

**ТЕРМИЗ ДАВЛАТ УНИВЕРСИТЕТИ  
ҲУЗУРИДАГИ ИЛМИЙ ДАРАЖА БЕРУВЧИ  
PhD.03/31.03.2021.Fil.78.04 РАҚАМЛИ ИЛМИЙ КЕНГАШ**

---

**ТЕРМИЗ ДАВЛАТ УНИВЕРСИТЕТИ**

**БЕГМАТОВА ГУЛЧЕХРА ХОЛМУРОТОВНА**

**ЎЗБЕК МИЛЛИЙ КОРПУСИДА ИДИОМАЛАР БАЗАСИНИ ЯРАТИШ**

**10.00.01 – Ўзбек тили**

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

**Термиз – 2021**

**Фалсафа доктори (PhD) диссертацияси автореферати мундарижаси**

**Оглавление автореферата диссертации доктора философии (PhD) по  
филологическим наукам**

**Contents of dissertation abstract of doctor of Philosophy (PhD) in  
philological sciences**

**Бегматова Гулчехра Холмуродовна**

Ўзбек миллий корпусида идиомалар базасини яратиш ..... 3

**Бегматова Гулчехра Холмуродовна**

Создание базы идиом в национальном корпусе узбекского языка..... 23

**Begmatova Gulchekhra Xolmurotovna**

Creating a base in the National Corpus of the Uzbek language ..... 43

**Эълон қилинган ишлар рўйхати**

Список опубликованных работ

List of published works..... 47

**ТЕРМИЗ ДАВЛАТ УНИВЕРСИТЕТИ**  
**ҲУЗУРИДАГИ ИЛМИЙ ДАРАЖА БЕРУВЧИ**  
**PhD.03/31.03.2021.Fil.78.04 РАҚАМЛИ ИЛМИЙ КЕНГАШ**

---

**ТЕРМИЗ ДАВЛАТ УНИВЕРСИТЕТИ**

**БЕГМАТОВА ГУЛЧЕХРА ХОЛМУРОТОВНА**

**ЎЗБЕК МИЛЛИЙ КОРПУСИДА ИДИОМАЛАР БАЗАСИНИ ЯРАТИШ**

**10.00.01 – Ўзбек тили**

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

**Термиз – 2021**

Фалсафа доктори (PhD) диссертацияси мавзуси Ўзбекистон Республикаси Вазирлар Маҳкамаси ҳузуридаги Олий аттестация комиссиясида B2019.4.PhD/Fil1026 рақам билан рўйхатга олинган.

Диссертация Термиз давлат университетиде бажарилган.

Диссертация автореферати уч тилда (Ўзбек, рус, инглиз (резюме)) Термиз давлат университети веб-саҳифасининг [www.tersu.uz](http://www.tersu.uz) ҳамда "Ziyounet" ахборот-таълим порталининг [www.ziyounet.uz](http://www.ziyounet.uz) манзилларига жойлаштирилган.

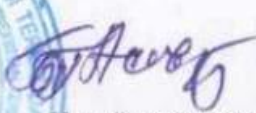
<b>Илмий раҳбар:</b>	<b>Махмаранмова Шохисга Тухташевна</b> филология фанлари доктори, доцент
<b>Расмий оппонентлар:</b>	<b>Каримов Сулол Амирович</b> филология фанлари доктори, профессор <b>Маматов Абди Эшонкулович</b> филология фанлари доктори, профессор
<b>Етақчи ташкилот:</b>	<b>Алишер Навоий номидаги Тошкент давлат Ўзбек тили ва адабиёти университети</b>


Диссертация ҳимояси Термиз давлат университети ҳузуридаги илмий даражалар берувчи PhD.03/31.03.2021.Fil.78.04 рақамли Илмий кенгашнинг 2021 йил «26» июль соат 10.00 даги мажлисида бўлиб ўтади. (Манзил: 190111, Термиз ш., Баркамол авлод кўчаси, 43-уй. Тел: (876) 221-74-55; факс: (876) 221-71-17; e-mail: termizdu@mail.uz).


Диссертация билан Термиз давлат университети Ахборот-ресурс марказида танишиш мумкин (1 рақам билан рўйхатга олинган). Манзил: 190111, Термиз ш., Баркамол авлод кўчаси, 43-уй. Тел: (876) 221-74-55.

Диссертация автореферати 2021 йил «15» июль куни тарқатилди.  
(2021 йил «15» июль даги 1 рақамли реестр баённомаси).



  
**Б. Умуркулов**  
Илмий даражалар берувчи илмий кенгаш раиси, ф.ф.д., доцент

  
**А. Эшмуминов**  
Илмий даражалар берувчи илмий кенгаш котиби, ф.ф.д.

  
**Ш. Ёкубов**  
Илмий даражалар берувчи илмий кенгаш қошидаги илмий семинар раиси ўринбосари, ф.ф.н., доцент

## КИРИШ (фалсафа доктори (PhD) диссертацияси аннотацияси)

**Тадқиқот мавзусининг долзарблиги ва зарурати.** Жаҳон фразеографик амалиётида идиоматик бирликларни таснифлаш муаммоси марказий масалалардан саналади. Тилнинг системавий унсури бўлган мазкур лингвомаданий бирликлар тадқиқи айниқса, лексикология, лексикография, корпус лингвистикасида тез-тез кўзга ташланмоқда. Гарчи XXI аср илм-фани бу борада электрон корпуслар мисолида кенг тадқиқ имконини яратган бўлса-да, идиомаларнинг миллий лисоний табиати уларни луғатлаштириш ҳамда миллий корпусда бериш ишларини бир қадар мураккаблаштиради. Бинобарин, идиоматик қурилмаларни қайта ишланувчи, лингвистик таснифланган, машинада ўқиладиган металингвистик маълумотлар форматида тақдим этиш муаммоси ҳануз долзарблигича қолмоқда.

Дунё корпус лингвистикасида идиомаларни семантик, синтактик ва прагматик жиҳатдан текшириш, улар таркибидаги сўзларнинг ўзаро бирика олиш (collocation) имкониятини тадқиқ этиш, фразеологик бутунликларни модификацияланиш хусусиятига кўра идиома ёки ўзга типли турғун бирикма сифатида таснифлаш муаммоси тўлиқ ҳал этилганича йўқ. Қатор тилларнинг Миллий корпусида идиомаларни разметкалашга доир жадал амалиёт кечаётганига қарамай, матнни автоматик қайта ишлайдиган дастурларнинг таржима сифатини яхшилаш мақсадида идиоматик конструкцияларни моделлаштириш ўз ечимини кутаётган муҳим масалалардан бири бўлиб турибди.

Мамлакатимизнинг бугунги мавқеи, жаҳон ҳамжамиятида тутган ўрни, қолаверса, глобализм феномени билан боғлиқ турли омил давлат тили масаласига жиддийроқ ёндашувни кун тартибига қўйди. “Бугунги глобаллашув даврида ҳар бир халқ, ҳар қайси мустақил давлат ўз миллий манфаатини таъминлаш, бу борада аввало, ўз маданиятини, азалий қадриятларини, она тилини асраб-авайлаш ва ривожлантириш масаласига устувор аҳамият қаратиши табиийдир”<sup>1</sup>. Бинобарин, барча илмий йўналишлар қаторида тилшуносликда ҳам “...илмий ва ижодий изланишларни ҳар томонлама қўллаб-қувватлаш, улар учун зарур шарт-шароитлар яратиш вазифа”сининг<sup>2</sup> белгиланиши тилшуносликнинг қўллаб-қувватлаш анъанавий муаммолари, хусусан, жонли ўзбек тилининг “қаймоғи бўлган”<sup>3</sup>, миллий-маданий кодланган турғун бирликларни авлоддан-авлодга етказиш масаласининг янгича ечим топиши ҳамда уларга доир анъанавий муаммоларни ҳал қилишда ишончли натижаларга эришишни кафолатлайди.

<sup>1</sup> Ўзбекистон Республикаси Президентининг 2019 йил 21 октябрдаги ПФ-5850-сонли “Ўзбек тилининг давлат тили сифатидаги нуфузи ва мавқеини тубдан ошириш чора-тадбирлари тўғрисида”ги фармони // [www.lex.uz](http://www.lex.uz)

<sup>2</sup> Мирзиёев Ш.М. Эркин ва фаровон, демократик Ўзбекистон давлатини биргаликда барпо этамиз. Ўзбекистон Республикаси Президенти лавозимига киришиш тантанали маросимида бағишланган Олий Мажлис палаталарининг қўшма мажлисидаги нутқ. – Тошкент: Ўзбекистон, 2016. – Б.13.

<sup>3</sup> Саломов Ғ.Т. Рус тилидан ўзбек тилига мақол, матал ва идиомаларни таржима қилиш масаласига доир. - Тошкент, 1961. – Б. 3.

Ўзбекистон Республикаси Президентининг 2019 йил 21 октябрдаги “Ўзбек тилининг давлат тили сифатидаги нуфузи ва мавқеини тубдан ошириш чора-тадбирлари тўғрисида”ги ПФ-5850-сон Фармонидан келиб чиқиб, давлат тилининг ахборот ва коммуникация технологиялари, хусусан, интернет жаҳон ахборот тармоғида муносиб ўрин эгаллашини таъминлаш, ўзбек тилининг компьютер дастурларини яратиш, давлат тилининг илм-фан тили сифатида нуфузини ошириш ва Ўзбекистон Республикаси Президентининг 2020 йил 20 октябрдаги “Мамлакатимизда ўзбек тилини янада ривожлантириш ва тил сиёсатини такомиллаштириш чора-тадбирлари тўғрисида”ги ПФ-6084-сон, 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 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 Қарори, 2017 йил 13 сентябрдаги “Китоб маҳсулотларини нашр этиш ва тарқатиш тизимини ривожлантириш, китоб мутолааси ва китобхонлик маданиятини ошириш ҳамда тарғиб қилиш бўйича комплекс чора-тадбирлар дастури тўғрисида”ги ПҚ-3271-сон Қарори, Ўзбекистон Республикаси Президентининг 2020 йил 24 январдаги Олий Мажлисга Мурожаатномасида белгиланган таълим, илм, фаннинг барча соҳалари ва босқичларида ахборот технологияларини жорий этиш ҳақидаги таклифи ҳамда мазкур фаолиятга тегишли бошқа меъёрий-ҳуқуқий ҳужжатларда белгиланган вазифаларни амалга оширишда ушбу диссертация иши муайян даражада хизмат қилади.

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

**Муаммонинг ўрганилганлик даражаси.** Мавзунинг умумий методологик асосини Ф.Ф.Фортунатов, Ф.И.Буслаев, Ш.Балли, А.А.Шахматов, Е.Д.Поливанов, В.В.Виноградов, С.И.Абакумов, Г.К.Данилов, Г.О.Винокур, В.Н.Державин, А.И.Ефимов, А.Я.Рожанский, И.Е.Аничкин, Л.П.Смит, В.П.Жуков, В.Н.Телия, Н.М.Шанский, Н.Н.Амосов, А.М.Бабкин, В.М.Савицкий, М.Л.Ковшова, Е.Д.Балигина, О.В.Елисеева, М.А.Ганюшина, С.Е.Максимец, Т.В.Кара-Казарьян, О.В.Куманок, Р.Саидов, Г.Ниннин, Т.С.Талабов, Е.М.Маклакова, Н.У.Ворокова, Н.В.Щенникова, Н.В.Баско, В.М.Мокиенко, Е.М.Гирняк, Т.С.Борщевская, И.Е.Аничков, К.Н.Мочалина, Н.Елисеева, Н.Горобец, А.Ў.Аksoy, Л.Э.Бинович, ўзбек тилшуносларидан Ш.Раҳматуллаев, А.Ҳожиев, А.Маматов, М.Рўзикулова,

Б.Йўлдошев, Д.Худайбергана, Ш.Усмонова каби олимларнинг назарий қарашлари ташкил қилди.

Тадқиқотнинг корпус лингвистикаси билан боғлиқ назарий асослари сифатида В.П.Захаров, В.Плунгян, В.В.Мамонтова, О.В.Кукушкина, А.А.Поликарпов, Е.В.Суровцевалар, ўзбек тилшунослигида компьютер лингвистикаси, матнга лексикографик ишлов бериш ва лингвостатистик таҳлил этиш борасида эса А.Л.Юсупов, А.Пўлатов, С.А.Каримов, Б.Р.Менглиев, С.Мухамедов, М.Айимбетов, С.Мухамедова, Н.Абдурахмонова, Ш.Ҳамраева<sup>1</sup> ва бошқаларнинг назарий қарашларига таянилди<sup>2</sup>. Бу тадқиқотларда тил бирликларини лексикографик ва лингвостатистик тадқиқ этиш, ўзбек лингвистик корпусларини яратишнинг усул ва тамойиллари тавсия этилган бўлса-да, уларда лингвомаданий бирликлар, жумладан, идиомалар базасини яратиш масаласи кун тартибига қўйилмаган.

**Тадқиқотнинг диссертация бажарилган олий таълим илмий-тадқиқот ишлари режалари билан боғлиқлиги.** Тадқиқот Термиз давлат университети илмий-тадқиқот ишлари режаси таркибига кирувчи “Ўзбек тилшунослиги: бадий нутқ таҳлили, антропоцентрик назария, этнолингвистика, корпус тилшунослиги” мавзусидаги тадқиқот доирасида бажарилган.

**Тадқиқот мақсади** ўзбек миллий корпусида идиомалар базасини шакллантиришнинг илмий, назарий-амалий асосларини ёритиш орқали уларни тил корпусига киритиш ҳамда разметкалаш масалаларининг назарий ва амалий асосларини ишлаб чиқишдан иборат.

**Тадқиқотнинг вазифалари:**

идиомаларнинг миллий лисоний табиатини ўрганишга бағишланган методологик қарашларни ўрганиш;

ўзбек тилида идиомаларнинг лексикографик тадқиқ даражасини текшириш;

идиомалар таснифида тиллараро тафовутни текшириш;

тил таълимида идиомаларни ўрганишнинг оптимал усулларини ўрганиш;

---

<sup>1</sup> Ҳамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари. Филол.фан.бўйича фалсафа доктори (PhD)...дис. афтореф. – Бухоро, 2018.

<sup>2</sup> Мухаммедов С.А. Статистический анализ лексико-морфологической структуры узбекских газетных текстов: Автореф. дисс... канд. фил. наук.- Ташкент, 1980; Бабанаров А. Разработка принципов построения словарного обеспечения турецко-русского машинного перевода: Автореф. дисс... канд. фил. наук. - Л., 1981; Айымбетов М.К. Опыт лингвостатистического анализа лексики и морфологии каракалпакского публицистического текста: Автореф. дисс... канд. фил. наук.- Ташкент, 1987; Каримов С., 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; Пўлатов А. Компьютер лингвистикаси. – Тошкент: Akademnashr, 2011.; Норов А. Компьютер лингвистикаси асослари. – Қарши, 2017. – 136 б..

тил корпусларида фразеологик чатишмаларнинг – идиомаларнинг берилишини жаҳон корпус лингвистикаси мисолида ўрганиш;

ўзбек тилининг Миллий корпусида идиомаларнинг электрон тезкор қидирув базасини яратиш.

**Тадқиқотнинг объектини** ўзбек тили идиомалари ташкил қилади.

**Тадқиқотнинг предметини** ўзбек тилидаги идиомаларни миллий корпусга киритиш масалалари ташкил этади.

**Тадқиқот усуллари.** Тадқиқотда қиёсий, статистик, шунингдек, корпус лингвистикасининг латент таҳлил, ўзаро маълумот, матнларни бирлаштириш усулларидан фойдаланилди.

**Тадқиқотнинг илмий янгилиги** қуйидагилардан иборат:

ўзбек тили миллий корпусида идиомалар базасини шакллантиришда лексикографик асоснинг етарли эмаслиги, луғатларда идиомалар фразеологик birlikлардан ажратилмай тавсифланганлиги, изоҳларда аниқ тамойилга асосланмаганлиги аниқланган;

коллокациялар (турғун бирикмалар)нинг ўзбек ва хорижий (рус, инглиз) тилларида морфологик анализаторлар функционал имкониятининг лингвистик таъминот жиҳатидан фарқли, семантик разметкалаш ҳамда математик моделлаштириш усуллари жиҳатидан эса муштарак эканлиги аниқланган;

латент-семантик таҳлил, ўзаро алоқадор ахборотни баҳолаш, параллел матнларни бирлаштириш усуллари асосида идиоматик қурилмаларни математик моделлаштиришнинг асослари очиб берилган;

Ўзбек Миллий Корпусида идиомаларни қидириш, танлаш ва вариантлари билан кўриш имконини берувчи, мазкур идиоматик birlikлар мазмуни, синонимлари ҳамда вариантлари, инглиз, рус ва баъзи туркий тиллардаги муқобилларини жамловчи ўзбек тили идиомалар онлайн базаси (Idioms.uz) веб-илоvasи яратилган.

**Тадқиқотнинг амалий натижалари** қуйидагилардан иборат:

тилнинг фразеологик тизимида ибораларнинг ўрни, лисоний моҳияти ва вазифаси; ҳосил бўлишнинг лингвомаданий асослари, ўзбек тили идиомаларини миллий корпусда беришнинг усул ва воситалари; ўзбек тили идиомаларининг нутқ маданиятини юксалтиришдаги ўрни ва уларнинг қўлланилиши, ўрганилишига доир илмий-назарий маълумотлар ҳозирги ўзбек тили (лексикология, фразеология, лексикография), лингвокультурология, корпус лингвистикаси, умумий тилшунослик каби фанлар бўйича яратиладиган дарслик ва ўқув қўлланмаларнинг мазмунан бойишига хизмат қилиши асосланган;

идиомаларнинг лингвистик шарҳига оид назарий хулосалар ўзбек идиоматикаси, лингвокультурологияси, корпус лингвистикаси бўйича тузилган луғатларнинг янада мукамаллашувига хизмат қилиши асосланган.

