалга ошириш йўллари» устувор йўналишига мувофиқ бажарилган.

Муаммонинг ўрганилганлик даражаси. Жаҳон ва ўзбек тилшунослигида синонимлар тадқиқи узоқ ривожланиш босқичига эга. Зеро, лисоний маънодошликка бағишланган қарашлар нафақат жаҳон тилшунослиги, туркология, қолаверса, Алишер Навоий ҳазратлари томонидан илгари сурилган назарияларда ҳам ўз ифодасини топган. Ўзбек тилшунослигида синонимлар тадқиқи А.Ҳожиев, Ш.Раҳматуллаев, Ё.Тожиев, Ҳ.Шамсиддинов, М.Вафоева, М.Ҳакимова²лар тадқиқотларида ўзининг нисбатан шаклланган назарий асосини топган бўлса-да, бу борада яратилган лексикографик ишлар ҳали бу соҳада қилинажак ишларнинг салмоғи катта эканлигини кўрсатади.

Жаҳон тилшунослигида корпус лингвистикаси ўтган асрнинг 60-йилларида ўрганиш объектига айланган. Корпус ҳақидаги илк фикрлар Р.Г.Пиатровский томонидан айтилган³ бўлса-да, бу соҳадаги мақсадли тадқиқотлар 40-йилларда Блумфильд, Фрайс ва Бонджерслар томонидан бошланган⁴. Н.Френсис ва Г.Кучера томонидан корпус тузиш принципларининг ишлаб чиқилиши,

² Ҳожиев А. Тилшунослик терминларининг изоҳли луғати. – Тошкент, 2002. –б.; Ё.Тожиев. Ўзбек тилида аффиксиал синонимия. – Тошкент: Университет, 1991; Шамсиддинов Ҳ. Сўзлар функционал-семантик синонимларининг шаклий ва маъновий табиати. – Тошкент: Университет, 1999; Шамсиддинов Ҳ. Ўзбек тилидасўзларнинг функционал-семантик синонимлари: Филол. фан докт. дис. автореф. – Тошкент, 1999. – 40 б.; Вафоева М. Ўзбек тилида фразеологик синонимлар ва уларнинг структурал-семантик таҳлили: филол. фан. номз. дисс. автореф. – Тошкент, 2009. – 26 б.; Ҳакимова М. Семасиология. – Тошкент, 2008.

³ Кутузов А.Б. Корпусная лингвистика. – (Электрон ресурс): Лицензия Creative commons Attribution Share-Alike 3.0 Unported (Электрон ресурс) - //lab314.brsu.by/kmp-lite/kmp-video/CL/CorporeLingva.pdf.

⁴ Блумфилд Л. Язык. – М.: «Прогресс», 1968. – 608 с.; Fries Ch.C. The structure of English. An introduction to the construction of English sentences. – L., 1969; Bongers H. The history and principles of Vocabulary control. – Woerden: WOCOPI, 1947.

Ж.Синклер томонидан дастлабки инглиз корпусининг яратилиши⁵ корпус лингвистикасида илк назарий ва амалий қадам бўлди. Рус тилшунослигида В.П.Захаров, А.Б.Кутузов, Е.В.Недошивина, В.В.Риков, В.Плунгян, О.В.Кукушкина, А.А.Поликарпов, Е.В.Суровцевалар корпус, унинг турлари, ўзига хос хусусияти, корпуснинг ижтимоий аҳамияти, корпус тузиш тамойиллари борасида тадқиқот олиб боришган. Ўзбек тилшунослигида компьютер лингвистикаси, матнга лексикографик ишлов бериш ва лингвостатистик таҳлил этиш борасида бирмунча тадқиқотлар амалга оширилган. А.Пўлатов, С.Муҳамедов, М.Айимбетов, С.Муҳамедова, С.Каримов, Г.Жуманазарова, А.Бабанаров, Д.Ўринбоева, Н.Абдурахмонова, А.Норов ва бошқаларнинг тадқиқотлари ана шундай ишлар сирасига киради⁶. Бу тадқиқотлар матни компьютер лингвистикаси ютуқлари ёрдамида лексикографик ва лингвостатистик тадқиқ этишнинг замонавий усуллари тавсия этган бўлса-да, уларда ўзбек лингвистик корпусларини яратиш масаласи кун тартибига қўйилмаган. Ш.Ҳамроева томонидан олиб борилган тадқиқотни мана шу йўналишдаги амалий қадам деса бўлади⁷. Ўзбек тилшунослигида бу борада қилинган ишларни Ўзбек тили Миллий корпусини яратишдаги илк қадам сифатида баҳолаш мумкин.

Тадқиқотнинг диссертация бажарилган олий таълим ёки илмий-тадқиқот муассасасининг илмий-тадқиқот ишлари режалари билан боғлиқлиги. Диссертация Термиз давлат университети Ўзбек тилшунослиги кафедрасининг 2017-2021 йилларга мўлжалланган “Ўзбек тилининг системавий тадқиқи” мавзусидаги илмий-тадқиқот ишлари режаси доирасида бажарилди.

⁵Френсис Н., Кучера Г. Вычислительный анализ современного американского варианта английского языка. – М., 1967.; Синклер Д. Предисловие к книге «Как использовать корпуса в преподавании иностранного языка»/ Д. Синклер [Электронный ресурс]. – Режим доступа: <http://www.ruscorpora.ru/corpora-infro.html>, свободный.; Кутузов А.Б. Корпусная лингвистика. – (Электрон ресурс): Лицензия Creative commons Attribution Share-Alike 3.0 Unported (Электрон ресурс)-/lab314.brsu.by/kmp-lite/kmp-video/CL/Corpore Lingva.pdf; Недошивина Е.В. Программы для работы с корпусами текстов: обзор основных корпусных менеджеров. Учебно-методическое пособие. – Санкт-Петербург. – 2006. – 26 с.; Рыков В.В. Курс лекций по корпусной лингвистике. URL: <http://rykov-cl.narod.ru/c.html>; Плунгян В. Зачем мы делаем Национальный корпус русского языка? «Отечественные записки» 2005, №2. http://magazines.russ.ru/oz/2005/2/2005_2_20-pr.html

⁶Муҳаммедов 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; Айимбетов М.К. Опыт лингвостатистического анализа лексики и морфологии каракалпакского публицистического текста: Автореф. дисс... канд. фил. наук. – Ташкент, 1987; Каримов С., Қаршиев 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.дис...автореф. – Тошкент, 2008; Муҳаммедова С. Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъминотияратиш. Методик қўлланма. – Тошкент, 2006; Ўринбоева Д.Б. Ўзбек фольклори матнларининг лингвостатистик тадқиқи. – Тошкент: «Фан», 2010; Жуманазарова Г.У. Фозил Йўлдош ўғли дostonлари тилининг лингвопоэтикаси: Фил. фан. док. дис...автореф. – Тошкент, 2017; Абдурахмонова Н.З. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти (Содда гаплар мисолида): Филол. фан. бўйича фалсафа доктори (PhD)...дис. афтореф. – Тошкент, 2018; Пўлатов А. Компьютер лингвистикаси. – Тошкент: Akademnashr, 2011; Норов А. Компьютер лингвистикаси асослари. – Қарши, 2017. – 136 б.

⁷Ҳамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари. Филол. фан. бўйича фалсафа доктори (PhD)...дис. афтореф. – Қарши, 2018

Тадқиқот мақсади синонимлар базасини шакллантиришнинг назарий асосларини ўрганиш, синонимлар разметкаси орқали синонимлар базасини такомиллаштириш тамойилларини ишлаб чиқишдир.

Тадқиқотнинг вазифалари:

жаҳон тилшунослиги, туркология ва ўзбек тилшунослигида синонимия ҳодисасининг ўрганилишини тадқиқ этиш;

ўзбек тилининг изоҳли луғатларида синонимлар тавсифи ва талқинини ўрганиш;

изоҳли луғатларда синонимлар шарҳидаги фарқларни кузатиш;

ўзбек тилининг синонимлар луғатида луғавий маънодош бирликларнинг изоҳланишини ўрганиш;

синонимлар базасини шакллантиришнинг лингвистик асосларини тадқиқ этиш;

лингвистик разметка турларини аниқлаш ва бунда семантик разметканинг аҳамияти ва моҳиятини текшириш;

семантик разметканинг синонимлар базасини тузишда қўлланадиган усул ва воситаларини ўрганиш, уларни амалиётга фаол тадбиқ этишни назарий асослаш;

синонимлар базасини тузиш тамойилларини ишлаб чиқиш;

семантик разметка хусусиятларини тавсифлаш.

Тадқиқотнинг объекти сифатидалисоний синонимлар ва уларнинг разметкаси танланган.

Тадқиқотнинг предметини ўзбек тили синонимлар базасини шакллантириш, синонимик бирликларни теглаш муаммолари ташкил этади.

Тадқиқотнинг усуллари. Тадқиқот мавзусини ёритишда таснифлаш, тавсифлаш, қиёслаш, статистик, когнитив таҳлил методларидан фойдаланилди.

Тадқиқотнинг илмий янгиллиги қуйидагилардан иборат:

синонимлар базасини шакллантиришда лисон – нутқ фарқланиши ва градуонимик (даражаланиш) тамойили лингвистик методология эканлиги асосланган;

семантик разметканинг лексема луғавий маъноси, маъно кўлами ва кўчим туридан кейинги поғонадаги босқичи синонимиктавсиф эканлиги аниқланган;

маънодош сўзлар базасини тузишда семантик разметканинг таснифий ўрни очиқ берилган;

ўзбек тилининг лисоний синонимларини теглаш алгоритми ишлаб чиқилган.

Тадқиқотнинг амалий натижалари қуйидагилардан иборат:

тадқиқот натижасида олинган хулосалар ўзбек тили миллий корпусининг синонимлар базасини тузиш тамойилларини ишлаб чиқиш учун ўзига хос манба эканлиги далилланган;

диссертацияда ўзбек тили синонимлари базасининг яратилганлиги лингвистик корпус тузувчилари учун муҳим маълумот бўлиб хизмат қилиши асосланган.

Тадқиқот натижаларининг ишончлилиги ўрганилган материалларнинг ўзбек тили табиатидан келиб чиққан ҳолда хулосалар қилишга ёрдам берганлиги, уларнинг асосли эканлиги, методологик мукамаллиги, амалда исботланган манбаларга таянилганлиги билан изоҳланади.

Тадқиқот натижаларининг илмий ва амалий аҳамияти ўзбек тили миллий корпусини яратишда синонимлар базасини яратишнинг назарий асослари ва алгоритминини ишлаб чиқиш билан белгиланади.

Тадқиқот натижаларининг жорий қилиниши. Ўзбек тили миллий корпусининг синоним сўзлар базасини ишлаб чиқиш жараёнида эришилган илмий натижалар асосида:

семантик разметканинг лексема луғавий маъноси, маъно кўлами ва кўчим туридан кейинги поғонадаги босқичи синонимиктавсиф эканлиги ҳақидаги илмий хулосалари асосида ўзбек тили синонимлари теглари лингвистик моделидан синонимлар базасининг on-line варианты фрагментини яратишда фойдаланилган (Олий ва ўрта махсус таълим вазирлигининг 2019 йил 21 майдаги 89-03-2113-сон маълумотномаси). Натижада “Ҳозирги ўзбек тили”, “Компьютер лингвистикаси” фанлари мазмунини такомиллаштириш ва ўқитиш сифатини оширишга эришилди;

маънодош сўзлар базасини тузишда семантик разметканинг таснифий ўрни эканлигига доир илмий натижаларидан Ўзбекистон Ёшлар иттифоқи Марказий Кенгаши давлатимиз раҳбари томонидан илгари сурилган 5 та муҳим ташаббуснинг учинчи ташаббуси доирасида бажарилаётган ишларда амалиётга жорий этилди (Ўзбекистон Ёшлар иттифоқи Марказий Кенгашининг 2019 йил 2 июлдаги 04-13/3033-сон далолатномаси). Натижада ёшларнинг билим олишга бўлган қизиқиши ҳамда интернет орқали ўзбек тилига оид билимларининг такомиллашига эришилди.

синонимлар базасини шакллантиришда лисон – нутқ фарқланиши ва градуонимик (даражаланиш) тамойили лингвистик методология эканлигини ақс эттиришга доир материалларидан Сурхондарё вилояти телерадио компаниясининг “Диёр янгиликлари” ва “Бир асар шарҳи” кўрсатувлари сценарийларини тайёрлашда фойдаланилган (Сурхондарё телерадиокомпаниясининг 2019 йил 9 апрелдаги 01-15/263-сон маълумотномаси). Натижада телекўрсатув ва радиоэшиттиришлар янги материаллар, илмий асосланган далиллар билан бойитилган.

