

**ЎЗБЕКИСТОН РЕСПУБЛИКАСИ ОЛИЙ ВА ЎРТА МАХСУС
ТАЪЛИМ ВАЗИРЛИГИ
АЛИШЕР НАВОИЙ НОМИДАГИ ТОШКЕНТ ДАВЛАТ ЎЗБЕК ТИЛИ
ВА АДАБИЁТ УНИВЕРСИТЕТИ**

Қўл ёзма ҳуқуқида

УДК: 81`33

ХАМРОЕВА ШАҲЛО МИРДЖОНОВА

**ЎЗБЕК ТИЛИ МОРФОЛОГИК АНАЛИЗАТОРИНИНГ ЛИНГВИСТИК
ТАЪМИНОТИ**

10.00.11 – Тил назарияси, амалий ва компьютер лингвистикаси

Филология фанлари доктори (DSc) илмий даражасини

олиш учун тақдим этилган

ДИССЕРТАЦИЯ

Илмий маслаҳатчи: Б.Р.Менглиев,
филология фанлари доктори, профессор

Тошкент - 2021

МУНДАРИЖА

КИРИШ	3
--------------------	----------

I БОБ. ЛИНГВИСТИК ТАЪМИНОТНИНГ АХБОРОТ-ҚИДИРУВ ТИЗИМЛАРИДАГИ ЎРНИ

1.1. Лингвистик таъминот воситаларининг умумий тавсифи, морфологик анализаторнинг тил корпуслари яратишдаги аҳамияти.....	19
1.2. Сўз шакллариининг морфологик анализ ва синтези.....	35
1.3. Лингвистик таъминотнинг тузилиши, таркиби ва вазифалари.....	49
Боб бўйича хулоса.....	59

II БОБ. МОРФОЛОГИК АНАЛИЗАТОР ТУЗИШНИНГ НАЗАРИЙ АСОСЛАРИ

2.1. Морфологик анализатор тузишнинг умумий тамойиллари.....	61
2.2. Морфологик анализаторнинг ахборот-қидирув тизимидаги аҳамияти...	77
2.3. Туркий тиллар морфологик анализаторларининг муштарак ва фарқли жиҳатлари.....	96
Боб бўйича хулоса.....	115

III БОБ. ЎЗБЕК ТИЛИ МОРФОАНАЛИЗАТОРИНИ ТУЗИШНИНГ ЛИНГВИСТИК АСОСЛАРИ

3.1. Ўзбек тили морфоанализатори бирликлари, тузилиши ва таркиби.....	117
3.2. Ўзбек тили морфоанализаторининг таҳлил бирликлари, сўзшакл ва морфемалар.....	142
3.3. Ўзбек тили морфоанализаторида морфотактика қоидалари.....	158
Боб бўйича хулоса.....	166

IV БОБ. ЎЗБЕК ТИЛИ МОРФОАНАЛИЗАТОРИ УЧУН СЎЗЛАРДА ҚЎШИМЧАЛАР ҚУРШОВИНИ ЛИНГВИСТИК МОДЕЛЛАШТИРИШ МАСАЛАЛАРИ

4.1. Сўзда луғавий шакл ясовчи, синтактик шакл ясовчи ва қўшимча кўринишидаги юкламалар қуршовини лингвистик моделлаштириш.....	169
4.2. Сўз ясовчи морфемаларнинг сўздаги қуршовини моделлаштириш.....	190
4.3. Сўз ясовчи қўшимчалар қуршовини моделлаштиришнинг айрим муаммолари	211
Боб бўйича хулоса.....	220

ХУЛОСА	222
---------------------	------------

Фойдаланилган адабиётлар рўйхати	224
---	------------

ИЛОВА	253
--------------------	------------

КИРИШ (докторлик (DSc) диссертацияси аннотацияси)

Тадқиқот мавзусининг долзарблиги ва зарурати. Жаҳон компьютер лингвистикасида автоматик морфологик таҳлил XX асрнинг 50-йилларидан табиий тилга ишлов бериш, математик лингвистика ва машина таржимаси амалиётида қўллана бошланди. Туркий тилларнинг морфологик анализаторларини ишлаб чиқишнинг амалий асосларига ўтган асрнинг 60-йилларидан эътибор қаратилди. Фандаги бу янгиланишлар ахборот технологияларини тилшуносликка татбиқ этиш билан боғлиқ истиқболли илмий йўналишларнинг вужудга келишига асос бўлди. Зеро, бу ахборот-қидирув тизими, корпус лингвистикаси ва машина таржимасида морфологик таҳлилнинг умумий тамойилларини ўрганиш, таҳлил тизимлари лингвистик таъминоти учун формал тил ва грамматикани яратиш тизимини ишлаб чиқишнинг заруриятини юзага келтирди. Бу эса сунъий интеллектнинг автоматик таржима, компьютер таҳлили ва таҳрири, тезаурус, электрон луғат борасидаги имкониятларининг амалий аҳамиятини белгилайди.

Дунё тилшунослигида XX аср охирида автоматик таржима сифатини яхшилаш, тилни лингвистик моделлаштириш, муайян тиллар доирасида сўзларни леммалаш назарияси, алгоритмининг яратиш каби масалалар кун тартибига қўйилди. Натижада ахборот-қидирув тизимида фаол қўлланувчи стеммер, орфографик корректор (имло муҳаррири) сингари иловалар ишлаб чиқилди. Бинобарин, тил корпусларини яратиш, уларни лингвистик аннотациялаш, матнни автоматик қайта ишлайдиган дастурлар – морфоанализатор, лемматизатор, стеммер, парсер, орфокорректорларни ишлаб чиқиш компьютер лингвистикасининг илмий-назарий муаммоларига айланди.

Ўзбек компьютер лингвистикаси соҳасида тил корпуси, нутқ синтезатори, автоматик таржима, сунъий интеллект, ўзбек тилини тушуниш ва қайта ишлаш борасида илмий изланишларга эътибор қаратилмоқда. Лекин

миллий тил морфологик анализаторининг дастурини тузиш масаласи илмий-назарий планда ўрганилмаган. Зеро, “...давлат тилининг софлигини сақлаш, уни бойитиб бориш ва аҳолининг нутқ маданиятини ошириш; давлат тилининг замонавий ахборот технологиялари ва коммуникацияларига фаол интеграциялашувини таъминлаш”¹ ҳозирги кунда ўзбек компьютер лингвистикаси олдида турган долзарб вазифалардан биридир. Шунга кўра, ўзбек тилига замонавий ахборот технологиялари воситасида ишлов берилишига эришиш, бунинг учун тил корпуслари, электрон таржимон, тезаурус, орфо корректор каби дастлабки автоматик ишлов воситалари ҳамда уларнинг лингвистик таъминотини яратиш сингари масалалар муҳим аҳамият касб этади.

Ўзбекистон Республикаси Президентининг 2016 йил 13 майдаги ПФ-4997-сон “Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университетини ташкил этиш тўғрисида”, 2017 йил 7 февралдаги ПФ-4947-сон “Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси тўғрисида”, 2019 йил 21 октябрдаги ПФ-5850-сон “Ўзбек тилининг давлат тили сифатидаги нуфузи ва мавқеини тубдан ошириш чора-тадбирлари тўғрисида”, 2020 йил 20 октябрдаги ПФ-6084-сон “Мамлакатимизда ўзбек тилини янада ривожлантириш ва тил сиёсатини такомиллаштириш чора-тадбирлари тўғрисида”ги фармонлари, 2017 йил 17 февралдаги ПҚ-2789-сон “Фанлар академияси фаолияти, илмий-тадқиқот ишларини ташкил этиш, бошқариш ва молиялаштиришни янада такомиллаштириш чора-тадбирлари тўғрисида”, 2019 йил 4 октябрдаги ПҚ-4479-сон “Ўзбекистон Республикасининг “Давлат тили ҳақида”ги Қонуни қабул қилинганлигининг ўттиз йиллигини кенг нишонлаш тўғрисида”ги қарорлари, 2020 йил 24 январдаги Олий Мажлисга Мурожаатномаси ҳамда

¹Ўзбекистон Республикаси Президенти Шавкат Мирзиёевнинг 2020 йил 202 октябрдаги «Мамлакатимизда ўзбек тилини янада ривожлантириш ва тил сиёсатини такомиллаштириш чора-тадбирлари тўғрисида»ги ПФ-6084-сон фармони // <https://lex.uz/docs/5058351>

бошқа меъёрий-ҳуқуқий ҳужжатларда белгиланган вазифаларни амалга оширишда ушбу тадқиқот муайян даражада хизмат қилади.