**Тадқиқот натижаларининг ишончлилиги** лингвистик таҳлил методологияси негизида шакллантирилганлиги, жумладан, идиомаларни ўрганишнинг жаҳон ва ўзбек луғатчилиги, фразеологияси, жаҳон корпус лингвистикаси кесимидаги ўрни, лингвомаданий мавқеи далилий



материаллар ёрдамида атрофлича очиб берилганлиги, назарий фикр ва хулосаларнинг амалиётга жорий этилганлиги, олинган натижаларнинг ваколатли ташкилотлар томонидан тасдиқланганлиги билан изоҳланади.

**Тадқиқот натижаларининг илмий ва амалий аҳамияти.** Тадқиқот натижаларининг илмий аҳамияти ўзбек тилшунослигида корпус лингвистикаси доирасидаги янги тадқиқотлар учун илмий-назарий манба бўла олиши билан белгиланади. Жумладан, идиомаларнинг лингвомаданий табиати, уларни тил корпусида лексик-семантик разметкаш, ўзбек тилидаги идиоматик қурилмаларнинг семантик-синтактик хусусиятларини таҳлилдан ўтказиш ҳамда уларнинг хорижий тиллардаги вариантларини миллий корпусда беришнинг самарадор усулларига доир назарий маълумотлар лексикология (семасиология), фразеология, лингвокультурология, корпус лингвистиканинг ривожига хизмат қилади.

Тадқиқот натижаларининг амалий аҳамияти эса ўзбек тили идиомалар базаси (Idioms.uz) веб-иловасининг мунтазам тўлдириб борилувчи базасидан илмий-амалий мақсадда фойдаланиш имконияти, тадқиқот материалларидан олий таълимда «Ҳозирги ўзбек адабий тили», «Лингвокультурология», «Корпус лингвистикаси» каби курсларни ўқитиш ҳамда шу фанларга доир дарслик ва ўқув қўлланмаларини тайёрлашда фойдаланиш мумкинлиги билан белгиланади.

**Тадқиқот натижаларининг жорий қилиниши.** Ўзбек тили миллий корпусида идиомалар базасини ишлаб чиқиш жараёнида эришилган илмий натижалар асосида:

туркий тиллар қатори ўзбек тили фразеологик тизимида мавжуд лексик-семантик бутунликларнинг ўрни, идиоматик конструкцияларнинг ўрганилишини ўзбек тили мисолидаги ҳолатини асослаш ва тасдиқлашга доир назарий маълумотлардан Ўзбекистон Республикаси Фанлар Академияси Қорақалпоғистон бўлими Қорақалпоқ гуманитар фанлар илмий-тадқиқот институти томонидан олиб борилган ФА-А1-Г007 «Қорақалпоқ нақл-мақоллари лингвистик тадқиқот объекти сифатида» мавзуларидаги фундаментал илмий лойиҳасида фойдаланилган (ЎзР Фанлар академияси Қорақалпоғистон бўлимининг 2021 йил 23 июндаги 17.01/127-сонли маълумотномаси). Натижада фразеологик бирикмаларнинг олам миллий манзарасини шакллантириши ва туркий тиллар қатори ўзбек тилидаги идиоматик қурилмаларнинг лингвомаданий табиатига оид билимларнинг мукамаллашувига эришилган;

идиоматик бирликлар тузилишининг умумий тамойиллари ҳамда унинг ахборот-қидирув тизимидаги аҳамиятини аниқлаш; туркий тиллар (қозоқ, қирғиз, озарбайжон, бошқирд, татар, туркман тили) қаторида ўзбек тили сўз ясалиш тизимида идиоматик бирликларнинг юзага келишида ўхшатиш ва “марказий сўз” ўрнида от ва сифат туркумидаги сўзнинг иштирокини эътироф этиш; идиоматик қурилмаларни лексикографик жиҳатдан идеографик таснифлашда улар таркибидаги локал компонентнинг нима орқали ёхуд қандай ифода топганини аниқлаш эвазига тил корпусида маълумотларнинг тезкор қидирувига эришиш имкониятларини ёритиш;

идиомаларнинг маълум ҳудудда қўлланилувчи, миллий хосланган ва амалий характерга эга луғавий бирликлар экани тўғрисидаги тадқиқ натижаларини компьютер тилшунослиги нуқтайи назаридан лингвистик тадқиқотлар сирасида илмий-таълимий жараёнга тадбиқ этиш ҳақидаги хулосалардан Ўзбекистон Республикаси Фанлар Академияси Қорақалпоғистон бўлими Қорақалпоқ гуманитар фанлар илмий-тадқиқот институтида бажарилган ФА-Ф1-Г003 “Ҳозирги қорақалпоқ тилида функционал сўз ясалиши” мавзусидаги амалий лойиҳасида фойдаланилган (ЎзР Фанлар академияси Қорақалпоғистон бўлимининг 2021 йил 22 февралдаги 35/1-сонли маълумотномаси). Натижада олий таълим соҳасида компьютер лингвистикасини ўрганиш ва лисоний компетенцияни ривожлантиришга доир қарашларнинг мукамаллашувига эришилган;

Идиомаларнинг ўзбек нутқ маданиятида тутган ўрни, лингвомаданий хоссаларининг жонли нутқда намоён бўлишига доир хулосалардан Сурхондарё вилояти телерадиокомпанияси кўрсатувларининг сценарийларини тайёрлашда фойдаланилган (2021 йил 17 майдаги 01-01/59-сон маълумотнома). Натижада юқори савияли баркамол авлодни вояга етиштириш, ёшлар орасида тил маданиятини шакллантириш ҳамда замонавий ахборот тизимида ўзбек тилининг имкониятларидан фойдаланиш ҳақидаги тарғибот ишлари мазмуни мукамаллашган.

**Тадқиқот натижаларнинг апробацияси.** Мазкур тадқиқот натижалари бўйича 2 та халқаро ва 3 та ОАК рўйхатидаги журналларда, 1 та республика илмий-амалий анжуманларида ва 1 та Scopus базасида индексладиган журналда эълон қилинган.

Тадқиқот натижаларининг эълон қилинганлиги. Диссертациянинг асосий мазмуни 1 та муаллифлик гувоҳномаси, 10 та илмий мақола (уларнинг 3 таси Ўзбекистон Республикаси ОАК тасарруф этган илмий журналлар ва 2 таси халқаро журналлар ва 2 таси халқаро конференция) 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 130 саҳифани ташкил этади.

## **ДИССЕРТАЦИЯНИНГ АСОСИЙ МАЗМУНИ**

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

Диссертациянинг “**Жаҳон ва ўзбек тилшунослигида идиоматик қурилмаларнинг лингвистик тавсифи**” деб номланган биринчи боби жаҳон тилшунослигида идиомаларга илмий-назарий ёндашувлар, ўзбек

тилшунослигида идиоматик конструкцияларга шарҳ ва улардаги тафовут, қиёсий тилшуносликда идиомаларнинг тавсифи, идиомаларнинг мавзуий гуруҳланишининг муҳокамасига оид масалаларни жамлайди.

Ҳар бир тил соҳиби ўз маданияти ва менталитетини намоён қилувчи тил бирликларига эга бўлиб, тилшунослар томонидан миллий хосланган тил бирликлари муаммоси XX асрнинг бошларида чуқур тадқиқ қилина бошланди. Профессор Ғ.Т.Саломов томонидан “тилнинг мўъжизаси, қаймоғи”<sup>1</sup> деб таърифланган идиомалар XX асрнинг 40-йилларида тадқиқотчилар диққатини тортди. Бу даврга келиб, жаҳон тилшунослигида фразеологик тадқиқотлар кўпроқ лексикографик жиҳатларга таянилиб амалга оширилди. Жумладан, фразеологизмларни луғатларда акс эттириш, этимологиясини изоҳлаш асосий ўринга олиб чиқилди. Проф. Б.Йўлдошевнинг фикрича, фразеологиянинг ривожланиш даври шартли равишда уч даврга бўлиб ўрганилди<sup>2</sup>:

**Биринчи давр** – XVIII аср ўрталаридан XX асрнинг 30-йилларини ўз ичига олади. Бу давр тадқиқотчилари орасида М.В. Ломоносов, В.Даль, А.А.Потебня, Ф.Ф.Фортунатов, Ф.И.Буслаев, Ш.Балли, А.А.Шахматовларнинг қарашлари машхур.

**Иккинчи давр** – XX асрнинг 30-50 йилларини ўз ичига олади, бу даврда Е.Д.Поливанов, В.В.Виноградов, С.И.Абакумов, Г.К.Данилов, Г.О.Винокур, В.Н. Державин, А.И.Ефимов, А.Я. Рожанский, И. Е. Аничкин кабиларнинг<sup>3</sup> илмий ғоялари кенг эътироф этилди.

**Учинчи давр** – XX асрнинг 60- йилларидан то ҳозирги кунгача бўлган даврни ўз ичига олади. Бу давр фразеологик тадқиқотлари янги методларнинг қўлланилиши, фразеологик назариянинг тез суръатлар билан ривожланиши ва кўплаб фразеолог мутахассисларнинг етишиб чиққанлиги билан характерланади<sup>4</sup>. Фразеологиянинг соҳа сифатида шаклланиши натижасида унинг таркибий қисмлари ҳам чуқур тадқиқ жараёнига тортилди. Хусусан, фразеологик чатишмалар – идиомалар кўплаб илмий изланишларнинг юзага келишига сабаб бўлди.

Фразеологияда турғун бирликларни ажратиш, таҳлил қилиш, ўрганиш ва тасниф қилишнинг компонент таҳлил, идентификация, аппликация ва комбинация методлари мавжуд. Аммо бу методларни идиоматик

<sup>1</sup> Саломов Ғ.Т. Рус тилидан ўзбек тилига мақол, матал ва идиомаларни таржима қилиш масаласига доир. - Тошкент, Ўзбекистон ССР Фанлар Академияси нашриёти, 1961. – Б. 3.

<sup>2</sup> Йўлдошев. Б. Ўзбек фразеологияси ва фразеографиясининг шаклланиши ҳамда тараққиёти. Монография. Тўлдирилган ва тузатилган иккинчи нашри. – Самарқанд: СамДУ нашри, 2013. – Б. 4

<sup>3</sup> Винокур Г.О. Избранные работы по русскому языку: сборник научных трудов / Г.О.Винокур; Российская Академия Наук, Институт русского языка им. В.В.Виноградова. – 2-е изд., стер. – Москва : ФЛИНТА, 2016. – 780 с. – (Стилистическое наследие).

<sup>4</sup> Бу ҳақда қаранг. Йўлдошев. Б. Ўзбек фразеологияси ва фразеографиясининг шаклланиши ҳамда тараққиёти. Монография. Тўлдирилган ва тузатилган иккинчи нашри. – Самарқанд: СамДУ нашри, 2013. – Б. 4; Ройзензон Л. И., Малиновский Е. А., Хаютин А. Д. Очерки по истории становления фразеологии как лингвистической дисциплины. – Самарқанд, 1975, - С. 5-70; Малиновский Е.А. История изучения проблем фразеологии в русском языкознании. – Ташкент: Фан, 1991. – С.5-92; Малиновский Е.А. Формирование русской фразеологии и фразеографии. АДД-Алматы, 1995. – С. 3-52; Малиновский Е.А.Формирование русской фразеологической теории в XX веке. Монография. – Самарқанд, 2007. – С.4-131.

бирликларга тўлиқ қўллаш ножоиз, бунга сабаб идиомаларнинг таркибий қисмларга бўлинмаслигидир.

XX асрнинг 40-йилларигача тилнинг фразеологик таркибига бағишланган қарашларни В.Виноградов, Ф.Фортунатов, А.А.Шахматов, А.Булаховский, И.Срезневский, А.Потебня, С.Абакумов ишларидагина учратиш мумкин эди<sup>1</sup>. Лекин кейинчалик бу соҳадаги ишлар салмоғи ошди. Бироқ уларнинг барчасида умумий қоида асосида иш кўриш, В.В.Виноградов, А.А.Шахматов таърифларидан фойдаланиш сезилиб қолди. Бу даврда амалга оширилган изланишларда фразеологик чатишманинг фразеологизм таркибига кириши ва айнан унинг ўзи эканлиги тўғрисидаги қарашлар устувор аҳамият касб этди. Қайд этиш керакки, бу давр жаҳон тилшунослигида фразеологик қатлам унсурлари орасида айнан идиомаларни ажратишга урғу берилди. Жумладан, С.Тарабрина хориж тилшунослари *идиома* деб атовчи, француз фразеологик бирлиги ўрнида қўлланилган ушбу терминни биринчи марта инглиз тилшуноси Л.П.Смит қўллаганини қайд этади<sup>2</sup>. Аммо Л.П.Смитнинг ўзи, шунингдек, В.П.Жуков, В.Н.Шанский ва бошқа олимлар ҳам турғун бирикмаларнинг барчасини фразеологияга оид деб ҳисоблашган.

Гарчи фразеология лингвистиканинг мустақил соҳаси сифатида ажралиб чиққанлигига қарамай, ҳозирда ҳам унга лексикологиянинг бир бўлими сифатида қаралади. Бироқ Н.Н.Амосов, А.М.Бабкин, А.И.Смирницкийлар фақат тегишли гуруҳларнигина бу бўлимга тааллуқли деб ҳисоблайдилар.

Рус тилшунослигида фразеологик бирликларни бир неча мезонларга кўра таснифлаш кузатилади: 1) фразеологик оборотларнинг таркибига кўра; 2) фразеологик оборотларнинг структурасига кўра таснифлаш.

Фразеологик оборотларни таркибига кўра таснифлашни Н.М.Шанский таклиф этган бўлиб, у фразеологик оборотларни икки гуруҳга бўлган: 1) фаол лексикага оид фразеологизмлар; 2) эскирган ёки диалектга оид фразеологизмлар<sup>3</sup>.

Ш.Балли фразеологизмларни икки асосий гуруҳга ажратган: фразеологик қатор ёки одатий фразеологик гуруҳ ва фразеологик бирлик. Ш.Баллига кўра, биринчи гуруҳ фразеологизмларининг қисмлари ўртасида бирмунча эркин бирикиш мавжуд, кейинги фразеологик гуруҳ таркибидаги бирликлар орасидаги боғланиш эса шу қадар кучлики, натижада улар орасидаги мазмун яхлит ҳолга келиб қолган<sup>4</sup>.

И.Ю.Третьякова фикрича: “Фразеологизмлар турли коммуникацион мақсадлар учун трансформацияга учраши мумкин, булар: мазмунни

<sup>1</sup> Фортунатов Ф.Ф. Сравнительное языковедение. Общий курс. В кн: «Избранные труды», т 1. – М., Учпедгиз, 1956. – С. 83.

<sup>2</sup> www.dissercat.com/ С. Ю.Тарабрина Семантика фразовых глаголов и глагольных единиц. Диссертация кандидата наук. 2003. Яна қаранг: Смит Л. Фразеология английского языка. – М., 1959. – С.10.

<sup>3</sup> Шанский Н.М. Фразеология современного русского языка. – М., 1963.

<sup>4</sup> Балли Ш. Французская стилистика. – М., 2001. – С. 89-90.

Яна қаранг: Колесникова А.А. Русские и английские идиомы мотивированные сценариями общеевропейской культуры( сравнительный анализ идиом со словами devil, hell, чёрт, ад и их квазисинонимами. Магистерская выпускная квалификационная работа. – С. Петербург, 2016. – С.6.

аниқлаштириш, маълум тил вазиятида фразеологизмни аниқлаштириш, субъектив баҳо бериш”<sup>1</sup>. Фразеологизмларни норматив қўллашда одатий қўлланилиш рўй беради, бунда одатийликдан ҳеч қандай оғишиш кузатилмайди, окказионал қўлланишда эса фразеологизм таркибига у ёки бу инновациялар киритилади. Бундай нормадан оғиш услубий приём ҳисобланади<sup>2</sup>.

Ўзбек ва жаҳон тилшунослигида фразеологиянинг ўрганилиш объекти кенг ва тор маънода тушунилади. Фразеологиянинг объектини тор маънода тушунувчилар ибораларни мақол ва маталлардан ажратиб изоҳламоқдалар<sup>3</sup>.

*Фразеологизм* тушунчасини кенг мазмунда тушунувчи биринчи гуруҳ олимлари (Л.П.Смит, В.П.Жуков, В.Н.Телия, Н.М.Шанский) барча турдаги барқарор бирикмаларни фразеологизмларга киритишса, Н.Н.Амосов, А.М.Бабкин, А.И.Смирницкийлар фақат маълум гуруҳни фразеологизмга мансуб деб биладилар. Академик В.В.Виноградов мақоллар, маталлар ва машҳур сўзлар (*крылатые слова*)ни фраземалар таркибига қўшмайди. Д.О.Добровольский идиома номи остида фақат фразеологик бирликларни келтирган бўлса, В.В.Виноградов таърифида фразеологизмлар уч гуруҳга бўлинган: А) фразеологик чатишмалар; б) фразеологик бутунликлар; в) фразеологик бирликлар.