Тадқиқот натижаларининг апробацияси. Мазкур тадқиқот натижалари бўйича 2 та халқаро ва 4 та республика илмий-амалий анжуманларида маърузалар қилинган.

Тадқиқот натижаларининг эълон қилинганлиги. Диссертациянинг асосий мазмуни муаллиф томонидан чоп этилган 1 та муаллифлик гувоҳномаси, 9 та илмий мақола (уларнинг 3 таси Ўзбекистон Республикаси ОАК тасарруф этган илмий журналларда) ва тезисларда ўз ифодасини топган.

Диссертациянинг тузилиши ва ҳажми. Диссертация кириш, уч асосий боб, умумий хулоса, фойдаланилган адабиётлар рўйхати ва иловадан иборат. Диссертациянинг умумий ҳажми 131 саҳифани ташкил этади.

ДИССЕРТАЦИЯНИНГ АСОСИЙ МАЗМУНИ

Кириш қисмида мавзунинг долзарблиги асосланган, тадқиқотнинг мақсад, вазифалари, объекти ва предмети тавсифланган, республика фан ва технологиялари ривожланишининг устувор йўналишларига мослиги кўрсатилган, илмий янгилиги ва амалий натижалари баён қилинган, натижаларни амалиётга жорий қилиш, нашр этилган ишлар ва диссертация тузилиши бўйича маълумотлар келтирилган.

Диссертациянинг **“Ўзбек тили миллий корпусида синонимларни беришнинг лингвистик асослари”** деб номланувчи I бобида дунё тилшунослигида синонимия ҳодисасининг ўрганилиши, ўзбек тилшунослигида синонимия ҳодисасининг тадқиқи, ўзбек тили изоҳли луғатларида синонимларнинг берилиши каби ҳолатлар ҳақида фикр билдирилган.

Қадимги юнон тилшунослигининг фалсафий давридан бошлаб файласуфлар Продик, Платон, Сукрот, Демокритнинг синонимия ҳодисасига муносабати ёритилган, шунингдек, Хитой, ҳинд, араб тилшуносликларида ҳам синонимия ҳодисасининг ўрганилиши акс этган. Синонимиянинг ўрганилишига боғлиқ иккинчи босқич XVII- XIX аср оралиғига алоқадор, бунда синонимларга оид кўплаб луғат чоп этилганлиги қайд қилинган. Ушбу даврда синонимия тадқиқи И.И.Давидов, Д.П.Калайдович, М.В.Ломоносов, И.Фонвизин ва бошқалар⁸ луғатчилик ҳамда синонимларни текшириш орқали муносабатини билдиришган.

С.Г.Бережан ХХI асргача синонимлар шунчаки нутқни бойитувчи восита сифатида ўрганилганини таъкидлаганлиги қайд қилинган. XIX асрнинг иккинчи ярми ҳамда XX асрнинг ўрталарига келиб синонимларни ўрганиш долзарблик касб этган. Семасиологиянинг фан сифатида эътироф этилиши эса олимларнинг ушбу тадқиқ объектига жиддий ёндашувларга туртки бўлди. Айтиш мумкинки, **“йигирманчи аср “синонимияни ўрганишнинг олтин асрига айланди”**⁹. XX асрнинг 50–70-йилларида фаннинг

⁸XVIII ва XIX асрларда синонимлар луғатини тузишда Д. И. Фонвизиннинг хизматлари улкан. Шунингдек, Н. Абрамовнинг 1890 (1-нашр), 1915 (4-н.), 1994 (5-н.) йилларда чоп этилган «Словарь русских синонимов и сходных по смыслу выражений» – М.: Русские словари, 1994. – 502 с.; 1840 йилда А. И. Галичининг «Словарь русских синоним или словослов, составленный редакцией нравственных сочинений» (226 словарных статей) издание в виде книги (I часть), П. Ф.Калайдовичнинг 1818 йилда нашр этилган «Опыт словаря русских синонимов» (77 групп синонимов) издание в виде книги (I часть), 1783 йилда Д. И. Фонвизин томонидан нашр этирилган «Опыт российского словословника» (32 группы синонимов) журнал «Собеседник Любителей Российского Слова» (ч. I, IV, X) асарларини эътироф этиш керак.

⁹Шумилова А.А. Истоки исследования синонимии в философских и лингвистических учениях. http://www.rusnauka.com/4_SND_2013/Philologia/3_127613.doc.htm; Горбачевич К.С. Русский синонимический словарь. [Ин-т лингвист. исслед. РАН]. – СПб: ИЛИ РАН, 1996. – 510 с.; Зимин В. И., Алекторова Л. П., Ким О. М., Колесников Н. П., Шанский В. Н. Учебный словарь синонимов русского языка (2800 синонимических рядов) – М., 1994; Жуков В. П., Сидоренко М. И., Шкляр В. Т. Словарь фразеологических синонимов русского языка. – М., 1987; Евгеньева А. П. Словарь синонимов. Справочное пособие. – М., 1975; Евгеньева А. П. Словарь синонимов русского языка. – Л.: Наука, 1970 – 1971. – Т. 1–2.; 1968 г. (1-изд, 1969 г. (2-и.), 1975 г. (4-и.), 1986 г. (5-и.), 1989 г. (6-и.), 1993 г. (7-и.), 2001 г. (11-и.) Александрова З. Е. «Словарь

ривожли лингвистикада синонимиянинг ривожига сабаб туғдирганлигини қайд қилиш мумкин.

Ўзбек тилшунослиги фанида синонимия ҳодисасига бағишланган тадқиқотлар асосан XX асрнинг 60-йилларига тўғри келган.

Ўзбек тилшунослигида синонимларни ўрганиш ўз тарихий тадрижига эга, уни шартли равишда қуйидаги босқичларга ажратиш мумкинлиги қайд қилинди:

- I босқич: XX асргача бўлган давр;
- II босқич: 1920-1930 йиллар;
- III босқич: 1950-1960 йиллар;
- IV босқич: 1960-1980 йиллар;
- V босқич: 1980-2000 йиллар;
- VI босқич: 2000ва ундан кейинги йиллар.

XX аср иккинчи ярмидан ўзбек тилшунослигида сўзлар (синтактик конструкциялар)даги синонимияга муносабат тараққий этиб, бу борадаги изланишлар XXI аср ўзбек тилшунослигида ҳам зинҳор сусаймаганлиги қайд қилинган.

Бобда ўзбек тилининг изоҳли луғатларида синонимларнинг таърифи ва тавсифи масаласига муносабат билдирилган. ЎТИЛ нинг икки нашрида синонимларга берилган изоҳлар шартли равишда қуйидаги гуруҳларга бўлинган:

- 1) изоҳи ва мисоллари айнан бўлган синонимлар;
- 2) изоҳи бутунлай янги, мисоллари айнан;
- 3) изоҳ қисман янги, мисоллар айнан;
- 4) изоҳ қисман янги, мисоллар қисман ўзгарган;
- 5) изоҳлар айнан, мисоллар янги;
- 6) изоҳлар ҳам, мисоллар ҳам янги;
- 7) олдингисида йўқ, янгисида бор;
- 8) синонимига ишора қилиш асосида айн. деб изоҳлаш (олдингисида ва кейингисида).

Юқорида келтирилган гуруҳларга ЎТИЛнинг икки нашрида келтирилган изоҳлар келтириб ўтилган.

Навбатдаги боб аввалида синонимларнинг тавсифига доир қуйидаги жадвални¹⁰ келтириб ўтмоқни жоиз деб билдик (1-жадвал). Зеро, корпусларнинг семантик разметкаси масаласида бу сингари таснифлар алоҳида аҳамият касб этади.

синонимов русского языка» – М.: Рус. яз. 2001. – 567 с. (11 000 синонимических рядов); Ключева В. Н. Краткий словарь синонимов русского языка. – М., 1956 г., 1961 г. (2-изд.); Павлов-Шишкин В. Д., Стефановский П. А. Учебный словарь синонимов русского литературного языка. – М., 1930 г., 1931 г. (2-изд.).

¹⁰<http://book-science.ru/humanities/linglang/lexicology/leksiko-semanticheskaja-paradigma-chast-3.html>

Синонимларнинг жадвалли таснифига доир намуна

Синонимик парадигма	Синонимларнинг лексик маъноси	Синонимларнинг семик таркиби				Синоним типлари		Синонимларнинг услубий хосланиши
		Интеграл семалар	Дифференциал семалар			семантик	структуравий	
			Миқдор билдирувчи семалар	Эмоционал-экспрессив коннотация	Услубий сема			
Енгил (доминанта)	‘тушуниш’, ‘ўрганиш’, ‘ўзлаштириш’ ‘ва’ ‘бажариш’ да ‘кийинчилик’ ‘туғдир’ ‘майдиган.’				нейтр.		Бир асосли	Барча услублар учун
Қийин эмас	‘етарли’ ‘даража’ да ‘енгил’		‘етарли’		нейтр.	идеографик	Ҳар хил асосли	Барча услублар учун
Оддий	‘алоҳида ақл, билим ва кўп меҳнат талаб қилмайдиган’		‘алоҳида’		нейтр.			Барча услублар учун
Элементар	‘жуда оддий’	«	‘жуда’		нейтр.	«	«	Илмий
Мураккаб эмас	‘заррача’ ‘машаққат’, ‘қийинчилик’нинг ‘мавжуд’ ‘эмас’лиги	‘Жуда’ ‘осон’	‘заррача’		Сўзлашув	Идеографик-услубий	«	Расмий, илмий, публицистик

“Ўзбек тили миллий корпусида синонимия ҳодисасини разметкаланишнинг назарий масалалари” деб номланувчи II бобда корпус лингвистикасининг тадрижий тараққиёти, терминологик аппарати, синонимик разметканинг ўзиги хос жиҳатлари тадқиқ қилинган.

Корпус лингвистикаси компьютер лингвистикасининг мустақил бўлими, лингвистик корпуслар (матн корпуслари)ни тузиш ва ундан фойдаланиш тамойилларини компьютер технологияларини қўллаган ҳолда ишлаб чиқиш билан шуғулланади¹¹. У XX аср 90-йилларининг ўрталаридан тилшуносликнинг алоҳида соҳаси сифатида ривожлана бошлаганлиги қайд қилинган. Корпус лингвистикаси матн корпуслари таҳлили билан шуғулланувчи фан сифатида Ғарбий Европа ва АҚШда 1960 йиллар охирида вужудга келди. Компьютер ва ахборот-коммуникатив технологияларининг шиддатли ривожини натижасида 1980 йиллар ўрталарида турли тил ва ҳажмдаги корпус лойиҳалари яратилди. Хусусан, Британия миллий корпусидан¹² илҳомланиб яратилган рус миллий корпуси ҳам ўз ичига 163 млн (бугунгача эса 500 млн.) тарихий ва замонавий қатламга тегишли сўзларни киритганлиги билан машҳур¹³. Ушбу корпус қуйидаги бўлимлардан иборат:

- 1) асосий корпус (XVIII-XXI асрга доир насрий матнлар);
- 2) синтактик корпус (ҳар бир гап учун морфологик ва синтактик курилманинг келтирилганлиги);
- 3) газета корпуси (1990-2000 йиллар ОАВ мақолалари);
- 4) параллел корпуслар (тегишли сўзнинг хорижий тилдан рус тилига ва рус тилидан хорижий тилга бўлган таржимаси);
- 5) диалектал матнлар корпуси;
- 6) шеърӣ матнлар корпуси;
- 7) рус тилини ўргатувчи корпус (содда мактаб ўқув дастурига асосланган);
- 8) оғзаки нутқ корпуси;
- 9) мультимедия материаллари корпуси (1930-2000 йиллар кинофильмлари жамланган);
- 10) русча урғу тарихининг корпуси (урғу тарихига оид матнлар)¹⁴.

Олимларнинг қайд этишича, XX асрнинг 90-йилларига келиб, яратилган корпусларнинг сони 600 дан ошган. Уларнинг яратилиш санаси йилларга нисбатан қуйидагича тақсимланган.

¹¹ Захаров В.П. Корпусная лингвистика. Учебно-методическое пособие. – Санкт-Петербург, 2005. – 48 с.

¹² Британия Миллий корпуси (British National Corpus – BNC)га тенг равишда, инглиз тили корпусларидан яна: инглиз тили Халқаро корпуси (International Corpus of English – ICE), инглиз тили лингвистик Банки (Bank of English), Ҳозирги замон Америкаинглиз тили Корпуси (Corpus of Contemporary American English – COCA) ва б.ларни ҳам келтириб ўтиш мумкин.

¹³ Майорова А.Д. Корпусная лингвистика: исторический и лингводидактический аспекты/ Педагогические Науки. Выпуск № 5(59) Май 2017.