Тадқиқотнинг республика фан ва технологиялар тараққиёти устувор йўналишларига мослиги. Тадқиқот республика фан ва технологиялари ривожланишининг I.“Ахборотлашган жамият ва демократик давлатни ижтимоий, ҳуқуқий, иқтисодий, маданий, маънавий-маърифий ривожлантириш, инновацион иқтисодиётни ривожлантириш” устувор йўналишига мувофиқ бажарилган.

Тадқиқот мавзуси бўйича хорижий илмий тадқиқотлар шарҳи². Жаҳон тилшунослигида автоматлаштирилган ахборот-қидирув тизимлари лингвистик таъминотини ўрганишга йўналтирилган илмий изланишлар етакчи илмий марказлар ва олий таълим муассасаларда, жумладан, Oxford University (Буюк Британия), Cambridge University (Буюк Британия), Humboldt Universitat (Германия), Boston University, University of New York, University of Pennsylvania, University of California (АҚШ); Indiana University (АҚШ), The University of Sorbonne (Франция), Universitat de Valencia (Испания), University of Zurich (Швейцария), Упсалла университети (Швеция), Kiofo University (Япония), Australian National University (Австралия), Истамбул техника университети (Туркия), Москва давлат университети (Россия), Москва давлат лингвистика университети (Россия), Санкт-Петербург давлат университети (Россия), Урал давлат педагогика университети, Харьков давлат педагогика университети (Украина), Baki Dövlət universiteti (Озарбайжон), Татаристон Республика Фанлар академияси амалий тилшунослик ва семиотика институти (Татаристон Республикаси, Россия), Қозон федерал университети (Татаристон Республикаси, Россия), В.Вернадский номидаги Таврия академияси Қрим федерал университети (Қрим Республикаси, Россия), шунингдек, Мирзо

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

Улуғбек номидаги Ўзбекистон Миллий университети, Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университети, Ал-Хоразмий номидаги Тошкент ахборот технологиялари университети ва унинг филиаллари, Бухоро давлат университети, Термиз давлат университети (Ўзбекистон)да олиб борилмоқда.

Дунё тилшунослигида ахборот-қидирув тизими(АҚТ)да лингвистик таъминот, матнга автоматик ишлов бериш технологияси ва воситалари, матннинг автоматик анализ/синтез дастурлари, морфологик анализаторлар тадқиқига оид олиб борилган изланишлардан қуйидаги илмий натижалар олинган: автоматлаштирилган ахборот-қидирув тизимлари лингвистик таъминоти тавсифланган, ахборот қидирув тизимининг асосий ва ёрдамчи воситалари, мультибазавий тил воситалари аниқланган (Cambridge University, Англия; University of Zurich, Швейцария); табиий тилга автоматик ишлов бериш (тил бирликларини анализ/синтез қилиш) хусусиятлари ёритилган (University of New Jersey, University of California, АҚШ; Boston University, New York University, АҚШ; Universitat de Valencia, Испания); табиий тилга компьютерда ишлов беришда трансдюссерлардан фойдаланиш усуллари аниқланган (Courant Institute of Mathematical Sciences, New York, АҚШ; Pereira – San Francisco, USA, АҚШ); матнга автоматик ишлов бериш технологияларининг асбоб-ускуналари ва воситалари ишлаб чиқилган (University of Tokyo, Япония; Australian National University, Австралия; Oxford University Language Centre, Англия); морфологик таҳлилнинг назарий асослари, умумий тавсифи, ишлаш тамойиллари, тилга оид грамматик категория, дастур учун лингвистик модуллар тавсифи ишлаб чиқилган (Kiofo University, Япония; Australian National University, Австралия; Истамбул техника университети, Туркия; Москва давлат университети, Москва давлат лингвистика университети, Санкт-Петербург давлат университети, Россия); матнга морфологик ва синтактик ишлов бериш модели ва дастурлари (Харьков давлат педагогика университети, Украина; Baki Dövlət universiteti,

Озарбайжон); туркий тиллар компьютер лингвистикасида табиий тилга ишлов бериш, морфологик анализатор тузиш масалалари тадқиқ қилинган, тил корпуси, морфоанализатор, овозли нутқни синтез қилиш воситалари, электрон таржимонлар ишлаб чиқилган (Татаристон Республика Фанлар академияси амалий тилшунослик ва семиотика институти, Қозон федерал университети, В.Вернадский номидаги Таврия академияси Қрим федерал университети, Россия); ўзбек-инглиз тили машина таржимасининг лингвистик таъминоти, ўзбек тили муаллифлик корпусини, тил корпусининг лингвистик базасини тузиш тамойиллари ва унинг лингвистик таъминоти, ўзбек тили birlikларини графематик таҳлил қилиш муаммолари, ўзбек тили атов birlikларини семантик теглашнинг назарий асослари, ўзбек-инглиз параллел корпуси инструментарийларини яратиш асослари ишлаб чиқилган (Мирзо Улуғбек номидаги ЎзМУ, Алишер Навоий номидаги ТошДЎТАУ, Ал-Хоразмий номидаги ТАТУ, Ўзбекистон).

Дунё компьютер лингвистикасида табиий тилга автоматик ишлов беришнинг куйидаги устувор йўналишларида илмий тадқиқотлар олиб борилмоқда: автоматлаштирилган ахборот-қидирув тизимлар лингвистик таъминотини ишлаб чиқиш; табиий тилга автоматик ишлов бериш (тил birlikларини анализ/синтез қилиш) хусусиятларини ёритиш; матнга автоматик ишлов бериш технологияларининг асбоб-ускуна ва воситаларини яратиш; морфологик таҳлилнинг назарий асослари, умумий тавсифи, ишлаш тамойиллари, тилга оид грамматик категория, дастур учун лингвистик модуллар ишлаб чиқиш; туркий тиллар компьютер лингвистикасида табиий тилга ишлов бериш, морфологик анализатор тузиш масалаларини тадқиқ қилиш, тил корпуси, морфоанализатор, овозли нутқни синтез қилиш воситалари, электрон таржимонлар ишлаб чиқиш.

Муаммонинг ўрганилганлик даражаси. Жаҳон тилшунослигида лингвистик таъминотнинг назарий, методологик муаммоларини ўрганиш ўтган асрнинг 60-йилларида бошланди: ахборот-қидирув тилисиз электрон

ҳисоблаш машинасининг машиналигича қолиб кетиши тан олинди, “механик” босқичдан “мантикий-лингвистик” босқичга, ахборот-қидирув тизими таракқиёт босқичига ўтилди³. Жаҳон компьютер лингвистикасида табиий тилга ишлов бериш муаммолари А.Кейлер, К.В.Линден, Н.Вард⁴, М.А.Моҳри⁵ томонидан ёритилди. Матнга автоматик ишлов бериш технология ва воситалари В.Ахо Альфред, С.Лам Моника, С.Рави, Д.Ульман Джеффри⁶, морфологик таҳлилнинг назарий асослари С.Ю.Толдова, А.А.Бонч-Осмоловская⁷, Т.Садиқов⁸ ишларида тавсифланди. Рус компьютер лингвистикасида О.В.Дереза, Д.А.Каютенко, А.С.Феногенова, М.Ҳақимовнинг морфологик анализатор, унинг умумий тавсифи, ишлаш принциплари, тилга оид грамматик категория, дастур учун лингвистик модуллар тавсифи ҳақидаги ишларини таъкидлаш ўринли⁹. Ахборот-қидирув тизимида лингвистик таъминотнинг аҳамияти А.Б.Антопольский, Ю.М.Арский, Г.Г.Белоногов, Б.А.Кузнецов, В.Н.Белоозеров, Р.Р.Мдивани, В.А.Глинский, А.И.Михайлов, В.А.Мишин, В.В.Морозов¹⁰ ишларида очиб берилган.

³ Соколов А.В. Автоматизация библиографического поиска / А.В.Соколов. – М.: Книга, 1981. – 167 с. – С.91.