Таъкидлаш зарурки, фразеологизм таркибидаги идиома ва идиоматик бирликларнинг изоҳи даврий аҳамиятга эга масала ҳисобланади. Жумладан, Ф.Ф.Фортунатов, В.М.Савицкий, М.Л.Ковшова, Е.Д.Балигина, О.В.Елисеева, М.А.Ганюшина, С.Е.Максимец, Т.В.Кара-Казарьян, О.В.Куманок, Р.Саидов, Г.Ниннин, Т.С.Талабов, Е.М.Маклакова, Н.У.Ворокова, Н.В.Щенникова, Н.В.Баско, В.М.Мокиенко, Е.М.Гирняк, Т.С.Боршевская, И.Е.Аничков, К.Н.Мочалина, Н.Елисеева, Н.Горобец каби тадқиқотчилар ўз ишларида фразеология ва идиоматика муносабатлари, менежмент ва иқтисодий терминологияда идиома, идиомаларнинг болалар томонидан қабул қилинишини нейролингвистик сатҳда ўрганишган, сиёсий ва иқтисодий аспектларда кўриб чиқишган ва бу олимларнинг барчасини фразеологизмни тор мазмунда тушунувчи гуруҳ олимлари қаторига қўшиш ўринлидир.

Шунингдек, тилшунос Ж.М.Ахмедова фразеологияни тор мазмунда тушуниш тарафдорлари бу чегарани фақат идиома (фразеологик чатишма) доирасида қабул қилишди, деган фикрларни билдиради<sup>4</sup>. Демак, фразеологияни барқарор бирикмаларни ўрганувчи соҳа, идиоматикани эса унинг таркибидаги бўлим сифатида тушуниш ўринли.

<sup>1</sup> Третьякова И.Ю. Окказиональные преобразования фразеологизмов с компонентом «душа» // Фразеологизм и слово в национальнокультурном дискурсе (лингвистический и лингвометодический аспекты): Междунар. науч.-практ. Конф., посвящ. юбилею д. ф. н, проф. А. М. Мелерович (Кострома, 20–22 марта 2008 г.). – М.: ООО Элпис, 2008. – С.270.

<sup>2</sup> Гирняк Е.М. Идиоматичность рекламного текста (На материале русской и китайской рекламы) Вестник Пермского университета Российская и зарубежная филология.2011.Вып.1(13). – С. 34

<sup>3</sup> Раҳматуллаев Ш. Нутқимиз кўрки. – Тошкент, 1970. – Б.30-31, 36.

<sup>4</sup><https://dissercat.com/> Ж.М. Ахмедова. Фразеология ахвахского языка: На материале ратлубского, цегобского и плянубского говоров. Диссертация кандидата наук. 2004.

В.Н.Телия фразеология объекти хусусида шундай ёзади: “Баъзи олимлар фразеология таркибига фақат идиомалар ва фразеологик бирликларни киритадилар, бошқалари мақоллар ва маталларни ҳам фразеологиянинг бир қисми деб ҳисоблайдилар. Бу бирликларнинг ҳаммасини бир хусусият бирлаштиради: бир неча сўздан иборат бўлиш ва мулоқотда уларга бўлган эҳтиёжнинг кучлилиги”<sup>1</sup>.

О.Ахманова идиоматик бирликларни ўрганувчи соҳани *идиоматика*, идиомаларнинг ўзини *идиоматизм*, *идиоматик ифода* деб атайди ҳамда бу бирликларнинг ингл. *idiom, idiomatic expression*, фр. *idiome, idiotisme*, нем. *idiom, idiotismus*, исп. *idiomatismo* сингари номланишини кўрсатади. Олима идиомаларни “...ўзининг синтактик ва семантик табиати билан тегишли тилнинг бетакрор хусусиятларини намоён этувчи, маълум халқнинг миллий ўзига хослигини кўрсатиб берувчи тил бирлиги” деб ҳисоблайди<sup>2</sup>.

Г.Х.Ахунзянов “фразеология” термини кўп қиррали ва кўп маъноли бўлганлиги учун уни “идиоматика”, “идиомалогия” терминлари билан алмаштириш кераклигини таъкидлайди<sup>3</sup>. Олим идиомаларни 1) яхлит маъноли идиомалар; 2) аналитик идиомаларга ажратади.

С. Жаъфаров турғун сўз бирикмаларини идиомалар, ҳикматли сўзлар, иборалар, оталар сўзи ва зарбулмасаллар каби беш гуруҳга ажратади<sup>4</sup>. С.Жаъфаровнинг таснифида камчиликлар ҳам мавжуд эди, чунки бу таснифда перифразалар ва тасвирий ифодалар аралаш ҳолда берилган. Бизнингча, С.Жаъфаров идиомани кенг маънода тушунувчилар қаторида, ўз талқинидаги таснифни юзага келтирган.

Идиомалар таснифи турли тилларнинг ички хусусиятларига кўра фарқланади. Бу уларнинг корпусда акс этишига бевосита таъсир кўрсатади. Масалан, Ф.П.Кауи ҳамда Р.Маккина идиомаларни семантик ва синтактик хусусиятларига кўра 4 гуруҳга бўлган: хусусий идиомалар (*to blow the guff*); метафорик идиома (*to beat one's breast*); ярим идиомалар (*a blind alley*); очик коллокациялар (*heavy rain*).<sup>5</sup>

Таникли инглиз тилшуноси М.Маккарти идиомаларни категориялаштиришда бир қанча терминларни қўллайди. Булар: 1) *opaque idioms* – тушунарсиз, ноаниқ идиома; умумий мазмуни компонентлар мазмунидан келиб чиқмайдиган идиомалар (*to kick the bucket*); 2) *semi-opaque idioms* – ярим тушунарсиз, ярим ноаниқ. Яъни қайта номлаш мумкин бўлган идиомалар (таққ.: *to pass the buck* ва *to pass the responsibility*); 3) *transparent idioms* – шаффоф идиома. Тушунишда қийинчилик туғдирмайдиган идиомалар (*to see the light*)<sup>6</sup> каби.

<sup>1</sup> Телия В.Н. Что такое фразеология. – М.: Наука, 1966.

<sup>2</sup> Ахманова О. С. Словарь лингвистических терминов. – М.: Сов. энциклопедия, 1966. – С. 161.

<sup>3</sup> Ахунзянов Г. Х. Идиомы (исследование на материале татарского языка): Автореф. дис. канд. филол. наук. – Казань, 1974.

<sup>4</sup> Жаъфаров С. Ҳозирги замон озарбайжон тили лексикаси (озарбайжон тилида). – Боку, 1958.

<sup>5</sup> Cowie A.P., Mackin R. Oxford Dictionary of Current Idiomatic English. - Vol. 2. Phrase, Clause and Sentence Idioms. - Oxford University Press, 1983.

<sup>6</sup> McCarthy M. Vocabulary. - Oxford University Press, 1990.

Таркиби вульгар бирликлардан таркиб топган идиомалар В.Буйнинг “Русская заветная идиоматика: весёлый словарь крылатых выражений”<sup>1</sup> луғатида учрайди ва унда 1000 га яқин сўкишга оид, ноодатий сўзлашувга хос барқарор бирликлар жамланган.

Ибораларни бадий асар ғоясига мувофиқ ўзгартириш мумкин. Ёзувчилар, одатда, ибораларни тасвир мақсадига мувофиқ танлаб қўллаш билангина кифояланиб қолмайдилар. Балки уларни қаҳрамонлар табиати, рухий ҳолати, ҳаёт тарзига мослаб ўзгартиради ва қайта ишлайди. Масалан, О.Аметова Юсуф Хос Ҳожибнинг тилдаги мавжуд, тайёр фразеологик иборалардан асар тилининг жонли, ширали чиқиши учун ўринли фойдалангани, кўпгина ҳолатларда ўзини янги фразеологик ибораларнинг ижодкори сифатида ҳам кўрсата олганини<sup>2</sup> қайд этади. Демак, иборалар қуйидаги ўзгаришларга учраши мумкин: 1) ибора таркибидаги айрим сўзлар алмаштирилади; 2) ибора таркиби кенгайди; 3) ибора таркиби қисқаради.<sup>3</sup> Бу қайта ишлаш натижасида иборалар: маънода янги талқиннинг юзага келиши; лексик таркибнинг ўзгариши; семантик-стилистик функцияларнинг кенгайиши; мазмунга янги мажозий ва образли маъноларнинг қўшилиши сингари “деформация”ларга учрайди<sup>4</sup>. Лекин бу ҳолат идиомаларда кузатилмайди, шу хусусияти билан у бошқа турдаги иборалардан фарқ қилади. Аммо Ж.Осмонова идиома компонентлари семантик ва синтактик яхлит бўлишига қарамай, шакл ясовчи қўшимчаларда ўзгариб келишини айтади. Масалан, *сирке+си суу котарбойт – сирке+м суу котарбойт* каби.<sup>5</sup> Кўринадики, идиомалар таркиби монолит (яхлит) бўлиб, уларнинг ўзига хос муҳим хусусияти иборалардаги каби ўзгаришларга учрамаслигидадир.

Инглиз тилшуноси Р.Мун инглиз идиомаларини функционал жиҳатдан беш гуруҳга ажратади: 1) ахборот идиомалари (*to catch sight of something, for sale*); 2) баҳолаш идиомалари (*kid's stuff, a different kettle of fish*); 3) вазият идиомалари (*long time no see, knock it off*); 4) модаллашган идиомалар (*I kid you not, you know what I mean*); 5) ташкилий идиомалар (*by the way, for instance*)<sup>6</sup>.

Идиомаларни ўрганган олимлардан яна бири А.Маккай идиомаларни алоҳида ўрганиб, икки қисмга ажратди: лексема даражасидаги идиомалар (lexemic stratum); семема даражасидаги идиомалар (sememic stratum). Лексема даражасидаги идиомаларга: – *come up with* типидagi феълли бирикмалар (phrasal verbs); – *to kick the bucket* типли фигуралар; биномлар (*here and there*); фразеологик сўз бирикмалари (*White House*); таркибли феълли идиомалар (*to baby-sit*). Олим семема даражасидаги идиомаларнинг номланишда аниқланадиган, семемаларга боғлиқ табиатини алоҳида

<sup>1</sup> <https://www.livelib.ru>>100 В.Буй. Русская заветная идиоматика: весёлый словарь крылатых выражений.

<sup>2</sup> Аметова О. “Кутадғу билик”да фразеологизмлар // Ўзбек тили ва адабиёти. №1. 2020. – Б. 114.

<sup>3</sup> Жумабоева С.М. Ўткир Ҳошимов асарларида фразеологизмлар. Битирув малакавий иши. – Тошкент: ЎзМУ, 2014. – Б.14.

<sup>4</sup> Шомаксудов А ва бошқ. Ўзбек тили стилистикаси. –Т., 1974. – Б. 70.

<sup>5</sup> Осмонова Ж. Идиомы в киргизском языке: Автореферат на соискание уч.степени канд. наук. – Фрунзе, 1969. – С.5.

<sup>6</sup> [https:// booknline.ru](https://booknline.ru). Петросян И.В. История изучения фразеологизмов и идиом.

белгилайди. Унингча, “маданий-прагмема” номи билан аталган ушбу идиомалар матн таҳлилида аниқланади ва уларнинг конвенционал характери белгиланади<sup>1</sup>.

Умумий ҳолда бугун тилшуносликда “Идиомалар табиатини очиб беришда қайси йўналиш ва усулни танлаш маъқул?”, “Уларни умумий билдирган мазмун асосида яхлит тушуниш керакми ёки миллий колорит, миллий-ментал хосланишга таяниб иш кўриш керакми?” деган саволларга жавоб изланмоқда. Идиомаларга тегишли мавжуд тавсифлар фразеологик бирликнинг идиома сифатига ўтиши учун қуйидаги хусусиятларга эга бўлишини исботлайди: фақат бир тилгагина хос бўлиб, бошқа тилга айнан таржима қилинмаслиги; миллий колоритга эга бўлиши; идиома маъноси компонентлари мазмунидан тубдан фарқ қилиши; икки ва ундан ортиқ сўз ва сўз бирикмасидан тузилиб, синтактик жиҳатдан бир бўлакка тенг бўлиши, гап таркибида бир сўз бирлигига тенг келиши; бошқа тилларда муқобилларининг мавжудлиги (аммо сўзма-сўз таржима қилиб бўлмаслиги аҳамиятлидир).

Ўзбек фразеологиясининг равнақ топишида Ш.Раҳматуллаев, А.Маматов, Б.Йўлдошев, А.Исаев, М.Рўзикулова, К.Бозорбоев, У.Рашидова, Ш.Ғаниева, Ш.Алмаматова, Б.Болтаева каби олимларнинг хизматлари улкан.

1960 йилларда фразеологик бирликларни таснифлаш масаласи жадаллашди. Жумладан, Ш.Раҳматуллаев, У.Турсунов, А.Маматов, Ф.Исҳоқов, К.Бозорбоев, Д.Абдуллаев, М.Вафоева, Ш.Алмаматова, М.Холиқова, Э.Умаров, Я.Пинхасов каби тилшуносларнинг ишларида айнан *идиома* термини келтирилган бўлмаса-да, моҳиятан ўзбек тили ибораларининг кенг кўламли тадқиқи амалга оширилди.

Ўзбек тилида идиоматикага оид тадқиқотлардан бири М.Рўзикулованинг “Ўзбек тилида идиомалар” мавзусидаги номзодлик диссертацияси (1966) ҳисобланиб, бу иш нисбатан анча олдинги ишлар сирасига киради. Унда идиоманинг табиати ва фразеологик бирликлар орасидаги ўрни ҳақида фикр юритилган, хулосалар қилинган. Ўзбек тилидаги фразеологик бирликлар маъно хусусиятларига кўра икки турга: а) фразеологик бутунлик; б) фразеологик чатишма каби турларга ажратилган.

XX асрнинг 60-йилларида диалектал фразеологизмлар соҳа тадқиқотчиларини ўзига жалб этди. Хусусан, Х.Дониёровнинг “Ўзбек тили “ж”лашган (кипчоқ) диалектининг фразеологияси устидан кузатишлар” номли<sup>2</sup> тадқиқотда фразеологизмлар ибора сифатида таҳлил қилинган, аммо идиоматик бирлик сифатида ўрганилмаган.

Бундан ташқари, диалектал фразеология масалалари бўйича Т.Содиқов, Ҳ.Узоқов, Н.Ражабов, М.Валиева, Б.Файзуллаев, А.Ишаев каби тилшунослар ўз ишларида Ўзбекистон ҳудудидаги турли вилоят шеваларининг диалектал

<sup>1</sup><https://www.dissercat.com/content/struktura-i-funktsii-idiom-na-materiale-substantivnykh-ustoichivyykh-slovosochetanii-i-slozhn> Ерохина Н.В. Структура и функция идиома; На материале субстантивных и устойчивых словосочетаний.

<sup>2</sup> Дониёров Х. Ўзбек тили “ж”лашган (кипчоқ) диалектининг фразеологияси устидан кузатишлар// Фразеология масалалари, СамДУ асарлари, янги серия, 106- чиқиши. – Самарқанд. – Б. 80-85.



фразеологиясини ўрганишган. Аммо профессор Э.Бегматовнинг фикрича, “халқ тили фразеологизми, айниқса, ўзбек диалектал фразеологияси ҳали чуқур тадқиқ қилингани йўқ”<sup>1</sup>. Фразеологиянинг таркибидаги бирлик сифатида идиомалар бу тадқиқотларда алоҳида эсланмаган.

Мазкур бобда фразеологизмларнинг туркий тиллардаги ўзига хосликлари фразеологик бирликлар ва идиомалар мисолида таҳлил қилинди.

**“Тилнинг миллий корпусида идиомалар базасини яратишда лексикографик асос”** деб номланган иккинчи бобда идиомаларнинг мавзуий гуруҳлари, жаҳон тиллари изоҳли луғатларида идиомаларнинг берилиши, тил корпуси ва лексикографик манба, жаҳон тиллари миллий корпусларида луғавий бирликларнинг берилиши (фразема ва идиомалар мисолида), идиоматик бирликларни Миллий корпусда акс эттиришнинг илмий-амалий аҳамияти масалалари ёритилди. Хусусан, идиомаларнинг тил тизимида пайдо бўлишига доир ҳолатларни тадқиқ этиш дунё тилшунослари томонидан ушбу турғун бирикмаларнинг турли мавзуий гуруҳларга ажратилишига олиб келган. Жумладан: диний асосларни ифодаловчи идиомалар (Муқаддас Қуръон оятлари, мусулмон эътиқодига мансуб шаърий қарашлардан таркиб топган идиомалар; инжил сюжетлари билан боғлиқ идиомалар – библеизмлар); муайян ижодкор фаолиятига алоқадор идиомалар; соматик тушунчаларни ифодаловчи идиомалар; “хайвон” семантик компонентли идиомалар; маиший буюмларни ифодаловчи идиомалар; таом, егулик номларини ифодаловчи идиомалар; мифологик асосга эга идиомалар.