¹⁴ Майорова А.Д. Корпусная лингвистика: исторический и лингводидактический аспекты/ Педагогические Науки. Выпуск № 5(59) Май 2017.

Бобда компьютер ва лисон ўртасида тўғридан-тўғри “ўзарочилик” тамойили мавжудлиги ҳам аниқланди; у *компьютер – дастур – билим – лисон/тил* концепцияда ҳам кўринади¹⁵. Корпус лингвистикасининг асосий тушунчаси – лингвистик корпус. Корпусга матнлар муаммога тегишлилик нуқтаи назаридан, яъни тилшунос тадқиқотчилар учун муҳим бўлган тематик алоқадорлик асосида танланади. Муаммолар соҳаси – тематик кўлам икки аспектдан иборат, ўзида тил ва нутққа доир матнларни жамлайди. Тил аспекти энг тадқиқотталаб соҳа бўлса, нутқ аспекти тил аспектига доир ҳолатларни акс эттирувчи контекстлардан таркиб топади. Матнларни ҳажмига кўра саралаш муаммоси эса корпус яратишнинг муҳим масалаларидан. Мазкур жиҳат корпуснинг *мукаммаллик* жиҳатини таъминлаб беради. Матнларни саралашюзасидан корпуслар қуйидаги турларга бўлинади:

- мувозанатга солинган корпуслар;
- мониторли корпуслар.

Матн бирликларига тегишли қўшимча лингвистик ва металингвистик ахборот (аннотация) – разметкани амалга ошириш зарурати юзага келади. Замонавий корпусларда метаразметка(матнга боғлиқ экстралингвистик маълумотлар)дан ташқари, лингвистик таърифни ифода этувчи: морфологик, синтактик, фонетик ва ҳ. турдаги разметка амалиёти амалга оширилади. Матн корпуси тушунчаси ўз ичига бир неча атамани жамлайди. Хусусан, корпус менеджери (корпус-менеджер) – корпус ичида тегишли маълумотларни излашга, фойдаланувчига тегишли маълумотлар юзасидан статистик ахборотни тақдим этувчи ҳамда истёъмолчилар оммасига уларга мақбул бўлган шаклдаги ахборотни узатишга ихтисослашган дастурларидан таркиб топувчи махсус қидирув тизим. Ўз навбатида, корпусдаги қидирув ўзига хос конкорданс – манбаси кўрсатилган контекстларда изланаётган лисоний бирлик тўғрисидаги барча қайдлар рўйхатини шакллантиради. Гарчи, олимлар бугунги кунгача корпус лингвистикасини назарий асослари етарлича шаклланмаган соҳа сифатида таснифласалар ҳам, умид борки, келажакда тадқиқотчилар корпусга шунчаки ахборот ресурси сифатида эмас, балки ўз методологик асослари ҳамда илмий-назарий аҳамиятга эга бўлган соҳа сифатида ёндаша бошлайдилар.

Шу ҳолатда анъанавий ва корпус лингвистиканинг ўзаро фарқлари қуйидагиларда кўринади (1.2-жадвал).

¹⁵ Шемякин Ю.И. Начала компьютерной лингвистики: Учеб. пособие. – М.: Изд-во МГОУ, А/О "Росвузнаука", 1992.

Анъанавий ва корпус лингвистиканинг ўзаро фарқлари

Хусусиятлар	Анъанавий лингвистика	Корпус лингвистикаси
Йўналганлик	Тил/лисон тадқиқи	Нутқ тадқиқи
Тадқиқ объектига нисбатан	Назария орқали нутқ ходисаларини тушунтириш	Матн корпуси
Усуллар	Квалитатив	Квантитатив
Матнга муносабат	Мавҳумлик	Физик мавжудлик

Корпуслар разметкаланганлик белгисига кўра ҳам муайян тоифаларга бўлинади. Бунда *разметка* атамаси билан ёндош ҳолда, бўлинишни ифодаловчи: *индексланган, индексланмаган, аннотацияланган, аннотацияланмаган, тегланган, тегланмаганлик* терминлари ҳам учрайди. Разметкаланган корпусда сўз ва жумлаларга разметканинг морфологик, синтактик, семантик, просодик ва ҳ. турини ифода этувчи лингвистик разметкага кўра *тегл*ар бириктирилади¹⁶.

Лингвистик разметканинг тамойиллари мавжуд, булар:

- 1) разметка схемаси/тузилишини тавсифлаш (асослаш);
- 2) лингвистик тушунчаларнинг умумқабул қилинган системаси;
- 3) таҳлилнинг фойдаланувчига маълум бўлган схемаси;
- 4) параметрларнинг асослантирилган тарзда киритилиши;
- 5) разметканинг анъанавий схемаси;
- 6) халқаро стандартларга амал қилиш¹⁷.

Семантик тег (рус., разметка) – сўз ёки сўз бирикманинг маълум бир семантик категория ёки ундан ҳам кичикроқ семантик гуруҳ(ЛСГ, семантик майдон ва тўда)га мансубликни билдирувчи, маънони махсуслаштирувчи белги, изоҳлар мажмуи¹⁸. Корпусларнинг семантик теглари сўз маъно(лар)и спецификаси, сўзнинг омонимлик, синонимлик билан боғлиқ изоҳлар мажмуини тузиш, сўзни категориялаш, унинг мавзувий гуруҳ, ЛСГ, семантик майдонга мансублигини белгилаш, деривацион характеристикаси, аташ маъноси каби белгиларни қамраб олади.

“Ўзбек тили миллий корпусида синоним сўзларни теглаш ва унинг дастурий таъминоти” деб номланган III бобда корпусда синоним сўзларни теглаш, унинг дастурий таъминоти, теглаш алгоритмлари ҳақида фикр билдирилган.

Синоним лексемалар бир қанча умумий семаларга эга бўлиши билан бирга, семантик жиҳатдан қуйидагича фарқланишларга ҳам эгаллиги аниқланди:

¹⁶ Захаров В.П., Богданова С.Ю. Корпусная лингвистика: учебник для студентов гуманитарных вузов. – Иркутск: ИГЛУ, 2011. – С. 25.

¹⁷ Захаров В.П., Богданова С.Ю. Корпусная лингвистика: учебник для студентов гуманитарных вузов. – Иркутск: ИГЛУ, 2011. – С. 45.

¹⁸ Захаров В.П., Богданова С.Ю. Корпусная лингвистика: учебник для студентов гуманитарных вузов. – Иркутск: ИГЛУ, 2011. – 161 с. – С. 48-49.

- 1) коннотатив маънодаги фарқлилик;
- 2) денотатив маънодаги фарқлилик;
- 3) услубий фарқ.

Изоҳ ва тавсифлар асосида синонимларнинг турлари ҳақида аниқ бир тўхтамга келиш лозим. Демак, синонимларнинг абсолют/мутлақ синоним (а); идеографик/семантик синоним (б); услубий синоним (в); контекстуал/матний синоним (г); фразеологик синоним (д) каби турларини ажратдик. Мавжуд изоҳли луғатлар тил бирлик(синоним)ларни теглаш учун база вазифасини бажаради. Бунинг учун мавжуд луғатлар қандай ишловдан ўтиши лозим, луғатлар қандай кўринишга келса, улардан автоматик теглашда фойдалана оламиз каби саволларга мавзунинг шу қисмида жавоб топамиз. Ўзбек тилининг мавжуд синонимик луғати ҳозирги ўзбек адабий тили учун характерли бўлган синонимик қаторларни ўз ичига олган. Аммо барча синонимик қаторларни қамраб ололмаган. Унга фақат лексик (лисоний) синонимларгина киритилган. Тил корпуслари эса контекстуал синонимларни ҳам қамраб олиши лозим. Энг муҳим иш – синонимик қаторни, доминантани аниқлаш ва уларни теглаш:

1) луғатнинг электрон матнида синонимик қатор ажратиб кўрсатилади: {куй, тарона}, {кифоя, етарли, етади, бўлади, бўлди, бас}. Мана шу чегара дастурлаш учун синонимик қатор аъзоларини белгилаб беради;

2) луғатнинг электрон матнида синонимик қатор доминантаси аниқланади ва асосий лемма белгиси билан ажратилади. Масалан: *кутмоқ*, *пойламоқ*. Ушбу синонимик қаторда *кутмоқ* доминанта, демак уни асосий лемма сифатида белгилаймиз: <<*кутмоқ*>>, <*пойламоқ*>, <<*кулги*>>, <*ханда*> каби. Умумий ҳолатда эса {<<*кутмоқ*>>, <*пойламоқ*>}, {<<*кулги*>>, <*ханда*>} шаклини олади.

Бобда электрон базага киритилаётган лингвистик маълумотлар қайси тамойиллар эвазига сараланиши аниқланди.

Саралаш тамойиллари синонимлар базасининг лингвистикқийматини белгилайди, бинобарин, у билан боғлиқ амалга ошириладиган тадқиқотлар самарадорлиги ҳам шунга мувофиқ белгиланади. Бунда бевосита қуйидагилар инобатга олинади:

- 1) синонимлар базасига киритилаётган бирликларнинг қайси қатламга мансублигини белгилаш;
- 2) базавий маълумотлар ҳажмини белгилаш;
- 3) базага киритилиши лозим бўлган матн манбаларининг тури, шакли, жанри ва б. хусусиятини белгилаш;
- 4) синонимик материалларни мазкур ҳодисанинг тилнинг қайси катламига тегишлилигига кўра саралаш;
- 5) киритилаётган бирликнинг частотасига эътибор бериш;
- 6) базага киритилувчи маълумотларнинг даврийлик жиҳатини белгилаш.

Маълумки, корпус материалларини саралаш амалиёти дастлаб Р.Г.Пиотровскийнинг 1965-1980 йиллар оралиғидаги тадқиқотлари эвазига мукаммаллаштирилган. Бу борада Л.Н.Засоринанинг частотали луғати ҳам назардан четда қолдирилмаган. Кўринадики, сараланаётган тўплам, унинг статистик кўрсаткичлари, даврий кесим бизнинг тадқиқ объектимиш – синонимлар базасини яратишда ҳам муҳим ўрин эгаллайди.

Корпус матнларининг асосий бирлиги сифатида сўз, сўз бирикмалари, ўзак(асос, леммалар) ҳамда жумлалар танланиши мумкин. Албатта, базанинг ҳажми уни яратувчиларнинг мақсадига кўра аниқланади, бу, одатда, тадқиқ предмети ўрнида танланган материалларнинг кўламига нисбатан белгиланади¹⁹.

Юқорида таъкидлаганимиз сингари, корпусга киритилаётган синонимлар базаси, даставвал, синонимик луғатларни тартибга солиш талабини ўртага ташлайди. Ҳолбуки, лексикографик асарларнинг тури уларнинг энциклопедик ва лингвистик хусусиятига нисбатан белгиланади. Бинобарин, энциклопедияда предмет, тушунча, ҳодиса ҳақида ахборот берилса, лингвистик луғатда сўзлар ҳақида маълумот берилади. Зеро, лингвистик луғат бир тилли, икки тилли, кўп тилли сингари турларга бўлинадики, кўп тилли луғат, одатда, таржима мақсадида тузилади. Бир тилли луғатда сўз маъноси шу тил асосида тушунтирилса, икки тилли луғатларда муайян сўзнинг маъноси икки тилдаги қиёси эвазига тушунтириб берилади.

Лингвистик луғатнинг бир тури ҳисобланган изоҳли луғат эса дескриптив ва норматив турга бўлинади. Дескриптив луғат тегишли соҳага оид сўзни тўлиқ изоҳлашга хосланган бўлади, уларда ҳатто релевант ҳолатга ҳам шарҳ берилади. Бундай луғатга ўзбек тилининг изоҳли луғатларини (1981, 2006 йилги нашрлари) мисол қилиш мумкин. Норматив луғатда эса, бундан фарқли, меъёрдан оғишган сўзлар таркибдан чиқарилади.

Синонимлар базасига киритилаётган маълумотлар сирасида ҳам худди шу масалага ойдинлик киритиб олиш талаб этилади. Яъни адабий тил ҳам, диалект, чегараланган лексика (жаргонлар) ҳам ўз меъёрига эга. Бу ҳолда ҳар бир бирлик лотинча Vs (*Versus*нинг қисқартмаси) белгиси билан аниқланади. Хусусан, *она* доминантаси ўз синонимик қаторига *ая* (Фарғона диалекти), *ача* (тожикча), *аба* (Сурхон диалекти), *она* (Самарқанд/Тошкент диалекти), *эна* (Сурхон, Қашқадарё, Самарқанд диалекти) сўзларини қабул қилади. Бу ўринда тузилаётган луғатнинг хусусиятидан қатъий назар, бизнингча, ушбу бирликларнинг барчаси синонимик базага киритилиши лозим.