⁴ Speech and Language Processing. Daniel S. Jurafsky and James H. Martin. Contributing writers: Andrew Kehler, Keith Vander Linden, Nigel Ward 2000y. Prentice Hall, Englewood Cliffs, New Jersey 07632. Pages: 950.

⁵ Mohri M.A. Finite-state transducers in language and speech processing. Computational Linguistics. – С. 269-312.

⁶ Ахо Альфред В., Лам Моника С, Сети Рави, Ульман Джеффри Д. Компиляторы: принципы, технологии и инструментарий, 2-е изд.: Пер. с англ. – М.: ООО “И.Д. Вильямс”, 2008. – 1184 с.

⁷ Толдова С.Ю., Бонч-Осмоловская А.А. Автоматический морфологический анализ // Фонд знаний «Ломоносов». М., 2011. www.lomonosov-fund.ru/enc/ru/encyclopedia:0127430

⁸ Садықов Т., Кочконбаева Б. Об оптимизации алгоритма морфологического анализа // Шестая Международная конференция по компьютерной обработке тюркских языков «Turklang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

⁹ Дереза О.В., Каютенко Д.А., Феногенова А.С. Автоматический морфологический анализ для русского языка сравнительный анализ систем // Шестая Международная конференция по компьютерной обработке тюркских языков «Turklang-2018». (Труды конференции) – Ташкент, 2018.; Хақимов М. Логико-лингвистические модели русского языка // Шестая Международная конференция по компьютерной обработке тюркских языков «Turklang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

¹⁰ Антопольский А.Б. Лингвистическое обеспечение АСНТИ: структура и проблема совместимости // НТИ. Сер.2, 1983. – №7. – С.17-25; Арский Ю.М. Информационные ресурсы для устойчивого развития общества / Ю.М.Арский, А.И.Черный // Междунар. форум по информ. и док. – 2000. – Т. 25, № 1. – С. 3-9.; Белоногов Г.Г. Автоматизированная обработка научно-технической информации. Лингвистические аспекты / Г.Г. Белоногов, Б.А. Кузнецов, А.П. Новоселов // Итоги науки и техники. Сер. Информатика / ВИНТИ. –1984. –Т.8. – 316; Белоозеров В.Н. Ведение и совершенствование Государственного рубрикатора НТИ (ГРНТИ) / В.Н. Белоозеров, В.И. Федосимов, Г.А.Кривенцова // Информ. продукты, процессы и технологии: НТИ-96: Материалы конф. – М., 1996. – С. 60-69.; Мдивани Р.Р. Лингвистическое обеспечение автоматизированной информационной системы по общественным наукам: состав и структура / Р.Р.Мдивани, В.А.Глинский // Вопр. информ. теории и практики. 1985. – №52. – С. 37-55.; Михайлов А.И. Проблемы лингвистического обеспечения АСНТИ / А.И. Михайлов // Вопр. информ. теории и практики. 1984.– № 51. – С. 5-11.; Мишин В.А. Лингвистическое обеспечение библиотечных автоматизированных систем / В.А. Мишин, Л. Санчес-

Морфоанализ хусусиятлари Д.Рудольф, морфоанализнинг амай аҳамияти А.Ермаков¹¹, матнга автоматик морфологик ва синтактик ишлов бериш модели ва дастурлари И.М.Ножов¹² томонидан ўрганилган.

Туркий компьютер лингвистикасида табиий тилнинг компьютер ишлови, жумладан, морфоанализатор тузиш масалалари атрофлича тадқиқ этилди. Бундай ишлар сирасига Г.Эрйигит, Е.Адали¹³, А.В.Дибо, А.В.Шеймович¹⁴, Ж.Сулейманов¹⁵, А.Р.Гатиатуллин, А.М.Баширов¹⁶, П.В.Желтов¹⁷, Н.А.Исраилова, П.С.Бакасова¹⁸, Н.А.Леонтьев¹⁹, В.В.Куканова, А.Ю.Каджиев²⁰ларнинг морфологик анализатор тузиш масалаларига бағишланган тадқиқотларини киритиш мумкин. Туркий тиллар морфологик

Мехидо // Сов. библиотекословение. 1981. – № 2. – С. 84-93.; Морозов В.В. Лингвистическое обеспечение автоматизированных систем научно-технической информации / В.В. Морозов, Б.Р. Певзнер. –М.: ИПКИР, 1987. – 76 с.

¹¹ Rodolfe D. Computational Linguistic Text Processing: Lexicon, Grammar, Parsing and Anaphora Resolution. Nova Science Publishers, Inc. – NewYork, 2008. – P. 4-5.; Ермаков А. Морфологический анализатор - основа поисковых систем // <https://www.kv.by/archive/index2004154301.htm>

¹² Ножов И.М. Морфологическая и синтаксическая обработка текста (модели и программы): диссертация канд. наук. – Москва, 2003. – 190 с.

¹³ Eryigit G., Adali E. An affix stripping morphological analyzer for Turkish // Proceedings of IASTED International Conference ARTIFICIAL INTELLIGENCE AND APPLICATIONS. February 16-18, 2004, Innsbruck, Austria. – P. 299-304.

¹⁴ Дыбо А.В., Шеймович А.В. Автоматический морфологический анализ для корпусов хакасского и древнетюркского языков / Научное обозрение саяно-алтая рецензируемый научный журнал Номер 2(08), 2014 (Периодичность – 2 раза в год. серия: Филология.Выпуск 2). С. 9-31.

¹⁵ Сулейманов Д.Ш., Гильмуллин Р.А., Гатауллин Р.Р. Морфологический анализатор татарского языка на основе двухуровневой модели морфологии / Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.; Сулейманов Д.Ш., Гатиатуллин А.Р. Модель татарской аффиксальной морфемы и ее реализация // Серия: Интеллект. Язык. Компьютер. – Вып.4. – Казань: Фэн. – 1996. – 113-127.

¹⁶ Гатиатуллин А.Р., Баширов А.М. Морфологический анализатор тюркских словоформ на базе структурно-функциональной модели тюркской морфемы // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.

¹⁷ Желтов П.В. Разработка морфологического анализатора чувашского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.

¹⁸ Исраилова Н.А., Бакасова П.С. Морфологический анализатор кыргызского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.

¹⁹ Леонтьев Н.А. Морфологический анализатор якутского языка // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с. – С. 276-279.

²⁰ Куканова В.В., Каджиев А.Ю. Алгоритм работы морфологического парсера калмыцкого языка // В сборнике: Писменото наследство и информационните технологии. E1Manuscript-2014 Материали от V международна научна конференция. Отговорни редактори В.А.Баранов, В.Желязкова, А.М.Лаврентьев. 2014. – С. 116-119.

таҳлил тизими ҳамда уларнинг моделлари тавсифи, таҳлилига бағишланган тадқиқотлар²¹ кўламининг кенглиги ушбу диссертациянинг назарий асосларини мустаҳкамлади. Шунингдек, морфоанализатор ишлаб чиқиш назарий ғоялари амалиётга жорий этилиб, Европа, рус, туркий тилларнинг морфоанализаторлари ишлаб чиқилди²². Ўзбек тилшунослигида компьютер лингвистикаси, матнга лексикографик ишлов бериш ва лингвостатистик таҳлил этиш борасида бирмунча тадқиқотлар амалга оширилган. М.Маҳмудов, М.Айимбетов, С.Каримов, Г.Жуманазарова, А.Бабанаров, Д.Ўринбоева, А.Норов, С.Муҳаммедов, Н.Абдурахмоноваларнинг тадқиқотларини ана шундай ишлар сифатида қайд этиш мумкин²³.

Ўзбек тили морфологик анализаторини тузиш махсус тадқиқот предмети бўлмаган, аммо айрим изланишларда масаланинг муайян қирралари ҳақида фикр билдирилган. Чунончи, ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш²⁴, ўзбек-инглиз тили машина

²¹ Kemal Oflazer. Two-level Description of Turkish Morphology. *Literary and Linguistic Computing*, – Vol. 9, No 2, – 1994.; Altintas K., Cicekli I. A morphological analyzer for Crimean Tatar // *Proceedings of the 10th Turkish Symposium on Artificial Intelligence and Neural Networks (TAINN'2001)*. – 2001. – P. 180-189.; Çağrı Çöltekin (2014) A Set of Open Source Tools for Turkish Natural Language Processing In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)* Ed. by N. Calzolari et al. 1079-1086.; G. Kessikbayeva and I. Cicekli, "Rule based morphological analyzer of Kazakh language," in *Proceedings of the 2014 Joint Meeting of SIGMORPHON and SIGFSM*. Baltimore, Maryland: ACL, June 2014, pp. 46-54.; Tantug C., Adali E., Oflazer K. Computer Analysis of the Turkmen Language Morphology, 5th International Conference on NLP (FinTAL 2006), Turku, Finland, 186-193, 2006.; Orhun M., Tantug C., Adali E. Rule Based Analysis of the Uyghur Nouns, *International Journal of Asian Lang. Proc.*,19(1), 33-44, 2009.