Ибора ёки турғун бирикмалар семантик жиҳатдан бўлинмас бирликни ташкил этгани учун лексикографияда матнга автоматик ишлов бериш жараёнида буни инобатга олиш масаласига алоҳида эътибор қаратмоқ зарур. Корпус материали асосида статистик метод орқали қайси сўзлар доим биргаликда қўлланиши, бунинг натижасида турғун бирикмага қанчалик алоқадор эканлигини аниқлаш мумкин<sup>2</sup>. Идиомаларда эса бу масъулият янада ошади, сабаби идиома нафақат турғун бирикма, балки семантик жиҳатдан компонентлар билдирган мазмундан фарқ қиладиган бирликдир. Маъноси ўз таркибидаги сўзлар маъноларига боғлиқ бўлмаган, яхлит ҳолда бир маъно берувчи, ихчам иборадир<sup>3</sup>. Шу боис таклиф этилаётган идиомалар базаси мазмуни, синоними, варианты, рус тили, инглиз тилидаги эквивалентлари ва бадий асар матни, сўзлашув услубидан олинган мисоллари билан келтирилиб, жонли тил вазияти очикланди. XIX-XX асрларда хорижда шаклланган бу принципни корпус лингвистикаси ўз методологиясига айлантирди. Зеро, корпус лингвистикасининг предмети – тил меъёри эмас, одатий, расмга айланган тилдир.

Идиомалар ўзбек тилидаги барча изоҳли луғатларда мавжуд, аммо уларда идиома сифатида эмас, балки фразеологик бирлик, ибора сифатида келтирилган. Идиомаларнинг рус тилидаги изоҳли луғатларда берилиши

<sup>1</sup> Бегматов Э. Ҳозирги ўзбек адабий тилининг лексик қатламлари. – Тошкент: Фан, 1985. – Б.15.

<sup>2</sup> Қаранг: Захаров В.П. Корпусная лингвистика. Учебно-методическое пособие.–Санкт-Петербург, 2005. – С. 48.

<sup>3</sup> ЎТИЛ. 1981 йил 2 жилдли. – Б. 169.

нисбатан мукаммал кўриниш ва афзалликларга эга, компьютер технологияларидан қисман бўлса-да фойдаланиш мавжуд бўлиб, баъзи луғатлар рус тили корпуси асосида тузилган ва маълумотларнинг тезкор қидирувига эга эканлиги билан таҳсинга сазовор. Идиомаларнинг инглиз тилидаги (American English, British English, Australian English) луғатларнинг мобил илова шаклидаги версияларида аудиовариантдаги, расмли, анимацион форматли талқини ҳам мавжудлиги бу борада инглиз тили идиомалар луғатларининг анча мукаммал эканлигини тасдиқлайди.

Ўзбек тили изоҳли луғатларидаги идиомаларни қуйидаги асосий мавзуй гуруҳларга бўлиш мақсадга мувофиқ. 1) “Инсон” семантик компонентли идиомалар; 2) табиий, антропоген ландшафт семантик компонентли; 3) топоним семантик компонентли; 4) “инсон-замон-макон” семантик компонентли; 5) мифоним семантик компонентли; 6) ижтимоий семантик компонентли; 7) фитоморф семантик компонентли; 8) зооморф семантик компонентли; 9) “предмет” семантик компонентли.

Идиомаларни миллий корпусда беришнинг илмий-амалий моҳияти қуйидагиларда намоён бўлади: ҳозирда идиомаларни тил корпусларида акс эттириш натижасида тилшунос, таржимон, ўқитувчи, журналист ва турли соҳадаги тадқиқотчиларга изланаётган бирликни тезкор ва қисқа муддатда топиш ва таҳлил қилиш имкониятини беради; коммуникация ва инфорацион воситалар ёрдамида ўқиш ва ўқитиш давр талаби сифатида янада такомиллашиб бормоқда. Бу эса миллий корпусда ёки тил корпусида жой олган маълумотлардан таълим жараёнларида ҳам тўғридан-тўғри фойдаланиш имконини беради; ўзбек тили дунёнинг бир неча давлатларида фан сифатида ўқитилишини ҳисобга олсак, идиомалар базасидан фойдаланиш орқали хорижлик қизиқувчиларда тилимизнинг миллий хосланган бирликлари билан танишиш, умуман тилшуносликнинг компьютерлашиши билан боғлиқ барча соҳаларида амалий қўллаш имкони вужудга келади.

Диссертациянинг **“Ўзбек тили миллий корпусида идиомалар базасини яратишнинг илмий-амалий асослари”** деб номланган учинчи бобида миллий корпусда луғавий бирликларни разметкалаш, идиомаларнинг мавзуй гуруҳланишида тиллараро тафовут (таълимий моҳият), идиомаларни автоматик аниқлаш методлари тўғрисидаги илмий-назарий маълумотлар атрофлича баён қилинди.

Идиоматик бирликларни разметкалаш масаласи компьютер лингвистикаси ва корпус лингвистикасининг асосий муаммоларидан бири ҳисобланади. Чунки разметка маълумотларни таҳрирлаш ва таҳлил қилиш воситаси ҳисобланади. Семантик разметканинг мақсади робот ёки компьютер учун тушунарли бўлган маълумотларни қидирув натижасида топиш ва тақдим этиш имконини беради.

Кўп ҳолларда маъновий бирлик бир неча компонентли шаклда келади. Мазкур қурилмаларнинг кетма-кетлик тартиби юқори бўлган ҳолларда

коллокация юзага келади<sup>1</sup>. Коллокацияда қурилманинг умумий маъноси компонентлар ташиган мустақил маъноларга тенг бўлмайди (*мехр булоғи* ≠ *мехр* + *булоқ*). Уларнинг орасида идиомалар (*илон ёғини ялаган, қўли очик*), топонимлар ва алоҳида олинган ономалар ҳам мавжуд бўлиб, бу ҳолда семантик разметкалаш амалиёти бироз мураккаблашади.<sup>2</sup>

Коллокация<sup>3</sup> - бу синтактик ва семантик жиҳатдан яхлитликка эга бўлган бирликлар бўлиб, лексик таркибига кўра қуйидаги турларга бўлинади: таркибий бўлмаган (*несоставные*); таркиби алмашмайдиган (*незаменяемые*); таркиби ўзгармайдиган (*неизменяемые*). Таркибий бўлмаган коллокациялар бу таркибий қисмларининг маъноси умумий мазмундан фарқ қилувчи бирликлардир. Буларга идиомалар ва идиоматик бирликлар киради. Таркиби алмашмайдиган (*незаменяемые*) коллокацияларга метафоралар киради.

Биз таклиф этаётган идиомалар базасида семантик разметка қуйидаги сабабларга кўра амалга оширилди: идиомаларнинг маълум бир қисмларга бўлинмаслиги; таркибидаги лексемаларнинг якка ҳолда мазмун англатмаслиги; маъноси битта луғавий қийматга тенглиги ҳамда шу жиҳати билан яхлит лексик-семантик бутунлик эканлиги ҳисобга олинди. Мазкур разметкалаш амалиётида жаҳон стандартларига ҳам амал қилинди (маълумки, 2011 йилда Google компанияси маълумотлар разметкасининг **Data Vocabulary** стандартини ишлаб чиқди. 1995 йилда кутубхона ва музейшунослик иши мутахассислари гуруҳи томонидан 15 асосий ва 3 та қўшимча хусусиятларга эга Dublin Core – Дублин ядроси (<http://dublincore/core/org/>) ишлаб чиқилган).

Турли типдаги идиомаларнинг қисмлари дистинктор (фонемалар), фиксатор (морфемалар), номинатор (сўз ва бирикмалар), коммуникаторлар (ифодалар) таъсирида турли лингвистик мақом касб этади. Бу жиҳат улардаги семантик сиғим, семантик вариативликни шартлайди ҳамда идиоматик қурилмалардаги кумулятив, эвфемик, эзотерик, тавсифловчилик, киноявийлик вазифаларини фаоллаштиради. Идиомалар структур хусусияти уларнинг лисоний бирликлар маъно мундарижасининг транспозицияси билан боғлиқ равишда юзага чиққанидан,<sup>4</sup> семантик разметка бир қадар мураккаблашади.

Ўзбек тилининг Миллий корпусида идиомалар базасини яратиш ўзбек тилидаги идиоматик қурилмаларнинг бир вақтнинг ўзида инглиз, рус ва бошқа тиллардаги муқобилини беришни шартлайди, ваҳоланки, идиомаларни ўзга тилга таржима қилиш уларга хос бўлган лисоний хусусиятларни инобатга олмаслик билан баробардир.

Ушбу бобда, шунингдек, идиомаларни автоматик аниқлаш методларига доир қарашлар муҳокама қилинди. Хусусан: латент-семантик таҳлил (Latent semantic analysis); параллел матнларни тенглаштириш (Bitext word alignment);

<sup>1</sup> Боярский К. К. Введение в компьютерную лингвистику. – СПб: НИУ ИТМО, 2013. – 72 с.

<sup>2</sup> Эшмўминов А. А. Ўзбек тили миллий корпусининг синонимлар базаси. Филол. фан. бўйича фалсафа доктори (PhD) дис.автореферати. – Қарши, 2019.

<sup>3</sup> <https://ru/m.wikipedia.org.> wiki>.

<sup>4</sup> <https://www.dissercat.com/content/struktura-i-funktsii-idiom-na-materiale-substantivnykh-ustoichivyykh-slovosochetanii-i-slozhn>.

ўзаро алоқадор ахборотни баҳолаш (Mutual informatsion)<sup>1</sup>. Жумладан, матнларни параллел тенглаштиришда энтропиянинг (юн.έν «ичида» +τροπή «айланиш») мавжуд-мавжуд эмаслиги ҳисобга олинади.

Энтропияни аниқлаш қуйидаги тартибда амалга оширилади:

$$H(T_s|s) = - \sum_{t \in T_s} P(t|s) \log P(t|s)$$

Энтропия идиомалар мавжудлигига ишора қилади.

PDA:

$$pda(S) = x = \frac{\sum_{s \in S} \sum_{d \in D_s} align\_freq(s,d)}{\sum_{s \in S} \sum_{t \in T_s} align\_freq(s,t)}$$

PDАнинг мавжудлиги текшириладиган бирикманинг ўз маънода келаётганлигини исботлайди.

Ўзаро маълумот коэффициенти аниқлаш даврида текшириладиган сўз бирикмаси ва сўз бирикмаларининг МІ (*mutual information*)си уларнинг синоним, антоним, голоним-мероним<sup>2</sup>, гипероним-гипоним муносабатига асосан таққосланади<sup>3</sup>. Жумладан, “O’zbek Milliy Korpusida idiomalar bazasi deponentlanadigan materiallari”<sup>4</sup>да бу маълумот қуйидаги кўринишга эга (дастур тури: Web илова (web application)):

- PHP v7.2 (Laravel Framework v6.2)
  - HTML
  - SASS/CSS (Bootstrap v4)
  - Javascript (Vuejs v2.5, JQuery v3.4)
- Қўлланилган маълумотлар базаси:
- MySQL v5.5

Матн жуфтликларини разметкалаш ўзига хос жуда фойдали лингвистик манба ҳисобланади. Уни қўл билан бажариш анча қиммат ва машаққатлидир. Шунинг учун тадқиқотчилар матнни автоматик бирлаштиришнинг турли усулларини ишлаб чиқмоқдалар. Ушбу усулларни икки гуруҳга бўлиш мумкин: қўшимча лингвистик билимларни жалб қилишга асосланган усуллар, масалан (Kayand Roscheisen, 1993, Kit *et al.*) ва статистик усуллар (Brown *et al.*, 1991; Gale and Church, 1991; McEnery and Oakes, 1996; Mikhailov). Бунда фарқ асосан ишлатилган маълумотлар турига боғлиқ бўлади.

Анъанага кўра параллел матнларни бирлаштириш учун статистик усуллардан фойдаланилади. Улар бир тилдаги матнлар узунлиги ва бошқа тилдаги матнлар узунлиги ўртасида сўзлар ёки белгилар билан ўлчанадиган

<sup>1</sup> Шильнов, А. Г., Миллер, Д., Митчелл, П. Д., & Смокотин, В. М. (2020). Сравнительный анализ идиом, представленных в российских учебниках английского языка, и идиом, представленных в современных аутентичных источниках. *Язык и культура*, (49), 157-174. <https://doi.org/10.17223/19996195/49/10>

<sup>2</sup> Масалан, *translation theory* термини таржима назариясининг кичик бўлимларига оид терминларга нисбатан голоним, *partial theories of translation* термини эса *translation theory* га нисбатан мероним ҳисобланади. Қаранг: [https://studwood.ru/712345/literatura/giperonimy\\_giponimy](https://studwood.ru/712345/literatura/giperonimy_giponimy).

<sup>3</sup> <https://www.slideshare.net/msucsai/ss-5430928>.

<sup>4</sup> O’zbek Milliy Korpusida idiomalar bazasi deponentlanadigan materiallari. Huquq egasi Begmatova Gulchehra Xolmurodovna. Mualliflar: Begmatova Gulchehra Xolmurodovna. Ulug’ov Anvar Abduraimovich. – 2019.

тахминий мувофиқликка асосланади. Шунинг учун, битта тилдаги матнларнинг узунлигига қараб, бошқа тилда матннинг мос парчалари қаердан бошланиши кераклигини ҳисоблаш лозим. Фақатгина нисбатан кўпроқ мос келадиган мувофиқликни танлаш қолади. Бу усуллар фақат сўзлар сони ёки белгилар сонини ҳисоблашга асосланган тезкор усуллар ҳисобланади. Бироқ бу усулларнинг ҳам камчиликлари бор. Таржима сезиларли даражада қисқа ёки узун маълумотдан иборат бўлиши биланоқ (мувофиқликнинг бир нечта вариантлари юзага чиққанда) мавжуд алгоритмлар маълумотларнинг ниҳоятда чекланганлиги сабабли қарор қабул қила олмайди.

**Натижаларни тақдим этиш муаммоси.** Йирик корпусларда илгари аҳамиятсиз бўлган ўзига хос муаммо мавжуд: сўровларни қидириш юзлаб ёки ҳатто минглаб натижаларни келтириб чиқариши мумкин, бироқ уларни чекланган вақт ичида кўриш имконсиздир. Ушбу муаммони ҳал қилишда қидирув натижаларини гуруҳлаш ва уларни автоматик равишда кластерлаш ёки уларнинг аҳамиятини статистик баҳолаш билан энг барқарор ибора (коллокациялар)ларни қайта ишлашга имкон берадиган тизимлар ишлаб чиқилмоқда.<sup>1</sup>

Тил корпуси негизида: сатрлар – идиомалар, устунлар – ҳужжатлар кўринишидаги сийрак матрица курилади. Яъни

$$t \frac{T}{i} \rightarrow \begin{bmatrix} x_{1,1} & \dots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{m,1} & \dots & x_{m,n} \end{bmatrix}$$

$t_i^T t_p$  — икки идиома (терма) ўртасидаги корреляция.

**Корпусдан коллокациялар** (сўз бирикмасига тенг яхлит синтактик курилмалар)ни қидиришга мўлжалланган вазифалар. Ушбу турдаги вазифаларда коллокациянинг корпус қидирувидан фойдаланилади: сўз бирикмалари уларнинг қўлланилиш частотаси ва коллокатив белгининг даражасига кўра, шунингдек, матннинг маълум қисмини танлаш билан, ҳам нутқнинг барча қисмларига нисбатан бир вақтнинг ўзида автоматик равишда танланади. Қидирувда матннинг жанровий хусусиятини эътиборга олиш лозим: коллокацияларнинг фаоллик даражаси матннинг турига қараб ўзгариб боришини талабалар амалда кузатишлари шарт.

Биз тақлиф этаётган базанинг ютуқли имкониятлари: идиоматик бирликлар турли йилларда яратилган луғатлардаги фразеологик бирликлар орасидан ажратиб олингани; идиомаларни йиғишда асосан бадий адабиёт матнларига таянилгани ва уларга идиоматик бирликнинг маъновий куршови ҳамда миллий-маданий тажрибадан келиб чиқиб баҳо берилгани; идиомалар диалектал бирликлар, эвфемизмлар орасидан ажратиб келтирилгани; идиомаларнинг инглиз тили, рус тили, баъзи туркий тиллардаги муқобиллари

<sup>1</sup> Корпусная лингвистика как раздел языкознания. — Текст : электронный // Myfilology.ru – информационный филологический ресурс : [сайт]. – URL: <https://myfilology.ru/177/korpusnaya-lingvistika-kak-razdel-yazykoznaneya/> (дата обращения: 5.02.2021).

тақдим этилгани; идиоматик бирликлар вариантлари, синонимлари билан берилгани (қисман бўлса ҳам); идиомаларнинг махсус веб-илова (idioms.uz) онлайн базаси яратилгани; онлайн база (idioms.uz) электрон почта орқали фикр ва мулоҳазаларни тўғридан-тўғри веб-илованинг ўзида олиш имконига эгаллиги билан белгиланади.

## ХУЛОСА

1. Ҳар бир миллатнинг ўз идиомалари мавжуд бўлиб, улар бошқа тилларга айнан таржима қилинмаслиги, аммо бошқа тилларда шундай ҳолатларга мос муқобиллари мавжудлиги билан аҳамиятлидир.

2. Идиомалар миллий хосланган иборалардир, аммо улар ибора каби таркибий ўзгартириш, айрим сўзларни алмаштириш, кенгайтириш ва ибора таркибини қисқартириш каби “деформация”га учрамаслиги билан бошқа фразеологизмлардан фарқланади.

3. Идиомаларнинг пайдо бўлишида маълум тарихий, ижтимоий ва иқтисодий ҳодиса прототип воқеа вазифасини бажаради ва кейинчалик воқеа унутулиб, идиома шу тарзда қотиб нутқ жараёнида ишлатилади.