Тезаруслар – тегишли билим, фаолият, соҳасига доир тушунча, изоҳ ва терминни ўзида жамлайди. Тезаруслар соҳавий алоқадорлик эвазига тузилади.

¹⁹ Захаров В.П., Богданова С.Ю. 3-38 Корпусная лингвистика: учебник для студентов гуманитарных вузов. – Иркутск: ИГЛУ, 2011. – 161 с.

Идеографик луғатда сўзлар алифбо тартибида эмас, лексик маъносига кўра тартибланади; алифбо тартибида тузилган луғат орқали фойдаланувчи муайян сўз ҳақида маълумотга эга бўлса, идеографик луғат ёрдамида фойдаланувчи тегишли маъно ҳақида билимга, масалан, муайян маъно/мазмунни қайси сўз ёрдамида ифода этиш мумкинлиги ҳақидаги ахборотга эга бўлади.

Умумий ҳолда, луғатлар этимологик, сўзлашгич, диалектал, орфографик, поэтик лексикага оид, частотали, тескари ва ҳ. турларга бўлинади. Улардан частотали луғат устида қисман тўхталиш мумкин. Негаки, синонимлар базасига киритилиши лозим бўлган аксар лисоний бирлик, дастлаб, истеъмол кўламига кўра сараланиши лозим. Юқорида лисоний бирликларни синонимлар базасига киритиш тамойиллари сирасида ҳам бу борада тўхталган эдик. Бинобарин, частотали луғатни мукаммал тузиш эвазига синонимик қаторнинг мукаммал тузилиши, пардигмага тўғри тартиб бериш, матнни тўғри таснифлашга эришиш мумкин. Албатта, бу вазиятда лемматизация, лисоний омонимия ҳодисасини ҳам назардан четда қолдирмаслик лозим бўлади. Зеро, синонимларни саралашда частотали луғатга мурожаат қилишнинг асосий сабабидан бири – доминантани тўғри белгилаш, тегишли синоним изоҳида хатоликларга йўл қўймаслик. Шу ўринда частотали луғатлар: 1) леммалаштирилган; 2) леммалаштирилмаган турга бўлинишини инобатга олиш керак.

Кўп ҳолларда маъновий бирлик бир неча узвлибўлиб келади. Мазкур қурилманинг кетма-кетлик тартиби юқори бўлган ҳолда коллокация²⁰ юзага келади²¹. Табиийки, коллокацияда қурилманинг умумий маъноси компонентлар ташиган мустақил маъноларга тенг бўлмайди (*мехр булоғи* $\neq$ *мехр* + *булоқ*). Уларнинг орасида идиомалар (*илон ёғини ялаган, қўли очик*), топонимлар ва алоҳида олинган ономалар ҳам мавжуд, бу ҳолда семантик разметкалаш амалиёти бироз мураккаблашади. Корпусда қўлланувчи конкорданс эвазига мазкур вазиятда сўзлар маъносини ойдинлаштириш имконини эса, ўйлаб кўриш керак бўлади. Зеро, конкордансга мўлжалланган қатор дастурларнинг имконияти юқоридаги вазиятга ечим топиб беради. Шу ўринда корпус лингвистикаси учун муҳим бўлган янги тоифадаги интеграциявий коллокациялар луғатини яратиш зарурати ҳақида қисқача тўхталиш зарур. Бундай луғат турғун бирикмаларнинг изоҳига бағишланган, унда фразема ва бошқа тоифадаги турғун қурилмаларга изоҳ берилади. Бу турдаги луғат ҳам, асосан, частотани аниқлаш тамойили асосида тузилади. Коллокациялар луғатининг объектини ташкил этувчи бирикмалар биграмма²² кўринишида тадқиқ этилади. Яъни тегишли сўзнинг ассоциатив хусусияти тўғрисида у мавжуд бўлган маъно кўламига мос тарзда хулосалар берилади. Буни қуйидаги жадвалда кузатишимиз мумкин (III.4- жадвал):

²⁰коллокат (collocates) – таҳлил қилинаётган сўзнинг ўзи боғланган компонент билан бирга келиши.

²¹Боярский К. К. Введение в компьютерную лингвистику. Учебное пособие. – СПб: НИУ ИТМО, 2013. – 72 с.

²²Таҳлил этилаётган сўз билан боғланган ўнг ва чап тарафдаги компонентларни инобатга олингандаги чизикли таркиб.

Компонент	Лемма	Компонент	Маъноси	Денотат	Синоним
<i>Беи</i>	<i>юз</i>		Микдор белгиси	Микдор	
	<i>юз</i>	бора	Микдор белгиси	Микдор	
<i>Гул</i>	<i>юз</i>		сифат (инсоний хусусият)	гўзал	+
<i>Анор</i>	<i>юз</i>	+ли	сифат (инсоний хусусият)	Гўзал/истарали	+
<i>Ширмой</i>	<i>юз</i>	+ли	сифат (инсоний хусусият)	Гўзал/истарали	+

Бобда ўзбек тили миллий корпусининг синоним сўзлар базаси дастурининг ишлаш принципи ва уларга лингвистик маълумотларнинг киритилиши таҳлил қилинди. Сўзларни лингвистик базага киритишнинг қуйидаги кетма-кетлиги таклиф қилинди:

Сўз киритилади
Сўз туркуми
Маъноси
Этимологияси
Ясалиши
Таркиби
Услуги
Тузилиши
Рус тилидаги муқобили
Инглиз тилидаги муқобили
Даражаланиш қатори
Ижобий салбийлигига кўра
Ишлатилиш микдорига кўра
Омоними
Пароними
Антоними
Ишлатилиши (мисоли)

Синонимлар базасига киритилиши лозим бўлган аксар лисоний бирлик, дастлаб, истеъмол кўламига кўра сараланиши лозим. Бинобарин, бу борада частотали луғатни саводхонлик тузиш эвазига синонимик қаторнинг мукамал тузилиши, пардигмага тўғри тартиб бериш, матнни тўғри таснифлашга эришиш мумкин. Албатта, бу вазиятда лемматизация, лисоний омонимия ҳодисасини ҳам назардан четда қолдирмаслик лозим. Синонимлар базасини шакллантиришда частотали луғатнинг ўрни доминанта ёхуд парадигма иштирокчисининг маъно кўламини, услубий хусусиятини аниқлаш имконини беради.

Бу ёндашув ва тамойиллар, пировардида, ўзбек тили синонимлар луғатини қайта ишлаш, тўлдириш амалиётини амалга ошириш заруратини туғдиради.

Негаки, синонимларга оид лексикографик асарлар мукаммаллаштирилмай туриб, синонимларнинг семантик разметкасини беҳато амалга ошириб бўлмайди. Ушбу тадқиқотимизда эса мавжуд материаллардан унумли фойдаланган ҳолда, ўзбек тили синонимлари базасини яратишнинг назарий ва амалий асосини ёритишга ҳаракат қилдик. Ваҳоланки, бугунги кунда биз томонимиздан ишлаб чиқилиб, фойдаланиш учун тақдим этилган “Ўзбек тили синонимлари базаси” келажакда мукаммаллаштирилиб, бойитилиб борадики, бу ҳам кенг фойдаланувчилар оммаси ҳамда лингвистик тадқиқотлар учун зарур.

ХУЛОСА

1. Синонимия ўрганилишининг I босқичи антик даврга, II босқичи XVII-XIX асрларга тўғри келади. Бу даврда кўплаб синонимлар луғати яратилган. XX асрга келиб синонимларни ўрганишда семантик, структур, прагматик ёндашувлар кузатилди. XXI асрга қадар синонимлар нутқни бойитувчи восита сифатида ўрганилган бўлса, XIX асрнинг II ярми ҳамда XX асрнинг ўрталарига келиб синонимларни ўрганиш долзарблик касб этди. XX аср синонимияни ўрганишнинг «олтин асри»га айланди. XXI аср синонимиянинг қайта туғилиш даври сифатида эътироф этилади. Бу даврда синонимия когнитив, психологик, корпус таҳлил аспектларида текширила бошлади.

2. Туркологияда синонимларни ўрганишда Маҳмуд Кошғарий, Алишер Навоий ҳазратларининг хизмати бекиёс. Синонимлар - ўтмишда бўлгани сингари ҳозир ҳам стилистиканинг ўзаги, унинг жони ва қони. Ўзбек тилшунослигида синонимия ҳодисасига бағишланган тадқиқотлар, асосан, XX асрнинг 60-йилларига тўғри келади. XX асрнинг 80-йилларида синонимиянинг лексик маънодошлик, фразеологик маънодошлик, лексик-фразеологик маънодошликдан иборат уч тури ажратилди. Ўзбек тилшунослигига структур таҳлилнинг кириб келиши синонимларни ўрганишнинг такомиллашувига сабаб бўлди. XX асрнинг 70-йилларидан бошлаб “маъноси айнан бир” меъёридан четланиб, умумий ва хусусий томонларига эътибор берилди.

3. Ўзбек тили изоҳли луғатларининг турли нашрлари, “Адабиётшунослик терминларининг русча-ўзбекча изоҳли луғати”, “Лингвистик терминларнинг изоҳли луғати”, ЎТИЛнинг 2006 йилдаги нашрида синонимларга берилган шарҳларни танқидий кўриб чиқиш шуни кўрсатдики, синонимларни изоҳлашда турли тамойилларга амал қилинган. ЎТИЛнинг икки нашрида синонимларга берилган изоҳлар шартли равишда (1) изоҳи ва мисоллари айнан бўлган синонимлар, (2) изоҳи бутунлай янги, мисоллари айнан, (3) изоҳ қисман янги, мисоллар айнан, (4) изоҳ қисман янги, мисоллар қисман ўзгарган, (5) изоҳлар айнан, мисоллар янги, (6) изоҳ ҳам, мисол ҳам янги, (7) олдингисида йўқ, янгисида мавжуд синонимлар, (8) синонимига ишора қилиш асосида *айн.* деб изоҳланган синонимлар каби гуруҳларга бўлинди.

4. Юқоридаги тасниф натижаси шуни кўрсатадики, ўзбек тилида яратилган синонимлар луғатлари келгусида синонимларни лексик-семантик разметкалашда тўлақонли лексикографик база сифатида хизмат қилолмайди. Бу эса разметкалаш жараёнининг аниқлигига таъсир этади. Шунга қарамай, улардаги синонимларга берилган шарҳ ва тавсифдан асос манба сифатида фойдаланиш лозим.

5. Корпусга синонимлар базасини киритиш мавжуд маълумотларни матн ҳолида тақдим эта олиш; корпуснинг ҳажмига кўра имкон даражада кўпроқ ахборот бериш; бир марта яратилган корпус маълумотларидан такроран, турли масалаларни ҳал қилишда фойдаланиш имконини беради. Бунинг учун ўзбек тили синонимлар электрон базасини тузишнинг тамойилларини ишлаб чиқиш, бу борадаги муаммоларни бартараф этиш лозим. Корпусни ташкил этувчи бирликларни аниқлаш, турли лисоний параметрлар бўйича стандарт синонимлар разметкаси ишлаб чиқиш (семантик, жанрий, услубий, морфологик разметка кўзда тутилмоқда) асосий масалалардан бири.

6. Семантик теглаш муаммоларни кузатиш синонимларни разметкалашнинг куйидаги тамойиллари мавжудлигини кўрсатди: разметка схемаси/тузилишини тавсифлаш (1); лингвистик тушунчаларнинг умумқабул қилинган системасини тузиш (2); таҳлилнинг фойдаланувчига маълум бўлган схемасини тузиш (3); синонимларни разметкалаш параметрларини танлаш (4); разметканинг анъанавий ва халқаро стандартлари схемасига амал қилиш (5).

7. Ўзбек тили Миллий корпусининг семантик разметкасини ишлаб чиқиш жараёнида изоҳли луғатларни қайта кўриб чиқиш; луғатнинг лексик бирликлар сигимини кенгайтириш; семантик синфлар структурасини мукаммаллаштириш; сўз ясовчилик белгиларини қўшиш талаб қилинади. Кўпинча, синонимияни изоҳлашда маънодаги тўлиқ ёки қисман айнанлик, бири ўрнида иккинчисини қўллаш имкониятининг мавжудлиги кабилар ҳисобга олинади. Синонимларнинг абсолют/мутлақ синоним (а); идеографик/семантик синоним (б); услубий синоним (в); контекстуал/матний синоним (г); фразеологик синоним (д) каби турлари ажратилади.