²² Морфологический анализатор текста на русском языке *mystem* [Электронный ресурс]; // Компания Яндекс [сайт] 2003-2013. URL: <http://company.yandex.ru/technologies/mystem/>; <http://tamgasoft.kg/morfo/>; <http://tatmorphon.pythonanywhere.com/>

²³ Маҳмудов М.А., Пиотровская А.А., Садыков Т. Система машинного анализа и синтеза тюркской словоформы // *Переработка текста методами инженерной лингвистики*. – Минск, 1982.; Муҳаммедов С.А. Статистический анализ лексико-морфологической структуры узбекских газетных текстов: Автореф. дисс... канд. фил. наук. – Ташкент, 1980.; Бабанаров А. Разработка принципов построения словарного обеспечения турецко-русского машинного перевода: Автореф. дисс... канд. фил. наук. – Л., 1981.; Айимбетов М.К. Опыт лингвостатистического анализа лексики и морфологии каракалпакского публицистического текста: Автореф. дисс... канд. фил. наук. – Ташкент, 1987.; Каримов С., Қаршиев А., Исроилова Г. Абдулла Қаҳхор асарлари тилининг луғати. Алфавитли луғат. Частотали луғат. Терс луғат. – Тошкент, 2007.; Ризаев С. Ўзбек тилининг лингвостатистик тадқиқи: Фил. фан. док. дис...автореф. – Тошкент, 2008.; Пўлатов А. Компьютер лингвистикаси. – Тошкент: Akadempnashr, 2011.; Норов А. Компьютер лингвистикаси асослари. – Қарши, 2017. – 136 б. Жуманазарова Г.У. Фозил Йўлдош ўғли дostonлари тилининг лингвопoeтикаси: Фил. фан. док. дис...автореф. – Тошкент, 2017.; Ўринбоева Д.Б. Ўзбек фольклори матнларининг лингвостатистик тадқиқи. – Тошкент: Фан, 2010.; Ўринбаева Д. Халқ оғзаки ижоди: жанрий-лисоний ва лингвостатистик тадқиқ муаммолари: фил. фан. док. дисс. автореф. – Самарқанд, 2019. – 74 б.

²⁴ Муҳаммедова С. Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. Методик қўлланма. – Тошкент, 2006.;

таржимасининг лингвистик таъминоти²⁵, ўзбек тили муаллифлик корпусини тузиш тамойиллари²⁶, ўзбек тили бирликларини графематик таҳлил қилиш муаммолари²⁷, тил корпуси лингвистик базасини тузиш тамойиллари²⁸ монографик планда ўрганилган. Шунингдек, ўзбек тили атов бирликларини семантик теглашнинг лингвистик асослари²⁹, ўзбек-инглиз параллел корпуси тузиш муаммолари³⁰, ўзбек тилидаги отларни автоматик таҳлил қилиш³¹, сўз ясалишининг формал моделлари³² ҳақида қатор мақолалар нашр этилган. Ўзбек тилида от ва сифат сўз туркуми морфологик анализини автоматлаштиришнинг дастурий томонлари ўрганилган³³. Бу тадқиқотлар компьютер лингвистикаси муаммоларини тадқиқ этганлиги билан долзарблик касб этган, аммо ўзбек тили автоматик морфологик таҳлил дастури – морфоанализатор ва унинг лингвистик таъминотини яратиш тамойиллари масаласи кун тартибига махсус қўйилмаган. Диссертацияни тайёрлаш жараёнида юқорида санаб ўтилган илмий изланишлар атрофлича ўрганилди,

²⁵ 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. – 52 б.

²⁶ Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол. фан. бўйича фалсафа доктори (PhD)...диссер. – Қарши, 2018. – 250 б.

²⁷ Абжалова М.А. Ўзбек тилидаги матнларни таҳрир ва таҳлил қилувчи дастурнинг лингвистик модуллари (Расмий ва илмий услубдаги матнлар таҳрири дастури учун): Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Фарғона, 2019. – 164 б.; Абжалова М. Матнларга автолингвистик ишлов бериш тизимлари // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

²⁸ Эшмўминов А.А. Ўзбек тили миллий корпусининг синоним сўзлар базаси: Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Қарши, 2019. – 140 б.;

²⁹ Akhmedova D.B., Mengliev B.R. Semantic Tag Categories in Corpus Linguistics: Experience and Examination International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-3S, October 2019. – P. 208-212.

³⁰ Karimov R., Mengliev B. Theoretical fundamentals of uzbek-english parallel corpus / Journal of critical reviews. ISSN- 2394-5125. – VOL 7, ISSUE 17, 2020. – P. 73-76.; Karimov R.A., Mengliev B.R. The Role of the Parallel Corpus in Linguistics, the Importance and the Possibilities of Interpretation International Journal of engineering and Advanced Technology (IJEAT). ISSN: 2249 – 8958, Volume-8, Issue-5S3 July 2019. – P. 388-391.

³¹ Орхун М. Computational analysis of uzbek nouns / Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

³² Турсунов А. Вопросы словообразования в формальных моделях тюркских языков (на примере узбекского языка) // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». Труды конференции) – Ташкент, 2018. – 320 с.

³³ Болтаев Т.Б., Ибрагимов С.И. О проекте программной системы морфологического анализа узбекского языка // <http://buxdu.uz/index.php/uz/>; Болтаев Т.Б., Кузьминов Т.В., Поттосин И.В. О структурном конструирование программ и инструментах его поддержки // Среда программирования: методы и инструменты. Новосибирск, 1992, С22-37.; The Structured Constructing as a Discipline of Safe Programming and Instruments Supporting It / Aniskov M.I., Boltaev T.B., Kochetov D.V. at al // Instrumental Congress on Computer Systems and Applied Mathematics CSAM'93. St-Petersburg. July 19-23.

зарур ўринларда уларга муносабат билдирилди ва улардан тадқиқотда фойдаланилди.

Тадқиқотнинг диссертация бажарилган олий таълим ёки илмий-тадқиқот муассасасининг илмий-тадқиқот ишлари режалари билан боғлиқлиги. Диссертация Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университетининг АМ-ФЗ-201908172 рақамли “Ўзбек тили таълимий корпусини яратиш” амалий лойиҳаси доирасида бажарилган.

Тадқиқотнинг мақсади ўзбек тили морфологик анализаторининг лингвистик таъминоти назарий асосларини ва моделларини яратишдан иборат.

Тадқиқотнинг вазифалари куйидагилардан иборат:

лингвистик таъминот воситаларининг умумий тавсифи ҳамда унинг морфологик анализатор ва тил корпуслари яратишдаги аҳамиятини ўрганиш;

тил бирликларининг автоматик морфологик анализ/синтез жараёни воситаларини ўрганиш;

лингвистик таъминотнинг тузилиши, таркиби ва вазифаларини тадқиқ этиш;

морфологик анализатор тузишнинг умумий тамойиллари ҳамда унинг ахборот-кидирув тизимидаги аҳамиятини аниқлаш;

туркий тиллар морфологик анализаторларининг муштарак ва фарқли жиҳатларини тадқиқ этиш;

ўзбек тили морфоанализатори бирликлари, тузилиши, таркиби ва таҳлил бирликларини аниқлаш;

сўзда луғавий шакл ясовчи, синтактик шакл ясовчи, сўз ясовчи морфемалар ҳамда кўшимча кўринишидаги юкламалар қуршовини лингвистик моделлаштириш.

Тадқиқотнинг объекти сифатида ўзбек тили морфологик анализаторининг лингвистик таъминоти танланган.

Тадқиқотнинг предметини лингвистик таъминот бирликларини моделлаштириш ҳамда тизимлаштириш асослари ташкил этади.