4. Идиоматик ифодалар ўзбек тилидаги барча изохли луғатларда мавжуд, аммо бу луғатларда улар идиома сифатида эмас, балки фразеологик бирлик, ибора сифатида шарҳланган.

5. Тил корпусидаги идиомалар базаси таълимга, айниқса тил таълимига доир қуйидаги типик хатоларни бартараф этиш имконини беради: тақдим этилган сўзлар қаторидан сўзларни тўғри танлашни билмаслик; турғун бирикмаларни (ўз ва кўчма маъноли), фразеологик бирликларни билмаслик, машҳур иборалардан беҳабарлик.

6. Биз таклиф этаётган идиомалар базасида семантик разметка идиомаларнинг маълум бир қисмларга бўлинмаслиги; таркибидаги лексемаларнинг барчаси яқка ҳолда мазмун англатмаслиги ва албатта кўчма маънода келиши; маъноси фразеологик бутунликлар сингари битта луғавий қийматга тенглиги ҳамда шу жиҳати билан яхлит лексик-семантик бутунлик эканлиги ҳисобга олиниб амалга оширилди. Мазкур разметкалаш амалиётида жаҳон стандартларига амал қилинди.

7. Биз таклиф этаётган базанинг: идиомалар диалектал бирликлар, эвфемизмлар орасидан ажратиб келтирилганлиги; идиомаларнинг инглиз тили, рус тили, баъзи туркий тиллардаги муқобиллари тақдим этилганлиги; идиоматик бирликлар вариантлари, синонимлари билан берилганлиги (тўлик бўлмаса ҳам); идиомаларнинг махсус веб-илова (idioms.uz) онлайн базаси яратилганлиги; онлайн база (idioms.uz) электрон почта орқали фикр ва мулоҳазаларни тўғридан –тўғри веб-илованинг ўзида олиш имконига эгаллиги мазкур тадқиқотнинг илмий янгилигини белгилайди.

**НАУЧНЫЙ СОВЕТ PhD.03/31.03.2021.Fil.78.04 ПО ПРИСУЖДЕНИЮ  
УЧЕНЫХ СТЕПЕНЕЙ ПРИ ТЕРМЕЗСКОМ ГОСУДАРСТВЕННОМ  
УНИВЕРСИТЕТЕ**

---

**ТЕРМЕЗСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**

**БЕГМАТОВА ГУЛЧЕХРА ХОЛМУРОТОВНА**

**СОЗДАНИЕ БАЗЫ ИДИОМ В НАЦИОНАЛЬНОМ КОРПУСЕ  
УЗБЕКСКОГО ЯЗЫКА**

**10.00.01 – Узбекский язык**

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

**Термез – 2021**

Тема диссертации доктора философии (PhD) по филологическим наукам зарегистрирована в Высшей аттестационной комиссии при Кабинете Министров Республики Узбекистан за B2019.4.PhD/Fil1026.

Диссертация выполнена в Термезском государственном университете.

Автореферат диссертации на трех языках (узбекский, русский, английский (резюме)) размещен на веб-странице Научного совета sandu.uz и Информационно-образовательном портале «ZiyoNet» www.ziynet.uz.

**Научный руководитель:** Махмаримова Шохиста Тухташевна  
доктор филологических наук, доцент

**Официальные оппоненты:** Каримов Суюн Амирович  
доктор филологических наук, профессор  
Маматов Абди Эшонкулович  
доктор филологических наук, профессор

**Ведущая организация:** Ташкентский государственный Университет  
узбекского языка и литературы имени Алишера  
Навои

Защита диссертации состоится « 26 » июля 2021 года в 10<sup>00</sup> часов на заседании Научного совета PhD.03/31.03.2021.Fil.78.04 при Термезском государственном университете. (Адрес: 190111, г. Термез, улица Баркамол авлод, 43. Тел.: по адресу: (876) 221-74-55; факс: (876) 221-71-17; e-mail: termizdu@umail.uz)

С диссертацией можно ознакомиться в Информационно-ресурсном центре Термезском государственном университете (зарегистрированной за № 9). Адрес: 190111, г. Термез, улица Баркамол авлод, 43. Тел.: по адресу: (876) 221-74-55;

Автореферат диссертации разослан « 15 » июля 2021 г.  
(Реестр протокола рассылки № 9 от « 15 » июля 2021 г.)



**Б. Умуркулов**  
Председатель Научного совета по  
присуждению ученых степеней,  
доктор филол. наук, доцент

**А. Эшмуминов**  
Ученый секретарь Научного совета по  
присуждению ученых степеней,  
доктор филологии по филол. наук

**Ш. Ёкубов**  
Заместитель председателя Научного  
семинара при Научном совете по  
присуждению ученых степеней,  
кандидат филол. наук, доцент



## **ВВЕДЕНИЕ (аннотация диссертации доктора философии (PhD) по филологическим наукам)**

**Актуальность и востребованность темы диссертации.** В мировой фразеографической практике проблема классификации идиоматических единиц считается одной из центральных. Исследование этих лингвокультурных единиц, которые являются системной единицей языка, часто встречаются особенно в лексикологии, лексикографии, корпусной лингвистике. Хотя наука XXI века позволяет широко изучать этот вопрос на примере электронных корпусов, национальная лингвистическая природа идиом усложняет работу в составлении словаря и представлении их в национальном корпусе. Следовательно, проблема представления идиоматических устройств в обработанном, лингвистически классифицированном, машиночитаемом формате металингвистических данных остается актуальной.

В соответствии с направлениями мировой лингводидактики, языковое образование осуществляется в первую очередь за счет освоения словарного содержания и тем самым за счет формирования знаний о национальном мировоззрении соответствующего языкового сообщества. В связи с этим невозможность перевода идиом на другие языки, выражающие дух нации, имеющие прагматическое значение, составляющие сущность внутренней формы языка, еще больше обостряет вопрос классификации идиом в корпусе национальных языков. В настоящее время, наряду с развитыми странами мира, практика присваивания и классификации идиом набирает обороты в некоторых тюркских языковых корпусах: это позволит протестировать предложенные теоретические модели идиом, определить семантико-синтаксическую сущность устойчивых соединений, несколько облегчить процесс машинного перевода и автоматического редактирования текстов.

В сегодняшнем времени Узбекистан вступил в новый этап своего развития – период национального продвижения. Нынешнее положение нашей страны, ее роль в мировом сообществе, а также различные факторы, связанные с феноменом глобализма, ставят на повестку дня более серьезный подход к вопросу о государственном языке. Поскольку, «Является естественным, что в нынешний период глобализации каждый народ, каждое независимое государство придает приоритетное значение обеспечению своих национальных интересов, прежде всего, сохранению и развитию своей культуры, исконных ценностей, родного языка»<sup>1</sup>. Поэтому определение «... задачи всестороннего сопровождения научных и творческих исследований, создания для них необходимых условий»<sup>2</sup> в языкознании, наряду со всеми другими научными направлениями, является одной из многих традиционных

---

<sup>1</sup> Ўзбекистон Республикаси Президентининг 2019 йил 21 октябрдаги ПФ-5850-сонли “Ўзбек тилининг давлат тили сифатидаги нуфузи ва мавқеини тубдан ошириш чора-тадбирлари тўғрисида”ги фармони // [www.lex.uz](http://www.lex.uz)

<sup>2</sup> Мирзиёев Ш.М. Эркин ва фаровон, демократик Ўзбекистон давлатини биргаликда барпо этамиз. Ўзбекистон Республикаси Президенти лавозимида киришиш тантанали маросимида бағишланган Олий Мажлис палаталарининг кўшма мажлисидаги нутқ. – Тошкент: Ўзбекистон, 2016. – Б.13.

проблем языкознания, в частности, это гарантирует новое решение проблемы передачи от поколения к поколению национально-культурных кодовых единиц, составляющих «сливки»<sup>1</sup> живого узбекского языка, а также достижение надежных результатов в решении традиционных проблем, связанных с ними.

В соответствии с Указом Президента Республики Узбекистан от 21 октября 2019 года № УП-5850 «О мерах по кардинальному повышению роли и авторитета узбекского языка в качестве государственного языка», обеспечении занимания достойного места государственный язык в мировой информационной сети интернет, в частности информационно-коммуникационных технологий, создания компьютерных программ узбекского языка, повышение авторитета государственного языка как языка науки и Указ Президента Республики Узбекистан от 20 октября 2020 года УП-6084 «О мерах по дальнейшему развитию узбекского языка и совершенствованию языковой политики в нашей стране» а также № УП-4997 от 13 мая 2016 года «О создании Ташкентского государственного университета узбекского языка и литературы имени Алишера Навои», Указ № УП-4947 от 7 февраля 2017 года «О Стратегии действий по дальнейшему развитию Республики Узбекистан», Постановление № ПП-2789 от 17 февраля 2017 года «О мерах по дальнейшему совершенствованию деятельности академии наук, организации, управлению и финансированию научно-исследовательской деятельности», Постановление № ПП-3271 от 13 сентября 2017 года «О программе комплексных мер по развитию системы издания и распространения книжной продукции, повышению культуры чтения», Предложение Президента Республики Узбекистан о внедрении информационных технологий во все сферы и этапы образования и науки, указанное в Послании к Олий Мажлису в 2020 году 24 января, а также данная диссертационная работа в определенной степени служит выполнению задач, поставленных в других нормативно-правовых актах, связанных с этой деятельностью.

**Соответствие актуальности исследования с учетом приоритетов развития науки и техники республики.** Диссертация выполнена в соответствии с учетом приоритетов развития науки и техники республики I. «Формирование системы инновационных идей и способов их реализации в социальном, правовом, экономическом, культурном, духовном и образовательном развитии информированного общества и демократического государства».

**Уровень изучения проблемы.** Теоретические взгляды таких ученых, как Ф.Ф.Фортунатов, Ф.И.Буслаев, Ш.Балли, А.А.Шахматов, Е.Д.Поливанов, В.В.Виноградов, С.И.Абакумов, Г.К.Данилов, Г.О.Винокур, В.Н.Державин, А.И.Ефимов, А.Я.Рожанский, И.Е.Аничкин, Л.П.Смит, В.П.Жуков, В.Н.Телия, Н.М.Шанский, Н.Н.Амосов, А.М.Бабкин, В.М.Савицкий,

---

<sup>1</sup> Саломов Ғ.Т. Рус тилидан ўзбек тилига мақол, матал ва идиомаларни таржима қилиш масаласига доир. – Тошкент, 1961. – Б. 3.

М.Л.Ковшова, Е.Д.Балыгина, О.В.Елисеева, М.А.Ганюшина, С.Е.Максимец, Т.В.Кара-Казарьян, О.В.Куманок, Р.Саидов, Г.Ниннин, Т.С.Талабов, Е.М.Маклакова, Н.У.Ворокова, Н.В.Щенникова, Н.В.Баско, В.М.Мокиенко, Е.М.Гирняк, Т.С.Борщевская, И.Е.Аничков, К.Н.Мочалина, Н.Елисеева, Н.Горобец, А.Ö.Aksoy, Л.Э.Бинович, узбекских лингвистов как Ш.Рахматуллаев, А.Ходжиев, А.Маматов, М.Рузикулова, Б.Юлдашев, Д.Худайбергана, Ш.Усманова составили общую методологическую основу темы.

Теоретической основой для исследования, связанного с корпусной лингвистикой послужили взгляды В.П.Захарова, В.Плунгяна, В.В.Мамонтовой, О.В.Кукушкиной, А.А.Поликарпова, Е.В.Суровцевой, также, лексическая обработка текста и лингвостатистический анализ текста основывался на теоретических воззрениях учёных, таких как А.Л.Юсупов, А.Пулатов, С.А.Каримов, Б.Р.Менглиев, С.Мухамедов, М.Айимбетов, С.Мухамедова, Н.Абдурахмонова, Ш.Хамраева<sup>1</sup> и других.<sup>2</sup> Несмотря на лексикографическое и лингвостатистическое изучение языковых единиц в данных исследованиях, предложенные методы и принципы создания узбекского лингвистического корпуса не затрагивают проблему создания базы лингвокультурных единиц, включая идиом.

**Связь дисциплинарной деятельности с планом исследований, проводимым высшим учебным заведением или исследовательским институтом.** Исследование проводилось в рамках научно-исследовательского плана Термезского государственного университета по теме «Узбекское языкознание: анализ художественной речи, антропоцентрическая теория, этнолингвистика, корпусная лингвистика».

**Целью исследования** является разработка теоретических и практических основ введения и обозначения идиом в узбекском национальном корпусе, освещение научных, теоретических и практических основ вопросов разметки, принципов совершенствования базы идиом посредством разметки различных типов коллокаций.

---

<sup>1</sup> Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол.фан.бўйича фалсафа доктори (PhD)...дис. афтореф. – Қарши, 2018.

<sup>2</sup>Мухаммедов С.А. Статистический анализ лексико-морфологической структуры узбекских газетных текстов: Автореф. дисс... канд. фил. наук. – Ташкент, 1980; Бабанаров А. Разработка принципов построения словарного обеспечения турецко-русского машинного перевода: Автореф. дисс... канд. фил. наук. - Л., 1981; Айымбетов М.К. Опыт лингвостатистического анализа лексики и морфологии каракалпакского публицистического текста: Автореф. дисс... канд. фил. наук. – Ташкент, 1987; Каримов С., Қаршиев А., Исроилова Г. Абдулла Қаҳҳор асарлари тилининг луғати. Алфавитли луғат. Частотали луғат. Терс луғат. – Тошкент, 2007; Қўнгуров Р., Каримов С. Зулфия поэзияси тилининг луғати. – Тошкент: Ўқитувчи, 1981; Қўнгуров Р., Каримов С. Зулфия поэзияси тилининг луғати. – Тошкент: Ўқитувчи, 1981; Ризаев С. Ўзбек тилининг лингвостатистик тадқиқи: Фил.фан.док.дис...афтореф. – 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...афтореф. – 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.

### **Задачи исследования:**

исследование методологических воззрений к изучению национального лингвистического характера идиомы;

проверить уровень лексикографической изученности идиом узбекского языка;

проверка межъязыковых различий в классификации идиом;

исследование оптимальных методов изучения идиом в обучении языкам;

исследование предоставления фразеологических помесей – идиом в языковых корпусах на примере лингвистики мирового корпуса;

создание электронной базы идиом в Национальном корпусе узбекского языка.

**Объектом исследования** являются идиомы узбекского языка.

**Предметом исследования** является создание базы идиом узбекского языка в национальном корпусе.

**Методы исследования.** В исследовании использовались такие методы как сравнительный, статистический а также латентный анализ корпусной лингвистики, перекрестные данные, объединение текста.

**Научная новизна исследования заключается в следующем:**

систематизировано содержание исследования и хронология идиоматических сочетаний в мировом языкознании, тюркологии и узбекском языкознании;

выявлено недостаточность лексикографических основ при формировании базы идиом в национальном корпусе узбекского языка, нераздельное описание идиом от фразеологических сочетаний в словарях, неотчетливый принцип основания в комментариях;

выявлены межъязыковые различия в тематической группировке идиом;

обосновано устаревание иллюстративной информации во фразеологических словарях и необходимости внедрения практики использования электронных корпусов при их обновлении;

описаны методы и средства математического анализа семантической разметки и автоматического определения коллокаций;

обосновано преимущества предлагаемого веб-приложения Idioms.uz.

**Практические результаты исследования заключается в следующем:**

Обоснованы роль, языковая сущность и функция оборотов во фразеологической системе языка; формирование лингвокультурологических основ, методов и средств представления идиом узбекского языка в национальном корпусе; роль идиом узбекского языка в развитии культуры речи и их применение; научно-теоретические данные по исследованию идиом послужило для обогащения содержания учебников и учебных пособий, созданных по таким дисциплинам, как современный узбекский язык (лексикология, фразеология, лексикография), лингвокультурология, корпусная лингвистика, общее языкознание.

Основаны теоретические выводы о лингвистической природе идиом, которые служат дальнейшему совершенствованию словарей по узбекской идиоматике, лингвокультурологии, корпусной лингвистики.

**Достоверность результатов исследования** сформулирована на основе методологии лингвистического анализа, в частности, с помощью теоретических источников подробно разъясняется роль изучения идиом в мировой и узбекской лексикографии, фразеологии, мировой корпусной лингвистики и исходя из этого, лингвокультурная позиция уточняется тщательным исследованием, основанным на доказательствах. Это также определяется тем, что теоретические суждения и выводы внедрены в практику, полученные результаты подтверждаются компетентными органами.

**Научно-практическое значение результатов исследований.** Научное значение результатов исследования определяется тем, что данные результаты могут стать научным источником для новых исследований, проводимых в области корпусной лингвистики в узбекском языкознании;

в частности, лингвокультурологическая природа идиом, теоретические данные о лексико-семантической разметки в языковом корпусе способствует развитию лексикологии (семасиологии), фразеологии, лингвокультурологии, корпусной лингвистики.

Разработка веб-приложения и онлайн-формата базы данных идиом узбекского языка, в свою очередь, расширяет возможности использования регулярно пополняемой базы данных этих единиц в научных и практических целях. Научная и практическая значимость результатов исследования приводит к совершенствованию теоретических взглядов к анализу семантических и синтаксических особенностей идиоматических конструкций в узбекском языке и эффективных методов представления их вариантов на иностранных языках в национальном корпусе.