8. Контекстуал (матний) синонимларни нутқий синоним номи остида алоҳида гуруҳ сифатида ажратиш лозим. Чунки синонимларни теглаш жараёнида луғатда синоним сифатида ажратилган ва луғатда ифодаланмаган синонимларни теглаш турли йўл билан амалга оширилади. Корпус учун синонимларни теглашда синоним луғати лингвистик таъминот вазифасини ўташини инобатга олсак, луғатда акс этган синонимларни теглаш бошқа йўл билан, луғатда акс этмаган, фақат маълум нутқий вазиятлардагина синонимликни юзага чиқарадиган (нутқий) синонимларни теглаш бошқа йўл билан амалга оширилади.

9. Мавжуд изоҳли луғатлар тил бирлик(синоним)ларни теглаш учун база вазифасини бажаради. Ўзбек тилининг мавжуд синонимик луғати

ҳозирги ўзбек адабий тили учун характерли бўлган синонимик қаторларни ўз ичига олган, аммо барча синонимик қаторларни қамраб ололмаган. Унга фақат лексик (лисоний) синонимларгина киритилган. Тил корпуслари эса контекстуал синонимларни ҳам қамраб олиши лозим. Демак, ҳатто лисоний синонимларни теглашда ҳам мавжуд синонимлар луғати тўла маънода манба бўлолмайди, аммо ҳозирча ундан бошқа луғатга таяна олмаймиз.

10. Ўзбек тили лисоний синонимларини теглаш учун мавжуд синонимлар луғатини қайта ишлаш лозим. Синонимлар корпуси яратилса, янги синонимлар луғати тузилади. Корпус даврий равишда тўлдирилиб борилса, электрон синонимлар луғати ҳам янгиланиб бораверади. Қайта ишланган луғат дастурчи учун дастурий таъминот вазифасини бажара олади. Синонимларга оид лексикографик асарлар мукаммаллаштирилмай туриб, синонимларнинг семантик разметкасини беҳато амалга ошириб бўлмайди.

SCIENTIFIC COUNCIL FOR AWARDING SCIENTIFIC DEGREES
PhD.30.05. 2018.Fil.70.01 AT KARSHI STATE UNIVERSITY
TERMEZ STATE UNIVERSITY

ESHMUMINOV ASKAR ALLAMURADOVICH

SYNONYM WORDS BASE OF UZBEK NATIONAL CORPUS

10.00.01 – Uzbek Language

DISSERTATION ABSTRACT FOR DOCTOR OF PHILOSOPHY (PHD)
ON PHILOLOGICAL SCIENCES

Karshi – 2019

The theme of PhD dissertation is registered by Supreme Attestation Commission at the Cabinet of Ministry of the Republic of Uzbekistan under the number №B2019.1.PhD/Phil 700.

The dissertation has been prepared at Termez State University.

The abstract of PhD dissertation is posted in three (Uzbek, English, Russian (resume)) languages on the website of "ZiyoNet" information and educational portal, www.karshidu.uz, www.ziynet.uz.

Supervisor: **Mengliev Bakhtiyor Rajabovich**
Doctor of Philological Sciences, professor

Official opponents: **Pardaev Azamat Bahronovich**
Doctor of Philological Sciences
Abjalova Manzura Abdurashetovna
Doctor of Philosophy on philology (PhD)

Leading organization: **Bukhara State University**

The dissertation defense will be held at the meeting of Academic Council under Qarshi State University, PhD, PhD, 30.05.2018.Fil.70.01, 5 XII, 2019 (Address: 180103, Qarshi Street, 17. Kuchabog Street Tel.: (0 375) 225-34-13; fax: (0375) 221-00-56; e-mail: rozidu@umail.uz). Karshi State University, the conference hall of faculty of Uzbek philology.

The dissertation could be reviewed in the information-resource center of Karshi State University (registration number 19). Address: 180103, Karshi, Kuchabag str., 17. Tel.: (0 375) 225-34-13; fax: (0375) 221-00-56; e-mail: nazidu@umail.uz).

The abstract of dissertation is distributed on: XI «22», 2019.
(Protocol at the register № 18 on XI «22», 2019.)



N.N.Shodmonov

Chairman of Scientific Council awarding
scientific degrees, Doctor of Philological sciences

G.N.Tojjeva

Secretary of Scientific Council awarding scientific
degrees, Doctor of Philological Sciences (PhD)

D.Turaev

Chairman of Scientific Seminar at the Scientific
Council awarding scientific degrees, Doctor
of Philological sciences, professor

INTRODUCTION (abstract of Doctor of philosophy (PhD) dissertation)

The actuality and necessity of the research topic. In world linguistics, scientific attitudes towards synonyms have their stages of development, mostly since the 19th century. The issue of creating an electronic database of synonyms with different and often contradictory interpretations, incorporating them into the national language corpus has been addressed only in the world's largest language corpus. In the development of corpus linguistics, many positive aspects of the humanities and technical sciences are combined. Currently, theoretical works fulfilled in this area can serve as a basis for the compilers of the national language corpus, which is emerging theoretically.

Lexicographic forms of the national corpus have been studied perfectly in world linguistics of the 21st century and up to this time, particularly, the dictionary of synonyms as well. Today, when corpus linguistics is raising up to the perfect scientific and theoretical level, these dictionaries play an essential role in the semantic content of the corpus. Intending to improve the quality of automatic translation, linguistic modelling of the language, the theory and algorithm of word lemming, as well as the electronification of the language, it has become a topical issue in world linguistics. Creating a variety of enclosures, expanding the size of the existing compounds, and developing automated text processing (tagger, parser) is a matter of linguistics.

During the years of independence, the process of turning the Uzbek language into information technology has been accelerated. One of the urgent problems facing the Uzbek linguistics is the critical task of the present time, the fact that the Uzbek language has not reached the level of the Internet language and this process is going very slowly. Stressing the importance of linguistic databases for the creation of the National Corpus of Uzbek Language is one of the crucial challenges facing Uzbek linguists, «Ensuring that information and communication technologies, in particular, the Internet have a decent place in the global information network, the creation of computer software for Uzbek language»¹. Note that in this context, corpus linguistics as a priority area of applied linguistics, which should be based on the linguistic corpus of applied research carried out today in Uzbek linguistics. The study of modern linguistic principles, such as the creation of synonyms databases for the Uzbek language corpus, the development of linguistic foundations of synonymised, is becoming increasingly important.

The results of the given dissertation will be used in realization of the tasks defined in the Decree 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, Presidential Decree PF-4947 «On the Strategy of Actions for the Further Development of the Republic of Uzbekistan» Decree issued on February 7, 2017, Decree No. PF-5850 of the President of the Republic of Uzbekistan dated October 21, 2019 «On measures to

¹Ўзбекистон Республикаси Президентининг “Ўзбек тилининг давлат тили сифатидаги нуфузи ва мавқеини тубдан ошириш чора-тадбирлари тўғрисида”ги ПФ-5850-сон Фармони. [www. http://lex.uz/docs/4561730](http://lex.uz/docs/4561730)

radically enhance the status and prestige of the Uzbek language as a state language», Decision of the President of the Republic of Uzbekistan No. PQ -2789 «About Measures for Further Improvement of the Furthering of Higher Education» dated February 17, 2017, Presidential decree PQ-4479 «About celebration of the 30th Anniversary of the Law “About the State Language”» issued on October 4, 2019.

The relevance of the research to the priority areas of science and technology development in the Republic of Uzbekistan. The study was carried out under the priority area of development of science and technology in the Republic of Uzbekistan «Formation and implementation of innovative ideas system in the social, economic, legal, cultural and moral education of the information society and democratic state».

The extent of the problem studied. In world and Uzbek linguistics, synonyms research has a long development phase. After all, the views on linguistic spirituality are reflected not only in world linguistics, in Turkology, but also the theories put forward by Alisher Navoi. Although the study of synonyms in Uzbek linguistics has found a relatively well-established theoretical basis in the studies of A. Hodjiev, Sh.Rakhmatullaev, Y.Tojiev, H.Shamsiddinov, M.Vafoeva, M.Hakimova², lexicographical work in this area still shows a great deal of work to be done.

In world linguistics, corpus linguistics became a subject of study in the 1960s. Although R.G. Piotrowski made the earliest thoughts on the corpus³, targeted research in this area was initiated by Bloomfield, Frays, and Bonders⁴ in the 40's. The development of the principles of corpus formation by N. Francis and G.Kuchera, the creation⁵ of the first English corpus by J. Sinkler was the first theoretical and practical step in corpus linguistics.

In Russian linguistics V.P. Zakharov, A.B.Kutuzov, E.V.Nedoshivina, V.V.Rikov, V.Plungyan, O.V.Kukushkina, A.A.Polikarpov, E.V.Surovtseva studied corpus types, characteristics, the social significance of the corpus, and the

²Хожиев А. Тилшунослик терминларининг изохли луғати. – Тошкент, 2002. –б.; Ё.Тожиев. Ўзбек тилида аффиксиал синонимия. – Тошкент: Университет, 1991.; Шамсиддинов Х. Сўзлар функционал-семантик синонимларининг шаклий ва маъновий табиати. – Тошкент: Университет, 1999.; Шамсиддинов Х. Ўзбек тилида сўзларнинг функционал-семантик синонимлари: Филол. фан докт. дис. автореф. – Тошкент, 1999. – 40 б.; Вафоева М. Ўзбек тилида фразеологик синонимлар ва уларнинг структурал-семантик таҳлили: филол. фан. номз. дисс. автореф. – Тошкент, 2009. – 26 б.; Ҳакимова М. Семасиология. – Тошкент, 2008.

³Кутузов А.Б. Корпусная лингвистика. – (Электрон ресурс): Лицензия Creative commons Attribution Share-Alike 3.0 Unported (Электрон ресурс) - //lab314.brsu.by/kmp-lite/kmp-video/CL/CorporeLingva.pdf.

⁴Блумфилд Л. Язык. – М.: «Прогресс», 1968. – 608 с.; Fries Ch.C. The structure of English. An introduction to the construction of English sentences. – L., 1969; Bongers H. The history and principles of Vocabulary control. – Woerden: WOCOPI, 1947.

⁵Френсис Н., Кучера Г. Вычислительный анализ современного американского варианта английского языка. – М., 1967.; Синклер Д. Предисловие к книге «Как использовать корпуса в преподавании иностранного языка»/ Д. Синклер [Электронный ресурс]. – Режим доступа: <http://www.ruscorpora.ru/corpora-info.html>, свободный.; Кутузов А.Б. Корпусная лингвистика. – (Электрон ресурс): Лицензия Creative commons Attribution Share-Alike 3.0 Unported (Электрон ресурс) - //lab314.brsu.by/kmp-lite/kmp-video/CL/ CorporeLingva.pdf; Недошивина Е.В. Программы для работы с корпусами текстов: обзор основных корпусных менеджеров. Учебно-методическое пособие. – Санкт-Петербург. – 2006. 26 с.; Рыков В.В. Курс лекций по корпусной лингвистике. URL: <http://rykov-cl.narod.ru/c.html>; Плунгян В. Зачем мы делаем Национальный корпус русского языка? «Отечественные записки» 2005, №2. http://magazines.russ.ru/oz/2005/2/2005_2_20-pr.html

principles of corpus formation. Much research has been done in Uzbek linguistics on computer linguistics, lexicographic processing of text, and linguistic analysis. These are the works of A. Pulatov, S.Muhamedov, M.Aimbetov, S.Muhamedova, S.Karimov, G.Jumanazarova, A.Babanov, D.Urunbayeva, N.Abdurahmonova, A.Norov and others⁶. Although these studies have suggested modern methods of lexicographic and linguistic studies using textual computer linguistics, they are not on the agenda. The research varied out by Sh. Khamraeva is a practical step in this filed⁷. In Uzbek linguistics, the work done in this regard can be regarded as the first step in establishing the National Corpus of the Uzbek language.

Connection of the research with the research plans of a higher education institution or research institution The dissertation was carried out within the framework of the project of research works of the Department of Uzbek linguistics at Termez State University for 2017-2021 on «Systematic study of the Uzbek language».

The aim of the research is to study the theoretical basis for the formation of a database of synonyms and to develop principles for improving the synonyms database using synonyms mapping.