Тадқиқотнинг усуллари. Тадқиқот мавзусини ёритишда таснифлаш, тавсифлаш, қиёслаш, статистика, моделлаштириш усулларидан фойдаланилди.

Тадқиқотнинг илмий янгилиги қуйидагилардан иборат:

Ўзбек тилида ахборот қидирув тизимининг лингвистик таъминоти таркиби, вазифаларини аниқлаш орқали орфо корректор, лемматизатор, морфоанализатор, парсер каби лингвистик анализаторларининг хусусий функционал аспекти аниқланган;

ахборот-қидирув тизимида морфологик анализаторнинг матнни лингвистик аннотациялаш ҳамда машина таржимасидаги табиий тилга автоматик ишлов беришнинг асосий воситаси эканлиги далилланган;

туркий тилларда ишлаб чиқилган морфологик анализаторлар функционал имкониятининг лингвистик таъминот жиҳатидан фарқли, тилга автоматик ишлов беришнинг морфологик таҳлил усуллари жиҳатидан эса муштарак эканлиги аниқланган;

Ўзбек тилидаги имловий ўзгаришлар туфайли сўз шакллариининг умумий қоидаларидан фарқланувчи сўзлар учун морфотактиканинги комбинатор ва позицион лингвистик таъминоти дастури ишлаб чиқилган;

Ўзбек тили бирликларини автоматик морфологик таҳлил қилиш учун сўз ясовчи, луғавий шакл ясовчи, синтактик шакл ясовчи морфемалар ҳамда қўшимча шаклидаги юктамаларнинг сўзга бирикиш позициялари дастурий таъминот алгоритми учун лингвистик моделлаштирилган.

Тадқиқотнинг амалий натижалари қуйидагилардан иборат:

тил бирликлари автоматик морфологик анализ/синтез жараёни воситалари таҳлил қилинган, морфологик анализатор тузишнинг умумий тамойиллари аниқланган, унинг ахборот-қидирув тизими, тил корпуслари, машина таржимаси ҳамда компьютер лексикографиясидаги аҳамияти далилланган;

туркий тиллар морфологик анализаторларининг муштарак ва фарқли жиҳатлари кўрсатилган, ўзбек тили морфоанализатори бирликлари, тузилиши, таркиби тавсифланган, тадқиқот иловасида таҳлил бирликлари рўйхати шакллантирилган;

ўзбек тили морфологик анализатори ва унинг лингвистик таъминотини ишлаб чиқиш бўйича таклиф берилган;

морфоанализатор учун сўз ясовчи, луғавий шакл ясовчи, синтактик шакл ясовчи морфемалар ва қўшимча кўринишидаги юкламалар қуршови дастурлаш учун лингвистик моделлаштирилган.

Тадқиқот натижаларининг ишончлилиги ўрганилган материалларнинг ўзбек тили табиатидан келиб чиққан ҳолда хулосалар қилишга ёрдам берганлиги, уларнинг асосли эканлиги, методологик мукамаллиги, ўзбек тили морфологик анализаторини тузиш тамойилларини ишлаб чиқишда амалда ишотланган манбаларга таянилганлиги билан изоҳланади.

Тадқиқот натижаларининг илмий ва амалий аҳамияти. Тадқиқот натижаларининг илмий аҳамияти ўзбек тили морфологик анализатори яратишнинг назарий асослари, унинг лингвистик таъминотини ишлаб чиқиш доирасидаги янги тадқиқотлар учун илмий-назарий манба бўла олиши, жумладан, тил бирликларининг автоматик морфологик анализ/синтез жараёни воситалари бўйича назарий маълумотлар бериши билан белгиланади.

Тадқиқот натижаларининг амалий аҳамияти амалий тилшунослик, компьютер лингвистикаси, корпус лингвистикаси фанлари ўқитилиш жараёнида дастур, режа тузиш, мавзуларни баён этишда манба вазифасини ўташи, морфологик анализатор тузишнинг умумий тамойиллари, унинг ахборот-қидирув тизими, тил корпуслари, машина таржимаси ҳамда компьютер лексикографиясининг ривожига хизмат қилиши, ўзбек тилидаги матнларни морфологик аннотациялашда қўлланилиши билан изоҳланади.

Тадқиқот натижаларининг жорий қилиниши. Ўзбек тили морфологик анализатори лингвистик таъминоти, морфологик анализатор тузишнинг умумий тамойиллари, морфологик анализаторнинг ахборот-қидирув тизимидаги аҳамияти, ўзбек тили морфоанализатори бирликлари, тузилиши, таркибига доир хулосалардан:

Ўзбек тили морфоанализатори бирликлари; тузилиши, таркиби ва таҳлил бирликларини аниқлаш, сўз ясовчи, луғавий шакл ясовчи, синтактик шакл ясовчи ва кўшимча кўринишидаги юкламалар куршовини лингвистик моделлаштиришга оид натижалардан Ўзбекистон Республикаси Фанлар Академияси Қорақалпоғистон бўлими Қорақалпоқ гуманитар фанлар илмий-тадқиқот институти томонидан олиб борилган ФА-Ф1-Г003 «Ҳозирги қорақалпоқ тилида функционал сўз ясалиши» мавзусидаги фундаментал илмий лойиҳада фойдаланилган (Олий ва ўрта махсус таълим вазирлигининг 2020 йил 20 ноябрдаги 89-03-4852-сонли маълумотномаси). Натижада ҳозирги қорақалпоқ тилида функционал сўз ясалиши моделлари мукамаллаштирилган; қорақалпоқ тилидаги шакл ясовчи морфема, кўшимча кўринишидаги юкламалар омонимиясини автоматик фарқлаш тамойиллари ишлаб чиқилган;

ахборот-қидирув тизимида ўзбек тили бирликларини қўллашнинг аҳамияти, ахборот-қидирув тизимининг тил маданиятини шакллантиришдаги ўрни ҳақидаги хулосалардан 2017-2020 йилларда Андижон давлат университетида бажарилган “ОТФ-1-18/ Оммавий лисоний маданиятни шакллантириш методлари ва методологиясини ишлаб чиқиш” давлат-фундаментал тадқиқотлар дастурлари доирасидаги лойиҳада фойдаланилган. (Олий ва ўрта махсус таълим вазирлигининг 2020 йил 20 ноябрдаги 89-03-4852-сонли маълумотномаси). Натижада оммавий лисоний маданиятни шакллантириш методлари ва методологиясини ишлаб чиқишда ахборот-қидирув тизимида ўзбек тили бирликларини кенг қўллаш тамойиллари ишлаб

чиқилган, ахборот-қидирув тизимининг тил маданиятини шакллантиришдаги ўрнини асослашга эришилган;

сўзда бошқа грамматик шакллар билан омонимия ҳосил қилувчи луғавий шакл ясовчи ва синтактик шакл ясовчи морфемалар; контекстда сўз ясовчи морфеманинг омонимик ҳамда полифункционаллик хусусиятини аниқлаш учун сўздаги қуршовни моделлаштириш натижаларидан Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университетида бажарилган I-OT-2019-42 рақамли «Ўзбек ва инглиз тилларининг электрон (Инсон қиёфаси, феъл-атвори, табиат ва миллий тимсоллар тасвири) поэтик луғатини яратиш» мавзусидаги фундаментал илмий лойиҳада фойдаланилган (Олий ва ўрта махсус таълим вазирлигининг 2020 йил 20 ноябрдаги 89-03-4852-сонли маълумотномаси). Натижада ўзбек ва инглиз тилларида инсон қиёфаси, феъл-атвори, табиат ва миллий тимсоллар тасвирини билдирувчи сўзларнинг контекстдаги омонимиясини автоматик аниқлашга эришилган;

ўзбек тили морфологик анализаторининг лингвистик таъминоти, морфологик анализатор тузишнинг умумий тамойиллари, ўзбек тили морфоанализатори бирликлари, тузилиши, таркиби ва таҳлил бирликлари, морфотактика қоидаларини тузиш натижаларидан “Ўзбек тилининг миллий ва муаллифлик корпуси” учун ЭҲМ дастури”ни тузишда фойдаланилган (Ўзбекистон Республикаси Интеллектуал мулк агентлиги томонидан 31.10.2018 санасида берилган № DGU 05735 рақамли гувоҳнома). Натижада ўзбек тили миллий ва муаллифлик корпуси лингвистик таъминоти ишлаб чиқилган, тил корпуси бирликлари, таҳлил бирликлари, теглар тизими сираси тўлдирилган;