Практическая значимость результатов исследования определяется тем, что материалы исследования могут быть использованы для преподавания, а также для подготовки учебников и методических пособий по таким дисциплинам, как «Современный литературный узбекский язык», «Лингвокультурология», «Корпусная лингвистика» в высших учебных заведениях

**Внедрение результатов исследования.** Материалы о роли лексико-семантической целостности во фразеологической системе узбекского языка, а также тюркских языков были использованы в проекте фундаментального исследования № PZ-20170927147 на тему «Изучение древнетюркской письменности и фольклора до XIII века», под руководством доцента Ташкентского государственного университета узбекского языка и литературы имени Алишера Навои А.Тилова (справка Ташкентского государственного университета узбекского языка и литературы имени Алишера Навои № 04/1-609 от 7 октября 2021 года). В результате достигнуто совершенствования знаний о национальной картины мира и лингвокультурной природе идиоматических конструкций в узбекском языке, а также в тюркских языках;

выводы о результатах исследования о том, что идиомы являются словарными единицами, которые используются в определенном регионе,

имеют национальную специфику и практический характер, по применению к научно-образовательному процессу в области языкознания с точки зрения компьютерной лингвистики, также, по определению общих принципов устройства идиоматических единиц и их значений в информационно-поисковой системе; признании сходства в встречаемости идиоматических единиц в системе словообразования узбекского языка среди тюркских языков (казахский, киргизский, азербайджанский, башкирский, татарский, туркменский) и участия слова в категории существительных и прилагательных вместо «центрального слова»; определению возможности достижения быстрого поиска информации в языковом корпусе путем определения того, что или как выражается локальный компонент содержания при лексикографической классификации идиоматических конструкций были использованы в научно-практическом проекте выполненном в Каракалпакском научно-исследовательском институте гуманитарных наук Каракалпакского отделения Академии наук Республики Узбекистан № FA-F1-G003 на тему «Функциональное словообразование в современном каракалпакском языке» (Справка Каракалпакского отделения Академии наук Республики Узбекистан № 35/1 от 22 февраля 2021 года). В результате достигнуто совершенствование знаний в подходе к изучению компьютерной лингвистики и взглядов на развитие лингвистической компетенции в области высшего образования;

разработана база данных идиом узбекского языка; теги узбекских идиом были применены для создания онлайн-версии базы данных идиом на основе лингвистической модели (свидетельство Агентства интеллектуальной собственности при Министерстве юстиции Республики Узбекистан № DGU 08092 от 5 мая 2020 года). Результатом является онлайн-версия базы данных идиом.

Роль идиом в узбекской культуре речи, выводы о проявлении лингвокультурных особенностей в живой речи использованы при составлении сценариев передач для Сурхандарьинской областной телерадиокомпании (справка Сурхандарьинской телерадиокомпании от 17 мая 2021 года № 01-01/59).

**Апробация результатов исследования.** Результаты данного исследования были опубликованы в 2 международных и 3 научных журналах, включенных в список ВАК, 1 республиканской научно-практической конференции и 1 в журнале, индексируемом в базе Scopus.

**Публикация результатов исследований.** Основное содержание диссертации отражено в 1 авторском свидетельстве, 10 научных статьях (3 из них - научные журналы, включенных в список ВАК Республики Узбекистан и 2 международных журнала и 2 международных конференциях) и тезисах.

**Объем и структура диссертации.** Диссертация состоит из введения, трёх основных глав, заключения и списка использованной литературы. Общий объем диссертации составляет 130 страницы.

## ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

**Во вводной части** изложена актуальность и необходимость исследования, описаны цель, задачи, объект и предмет исследования, показана его соответствие приоритетным направлениям развития науки и технологий Республики, изложены научные новшества, практические результаты исследования, раскрыта научно-практическая значимость полученных результатов, приведены сведения о внедрении результатов исследований, структуре и объеме диссертации.

В первой главе диссертации, озаглавленной «**Лингвистическое описание идиоматических конструкций в мировой и узбекской лингвистике**», представлены научные и теоретические подходы к идиомам в мировой лингвистике, комментарии к идиоматическим конструкциям в узбекском языкознании и их различию, описание идиом в сравнительном языкознании, обсуждение тематических классификаций идиом.

Каждый владелец языка имеет свои собственные языковые единицы, отражающие их культуру и менталитет, и проблема национализированных языковых единиц стала углубленно изучаться лингвистами в начале двадцатого века. Идиомы, описанные профессором Г.Т. Саломовым как «чудо языка, сливки»,<sup>1</sup> привлекли внимание исследователей еще в 40-х годах XX века. К этому времени, фразеологические исследования в мировой лингвистике стали более лексикографическими. В частности, на первый план выдвинуто отражение фразеологии в словарях, трактовка ее этимологии. По словам проф. Б. Юлдашева, период развития фразеологии был изучен условно разделив на три периода<sup>2</sup>:

**Первый период** – с середины XVIII века до 30-х годов XX века. Среди исследователей этого периода известны взгляды М.В. Ломоносова, В. Даля, А. А. Потебни, Ф. Ф. Фортунатова, Ф. И. Буслаева, Ш. Балли, А. А. Шахматова.

**Второй период** – 30-е и 50-е годы XX века, в этот период широкое признание получили научные идеи Е.Д. Поливанова, В.В. Виноградова, С.И. Абакумова, Г.К. Данилова, Г.О. Винокур, В. Державина, А.И. Ефимова, А.Я. Рожанского, И. Е. Аничкина и др.<sup>3</sup>

**Третий период** – охватывает период с 1960-х годов до наших дней. Фразеологические исследования этого периода характеризуются применением новых методов, бурным развитием теории фразеологии и появлением множества специалистов фразеологии<sup>4</sup>. В результате

---

<sup>1</sup>Саломов Г.Т. Рус тилидан ўзбек тилига мақол, матал ва идиомаларни таржима қилиш масаласига доир. - Тошкент, Ўзбекистон ССР Фанлар Академияси нашриёти, 1961. – Б. 3.

<sup>2</sup>Йўлдошев. Б. Ўзбек фразеологияси ва фразеографиясининг шаклланиши ҳамда тараққиёти. Монография. Тўлдирилган ва тузатилган иккинчи нашри. – Самарқанд: СамДУ нашри, 2013. – Б. 4

<sup>3</sup>Винокур Г.О. Избранные работы по русскому языку: сборник научных трудов / Г.О.Винокур; Российская Академия Наук, Институт русского языка им. В.В.Виноградова. – 2-е изд., стер. – Москва: ФЛИНТА, 2016. – 780 с. – (Стилистическое наследие).

<sup>4</sup>Бу ҳақда қаранг. Йўлдошев. Б. Ўзбек фразеологияси ва фразеографиясининг шаклланиши ҳамда тараққиёти. Монография. Тўлдирилган ва тузатилган иккинчи нашри.- Самарқанд: СамДУ, 2013. – Б. 4; Ройзензон Л. И., Малиновский Е. А., Хаютин А. Д. Очерки по истории становления фразеологии как

формирования фразеологии как сферы ее компоненты также были привлечены к процессу углубленного исследования. В частности, фразеологические скрещивания – идиомы привели к появлению большого количества научных исследований.

Во фразеологии существуют методы компонентного анализа, идентификации, аппликации и комбинирования для разделения, анализа, изучения и классификации фразеологических единиц. Однако невозможно полностью применить эти методы к идиоматическим единицам, потому что идиомы не разделены на компоненты.

До 40-х годов XX века взгляды на фразеологическую структуру языка можно было найти только в трудах В. Виноградова, Ф. Фортунатова, А.А. Шахматова, А. Булаховского, И. Срезневского, А. Потебни, С. Абакумова<sup>1</sup>. Однако затем вес работы в этой сфере увеличился. Тем не менее во всех чувствовалась работа по общему правилу, пользоваться определениями В. В. Виноградова, А. А. Шахматова. В исследованиях, проведенных в этот период, первостепенное значение имела идея о том, что фразеологическая скрещивание является частью фразеологии и что она сама по себе имеет первостепенное значение. Следует отметить, что в мировой лингвистике этого периода упор делался на выделение идиом среди элементов фразеологического пласта. В частности, С. Тарабрина отмечает, что термин впервые был использован английским лингвистом Л. П. Смитом, вместо французского фразеологизма, который иностранные лингвисты называют *идиомами*<sup>2</sup>. Но сам Л.П. Смит, а также В.П. Жуков, В.Н. Шанский и другие ученые считали все устойчивые соединения фразеологическими.

Хотя фразеология возникла как самостоятельный раздел лингвистики, она по-прежнему рассматривается как раздел лексикологии. Однако Н. Н. Амосов, А. М. Бабкин, А. И. Смирницкий считают, что в этот раздел входят только соответствующие группы.

В русском языкознании существует классификация фразеологических единиц по нескольким критериям: 1) классификация по составу фразеологического оборота; 2) классификация по структуре фразеологического оборота.

Классификация фразеологического оборота по их составу была предложена Н.М. Шанским, который разделил фразеологические обороты на две группы: 1) фразеологизмы, относящиеся к активной лексике; 2) фразеологизмы, относящиеся к устаревшим или диалектическим<sup>3</sup>.

---

лингвистической дисциплины. – Самарканд, 1975. – С. 5-70; Малиновский Е.А. История изучения проблем фразеологии в русском языкознании. – Ташкент: Фан, 1991. – С. 5-92; Малиновский Е.А. Формирование русской фразеологии и фразеографии. – Алматы, АДД, 1995. – С. 3-52; Малиновский Е.А. Формирование русской фразеологической теории в XX веке. Монография. – Самарканд, 2007. – С. 4-131.

<sup>1</sup>Фортунатов Ф.Ф. Сравнительное языковедение. Общий курс. В кн: «Избранные труды», т 1. – М., Учпедгиз, 1956. – С. 83.

<sup>2</sup>[www.dissercat.com/](http://www.dissercat.com/) С. Ю.Тарабрина Семантика фразовых глаголов и глагольных единиц. Диссертация кандидата наук. 2003. Яна қараңг: Смит Л. Фразеология английского языка. – М., 1959. – С.10.

<sup>3</sup>Шанский Н.М. Фразеология современного русского языка. – М., 1963.



Ш.Балли разделил фразеологизмы на две основные группы: фразеологический ряд или обычная фразеологическая группа и фразеологическая единица. По мнению Ш. Балли, существует некоторая свободная связь между частями фразеологизмов первой группы, а связь между единицами фразеологии следующей группы настолько сильна, что содержание между ними стало единым целым<sup>1</sup>.

По словам И.Ю. Третьяковой: «Фразеологизмы трансформируются для различных коммуникативных целей, в том числе: определение содержания, определение фразеологизма в данной языковой ситуации, субъективная оценка».<sup>2</sup> В нормативном употреблении фразеологизмы есть нормальное применение, в котором нет отклонений от нормы, а в эпизодическое употребление фразеологизмов вносятся некоторые инновации. Отклонение от такой нормы – методический приём.<sup>3</sup>

В узбекском и мировом языкознании объект изучения фразеологии понимается в широком и узком смысле. Те, кто понимает объект фразеологии в узком смысле интерпретируют фразы отдельно от пословиц и притч.<sup>4</sup>

Если первая группа ученых (Л.П. Смит, В.П. Жуков, В.Н. Телия, Н.М. Шанский), понимавших понятие *фразеологизм* в широком смысле, включала во фразеологию все виды устойчивых соединений, то Н.Н. Амосов, А.М. Бабкин, А.И. Смирницкий рассматривают только определенную группу принадлежащему к фразеологизму. Академик В.В. Виноградов не включает в фраземы пословицы, поговорки и известные слова (крылатые слова). Если под названием идиома Д.О.Добровольский цитировал только фразеологические единицы фразеологизмы, то в определении В. В. Виноградова фразеологизмы делятся на три группы: А) фразеологические скрещивания; б) фразеологическая целостность; в) фразеологические единицы.

Следует отметить, что интерпретация идиом и идиоматических единиц во фразеологии имеет периодическое значение. В частности, такие исследователи, как Ф.Ф. Фортунатов, В.М. Савицкий, М.Л. Ковшова, Е.Д. Балигина, О.В. Елисеева, М.А. Ганюшина, С.Е. Максимец, Т.В. Кара-Казарян, О.В. Куманок, Р.Саидов, Г.Ниннин, Т.С. Талабов, Е.М. Маклакова, Н.В. Н.В. Баско, В.М. Мокиенко, Е.М. Гирняк, Т.С. Борщевская, И.Е. Аничков, К.Н. Мочалина, Н.Елисеева, Н.Горобец, в своих работах о соотношении фразеологии и идиоматики, управленческой и экономической терминологии идиомы, которые изучали восприятие идиом детьми на

<sup>1</sup>Балли Ш. Французская стилистика. – М., 2001. – С. 89-90.

Яна қаранг: Колесникова А.А. Русские и английские идиомы мотивированные сценариями общеевропейской культуры( сравнительный анализ идиом со словами devil, hell, чёрт, ад и их квазисинонимами. Магистерская выпускная квалификационная работа. – С. Петербург, 2016. – С.6.

<sup>2</sup>Третьякова И.Ю. Оказиональные преобразования фразеологизмов с компонентом «душа» // Фразеологизм и слово в национальнокультурном дискурсе (лингвистический и лингвометодический аспекты): Междунар. науч.-практ. Конф., посвящ. юбилею д. ф. н, проф. А. М. Мелерович (Кострома, 20–22 марта 2008 г.). – М.: ООО Элпис, 2008. – С. 270.

<sup>3</sup>Гирняк Е.М. Идиоматичность рекламного текста (На материале русской и китайской рекламы) Вестник Пермского университета Российская и зарубежная филология. 2011. Вып.1(13). – С. 34.

<sup>4</sup>Раҳматуллаев Ш. Нутқимиз кўрки. – Тошкент, 1970. – Б.30-31, 36.

нейролингвистическом уровне, рассматривали их в политическом и экономическом аспектах и целесообразно включать в группу ученых понимающих фразеологизм в узком смысле.

Языковед Ж.М.Ахмедова также утверждает, что сторонники узкого понимания фразеологии приняли эту границу только в контексте идиомы (фразеологических скрещиваний)<sup>1</sup>. Таким образом, фразеологию уместно понимать, как область изучения устойчивых соединений, а область идиоматики как часть ее структуры.

По поводу предмета фразеологии В.Н.Телия пишет: «Некоторые ученые включают во фразеологию только идиомы и фразеологические единицы, а другие считают пословицы и поговорки частью фразеологии. У всех этих единиц есть одна общая черта: они состоят из нескольких слов, и потребность в них в общении очень велика».<sup>2</sup>

О. Ахманова называет область изучения идиоматических единиц *идиоматикой*, сами идиомы *идиоматизмом*, *идиоматическим выражением* и показывает названия этих единиц англ. *idiom, idiomatic expression*, фр. *idiome, idiotisme*, нем. *idiom, idiotismus*, исп. *Idiomatismo*. Ученая считает идиомы «... языковой единицей, которая по своей синтаксической и семантической природе отражает уникальные особенности языка, отражая национальную идентичность конкретного народа»<sup>3</sup>.

Г.Х. Ахунзянов подчеркивает, что термин «фразеология» многогранен и многозначен, поэтому его следует заменить терминами «идиоматика», «идиомалогия»<sup>4</sup>. Ученые делят идиомы на 1) целостные идиомы; 2) аналитические идиомы.

В монографии «Современная азербайджанская лексика» С. Джафаров делит устойчивые словосочетания на пять групп: идиомы, мудрые изречения, словосочетания, пословицы и поговорки<sup>5</sup>. Были и недостатки в классификации С. Джафарова, так как в этой классификации перифразы и описательные выражения даны в смешении. На наш взгляд, С. Джафаров из числа тех, кто понимает идиому в самом широком смысле, создал классификацию в своей интерпретации.

Классификация идиом варьируется в зависимости от внутренних характеристик разных языков. Это напрямую влияет на их отражение в корпусе. Например, Ф.П. Кауи и Р. Маккина разделили идиомы на 4 группы в соответствии с их семантическими и синтаксическими особенностями: специальные идиомы (*to blow the guff*); метафорическая идиома (*to beat one's breast*); полуидиомы (*a blind alley*); открытые коллокации (*heavy rain*).<sup>6</sup>

<sup>1</sup> <https://dissercat.com/> Ж.М. Ахмедова. Фразеология ахвахского языка: На материале ратлубского, цегобского и тлянубского говоров. Диссертация кандидата наук. 2004.

<sup>2</sup> Телия В.Н. Что такое фразеология. – М.: Наука, 1966.

<sup>3</sup> Ахманова О. С. Словарь лингвистических терминов. – М.: Сов. энциклопедия, 1966. – С. 161.

<sup>4</sup> Ахунзянов Г. Х. Идиомы (исследование на материале татарского языка): Автореф. дис. канд. филол. наук. – Казань, 1974.

<sup>5</sup> Жаьфаров С. Хозирги замон озарбайжон тили лексикаси. (озарбайжон тилида). – Боку, 1958.

<sup>6</sup> Cowie A.P., Mackin R. Oxford Dictionary of Current Idiomatic English. - Vol. 2. Phrase, Clause and Sentence Idioms. - Oxford University Press, 1983.

Известный английский лингвист М.Маккарти использует несколько терминов для классификации идиом. Это: 1) *opaque idioms* – неясные, расплывчатые идиомы; идиомы, которые не проистекают из содержания общих компонентов содержания (*to kick the bucket*); 2) *semi-opaque idioms* – наполовину непонятные, наполовину расплывчатые. То есть идиомы, которые можно переименовать (сравн.: *to pass the buck* и *to pass the responsibility*); 3) *transparent idioms* – прозрачные идиомы. Как идиомы (*to see the light*)<sup>1</sup> которые нетрудно понять.