Research objectives:

study of synonymy phenomena in world linguistics, turkology and Uzbek linguistics;

to study the description and interpretation of synonyms in the Uzbek dictionaries;

observation of differences in the analysis of synonyms in annotated dictionaries;

study of the understanding of lexical spiritual units in the Uzbek synonym dictionary;

to study linguistic bases of formation of synonyms database;

to identify the types of linguistic coefficients, and to check the significance and significance of semantic assimilation;

study of the methods and tools used in the creation of synonyms databases, theoretical justification for their active implementation in practice;

⁶Мухаммедов 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., Қаршиев 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. дис... автореф. - Тошкент, 2008; Мухаммедова С. Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. Методик қўлланма. - Тошкент, 2006; Ўринбоева Д.Б. Ўзбек фольклори матнларининг лингвостатистик тадқиқи. – Тошкент: «Фан», 2010; Жуманазарова Г.У. Фозил Йўлдош ўғли дostonлари тилининг лингвопоэтикаси: Фил. фан. док. дис...автореф. - Тошкент, 2017; Абдурахмонова Н.З. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти (Содда гаплар мисолида): Филол. фан. бўйича фалсафа доктори (PhD)...дис. афтореф. – Тошкент, 2018; Пулатов А. Компьютер лингвистикаси. – Тошкент: Akademnashr, 2011.; Норов А. Компьютер лингвистикаси асослари. – Қарши, 2017. – 136 б.

⁷ Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари. Филол.фан.бўйича фалсафа доктори (PhD)...дис. афтореф. – Қарши, 2018

working out the principles of creating a database of synonyms;
to describe the properties of semantic routers.

Objects of the study were linguistic synonyms and their terms.

The subject of the research is the formation of the Uzbek language synonyms, the problems of tapping synonymic units.

Methods of research. Purposes of classification, description, comparison, statistical, cognitive analysis were used to cover the research topics.

Scientific novelty of the research is:

linguistics in the formation of a synonym database is based on the fact that the principle of speech differentiation and graduonimic (linguistic) is a linguistic methodology;

it has been established that the lexeme meaning of the semantic razor is synonymous with the lexeme meaning, the meaning of the definition, and the post-migration phase;

the classical role of semantic charts in the creation of a database of meaningful words is revealed;

An algorithm for touching linguistic synonyms of the Uzbek language has been developed.

The practical implications of this study are:

the conclusions drawn from the research are a specific resource for developing principles for the creation of synonyms databases of the Uzbek national corpus;

The dissertation is based on the creation of a base of synonyms of the Uzbek language for the linguistic corpus developers.

Reliability of the results of the research is since the studied materials have helped to make conclusions based on the nature of the Uzbek language, their validity, methodological excellence, and reliance on proven sources.

The scientific and practical significance of the results of the research is determined by the development of theoretical foundations and algorithms for the creation of a synonym database for the creation of the national language of the Uzbek language.

Implementation of research results. Based on the scientific results obtained during the development of the synonym words database of the Uzbek language:

the linguistic model of Uzbek synonyms has been used to create an on-line version of the synonym database based on scientific conclusions that the lexeme meaning, scale and post-migration phase of the semantic tagging (indexing) are synonymic (Certificate No. 89-03-2113 of the Ministry of Higher and Secondary Special Education dated May 21, 2019). As a result, the quality of teaching and improving the content of “Modern Uzbek language” and “Computer linguistics” subjects has been achieved;

The Central Council of the Youth Union of Uzbekistan has been put into practice in the framework of the third initiative of the five key actions of the President of the Republic of Uzbekistan, from the scientific results of the classification of semantic arrangements in the formation of the base of meaning

words. (Protocol 04-13/3033-of the Central Council of the Youth Union of Uzbekistan). As a result, the interest of young people in learning and improving their knowledge of the Uzbek language through the Internet was achieved.

Materials reflecting the fact that the language is a linguistic methodology of linguistics - speech distinction and grading (syllabus) in the formation of synonyms databases was used in preparation the sceneries of the TV programs “Диёр янгиликлари” (Country news) and “Бир асар шарҳи” (Review of a certain literary work) (Certificate No. 01-15/263 of the teleradiocompany of the Surkhandarya region dated April 9, 2019). As a result, television and radio programs are enriched with new material, scientific evidence.

Approbation of research results. Following the results of this research were made presentations at two international and four republican scientific-practical conferences.

Publication of the study results. The main content of the thesis is contained in 1 author's certificate, nine scientific articles (3 of them are published in the journals recommended by the Higher Attestation Committee of the Republic of Uzbekistan) and abstracts.

Structure and volume of the dissertation. The dissertation consists of an introduction, three main chapters, an overview, a list of references and an appendix. The total volume of the dissertation is 131 pages.

BASIC CONTENT OF DISSERTATION

The introductory part is based on the urgency of the topic, describes the purpose, objectives, object and subject of the research, demonstrates the correspondence with the priority areas of science and technology development in the country, presents the scientific novelty and practical results, the results of which are published.

Chapter I of the dissertation, entitled «**Linguistic Basis of Synonyms in the National Corpus of Uzbek Language**», explores the phenomenon of synonymy in world linguistics, the study of synonymy phenomena in Uzbek linguistics, and the use of synonyms in Uzbek-language dictionaries.

From the ancient Greek philosophical period, philosophers Prodic, Plato, Socrates, Democritus have been described in relation to the synonymy phenomenon, as well as the study of synonymy in Chinese, Indian, and Arabic linguistics. The second stage related to the study of synonyms dates back to the 18th and 19th centuries when numerous dictionaries of synonyms were published. During this period, I. Davidov, D.P. Kalaydovich, M.V. Lomonosov, I. Fonvizin and others⁸ commented on the synonymy study by examining dictionaries and synonyms.

⁸ In compiling a dictionary of synonyms in the 18th and 19th centuries D.I. Fonvizin has contributed much. See also N. Abramov's «Словарь русских синонимов и сходных по смыслу выражений» published in 1890 (1st edition), 1915 (4th ed.), 1994 (5th ed.) – М.: Russkiye Slovari, 1994. – 502 p.; A.I. Galich's Dictionary «Словарь русских синоним или сословов, составленный редакцией нравственных сочинений» published in 1840 (226 lexicographical articles), F. Kalaydovich «Опыт словаря русских синонимов» published in 1818 (77 group of synonyms) in the for of book (Part I). It is necessary to acknowledge the works of «Опыт российского

It is noted that S.G. Berejan argued that until the XXIst century, synonyms were simply explored as enriching speech. In the second half of the XIXth and mid-XXth centuries, the study of synonyms became relevant. The recognition of semasiology as science has prompted scholars to take this approach seriously. We can say that the twentieth century has become «the golden age of synonymy learning»⁹. It is worth noting that the development of science in the 50-70s of the 20th century gave rise to synonyms in linguistics.

Studies in the synonymy phenomenon in Uzbek science dates back to the 1960s. The study of synonyms in Uzbek linguistics has its own historical evolution, which can be conditionally divided into the following stages:

- Stage I: Up to the XXth Century;
- Stage II: 1920-1930s;
- Stage III 1950-1960s;
- Stage IV: 1960-1980s;
- Stage V: 1980-2000s;
- Stage VI: 2000 and beyond.

In the second half of the twentieth century, Uzbek linguistics developed an attitude towards synonyms in words (syntactic constructions), and studies in this area did not slow down even in the XXIst century Uzbek linguistics.

The chapter deals with the definition and description of synonyms in the explanatory dictionaries of the Uzbek language. The comments on synonyms in the two editions of the O'Leary are conventionally divided into the following groups:

- 1) synonyms with clear explanations and examples;
- 2) the commentary is completely new, just the examples;
- 3) the comments are partly new, the examples exactly;
- 4) the comments are partly new, the examples partially changed;
- 5) the comments are just the examples;
- 6) Both the comments and the examples are new;
- 7) not the previous one, but the new one;
- 8) interpreting as «exactly» based on the synonym (previous and next)

The above groups are commented on in the two editions of the Explanatory dictionary of Uzbek language.

словника» published by L.I. Fonvizin in 1783 (32 groups synonyms) and the journal Собеседник Любителей Российского Слова» (part I, IV, X).

⁹ Шумилова А.А. Истоки исследования синонимии в философских и лингвистических учениях. http://www.rusnauka.com/4_SND_2013/Philologia/3_127613.doc.htm; Горбачевич К.С. Русский синонимический словарь. [Ин-т лингвист. исслед. РАН]. – СПб: ИЛИ РАН, 1996. – 510 с.; Зимин В. И., Алекторова Л. П., Ким О. М., Колесников Н. П., Шанский В. Н. Учебный словарь синонимов русского языка (2800 синонимических рядов) – М., 1994; Жуков В. П., Сидоренко М. И., Шкляров В. Т. Словарь фразеологических синонимов русского языка. – М., 1987; Евгеньева А. П. Словарь синонимов. Справочное пособие. – М., 1975; Евгеньева А. П. Словарь синонимов русского языка. – Л.: Наука, 1970 – 1971. – Т. 1–2.; 1968 г. (1-изд, 1969 г. (2-и.), 1975 г. (4-и.), 1986 г. (5-и.), 1989 г. (6-и.), 1993 г. (7-и.), 2001 г. (11-и.) Александрова З. Е. «Словарь синонимов русского языка» – М.: Рус. яз. 2001. – 567 с. (11 000 синонимических рядов); Ключева В. Н. Краткий словарь синонимов русского языка. – М., 1956 г., 1961 г. (2-изд.); Павлов-Шишкин В. Д., Стефановский П. А. Учебный словарь синонимов русского литературного языка. – М., 1930 г., 1931 г. (2-изд.).

At the beginning of the next chapter, we would like to make the following table ¹⁰ describing synonyms (Table 1). After all, such classifications are particularly important when it comes to semantic circuitry of corpus.

I .1. Table

Example of a table classification of synonyms

Synonymic paradigm	The lexical meaning of synonyms	Semantic structure of synonyms				Types ofsynonyms		Methodological characterization of synonyms
		Integral semesters	Differential semes			semantic	struktural	
			Quantifying semesters	Emotional expressive connotations	Methodical semes			
Енгил (доминанта)	‘тушуниш’, ‘ўрганиш’, ‘ўзлашти-риш’ ‘ва’ ‘бажариш’ да ‘қийин-чилик’ ‘туғдир’ ‘майдиган.’				нейтр.		Бир асосли	Барча услублар учун
Қийин эмас	‘етарли’ ‘даража’ да ‘енгил’		‘етарли’		нейтр.	идеогра-фик	Ҳар хил асосли	Барча услуб-лар учун
Оддий	‘алоҳида ақл, билим ва кўп меҳ-нат талаб қилмайдиган’		‘алоҳида’		нейтр.			Барча услуб-лар учун
Элемен-тар	‘жуда оддий’	«	‘жуда’		нейтр.	«	«	Илмий
Мурак-каб эмас	‘заррача’ ‘машаққат’, ‘қийинчи-лик’нинг ‘мавжуд’ ‘эмас’лиги	‘Жуда’ ‘осон’	‘заррача’		Сўзла-шув	Идео-график-услубий	«	Расмий, илмий, публи-цистик

¹⁰<http://book-science.ru/humanities/linglang/lexicology/leksiko-semanticheskaja-paradigma-chast-3.html>

Chapter II, entitled «**Theoretical Problems of Dispersion of Synonyms in the National Corpus of the Uzbek Language**», explores the evolution of corpus linguistics, the terminological apparatus, and the peculiarities of synonymic intelligence.

Corpus linguistics is an independent division of computer linguistics, developing the principles of creating and using linguistic corpus (text corpus) using computer technology¹¹. It has been noted that in the mid-1990s it began to develop as a separate area of linguistics. Corpus linguistics as a science textual analysis appeared in Western Europe and the USA in the late 1960s. In the mid-1980s, as a result of the rapid development of computer and information and communication technologies, corpus projects of various languages and sizes were created. In particular, the Russian National Corpus¹², inspired by the British National Corpus, is also known for incorporating 163 million (up to 500 million today) words into the historical and modern strata¹³. This body consists of the following sections:

- 1) The main building (prose texts of the 18th and 21st centuries);
- 2) syntactic body (the morphological and syntactic device for each sentence);
- 3) Newspaper (1990-2000 media articles);
- 4) parallel cases (translation of the corresponding word from Russian into Russian and from Russian into foreign language);
- 5) the body of dialectal texts;
- 6) corpus of poetic texts;
- 7) Russian language learning corpus (based on simple school curriculum);
- 8) spoken language;
- 9) multimedia material (1930-2000 film collection);
- 10) Russian accent history (texts on accent history)¹⁴.

Scientists say that by the 1990s, the number of the corpus had been more than 600. The date of their creation is divided into years as follows.

The chapter also reveals the principle of «interaction» between the computer and the lens; it also appears in the concept of computer - software - knowledge - linguistics/language¹⁵. The basic notion of corpus linguistics is the linguistic corpus. The texts to the corpus are chosen from the point of view of their relevance

¹¹ Захаров В.П. Корпусная лингвистика. Учебно-методическое пособие. – Санкт-Петербург, 2005. – 48 с.