сўз ясовчи қўшимчаларнинг қуршовдаги маъносини аниқлаш учун таклиф этилган қўшимчалар аранжировкаси, сўз ясовчи қўшимчаларнинг сўздаги позициясига қараб маъносининг ўзгариши хусусидаги хулосалардан 5220100 – Филология ва тилларни ўқитиш (ўзбек тили) бакалавр таълим йўналиши талабаларига мўлжалланган “Ҳозирги ўзбек тили” дарслигининг

“Ясама сўзларнинг лексемалашувчи” деб номланган 13-параграфиди ўзбек тили лексикасидаги лисоний ясама сўз ҳамда соф нутқий ясама сўзларнинг ясаиш структурасини тавсифлашда фойдаланилган (Олий ва ўрта махсус таълим вазирлигининг 2020 йил 20 ноябрдаги 89-03-4852-сонли маълумотномаси). Натижада ўзбек тили лексикасидаги лисоний ясама сўз ҳамда соф нутқий ясама сўзларнинг ясаиш структураси тавсифи мукаммаллашган;

ўзбек тили морфологик анализаторининг табиий тилга автоматик ишлов бериш усуллари: стемминг, парсинг, таггинг натижаларидан 5A120106 – Компьютер лингвистикаси магистратура мутахассислигига мўлжалланган “Korpus lingvistikasi: korpus tuzish va undan foydalanish”. Amaliy mashg`ulot uchun qo`llanma”сида фойдаланилган (Олий ва ўрта махсус таълим вазирлигининг 2020 йил 20 ноябрдаги 89-03-4852-сонли маълумотномаси). Натижада “Корпус лингвистикаси: корпус тузиш ва ундан фойдаланиш” фанини ўқитишда адабиёт ва ахборот манбаларилар сираси бойитилган;

оммавий саводхонликни ошириш, ёшлар орасида тил маданиятини шакллантириш ҳамда замонавий ахборот тизимида ўзбек тилининг имкониятларидан фойдаланишда, интернетда мавжуд бўлган автоматик таржимон, кирилл-лотин конвентори, компьютер таҳлили, таҳрири, тил корпуси, электрон луғат ва тезаурусларнинг аҳамияти ҳақидаги хулосалардан Бухоро вилояти телерадиокомпанияси «Бухоро» телеканалининг «Ассалом, Бухоро» телекўрсатувида фойдаланилган (Бухоро вилояти телерадиокомпаниясининг 2020 йил 12 ноябрдаги № 1/296-сонли маълумотномаси). Натижада юқори савияли баркамол авлодни вояга етиштириш, ёшлар орасида тил маданиятини шакллантириш ҳамда замонавий ахборот тизимида ўзбек тилининг имкониятларидан фойдаланиш ҳақидаги тарғибот ишлари мазмуни мукаммаллашган.

Тадқиқот натижаларининг апробацияси. Мазкур тадқиқот натижалари 12 та халқаро, 18 та республика анжуманида муҳокамадан

Ўтказилган ҳамда қуйидаги профилларда ҳаволага эришилган: orcid.org/0000-0002-5429-4708, www.researchgate.net/profile/Shahlo_Khamroeva, <https://scholar.google.com/citations>.

Тадқиқот натижаларининг эълон қилиниши. Диссертация мавзуси бўйича 54 та илмий иш нашр эттирилган, жумладан, 2 та монография, Ўзбекистон Республикаси Вазирлар Маҳкамаси ҳузуридаги Олий аттестация комиссиясининг докторлик диссертациялари асосий илмий натижаларини чоп этиш учун тавсия этилган илмий нашрларда 11 та илмий мақола, улардан 4 таси хорижий журналларда чоп қилинган.

Диссертациянинг ҳажми ва тузилиши. Диссертация кириш, тўрт боб, хулоса, фойдаланилган адабиётлар рўйхати ва иловадан иборат. Ҳажми 264 саҳифани ташкил этади.

I БОБ. ЛИНГВИСТИК ТАЪМИНОТНИНГ АХБОРОТ-ҚИДИРУВ ТИЗИМЛАРИДАГИ ЎРНИ

1.1. Лингвистик таъминот воситаларининг умумий тавсифи, морфологик анализаторнинг тил корпуслари яратишдаги аҳамияти

Ҳозирги ахборот технологияларида инсон ва ахборот тизимлари нутқий мулоқоти(коммуникатив ўзаро таъсир)нинг асосий аспектларини қамраб олувчи назарий модел мавжуд эмас. Аммо коммуникатив ўзаро таъсир аспектида қатор фундаментал тадқиқотлар бажарилмоқда³⁴.

Маълумки, лингвистик таъминот тушунчасига турли ёндашув мавжудки, улар орасида энг машҳури классик ёндашув ҳисобланади. Бундай ёндашувда лингвистик таъминот сирасига ахборот-қидирув тили, таснифлаш ва вербал тиллар мажмуи киритилади. Одатда, “лингвистик таъминот” тушунчаси остида библиографик маълумотлар тили ҳам англашилади. Ушбу атамага лингвистик ёндашув ҳам мавжуд, у ахборот тизимларига лингвистик нуқтаи назар асосида қарашдан келиб чиқади, асосан, амалий ва компьютер лингвистикаси мутахассислари томонидан ривожлантирилади. Шу ёндашувдан келиб чиқиб, лингвистик таъминотга табиий тил матнларини автоматик қайта ишлаш (сўров тили ва қидирув)да қўлланадиган воситалар мажмуи, тил процессори деб қараш тўғрироқ бўлади³⁵.

Ёндашувлар орасида нисбатан умумийси – семантик ёндашув; у тил ҳақидаги классик-семиотик қарашлар (тил – турли рамзлар мажмуи, табиийки, алифбодан бошлаб) асосида шаклланган. Бундай ёндашув лингвистик таъминот ахборотига маълумот ва унинг интерпретацияси сифатида қарашдан келиб чиққан. Бунда алифбони кодлаш воситаси, маълумотни акс эттириш форматида лингвистик таъминот сифатида қараш ўринли, аммо дастурлаш тилини бунга аралаштириш мумкин эмас.

³⁴http://tspu.ru/res/informat/sist_seti_fmo/lekcii/lekcii-4.html#name3

³⁵Ўша манба.

Лингвистик таъминот – у ёки бу ижтимоий мулоқот доирасида тил воситаларининг адекват ишлашини таъминловчи белгилар мажмуи³⁶, ахборот технологиялари дастурий таъминотида ишлатилувчи тиллар, уларни яратиш усул/воситалари, қўллаш ва фойдаланиш комплекси. Лингвистик таъминот сирасига халқаро тасниф, рубрикатор, ахборот-қидирув тиллари тавсифи, луғат ва тезаурус киради. *Автоматлаштирилган тизимнинг лингвистик таъминоти* атамаси остида табиий тилдан умумистеъмол қўллаш ва эксплуатация қилиш жараёнида фойдаланиш учун формаллаштириш қоида ҳамда воситалари тўплами тушунилади³⁷. Адабиётларда лингвистик таъминотга берилган шарҳда “лингвистик таъминот инсон ва сунъий интеллект мулоқотини таъминлаш самарадорлигини ошириш учун маълумотни қайта ишлаш тизимини яратиш, фойдаланишнинг турли боскичида ишлатиладиган тил воситалари йиғиндисини ўзида намоён этади”,³⁸ деган умумий хулоса мавжуд.

Автоматлаштирилган тизимнинг лингвистик таъминоти автоматлаштирилган тизимда қўлланувчи, ахборотни автоматик таҳлил қилишни унумли даражага кўтариш мақсадида табиий тилнинг формал қоидаларини яратишга хизмат қилувчи тил воситаларидан иборат³⁹. Ахборот тизимларининг лингвистик таъминоти деганда ахборот объекти тузилиши, унинг ишлаши учун фойдаланиладиган, формаллашган, табиий тил ифодалядиган тушунча устида бажариладиган амаллар, унинг воситалари таснифи ва тўплами назарда тутилади. Автоматик лойиҳалаш воситаларининг умумий мажмуини унга нисбатан хусусий бўлган турли таъминотлар мажмуи ташкил этади. Бундай таъминотнинг математик (1), лингвистик (2), техник (3),

³⁶https://sociolinguistics.academic.ru/915/%D0%9B%D0%B8%D0%BD%D0%B3%D0%B2%D0%B8%D1%81%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%BE%D0%B5_%D0%BE%D0%B1%D0%B5%D1%81%D0%BF%D0%B5%D1%87%D0%B5%D0%BD%D0%B8%D0%B5

³⁷Ўша манба

³⁸Ахборотлар тизимлари ва технологиялари (С.С.Ғуломовнинг умумий таҳрири остида). – Тошкент: Шарқ, 2000. – Б. 56

³⁹Ўша манба. – Б. 59.