Идиомы, состоящие из вульгарных единиц, можно найти в словаре В.Буя «Русская заветная идиоматика: весёлый словарь крылатых выражений»<sup>2</sup>, который содержит около 1000 устойчивых единиц ругательной, странной речи.

Фразы можно менять по идее произведения искусства. Авторы обычно не довольствуются выбором и применением фраз в соответствии с назначением изображения. Возможно, он адаптирует и переделывает их в соответствии с характером, настроением и образом жизни главных героев. Например, О.Аметова отмечает, что Юсуф Хос Хаджиб использовал имеющиеся в языке готовые фразеологические выражения, чтобы сделать язык произведения живым и ярким, и во многих случаях смог проявить себя как создатель новых фразеологизмов<sup>3</sup>. Таким образом, фразы могут претерпевать следующие изменения: 1) заменяются некоторые слова во фразе; 2) структура фразы расширяется; 3) содержание фразы сокращается.<sup>4</sup> В результате такой обработки оборот встречаются: появление новой интерпретации в смысле; изменения лексического содержания; расширение семантико-стилистических функций; добавление новых образных и переносных значений к содержанию, это также называется "деформацией"<sup>5</sup>. Но это условие не соблюдается в идиомах, этим он отличается от других типов выражений. Однако, по словам Ж.Осмонова, несмотря на то, что компоненты идиомы являются семантическими и синтаксическими, они меняются в формообразующих аффиксах. Например, *сирке+си суу котарбойт* – *сирке+м суу котарбой*<sup>6</sup>. Судя по всему, состав идиом монолитен, и их важная особенность состоит в том, что они не меняются так сильно, как в выражениях.

Английский лингвист Р. Мун делит английские идиомы на пять функциональных групп: 1) информационные идиомы (*to catch sight of something, for sale*); 2) оценочные идиомы (*kid's stuff, a different kettle of fish*); 3) ситуативные идиомы (*long time no see, knock it off*); 4) модальные идиомы

<sup>1</sup> McCarthy M. Vocabulary. - Oxford University Press, 1990.

<sup>2</sup> <https://www.livelib.ru>>100 В.Буй. Русская заветная идиоматика: весёлый словарь крылатых выражений.

<sup>3</sup> Аметова О. "Кутадгу билик"да фразеологизмлар // Ўзбек тили ва адабиёти. №1. 2020. – Б. 114.

<sup>4</sup> Жумабоева С.М. Ўткир Ҳошимов асарларида фразеологизмлар. Битирув малакавий иши. – Тошкент: ЎзМУ, 2014. – Б.14.

<sup>5</sup> Шомаксудов А ва бошқ. Ўзбек тили стилистикаси. –Тошкент, 1974. – Б. 70.

<sup>6</sup> Осмонова Ж. Идиомы в киргизском языке: Автореферат на соискание уч.степени канд. наук. – Фрунзе, 1969. – С.5.

(*I kid you not, you know what I mean*); 5) организационные идиомы (*by the way, for instance*)<sup>1</sup>.

Другой ученый, изучавший идиомы, А.Маккай, изучал идиомы отдельно и разделил их на две части: идиомы лексемного уровня (*lexemic stratum*); идиомы уровня семемы (*sememic stratum*). Идиомы на уровне лексем включают: - *come up with* типовые фразовые глаголы; - *to kick the bucket* типовые фигуры; биномы (*here and there*); фразеологические идиомы (*White House*); идиомы составных глаголов (*to baby-sit*). Ученый подчеркивает семантико-зависимый характер идиом на уровне семемы, которая определяется в номенклатуре. По его словам, эти идиомы, получившие название «культурный-прагмема», выявляются при анализе текста и определяется их конвенциональный характер.<sup>2</sup>

Вообще в лингвистике сегодня «Какое направление и метод следует выбрать, чтобы раскрыть природу идиом?», «Следует ли понимать их как единое целое на основе общего содержания или они должны основываться на национальном колорите, национально-ментальной адаптации?» ответы на эти вопросы ищутся. Существующие описания идиом доказывают, что фразеологизм обладает следующими характеристиками для перехода к качеству идиомы: он уникален для одного языка и не может быть переведен точно на другой язык; иметь национальный колорит; кардинально отличается от содержания компонентов значения идиомы; состоит из двух или более слов и словосочетаний, синтаксически равных одной части, равной одному слову в предложении; наличие альтернатив на других языках (но важно, чтобы их нельзя было перевести буквально).

Вклад таких ученых, как Ш. Рахматуллаев, А. Маматов, Б. Юлдашев, А. Исаев, М.Рузикулова, К.Бозорбоев, У. Рашидова, Ш. Ганиева, Ш. Алмаматова, Б. Болтаева, в развитии узбекской фразеологии невосомая.

В 1960-е годы обострилась проблема классификации фразеологических единиц. В частности, хотя термин *идиома* не упоминается в работах таких лингвистов, как Ш. Рахматуллаев, У. Турсунов, А. Маматов, Ф. Исхаков, К. Бозорбоев, Д. Абдуллаев, М. Вафоева, Ш. Алмаматова, М. Холикова, Э. Умаров, Я. Пинхасов по сути, было проведено масштабное изучение фраз узбекского языка.

Одним из исследований идиоматики в узбекском языке является диссертация М.Рузикуловой «Идиомы в узбекском языке» (1966), которая является одной из ранних работ. В нем обсуждается природа идиомы и ее место между фразеологизмами и делаются выводы. Фразеологические единицы в узбекском языке делятся на два типа по признакам значения: а) фразеологическая целостность; б) фразеологическое скрещение.

В 1960-е годы диалектный фразеологизм привлек внимание исследователей. В частности, в исследовании под названием «Заметки о

<sup>1</sup> [https:// booknlime.ru](https://booknlime.ru). Петросян И.В.История изучения фразеологизмов и идиом.

<sup>2</sup><https://www.dissercat.com/content/struktura-i-funktsii-idiom-na-materiale-substantivnykh-ustoichivyykh-slovosochetanii-i-slozhn> Ерохина Н.В. Структура и функция идиом; На материале субстантивных и устойчивых словосочетаний.

«ж»фицированной (кипчакский) диалекте фразеологии узбекского языка» Х.Дониёрова<sup>1</sup> фразеология анализировалась как фраза, но не изучалась как идиоматическая единица.

Кроме того, лингвисты Т. Содиков, Х. Узаков, Н. Раджабов, М. Валиева, Б. Файзуллаев, А. Ишаев изучали диалектную фразеологию различных региональных диалектов в Узбекистане. Однако, по словам профессора Э.Бегматова, «фразеологизм народного языка, особенно узбекская диалектная фразеология, еще глубоко не изучена».<sup>2</sup> Идиомы как единица в структуре фразеологии в этих исследованиях специально не упоминаются.

В этой главе особенности фразеологизмы тюркских языков анализируются на примере фразеологических единиц и идиом.

Во второй главе, озаглавленной «**Лексикографическая основа для создания базы данных идиом в корпусе национального языка**», были затронуты вопросы научной и практической значимости представления в национальном корпусе тематические группы идиом, идиомы в толковых словарях мировых языков, языковой корпус и лексикографические источники, лексические единицы в мировых языковых корпусах (фраземы и идиомы), идиоматические единицы. В частности, изучение случаев, в которых идиомы появляются в языковой системе, привело лингвистов всего мира к разделению этих устойчивых соединений на различные тематические группы. В частности, идиомы, выражающие религиозные принципы (стихи Священного Корана, идиомы, состоящие из взглядов шариата, принадлежащих к мусульманской вере; идиомы, связанные с библейскими сюжетами - библеизмы); идиомы, относящиеся к определенной творческой деятельности; идиомы, представляющие соматические концепции; идиомы с «животным» смысловым компонентом; идиомы, представляющие предметы домашнего обихода; идиомы, представляющие названия еды, еды; идиомы на мифологической основе.

Поскольку выражения или фиксированные соединения образуют семантически неделимую единицу, необходимо уделить особое внимание вопросу учета этого в процессе автоматической обработки текста в лексикографии. На основе материала корпуса можно использовать статистический метод, чтобы определить, какие слова всегда используются вместе, в результате чего определяется, насколько они связаны со стабильным составом.<sup>3</sup> В идиомах, однако, эта ответственность еще больше усиливается, потому что идиома - это не только стабильное соединение, но и единица, которая семантически отличается от содержимого, выражаемого компонентами. Это лаконичная фраза, значение которой не зависит от значений слов в ней.<sup>4</sup> Таким образом, содержание предложенной базы идиом, синонимов, вариантов, русских эквивалентов, английских эквивалентов и в

<sup>1</sup> Дониёров Х. Ўзбек тили “ж”лашган (кипчоқ) диалектининг фразеологияси устидан кузатишлар // Фразеология масалалари. СамДУ асарлари, янги серия, 106- чиқиши. – Самарқанд. – Б. 80-85.

<sup>2</sup> Бегматов Э. Ҳозирги ўзбек адабий тилининг лексик қатламлари. – Тошкент: Фан, 1985. – Б.15.

<sup>3</sup> Қаранг: Захаров В.П. Корпусная лингвистика. Учебно-методическое пособие. – Санкт-Петербург, 2005. – С. 48.

<sup>4</sup> ЎТИЛ. 1981 йил 2 жилдли. – Б. 169.

тексте художественного произведения, с примерами из стиля речи, раскрывается ситуация живого языка. Корпусная лингвистика приняла в качестве своей методологии этот принцип, сформировавшийся за рубежом в XIX-XX веках. В конце концов, предметом корпусной лингвистики является не языковая норма, а обычный язык.

Идиомы присутствуют во всех толковых словарях узбекского языка, но даны они не как идиомы, а как фразеологизмы, словосочетания. Представление идиом в русскоязычных аннотированных словарях имеет относительно безупречный внешний вид и преимущества, есть хотя бы частичное использование компьютерных технологий, некоторые словари основаны на корпусе русского языка и заслуживают похвалы за быстрый поиск информации. Наличие аудиоверсии, картинки, анимированного формата интерпретации идиом в словарях английского (американский английский, британский английский, австралийский английский) в виде мобильных приложений подтверждает, что английские словари идиом в этом плане более совершенны.

Идиомы в узбекских словарях целесообразно разделить на следующие основные тематические группы: 1) идиомы со смысловой составляющей «человек»; 2) с природной, антропогенной смысловой составляющей ландшафта; 3) топоним с семантической составляющей; 4) семантическая составляющая «человек-время-пространство»; 5) мифоним со смысловой составляющей; 6) с социально-смысловой составляющей; 7) с фитоморфно-смысловой составляющей; 8) с зооморфным смысловым компонентом; 9) с смысловой составляющей «предмет».

Научная и практическая значимость презентации идиом в национальном корпусе отражается в том, что теперь, в результате отражения идиом в языковом корпусе, лингвисты, переводчики, учителя, журналисты и исследователи в различных областях могут быстро и быстро найти и проанализировать, чтение и обучение с использованием средств коммуникации и информации становится все более и более изощренным требованием времени. Это позволяет напрямую использовать информацию, содержащуюся в национальном или языковом корпусе, и в образовательном процессе; учитывая тот факт, что узбекский язык преподается как наука в нескольких странах мира, используя базу идиом для ознакомления с национальной принадлежностью нашего языка в иностранных интересах, в целом есть возможность практического применения во всех областях, связанных с компьютеризацией лингвистики.

В третьей главе диссертации, озаглавленной «**Научно-практические основы создания базы данных идиом в национальном корпусе узбекского языка**», обозначены разметки лексические единицы в национальном корпусе, межъязыковые различия в тематической группировке идиом (учебная сущность) подробно описаны научно-теоретические данные о методах автоматического определения идиом.

Проблема разметки идиоматических единиц - одна из основных проблем компьютерной лингвистики и корпусной лингвистики. Потому что разметка -

это инструмент для редактирования и анализа данных. Цель семантической разметки - найти и представить информацию, понятную роботу или компьютеру, в результатах поиска.

Во многих случаях смысловое единство состоит из нескольких компонентов. Совместная коллокация происходит, когда последовательность этих конструкций высока.<sup>1</sup> В словосочетании общее значение устройства не равно независимым значениям, которые несут компоненты (*мехр булоги* ≠ *мехр* + *булоқ*). К ним относятся идиомы (*илон ёгини ялаган, қўли очиқ*), топонимы и отдельные ономы, и в этом случае практика семантической разметки несколько усложняется.<sup>2</sup>

Коллокация<sup>3</sup> - это единица с синтаксической и семантической целостностью, которые по лексическому составу делятся на следующие типы: несоставный; незаменяемый состав; неизменяемые состав. Несоставные коллокации - это единицы, значение которых отличается от общего содержания. К ним относятся идиомы и идиоматические единицы. Незаменяемые коллокации включают метафоры.

Семантическая разметка на основе предлагаемых идиом была сделана по следующим причинам: тот факт, что идиомы не делятся на определенные части; содержащиеся в нем лексемы не означают только содержание; значение приравнивается к единому лексическому значению и в этом отношении представляет собой целостную лексико-семантическую целостность. В этой практике разметки также соблюдались мировые стандарты (как известно, в 2011 году Google разработал стандарт Data Vocabulary для разметки данных. В 1995 году группа библиотекарей и музейных ученых разработала Dublin Core (<http://dublincore/core/org/>) с 15 основными и 3 дополнительными функциями).

Части разных типов идиом приобретают разный лингвистический статус под влиянием различителей (фонем), фиксаторов (морфем), номинаторов (слов и словосочетаний), коммуникаторов (выражений). Этот аспект обуславливает их семантическую емкость, семантическую изменчивость и активизирует кумулятивные, эвфемистические, эзотерические, описательные, сатирические функции в идиоматических приемах. Поскольку структурная особенность идиом возникает в связи с транспозицией семантического содержания их языковых единиц,<sup>4</sup> семантическая разметка несколько усложняется.

Создание базы идиом в Национальном корпусе узбекского языка требует одновременного предоставления идиоматических конструкций на узбекском, английском, русском и других языках, однако перевод идиом на другой язык равносителен игнорированию языковых особенностей, которые для них не характерны.

<sup>1</sup> Боярский К. К. Введение в компьютерную лингвистику. – СПб: НИУ ИТМО, 2013. – 72 с.

<sup>2</sup> Эшмўминов А. А. Ўзбек тили миллий корпусининг синонимлар базаси. Филол. фан. бўйича фалсафа доктори (PhD) дис.автореферати. – Қарши, 2019.

<sup>3</sup> <https://ru/m.wikipedia.org.> wiki>.

<sup>4</sup> <https://www.dissercat.com/content/struktura-i-funksii-idiom-na-materiale-substantivnykh-ustoichivnykh-slovosochetanii-i-slozhn>.

В этой главе также обсуждаются взгляды на методы автоматического обнаружения идиом. В частности, латентно-семантический анализ (Latent semantic analysis); параллельное выравнивание текста (Bitext word alignment); оценка взаимосвязанной информации (Mutual information).<sup>1</sup> В частности, параллельное уравнение текстов учитывает наличие или отсутствие энтропии (греч. ἐν «Внутри» + τροπή «вращение»).

Энтропия определяется в следующем порядке:

$$H(T_s|s) = - \sum_{t \in T_s} P(t|s) \log P(t|s)$$

Энтропия означает наличие идиом.

PDA:

$$pda(S) = x = \frac{\sum_{s \in S} \sum_{d \in D_s} \text{align\_freq}(s,d)}{\sum_{s \in S} \sum_{t \in T_s} \text{align\_freq}(s,t)}$$

Наличие PDA доказывает, что тестируемое сочетание имеет собственное значение.

При определении коэффициента взаимной информации сравнивается MI (*mutual information*) исследуемой фразы и словосочетаний на основе их синонима, антонима, голонима-меронима<sup>2</sup>, гиперонима-гипонима<sup>3</sup>. В частности, в «Материалах, хранящихся в базе идиом в Узбекском национальном корпусе»<sup>4</sup> эта информация выглядит следующим образом: (тип программы: Веб-приложение (web application):

Используемые языки программирования:

- PHP v7.2 (Laravel Framework v6.2)
- HTML
- SASS/CSS (Bootstrap v4)
- Javascript (Vuejs v2.5, JQuery v3.4)

Применимая база данных:

- MySQL v5.5

Разметка текстовых пар - это сам по себе очень полезный лингвистический ресурс. Делать это вручную намного дороже и трудоемко. Поэтому исследователи разрабатывают разные способы автоматического комбинирования текста. Эти методы можно разделить на две группы: методы, основанные на привлечении дополнительных лингвистических знаний, например (Kayand Roscheisen, 1993, Kit et al.) и статистические методы (Brown et al., 1991; Gale and Church, 1991; McEnery and Oakes, 1996; Mikhailov). Разница в основном зависит от типа используемых данных.

<sup>1</sup>Шильнов, А. Г., Миллер, Д., Митчелл, П. Д., & Смокотин, В. М. (2020). Сравнительный анализ идиом, представленных в российских учебниках английского языка, и идиом, представленных в современных аутентичных источниках. *Язык и культура*, (49), 157-174. <https://doi.org/10.17223/19996195/49/10>

<sup>2</sup>Масалан, *translation theory* термини таржима назариясининг кичик бўлимларига оид терминларга нисбатан голоним, *partial theories of translation* термини эса *translation theory* га нисбатан мероним хисобланади. Қаранг: [https://studwood.ru/712345/literatura/giperonimy\\_giponimy](https://studwood.ru/712345/literatura/giperonimy_giponimy).