¹² Британия Миллий корпуси (British National Corpus – BNC)га тенг равишда, инглиз тили корпусларидан яна: инглиз тили Халқаро корпуси (International Corpus of English – ICE), инглиз тили лингвистик Банки (Bank of English), Ҳозирги замон Америкаинглиз тили Корпуси (Corpus of Contemporary American English – COCA) ва б. ларни ҳам келтириб ўтиш мумкин.

¹³ Майорова А.Д. Корпусная лингвистика: исторический и лингводидактический аспекты/ Педагогические Науки. Выпуск № 5(59) Май 2017

¹⁴ Майорова А.Д. Корпусная лингвистика: исторический и лингводидактический аспекты/ Педагогические Науки. Выпуск № 5(59) Май 2017.

¹⁵ Шемякин Ю.И. Начала компьютерной лингвистики: Учеб. пособие. – М.: Изд-во МГОУ, А/О "Росвузнаука", 1992.

to the problem, that is, thematic links that are important to linguistic researchers. Problem area - Thematic scope consists of two aspects, including language and speech texts. While the language aspect is the most researched area, the speech aspect is composed of contexts reflecting the context of the language aspect. And the problem of sorting text by volume is one of the most important issues in building a body. This aspect ensures that the body is perfect. Bodies of text sorting are subdivided into the following types:

- Balanced enclosures;
- monitor cases.

There is a need to implement additional linguistic and metallurgical information (annotation) - text units. In modern buildings, besides metarasyymetry (textual extrovalvistical data), there is a linguistic definition of morphological, syntactic, phonetic, and so on. This type of jacket is used. The notion of a text corpus contains several terms. In particular, the case manager is a special search engine that includes software that specializes in searching for relevant information within the body, providing statistical information on user-related information, and transferring information in a way that is convenient to consumers. In turn, the search on the body generates a list of all the records of the linguistic unit searched for in the context of the specific concordance. Although scholars still classify corpus linguistics as an area with the little theoretical basis, it is hoped that in the future researchers will approach the corpus not only as an information resource but also as a methodological basis and area of scientific and theoretical significance.

In this case, the differences between traditional and corpus linguistics can be seen in the following (Table 2).

I. Table 2.

Differences between traditional and corpus linguistics

Features	Traditional linguistics	Corpus linguistics
Direction	Language / Language Study	Speech Research
With respect to the object of study	Explanation of speech events through theory	Text corpus
Methods	Qualitative	Quantitative
Attitude to the text	Abstract	Physical Presence

Bodies are also classified according to their designation. There are also terms used to denote the division: indexed, indexed, annotated, annotated, tagged, untagged. Morphological, syntactic, semantic, prosodic etc. tags are attached according to the linguistic symbol that represents the type¹⁶.

¹⁶ Захаров В.П., Богданова С.Ю. Корпусная лингвистика: учебник для студентов гуманитарных вузов. – Иркутск: ИГЛУ, 2011. – С. 25.

There are the principles of linguistic tagging:

- 1) description (justification) of the scheme/structure of the tag;
- 2) a common system of linguistic concepts;
- 3) scheme of analysis which is familiar to user;
- 4) asymmetric input of parameters;
- 5) traditional schematic tag;
- 6) adherence to international standards¹⁷.

Semantic tag (Russ., разметка) is a set of meanings, comments, expressions that refer to a particular semantic category or smaller semantic group (LSG, semantic field, and group) of a word or phrase¹⁸. The semantic labels of the corpus include such words as the meaning (s) of the word, the creation of a set of comments related to the word survival, synonyms, categorization of the word, its thematic group, LSG, semantic domain, derivative characterization, and meaning.

Section III, entitled «**Synonym words and their software in the national language of the Uzbek language**», provides an overview of synonym words, its software, and algorithms for tagging.

Although synonymous lexemes have several common semesters, they have also been shown to have semantic differences as follows:

- 1) the difference in the connotative meaning;
- 2) the difference in the denotative meaning;
- 3) methodological difference.

Based on the explanations and descriptions, it is important to come to a clear understanding of the types of synonyms. Hence, the absolute / absolute synonym of synonyms (a); ideographical / semantic synonym (b); methodical synonym (v); contextual / logical synonym (g); we have distinguished such phrases as synonym (d). Existing annotated dictionaries serve as a base for tapping language units (synonyms). To do this, how should the existing dictionaries be processed, and how can we use them automatically when the dictionaries appear? We will find answers to such questions in this section of the topic. The existing synonymic dictionary of the Uzbek language contains synonyms that are characteristic of modern Uzbek literary language. But it does not cover all synonyms. It includes only lexical synonyms. The language corpus should also include contextual synonyms. The most important thing is to identify a synonym, a dominant, and to tag them:

1) Synonymic text is highlighted in the electronic text of the dictionary: {melody, song}, {enough, enough, enough, is, has, bass). This threshold specifies members of a synonym for programming;

¹⁷ Захаров В.П., Богданова С.Ю. Корпусная лингвистика: учебник для студентов гуманитарных вузов. – Иркутск: ИГЛУ, 2011. – С. 45.

¹⁸ Захаров В.П., Богданова С.Ю. Корпусная лингвистика: учебник для студентов гуманитарных вузов. – Иркутск: ИГЛУ, 2011. – 161 с. – С. 48-49.

2) In the electronic text of the dictionary, synonymic dominance is defined and separated by the main lemma symbol. For example *кутмоқ*, *пойламоқ*. In this synonymic line, the expectation is dominant, which means we define it as the main lemma: <<кутмоқ>>, <пойламоқ>, <<кулги>>, <ханда>. In general, it takes the form of {<<кутмоқ>>, <пойламоқ>}, {<<кулги>>, <ханда>}.

The chapter outlines the principles by which linguistic data entered into an electronic database is sorted out.

The selection principles determine the linguistic value of the synonyms database, and the effectiveness of related research is also determined. The following shall be taken into consideration directly:

- 1) to determine which layers of units are included in the synonyms database;
- 2) defining the size of the database;
- 3) the type, format, genre, etc. setting the property;
- 4) sorting of synonymic materials according to which language of the phenomenon they belong to;
- 5) Pay attention to the frequency of the unit being introduced;
- 6) Determine the frequency of data entered into the database.

It is well-known that the practice of sorting corpus materials was initially enhanced by the research of R.G. Piotrovsky between 1965 and 1980. The frequency dictionary of L.N. Zazorina in this regard is also not ignored. As you can see, the sampled collection, its statistical indices, periodic cross-section also play an important role in the creation of our research object - a synonym database.

The main unit of corpus texts may be words, phrases, cores (bases, lemmas), and sentences. Of course, the size of the database is determined by the purpose of its creators, which is usually determined by the amount of material selected in place of the subject of the study¹⁹.

As we have already mentioned, the synonyms that are being inserted into the body imply, first of all, the need to regulate synonyms. However, the type of lexicographic works is determined by their encyclopedic and linguistic features. Consequently, the encyclopedia provides information about the subject, the concept, the event, and the linguistic dictionary provides information about the words. After all, a linguistic dictionary is divided into one-language, bilingual, and multilingual types, and a multilingual dictionary is usually created for translation purposes. In a linguistic dictionary, the meaning of a word is explained in that language, whereas in a bilingual dictionary the meaning of a particular word is explained by comparing it to two languages.

An annotated dictionary, which is a type of linguistic dictionary, is classified as descriptive and normative. The Descriptive Dictionary is characterized by a

¹⁹ Захаров В.П., Богданова С.Ю. 3-38 Корпусная лингвистика: учебник для студентов гуманитарных вузов. – Иркутск: ИГЛУ, 2011. – 161 с.

complete interpretation of the word in the relevant field, which can even be interpreted as relevant. An example of this is an explanatory dictionary of the Uzbek language (1981, 2006 edition). In the normative dictionary, however, deviations from the norm are excluded.

The same information needs to be clarified in the synonym database. That is, the literary language, the dialect, the limited vocabulary (jargons) have their own norms. In this case, each unit is identified by the Latin symbol Vs. In particular, the *она* dominant takes on its synonymic word (*ая* (Ferghana dialect), *ача* (Tajik), *аба* (Surkhan dialect), *она* (Samarkand / Tashkent dialect), *эна* (Surkhan, Kashkadarya, Samarkand dialect). Regardless of the dictionary, we believe that all of these units should be synonymized.

Thesaurus is a term that includes the relevant knowledge, activity, field, understanding, interpretation and terminology. Thesaurus is formed by industry links.

In the ideological dictionary, words are arranged according to lexical meaning, not alphabetically; If a user has information about a particular word in an alphabetical dictionary, using an ideographic dictionary the user will have the knowledge of the relevant meaning, such as how to express a particular meaning/meaning.

In general, dictionaries are etymological, vocabulary, dialectal, orthographic, poetic vocabulary, frequency, reverse, etc. types. Of these, we can partially focus on the frequency dictionary. This is because most linguistic units that need to be included in the synonyms database must first be sorted by consumption. We have already talked about the principles of introducing linguistic units into synonyms. Consequently, it is possible to achieve a perfect synonymic structure, proper arrangement of the paradigm, and proper classification of the text, thanks to the perfect compilation of the frequency dictionary. Of course, lemmatization and laryngeal symmetry should not be ignored in this case. After all, one of the main reasons for referring to a frequency dictionary when selecting synonyms is to correctly identify the dominant, avoiding errors in the appropriate synonym. Frequency dictionaries are 1) limited; 2) it is important to take into account the non-lemma type.

In many cases, spiritual unity comes from several members. With the high sequence of the device, collocation²⁰ occurs²¹. Naturally, in the collocation, the total value of the device is not equal to the independent meanings that the components carry (*мехр булоғи* ≠ *мехр* + *булоқ*). These include idioms (*илон ёғини ялаган*, *қўли очик*), toponyms, and separately derived onomastics, making the semantic excision procedure a bit more complicated. It is necessary to consider

²⁰ коллокат (collocates) - when the analysed word comes together with the word's component which is bound with itself.

²¹ Боярский К. К. Введение в компьютерную лингвистику. Учебное пособие. – СПб: НИУ ИТМО, 2013. – 72 с.

the possibility of clarifying the meaning of the words in the present situation, due to the concordance used in the body. After all, the possibility of a number of concordance programs will address the situation. In this regard, it is necessary to briefly dwell on the need to create a new category of colloquial collocations that is important to corpus linguistics. Such a dictionary is devoted to the explanation of static compounds that explain the phrase and other types of static devices. This type of dictionary is also based on the principle of frequency determination. The compounds that form the object of the colloquial dictionary are studied in the form of bigramma²². That is, conclusions about the associative nature of the relevant word are made according to the extent of its meaning. This is illustrated by the following table (III. Table 4):

Component	Lemma	Component	Meaning	Denotat	Synonym
<i>Беш</i>	<i>Юз</i>		the amount of mark	Amount	
	<i>Юз</i>	бора	the amount of mark	Amount	
<i>Гул</i>	<i>Юз</i>		quality (human trait)	гўзал	+
<i>Анор</i>	<i>Юз</i>	+ли	quality (human trait)	Гўзал/истарали	+
<i>Ширмой</i>	<i>Юз</i>	+ли	quality (human trait)	Гўзал/истарали	+

The chapter analyzes the working principles of synonymic word database program of the Uzbek language and introduces linguistic information in them. The following sequence of words is included in the linguistic database:

The word is entered
The category of the word
Meaning
Etymology
Creation
Composition
Style
Structure
Russian alternative
Alternative in English
Range of degrees
According to the positive
By the amount of use

²² Linear structure, taking into account the left and right components associated with the analyzed word.

Homonym

Paronym

Antonym

Usage (example)

Most linguistic units that need to be included in the synonyms database must first be sorted by consumption. Consequently, it is possible to achieve a perfect synonymic structure, proper arrangement of the text, and proper classification of the text, thanks to the literacy dictionary of the frequency dictionary. Of course, lemmatization and laryngeal symmetry should not be ignored in this case. The role of the frequency dictionary in the formation of a synonym database allows determining the scope, methodological nature of the dominant or paradigm participant.

These approaches and principles eventually lead to the need for the practice of refining and filling the dictionary of Uzbek synonyms. This is because, without the perfection of the lexicographic works on synonyms, the semantic assimilation of synonyms cannot be done without fail. In our present research, we have tried to make use of the available materials and to elucidate the theoretical and practical basis for creating a database of Uzbek synonyms. However, the “Uzbek Synonyms Database”, developed and presented today by us, will be further enhanced and enriched, which is also required for a wide range of users and for linguistic research.