информацион (4), дастурий (5), методик (6), ташкилий (7) турлари ажратилади⁴⁰.

Дастурий таъминотнинг лингвистик таъминот қисми автоматик лойиҳалаш учун зарур бўлган маълум тил бирликлари (атама, аниқловчи, тилни формаллаштириш қодалари, тил бирлигини қисқартириш ва умумлаштириш) шакллари, тўпламини сақлайди. Ахборот технологиялари соҳасида қўлланувчи тиллар алгоритмик тил саналади ва улардан маълум бир ахборотни қайта ишлаш алгоритмини яратишда фойдаланилади; лингвистик таъминот шу тилга тегишли символ ва қодалар тизимини тўплаш орқали яратилади⁴¹. Қўллаш соҳасидан қатъи назар, ахборот тизими маълумотини қайта ишлашнинг деярли барча тизими таъминлаш турлари деб аталадиган таркибий қисми ўз ичига олади. Юқорида айтилганидек, улар дастурий, техник, ҳуқуқий, ахборот-ташкилий, математик ва лингвистик таъминотга ажратилади. *Ахборот таъминоти* ахборот тизимлари ички машина ахборот базасини яратишни таснифлаш, кодлаштириш тизими, ҳужжатлаштиришнинг унификациялашган тизими, ҳужжат айланмаси ва ҳужжат шакли услубини рационал ҳолга келтиришни ўз ичига олган ахборотни жойлаштириш, ташкил қилиш усул ва воситалар йиғиндиси саналади. *Дастурий таъминот* ҳисоблаш техникаси воситасида маълумотни қайта ишлаш тизимини яратиш ва фойдаланиш дастурий воситалари мажмуи ҳисобланади. *Техник таъминот* эса маълумотни қайта ишлаш тизимини фаоллаштиришда қўлланувчи техник воситалар комплекси, у маълумотни қайта ишловчи намунавий амалларни бажарувчи қуролмаларни ўз ичига олади. Шунингдек, *ҳуқуқий таъминот* ахборот тизимини яратиш, функциялаштиришни тартибга солувчи ҳуқуқий меъёрлар йиғиндиси сифатида намоён бўлади. Ахборотни қайта ишлашнинг ҳуқуқий таъминоти АТ буюртмаси ва тайёрловчи ўртасидаги шартномавий

⁴⁰<https://works.doklad.ru/view/IP7CrszcUzY.html>

⁴¹Митрофанов В.Г., Калачев О.Н., Схирладзе А.Г. и др. САПР в технологии машиностроения. Учеб. пособие. – Ярославль; Ярослав. гос. техн. ун-т, 1995. – 298с.

ўзаро муносабат меъёрий актининг ҳуқуқий тартибга солинишини ўз ичига олади.

Лингвистик дастурий таъминот (инг. *lingware* = linguistic + software) – табиий тилдаги аудиомаълумот, расм ва матнни таҳлил қилиш, қайта ишлаш, сақлашга оид дастур ва маълумотлар мажмуи. Лингвистик дастурий таъминотнинг қуйидаги турлари ажратилади⁴²:

- 1) электрон луғат: Викисловарь, Мультитран, АБВУД Lingvo;
- 2) орфоқорректор (ёки спеллчекер): Орфо, MS Word;
- 3) қидирув тизими;
- 4) машина таржимаси тизими: PROMT, Socrat, Apertium ва х.;
- 5) автоматлаштирилган таржима тизими, таржимон хотирасини бошқариш дастури: OmegaT, STAR Transit NXT, Trados, MetaTaxis, Deja Vu ва б.;
- 6) OCR: Finereader, CuneiForm, Tesseract, OCRopus;
- 7) нутқ анализи тизими: Dragon, IBM via voice;
- 8) нутқ синтези тизими: Агафон;
- 9) овозли таржима тизими (овозни аниқлаш, таниш ва синтез қилиш): Speereo.

Ахборот қидирув тили ва ахборот қидирув дастурлари. Ахборот қидирув дастурлари (АҚД) яратилган пайтдан бошлаб ахборот-қидирув тили (АҚТ) уларнинг энг асосий қисми ҳисобланган. Ўтган асрнинг 60-70-йилларга келиб АҚТни ишлаб чиқиш муаммоларига ҳам назарий, ҳам амалий жиҳатдан алоҳида эътибор қаратиш бошланган. Н.И.Гендина маълумотларига қараганда, Россияда АҚТ ҳақидаги энг кўп нашрлар 1974-1978-йилларга тўғри келади⁴³. Бу давр ишларида АҚТ табиий тиллар билан қиёсий ҳолда ўрганилган. Асосий эътибор аниқ АҚТни яратиш, қўллаш усуллари(индекслаш)га қаратилган. Ахборот-қидирув тили билан бирга

⁴²https://ru.wikipedia.org/wiki/%D0%9B%D0%B8%D0%BD%D0%B3%D0%B2%D0%B8%D1%81%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%BE%D0%B5_%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%BC%D0%BD%D0%BE%D0%B5_%D0%BE%D0%B1%D0%B5%D1%81%D0%BF%D0%B5%D1%87%D0%B5%D0%BD%D0%B8%D0%B5

⁴³https://studopedia.ru/6_11435_ponyatie-i-sostav-lingvisticheskogo-obespecheniya.html

лингвистик таъминот (ЛТ) атамаси кенг қўллана бошлаган. Бу атама (ЛТ) остида ахборот қидирувининг асосий вазифаси – ҳужжат мазмуни ва қидирувни белгилаб берадиган лисоний ҳамда мантиқий восита/усуллар мажмуи тушунилган⁴⁴. Уларга индекслаш усуллари, тезаурус яратиш методикаси, турли қоидалар, АҚТни яратиш ва қўллашнинг ёрдамчи воситалари киради. Дастурлаш тили ва АҚТнинг вазифалари батамом бошқа-бошқа. Шунинг учун бу икки тил орасидаги чегарани аниқ белгилаб олиш лозим.

Ахборот-қидирув тизимининг лингвистик таъминоти таркибига қуйидагилар киради⁴⁵:

1. Ахборот-қидирув тили (ҳужжат мазмуни ва қидирув тили):

- a) алифбо ва АҚТ микросинтаксиси;
- b) мазмун режаси(семантика)ни ифодалаш воситаси: луғат, тезаурус, ёрдамчи луғатлар (қисқартма ва ўлчов бирликлари тизими);
- c) классификатор;
- d) АҚТнинг граматик воситалари (қўшимча, ёрдамчи сўзлар);
- e) табиий тил маълумотини ифодалаш воситалари (матнни табиий тилда ифодалаш воситалари, кўптиллилиқ муаммоси ва ҳ.).

2. Турли тилларни мослаштиришни таъминловчи воситалар:

- 1) мослик жадвали;
- 2) конвентрлаш жадвали;
- 3) протоколлар.

3. Амалиёт воситалари:

- 1) индекслаш методикаси;
- 2) автоматик индекслаш алгоритми;
- 3) қидирув воситаси: (сўров тили, маълумотлар базасига автоматик сўров бериш воситаси, ягона буйруқ тили).

⁴⁴http://tsput.ru/res/informat/sist_seti_fmo/lekcii/lekcii-4.html#name3

⁴⁵https://studopedia.ru/6_11435_ponyatie-i-sostav-lingvisticheskogo-obespecheniya.html

4. Мазмуний мутаносиблик мезонлари.

5. Тузилиш тиллари ва маълумотларнинг метатавсифи.