<sup>3</sup> <https://www.slideshare.net/msucsai/ss-5430928>.

<sup>4</sup>О'zbek Milliy Korpusida idiomalar bazasi deponentlanadigan materiallari. Huquq egasi Begmatova Gulchehra Xolmurodovna. Mualliflar: Begmatova Gulchehra Xolmurodovna. Ulug'ov Anvar Abduraimovich. – 2019.



Традиционно для объединения параллельных текстов используются статистические методы. Они основаны на приблизительном соответствии между словами на одном языке и длиной текстов на другом языке, измеряемой в словах или символах. Поэтому, в зависимости от длины текстов на одном языке, необходимо рассчитать, где должны начинаться соответствующие части текста на другом языке. Остается только выбрать относительно более подходящую совместимость. Эти методы представляют собой быстрые методы, основанные только на подсчете количества слов или количества символов. Однако у этих методов есть и недостатки. Как только перевод состоит из значительно более коротких или более длинных данных (когда возникает несколько вариантов совместимости), существующие алгоритмы не могут принять решение, потому что данные чрезвычайно ограничены.

**Проблема представления результатов.** У крупных корпусов есть специфическая проблема, которая раньше была незначительной: поисковые запросы могут давать сотни или даже тысячи результатов, но их невозможно увидеть за ограниченный промежуток времени. Для решения этой проблемы разрабатываются системы, позволяющие обрабатывать наиболее устойчивые выражения (коллокации) путем группировки результатов поиска и автоматической их кластеризации или статистической оценки их важности.<sup>1</sup> На основе языкового корпуса: строки - идиомы, столбцы - разреженная матрица в виде документов. Это

$$t \frac{T}{i} \rightarrow \begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{m,1} & \cdots & x_{m,n} \end{bmatrix}$$

$t_i^T t_p$  — соотношение двух идиом (выбранная).

**Задачи по поиску коллокаций** (целостных синтаксических конструкций, равных фразе) **из корпуса.** В этих типах задач используется поиск по корпусу разговорных выражений: фразы выбираются автоматически в зависимости от частоты их использования и времени уровня разговорного характера, а также путем выделения определенной части текста, как для всех частей речи одновременно. При поиске следует учитывать жанровую природу текста: студенты должны на практике заметить, что уровень активности словосочетаний варьируется в зависимости от типа текста.

Предлагаемые нами успешные возможности базы: идиоматические единицы отличаются от фразеологизмов в словарях, созданных в разные годы; сборник идиом основан в основном на текстах художественной литературы и их оценке на основе семантической среды идиоматической единицы и национально-культурного опыта; идиомы отличаются от

<sup>1</sup> Корпусная лингвистика как раздел языкознания. — Текст: электронный // Myfilology.ru – информационный филологический ресурс: [сайт]. – URL: <https://myfilology.ru/177/korpusnaya-lingvistika-kak-razdel-yazykoznanija/> (дата обращения: 5.02.2021).

диалектных единиц, эвфемизмов; представлены варианты идиом в английском, русском и некоторых тюркских языках; варианты идиоматических единиц, заданные их синонимами (хотя и частично); создание специального веб-приложения идиом ([idioms.uz](http://idioms.uz)); ([idioms.uz](http://idioms.uz)) онлайн-база данных определяется возможностью получать мнения и отзывы по электронной почте непосредственно в самом веб-приложении.

## ОБЩЕЕ ЗАКЛЮЧЕНИЕ

1. У каждой нации есть свои собственные идиомы, которые важны в мировой лингвистике, потому что они являются национальными по своей природе, не переведены доподлинно на другие языки, но для других языков есть подходящие альтернативы таким случаям.

2. Идиомы - это национально адаптированные выражения во фразеологии, но идиомы отличаются от других фразеологий тем, что они не подвергаются «деформации», такой как структурное изменение, замена определенных слов, расширение и сокращение структуры фразеологизмов.

3. При возникновении идиом конкретное историческое, социальное и экономическое событие выступает в качестве события-прототипа, а затем событие забывается, и идиома, таким образом, используется в процессе застывшей речи.

4. Идиоматические выражения присутствуют во всех толковых словарях узбекского языка, но в этих словарях они трактуются не как идиомы, а как фразеологические единицы, словосочетания.

5. База идиом в языковом корпусе позволяет устранить следующие типичные ошибки в обучении, особенно в языковом обучении: незнание, как правильно подбирать слова из предоставленного списка слов; незнание фиксированных словосочетаний (собственных и переносных значений), фразеологических единиц, незнание популярных фраз.

6. Семантическая разметка на основе предлагаемых идиом не разделяет идиомы на определенные части; содержащиеся в нем лексемы не означают только содержание; Значение приравнивается к единому лексическому значению, как и фразеологизмы, и в этом отношении представляет собой целостную лексико-семантическую целостность. В этой практике маркировки соблюдались мировые стандарты.

7. База, которое мы предлагаем, заключается в следующем: идиомы отличаются от диалектных единиц, эвфемизмов; английский, русский и некоторые тюркские альтернативы идиом; варианты идиоматических единиц, заданные их синонимами (хотя и не полные); создание специального онлайн веб-приложения идиом ([idioms.uz](http://idioms.uz)); важность научной новизны данного исследования определяется тем, что онлайн-база данных ([idioms.uz](http://idioms.uz)) имеет возможность получать мнения и комментарии непосредственно через электронную почту в веб-приложении.

**SCIENTIFIC COUNCIL NUMBER PhD.03/31.03.2021.Fil.78.04  
ON AWARDING SCIENTIFIC DEGREES UNDER  
TERMEZ STATE UNIVERSITY**

---

**TERMEZ STATE UNIVERSITY**

**BEGMATOVA GULCHEKHRA KHOLMUROTOVNA**

**CREATION OF IDIOMAL DATABASE IN UZBEK NATIONAL CORPUS**

**10.00.01 – Uzbek language**

**DISSERTATION ABSTRACT FOR THE DOCTOR OF PHILOSOPHY (PhD) ON  
PHILOLOGICAL SCIENCES**

**Termez – 2021**

The theme of the dissertation for the Doctor of Philosophy (PhD) was registered at the Supreme Estimation Commission of the Cabinet of Ministers of the Republic of Uzbekistan under № 119.4.PhD/Fil1026.

The dissertation has been prepared at the Termez State University.

The abstract of the dissertation is posted in three languages (Uzbek, Russian, English (resume)), is placed on the website of Scientific Council ([www.tersu.uz](http://www.tersu.uz)) and on the website of 'Ziyonet' information and educational portal ([www.ziyonet.uz](http://www.ziyonet.uz)).

**Scientific supervisor:** **Makhmaraimova Shokhista Tukhtashevna**  
Doctor of Philological Sciences, Associate Professor

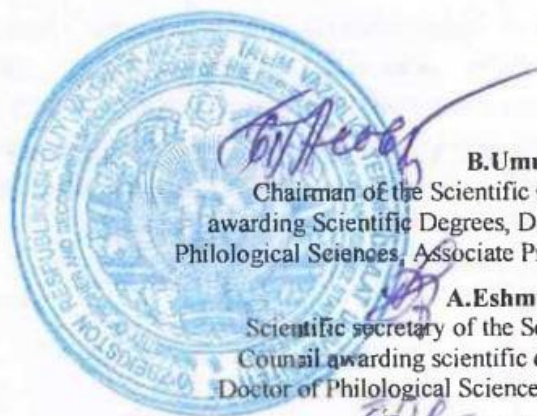
**Official opponents:** **Karimov Suyun Amirovich**  
Doctor of Philological Sciences, Professor  
**Mamatov Abdi Eshonkulovich**  
Doctor of Philological Sciences, Professor

**Leading organization:** **Tashkent State university of Uzbek language and literature named after Alisher Navai**

The defence of the dissertation will be held on July «26», 2021 at 10<sup>00</sup> at the meeting of the Scientific Council PhD.03/31.03.2021.Fil.78.04 at the Termez State University. Address: Termez city, Barkamol avlod street, 43. Phone: (876) 221-74-55; fax: (876) 221-71-17; e-mail: [termizdu@umail.uz](mailto:termizdu@umail.uz).

The dissertation can be found in the Information Resource Center of Termez State University (registered with the number 1). Address: Termez city, Barkamol avlod street, 43. Phone: (876) 221-74-55.

The abstract of dissertation was distributed on "15" July 2021.  
(Registry record № 1 dated "15" July 2021.)



**B.Umurkulov**  
Chairman of the Scientific Council  
awarding Scientific Degrees, Doctor of  
Philological Sciences, Associate Professor

**A.Eshmuminov**  
Scientific secretary of the Scientific  
Council awarding scientific degrees,  
Doctor of Philological Sciences (PhD)

**Sh.Yokubov**  
Vice-chairman of the Scientific Seminar at the  
Scientific Council Awarding Scientific degrees,  
Candidate of Philological Sciences, Associate Professor

## **INTRODUCTION**

**(abstract of the doctor of philosophy (PhD) dissertation)**

**Actuality and relevance of the research topic.** According to the principles of world linguodidactics, language education is carried out primarily through the acquisition of vocabulary content and thus the formation of knowledge about the national worldview of the relevant linguistic community. In this regard, the inability to translate idioms into other languages, in terms of national language corpora, further exacerbates the problem of classifying idioms which express the spirit of the nation, hold pragmatic value and contribute to the content of the internal form of language. Currently, the practice of classifying idioms is gaining momentum in some Turkic languages, as well as in the developed countries of the world which allows you to test the proposed theoretical models of idioms, to determine the semantic-syntactic view of fixed phrases, to facilitate the process of machine translation and automatic editing of texts.

**The purpose of the research** is to develop a theoretical and practical basis for the inclusion and tagging of idioms in the language, to advance principles for improving the database of idioms through tagging of different types of collocations by highlighting scientific, theoretical and practical bases of formation of idioms base in Uzbek national corpus.

**The aim of the study** is to develop the theoretical and practical foundations for the introduction and designation of idioms in the Uzbek national corpus based on the coverage of the scientific, theoretical and practical foundations of marking issues, the principles of improving the idiom base by marking various types of collocations.

**The object of the research** is the idioms of the Uzbek language.

**The scientific novelty of the research** is determined as follows:

the content of the research and the chronology of idiomatic combinations in world linguistics, Turkic studies and Uzbek linguistics are systematized;

an insufficient lexicographic base was revealed in the formation of a base of idioms in the national corpus of the Uzbek language, an inseparable description of idioms from phraseological combinations in dictionaries, an indistinct principle of foundation in comments;

revealed interlanguage differences in the thematic grouping of idioms;

justified the obsolescence of illustrative information in phraseological dictionaries and the need to introduce the practice of using electronic corpora when updating them;

methods and means of mathematical analysis of semantic marking and automatic collocation determination are described;

justified the advantages of the proposed web application Idioms.uz.

**Implementation of research results.** The role of lexical and semantic integrity in the phraseological system of the Uzbek language, as well as the Turkic languages was used in the project of fundamental research №. PZ-20170927147 on the topic "Study of ancient Turkic writing and folklore until the 13th century", under the guidance of associate professor of the Tashkent State University of

Uzbek language and literature named after Alisher Navoi A.Tilovov (reference No. 04 / 1-609 dated October 7, 2021 of the Tashkent State University of Uzbek Language and Literature named after Alisher Navoi). As a result, the formation of the national picture of phraseological combinations and the improvement of knowledge about the linguocultural nature of idiomatic constructions in the Uzbek language, as well as in the Turkic languages;

to determine the general principles of the structure of idiomatic units and their meaning in the information retrieval system; recognition of the similarity in the occurrence of idiomatic units in the word-formation system of the Uzbek language among the Turkic languages (Kazakh, Kyrgyz, Azerbaijani, Bashkir, Tatar, Turkmen) and the participation of a word in the category of nouns and adjectives instead of the "central word"; to emphasize the possibility of achieving a quick search for information in the linguistic corpus by determining what or how the local component of the content is expressed in the lexicographic classification of idiomatic constructions;

Conclusions on the application of the research results that idioms are vocabulary units that are used in a certain region, have a national specificity and practical nature, for application to the scientific and educational process in the field of linguistics from the point of view of computational linguistics were used in a scientific and practical project carried out at the Karakalpak Scientific Research Institute of Humanities of the Karakalpak Branch of the Academy of Sciences of the Republic of Uzbekistan No. FA-F1-G003 on the topic "Functional word formation in the modern Karakalpak language" (reference No. 35/1 dated February 22, 2021 from the Karakalpak branch of the Academy of Sciences of the Republic of Uzbekistan). As a result, an improvement has been achieved in the approach to the study of computational linguistics and views on the development of linguistic competence in the field of higher education;

a database of the idiom of the Uzbek language has been developed; tags of Uzbek idioms were used to create an online version of the idiom database based on the linguistic model (reference No. DGU 08092 dated May 5, 2020 of the Intellectual Property Agency under the Ministry of Justice of the Republic of Uzbekistan). The result is an online version of the idiom's database.

The role of idioms in the Uzbek culture of speech, conclusions about the manifestation of linguacultural features in live speech were used in the compilation of program scenarios for the Surkhandarya regional TV and radio company (reference No. 531 dated July 9, 2020 of the Surkhandarya TV and Radio Company).

**The structure and length of the dissertation.** The dissertation consists of an introduction, three main chapters, a general conclusion, a list of references and an appendix. The total length of the dissertation is 130 pages.

**ЭЪЛОН ҚИЛИНГАН ИШЛАР РЎЙХАТИ**  
**СПИСОК ОПУБЛИКОВАННЫХ РАБОТ**  
**LIST OF PUBLISHED WORKS**

**I бўлим (I часть; I part)**

1. Begmatova G. Scientific and theoretical approaches to idioms in world // *Academicia: An international Multidisciplinary Research Journal*. – India, 2020. Mart, Volume 10. Issue 3. – P.178-181. (Impact factor SJIF 2020=7.13) <http://www.saarj.com>

2. Begmatova G. On the etymology of idioms and phraseologisms. *Ilkogretim Online - Elementary Education Online*, 2020; 19 (4): P. 842-845. <http://ilkogretim-online.org>. doi:10.17051/ilkonline.

3. Begmatova G. Idioms in turkish languages. *Clarivate. International Scientific Journal ISJ Theoretical & Applied Science*. Philadelphiya, USA. Issue 06. Volume 86 published June 30, 2020. P. 368-371. <http://t-science.org/axivDOI/2020/06-86.html>.

4. Бегматова Г. Ўзбек миллий корпусида идиомаларни яратиш муаммолари // НамДУ илмий ахборотномаси. 2019, 7-сон. – Б. 211-215.(10.00.00; №26)

5. Бегматова Г. Идиома ва фразеологизмларнинг этимологияси ҳақида // Хоразм Маъмун академияси ахборотномаси. 2021 йил, махсус сон. – Б. 92-95. (10.00.00; №21)

6. Бегматова Г. Изоҳли луғатларда идиомаларнинг берилиши // Ўзбекистон Миллий ахборот агентлиги – ЎЗА. Илм-фан бўлими. – Тошкент, 2020. – Б. 196-207.

7. Бегматова Г. Х. Идиоматик бирликларни разметкалашнинг ўзига хос мураккабликлари / *Modern Scientific challenges and trends*. ISSUE 11(33). Part 2. Warsaw, Poland. November, 2020. P.313-316.

8. Бегматова Г. Идиомаларда миллийликнинг намоён бўлиши / Нутқ маданияти ва ўзбек тилшунослигининг долзарб муаммолари. Халқаро онлайн илмий- амалий конференция. – Андижон, 4 май, 2020. – Б.106-108.

9. Бегматова Г. Ареал лингвистика ва диалектология муаммолари / “O’zbekistonda ilmiy-amaliy tadqiqotlar” mavzusidagi 16-ko’p tarmoqli ilmiy konferensiya. – Toshkent, may, 2020 yil.

**II бўлим (II часть; II part)**

1. Begmatova G. Nutqda dialektal idiomalarning qo’llanilishi va tadqiqi / “O’zbekistonda ilmiy-amaliy tadqiqotlar” mavzusidagi 16-ko’p tarmoqli ilmiy konferensiya Toshkent, may,2020 yil.

2. Бегматова Г. Ўзбек миллий корпусида идиомалар базаси ва унинг техник таъминоти / “Filologiya masalalari” Respublika ilmiy amaliy onlayn konferensiyasi. – Andijon, 2020.

3. Бегматова Г. Жаҳон тиллари Миллий корпусларида луғавий бирликларнинг берилиши (ибора ва идиомалар мисолида) / “Ta’limga

kompetensiyaviy yondashuv: muammo va yechimlar”. Respublika 3-online ilmiy-amaliy konferensiyasi. – Andijon, avgust, 2020.

4. Бегматова Г. Ўзбек миллий корпусида идиомалар базаси. Веб-илова. Ўзбекистон Республикаси Адлия вазирлиги ҳузуридаги Интеллектуал мулк агентлигининг 2020 йил 5 майдаги №DGU 08092-сонли гувоҳномаси.

5. Бегматова Г. Ўзбек тили идиомаларининг кичик изоҳли луғати. – Термиз: ТерДУ нашр матбаа маркази нашриёти, 2021. – 34 б.



Автореферат Термиз давлат университетининг  
«Филологик тадқиқотлар: тил, адабиёт, таълим» илмий журнали таҳририятида  
таҳрирдан ўтказилди (9.07.2021 йил)