CONCLUSION

1. The first phase of the synonymy study is of the antiquity, the second phase is in the 17th and 19th centuries. During this time many dictionaries of synonyms were created. By the 20th century, semantic, structural, and pragmatic approaches to the study of synonyms were observed. Until the 21st century, synonyms have been explored as enriching speech, whereas in the second half of the 19th and mid-20th centuries, the study of synonyms became relevant. The 20th century has become a golden age in the study of synonymy. The 21st century is recognized as a rebirth of synonymy. During this period, synonyms began to be investigated in cognitive, psychological, and corpus aspects.

2. Hazrat Mahmud Kashgari and Alisher Navoi are invaluable in studying synonyms in Turkology. Synonyms are the core of stylistics, its soul and blood, as in the past. Studies of synonymy in Uzbek linguistics are mainly in the '60s of the 20th century. In the 1980s, three types of synonymy consisted of lexical meaning, phraseological meaning, lexical-phraseological spirituality. The introduction of structural analysis into Uzbek linguistics contributed to the development of synonyms. Since the 70s of the 20th century, attention has been drawn to both general and private aspects, deviating from the «same meaning» principle.

3. Different publications of the Uzbek-language dictionary, “Russian-Uzbek explanatory dictionary of literary terms”, “Explanatory dictionary of linguistic terms”, 2006, review of the synonyms commentary showed that different principles were used to interpret synonyms. In the two editions of O’Leary, the comments on the synonyms are conventionally (1) synonyms and examples, (2) the comments are completely new, (3) the comments are partly new, (4) the comments are new, the examples are partly new, 5) the comments are clear, examples are new, (6) both comments are new, and (7) are not the previous ones, but the synonyms that exist in the new ones are based on the reference to the synonym (8). divided into groups such as synonyms that are interpreted as.

4. The result of the above classification shows that dictionaries of synonyms created in Uzbek language will no longer serve as a full lexicographic database for lexical-semantic assimilation. This affects the accuracy of the process. Nevertheless, the commentary and description given in their synonyms should be used as the basis.

5. Entering a synonym database into the body will be able to present existing data in the text; provide as much information as possible with the size of the body; allows you to use the once-built corpus data to solve multiple issues. For this purpose, it is necessary to develop principles of creation of the electronic database of the Uzbek language synonyms and to solve these problems. One of the main tasks is to identify corpuscular units, to develop standard synonyms coefficients for various linguistic parameters (semantic, genre, methodological, morphological).

6. Observing semantic tagging problems shows the following principles for synonymizing: the schematic/layout description (1); (2) establishing a common system of linguistic concepts; use of analysis

7. Revision of Explanatory Glossaries in the Development of the Semantic Characteristics of the National Corpus of Uzbek Language; expand the vocabulary of lexical units; perfection of structure of semantic classes; you will be required to add words. Often, when interpreting synonyms, the meaning is entirely or partially identical, such as the possibility of using one or the other. Absolute / absolute synonym of synonyms (a); ideographical / semantic synonym (b); methodical synonym (v); contextual / logical synonym (g); such as phraseological synonyms (d).

8. Contextual (textual) synonyms should be separated into a separate group under the name of a speech synonym. Because when you use synonyms, synonyms that are synonymous with a dictionary and not expressed in a dictionary are done in different ways. Given that synonyms are synonymous with linguistic support when tapping synonyms for corpus, tagging synonyms displayed in the dictionary is another way, and synonyms that do not appear in the vocabulary, but only synonyms that appear in certain speech situations.

9. Existing annotated dictionaries serve as a base for tapping language units (synonyms). The existing synonymic dictionary of the Uzbek language contains synonyms that are characteristic of the modern Uzbek literary language but cannot contain all synonyms. It includes only lexical synonyms. The language corpus should also include contextual synonyms. Thus, even when tapping on linguistic synonyms, the dictionary of existing synonyms may not be a complete source, but so far we cannot rely on any other dictionary.

10. To touch the linguistic synonyms of the Uzbek language, it is necessary to revise the existing dictionary of synonyms. When synonyms are created, a new synonym dictionary is created. When the body is filled up periodically, the dictionary of electronic synonyms will be updated. The revised dictionary can serve as a programmer for the programmer. Without perfecting the lexicographical works of synonyms, the semantic mapping of synonyms cannot be successfully performed

**НАУЧНЫЙ СОВЕТ PhD.30.05.2018.Fil.70.01
ПО ПРИСУЖДЕНИЮ УЧЕНЫХ СТЕПЕНЕЙ ПРИ
КАРШИНСКОМ ГОСУДАРСТВЕННОМ УНИВЕРСИТЕТЕ**

ТЕРМЕЗСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ЭШМУМИНОВ АСҚАР АЛЛАМУРАДОВИЧ

**БАЗА СИНОНИМИЧНЫХ СЛОВ НАЦИОНАЛЬНОГО КОРПУСА
УЗБЕКСКОГО ЯЗЫКА**

10.00.01 – Узбекский язык

**АВТОРЕФЕРАТ ДИССЕРТАЦИИ ДОКТОРА ФИЛОСОФИИ (PhD)
ПО ФИЛОЛОГИЧЕСКИМ НАУКАМ**

Карши – 2019

Тема диссертации доктора философии (PhD) по филологическим наукам зарегистрирована в Высшей аттестационной комиссии при Кабинете Министров Республики Узбекистан за № В2019.1.PhD/Fil930.

Диссертация выполнена в Каршинском государственном университете.

Автореферат диссертации размещен на трех языках (узбекский, английский, русский (резюме)) в веб-странице Научного совета karshi.uz и в Информационном образовательном портале «ZiyoNet»

Научный руководитель:

Менглиев Бахтиёр Ражабович
доктор филологических наук, профессор

Официальные оппоненты:

Пардаев Азамат Бахранович
доктор филологических наук, профессор

Абжалова Манзура Абдурашатовна
доктора философии (PhD) по филологическим наукам

Ведущее учреждение:

Бухарский государственный университет

С диссертацией можно ознакомиться в Информационно-ресурсном центре Каршинского государственного университета (зарегистрирован под номером 29). (Адрес: 180103, город Карши, улица Кучабог, 17. Тел.: (0375) 225-34-13; факс: (0375) 221-00-56; e-mail: qarshidu@umail.uz). Каршинский государственный университет, зал заседаний факультета узбекской филологии.

Защита диссертации состоится «5.12» 2019 года в 10⁰⁰ часа на заседании Научного совета PhD.30.05.2018.Fil.70.01 при Каршинском государственном университете (Адрес: 180103, город Карши, улица Кучабог, 17. Тел.: (0375) 225-34-13; факс: (0375) 221-00-56; e-mail: qarshidu@umail.uz).

Автореферат диссертации разослан «22.11» 2019 года.
(протокол реестра № 18 от 22.11 2019 года).



Н.Н.Шодмонов

председатель Научного совета по присуждению научных степеней, доктор филологических наук

Г.Н.Тожиева

ученый секретарь Научного совета по присуждению ученых степеней, доктор филологических наук (PhD)

Д.Тураев

председатель научного семинара при Научном совете по присуждению ученых степеней, доктор филологических наук, профессор

ВВЕДЕНИЕ (аннотация диссертации доктора философии (PhD))

Актуальность и необходимость темы исследования. Научное отношение к синонимам в мировой лингвистике имеет свои стадии развития, в основном начиная с XIX века. Проблема создания электронной базы данных синонимов с различными и часто противоречивыми интерпретациями и включения их в корпус национальных языков решалась в рамках крупнейшем в мире языковом корпусе. В развитии корпусного языкознания сочетаются многие положительные аспекты гуманитарных и технических наук. В настоящее время теоретическая работа в этой области является основой для создания национального языкового корпуса.

В лексикографии мировой лингвистики начиная с XXI века и ранее прекрасно развиты лексикографические формы национального корпуса, в том числе словарь синонимов. Сегодня, когда корпусная лингвистика находится на научно-теоретическом уровне, эти словари играют важную роль в семантическом содержании корпуса. С целью улучшения качества автоматического перевода, лингвистического моделирования языка, теории и алгоритма леммы слова, а также электронизации языка для расширения доступа к многовековому национально-культурному наследию она стала актуальной проблемой мировой лингвистики. Создание различных корпусов, расширение размера существующих корпусов и разработка автоматизированной обработки текста (тэгеров, анализаторов) - это вопрос корпусной и компьютерной лингвистики.

Одной из актуальных проблем, стоящих перед узбекской лингвистикой, является актуальная задача современности - тот факт, что узбекский язык не достиг уровня Интернет-языка и работает очень плохо. Изучение современных лингвистических принципов, таких как создание баз синонимов для национального корпуса узбекского языка, разработка лингвистических основ синонимизации.

Соответствие исследования приоритетным направлениям развития науки и технологии республики. Диссертационное исследование выполнено в соответствии с приоритетными направлениями развития науки Республики Узбекистан – I. «Духовно-нравственное и культурное развитие демократического и правового общества, формирование инновационной экономики».

Целью исследования является изучение теоретических основ формирования баз данных синонимов и разработка принципов улучшения базы данных синонимов при помощи синонимической разметки.

Задачи исследования:

изучение феномена синонимии в мировой лингвистике, тюркологии и узбекской лингвистике;

изучить описание и толкование синонимов в узбекских словарях;

наблюдение за различиями в синонимах комментариев в аннотированных словарях;

изучение толкования лексических духовных единиц в словаре синонимов узбекского языка;

исследование лингвистических основ формирования базы синонимов;
определить типы лингвистических коэффициентов и проверить значимость и важность семантики;

изучение методов и средств, используемых при создании баз синонимов, теоретическое обоснование их активной реализации на практике;

разработка принципов создания базы данных синонимов;

описать свойства семантических роутеров.

Объектами исследования являются лингвистические синонимы и их разметка.

Предмет исследования - формирование баз синонимов узбекского языка, проблемы выявления синонимических единиц.

Методы исследования. Методы классификации, описания, сравнения, статистического, когнитивного анализа были использованы для охвата темы исследования.

Апробация результатов исследования. Результаты исследования в виде лекции и тезисов были апробированы на 6, в том числе 2 международных и 4 республиканских конференциях.

Опубликование результатов исследования. По теме диссертация было опубликовано всего 10 трудов. Из них 1 авторское свидетельство, 9 научных статей в изданиях, предложенных Высшей аттестационной комиссией Республики Узбекистан для публикации докторских диссертаций, из которых 3 в республиканских журналах.

ЭЪЛОН ҚИЛИНГАН ИШЛАР РЎЙХАТИ
СПИСОК ОПУБЛИКОВАННЫХ РАБОТ
LIST OF PUBLISHED WORKS

I бўлим (I часть; I part)

1. 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).
2. Эшмуминов А. Ўзбек тили миллий корпусининг синонимик базасини яратиш меъёрлари // Сўз санаъти, – Тошкент, 2019. №1. – Б. 29-34 (10.00.00 №31).
3. Эшмуминов А. Ўзбек миллий корпусини яратиш муаммолари // Илм сарчашмалари, – Урганч, 2019. №6. –Б. 117-121, (10.00.00 №3).
4. Эшмуминов А.А. Корпусда синоним сўзларни теглашнинг дастурий таъминоти масалалари // Konferencji Miedzynarodowej Naukowo-Praktycznej (on-line): Science, Research, Development #17. – Belgrade (Serbia), 2019, –Б. 21-25.
5. Эшмуминов А. Ўзбек миллий корпусини яратишда аждодлар тажрибаси. / «Адабий алоқалар ва маданиятларнинг ўзаро таъсири» Халқаро илмий-амалий анжуман материаллари. – Термиз, 2019. –Б. 287-289.
6. Эшмуминов А. Korpus lingvistikasi xususida // «Ўзбек тилшунослигининг долзарб масалалари» VII Республика илмий-амалий анжуман материаллари. – Тошкент, 2015. –Б. 199-200.
7. Эшмуминов А. Ўзбек тилининг изоҳли луғатларида синонимлар тавсифи ва талқини / «Lisoniy meros va ona tili taraqqiyoti» Республика илмий-амалий анжуман материаллари. – Тошкент, 2019. – Б. 92-94.

II бўлим (II часть; II part)

8. Эшмуминов А. Ўзбек тилшунослигида тил корпусларининг тузилиши зарурат ва эҳтиёж ҳақида // «Давр, Замон, Шахс» Республика илмий-амалий анжуман материаллари. – Термиз, 2017. – Б. 713-714.
9. Eshmuminov A. O'zbek tilining milliy korpusini tuzish - o'zbek tilshunosligining dolarb vazifalaridan biri // Magistrantlarning ilmiy xabarlar materiallari. – Termiz, 2015. – B. 29-30
10. Mengliyev B., Eshmuminov A., Shukurov A., Maxmaraimova Sh. "O'zbek tili milliy korpusining sinonim so'zlar bazasi" dasturining on-layn versiyasi uchun mualliflik guvohnomasi. – Toshkent, 04.04.2019. №DGU 06200.