6. Матнларга табиий тилда ишлов беришнинг ёрдамчи воситалари: морфологик ва синтактик таҳлилнинг дастурий-алгоритмик воситалари; матн маълумотларини семантик қайта ишлаш воситаси; материални автоматик таснифлаш алгоритми ва лингвистик воситаси; автоматик қидирув буйруқларини тузиш алгоритми, лингвистик воситалари; хатони автоматик аниқлаш воситаси; лингвистик маълумотлар банки, табиий тилдаги лексик (сўз), морфологик бирлик (асос, қўшимча, флектив шакл) ва синтактик бирлик (сўз бирикмаси) кабиларнинг машина луғатлари.

7. Лингвистик таъминотнинг таянч воситалари: машина луғатлари тузиш амаллари; машина луғатлари билан ишлаш жараёни тавсифи; лингвистик таъминотни ҳужжатлаштириш.

Сунъий интеллект ва инсон тили, умуман барча тилларнинг асосий вазифаси – ахборотни семантик жиҳатдан қайта ишлаш. Матн мазмунини тавсифлаш – назарий ва амалий тилшунослик муаммоларидан. Матн ёки маълумот мазмунини формал тавсифлашни кўзда тутадиган сунъий тиллар ахборот-қидирув тили дейилади. Сунъий тилларнинг асосий вазифаси табиий тиллардаги синонимия ва омонимия муаммоларини АҚТ учун ечишдан иборат бўлади.

В.А.Москович томонидан (Э.Сепирнинг табиий тиллар таснифига асосланиб) ишлаб чиқилган ахборот тилининг морфологик таснифи тилшуносларда катта қизиқиш уйғотди. Э.Сепир тилда англашиладиган барча тушунчани тўрт гуруҳга ажратади. Турли тилларда бу тасниф турлича (изоляция, агглютинация, фузия, символизация) бўлиши мумкин. Бундан ташқари, олим томонидан ўзакка аффиксал элемент қўшилишининг турли даражаси аниқланган. В.А.Москович морфологик тилнинг 8 турини аниқлаган; уларнинг ҳар бирига ҳам табиий тилдан, ҳам ахборот тилидан мисоллар келтирган, далиллаган. Шунингдек, лингвистик асосларга

таянадиган ахборот тили тузишни таҳлил қилиш усули ҳам соҳа олимларида қизиқиш уйғотган⁴⁶.

Ахборот қидирув тизими(АҚТ)нинг лингвистик таъминоти сирасига автоматик таржима луғатларда қўллашга мўлжалланган АҚТнинг асосий ва ёрдамчи воситалари, мультибазавий тил воситалари, диалогли қидирув режимининг тилга оид махсус воситалари ҳам киритилади⁴⁷. *Лингвистик таъминот* (лингвистическое обеспечение) атамасини В.В.Пономарёв рус тили луғатлари асосида шундай изоҳлайди⁴⁸: “лингвистический – а) языковедческий – относящийся к лингвистике; б) языковой – относящийся к языку. Обеспечение – то, что обеспечивает исполнение, реализацию. Средство – а) прием, способ действия для достижения чего-либо; б) орудие для осуществления какой-либо деятельности”. Таъминот – юзага чиқариш, амалга оширишни таъминловчи, восита эса а) нимагадир эришиш усул ва йўллари; б) бирор фаолиятни амалга ошириш қуроли) саналади.

Дастлаб ушбу тушунчани ўзбек тилида қайси атама билан номлаш тўғрироқ бўлиши таҳлил қилиниши лозим. Ўзбек тилининг изоҳли луғатида ушбу сўзлар қуйидагича изоҳланади⁴⁹:

ТАЪМИН *а.* (кафолат, таъминлаш; гаров; хавфсизлик; “омин” дейиш, омонлик бериш]: таъмин қилмоқ (*ёки* этмоқ) 1) керакли нарсаларни етказиб бермоқ, таъминламоқ. *Керакли маҳсулотлар билан таъмин этмоқ. – Отам учун икки кишини шаҳарга юбориб, моддий жиҳатдан таъмин қилиш оғирлик қиларди.* С.Айний, Эсдаликлар; 2) бирор нарсанинг амалга ошиши, рўёбга чиқиши, мухайё бўлиши учун зарур шароит яратмоқ. *Муваффақиятни таъмин қилмоқ. – Ҳамроҳларим ўз танки билан душман зотини абжақ қилган*

⁴⁶https://studopedia.ru/6_11435_ponyatie-i-sostav-lingvisticheskogo-obespecheniya.html

⁴⁷ Пономарев В.В. Лингвистическое обеспечение и социолингвистическая специфика проблемы автоиндексационной актуализации информационных систем: автореф. диссер. канд. филол. наук. – Москва, 2005. – 25 с.

⁴⁸ Ўша манба. – Б. 14.

⁴⁹ Ўзбек тилининг изоҳли луғати. 5 жилдли. IV жилд. – Тошкент: "Ўзбекистон миллий энциклопедияси" Давлат илмий нашриёти, 2006. – Б. 21-22.

ва галабани таъмин этган. Н.Сафаров, Оловли излар; 3) кониктирмақ, каноатлантирмақ; кифоя қилмақ, етарли бўлмақ. *Бунинг сабабини сўзлашдан илгари имом афандининг ўтган тарихига бир оз кўз ташлаб олишга мажбурмиз. Чунки бусиз ўқувчимизни таъмин қилиш мумкин эмасдир.* А.Қодирий, Меҳробдан чаён.

ТАЪМИНОТ [а. –“таъмин”с. кўпл.] 1 Керакли нарсаларни муҳайё қилиб етказиб бериш; яшаш, ишлаш учун зарур нарсаларга бўлган эҳтиёжни кондириш. *Ижтимоий таъминот. Давлат таъминоти. Таъминот бўлими. –Таъминот ҳам тўла йўлга қўйилмаган: гоҳи ёқилги етказиб берилмайди... “Ўзбекистон кўриклари”.*

2 Яшаш, ишлаш учун керакли пул, озиқ-овқат, кийим-кечак каби нарсалар мажмуи. *...лашкар таъминотини зиёратчиширнинг назр-ниёзларига юклаганлар.* К.Яшин, Ҳамза.

3 кўчма кам қўлл. Маънавий озуқа; далда; умид. *Отасининг бу сўзлари унинг [Отабекнинг] истиқболи учун яхши таъминотлар берар.* А.Қодирий, Ўткан кунлар.

Кўриниб турибдики, ўзбек тилида атама сифатида *таъмин* ва *таъминот* сўзларини синоним сифатида талқин қилинган. *Таъмин* сўзининг биринчи ва иккинчи маъноси бу сўзни “лингвистик таъминот” (лингвистическое обеспечение) бирикмаси таркибида атама сифатида қўллаш моҳиятан тўғри эканлигини кўрсатса, *таъминот* сўзи биринчи маъноси билан ушбу бирикма таркибига киришга ҳақли. Демак, ўзбек тилида атама сифатида *лингвистик таъминот* ва *лингвистик таъмин* атамаларини синоним сифатида қўллаш мумкин.

Жаҳон тилшунослигида лингвистик таъминотнинг назарий, методологик муаммоларини ўрганиш ўтган асрнинг 60-йилларида бошланди. Бу вақтга келиб ахборот-қидирув тилисиз электрон ҳисоблаш машинаси машиналигича қолиб кетиши тан олинди. А.В.Соколов таъбири билан айтганда, “механик” босқичдан “мантикий-лингвистик” босқичга, ахборот-

кидирув тизими тараққийёт босқичига ўтилди⁵⁰. Бу муаммонинг ўрганилишига ғарб олимлари, қатор рус олимлари ҳам киришди, чунки ахборот-қидирув тизими (аксарият олимларнинг яқдил фикрига кўра) миллий характерга эга⁵¹. Лингвистик таъминот муаммолари ахборот-қидирув тизими билан боғлиқ барча тадқиқотларда турли даражада ўз ечимини топган. Чунки лингвистик таъминот ахборот-қидирув тизими, маълумотлар базаси, электрон каталогларнинг ажралмас қисми саналади. Лингвистик таъминотнинг аҳамияти ҳақидаги батафсил маълумотларни А.Б.Антопольский⁵², Ю.М.Арский⁵³, Г.Г.Белоногов, Б.А.Кузнецов⁵⁴, В.Н.Белоозеров⁵⁵, Р.Р.Мдивани, В.А.Глинский⁵⁶, А.И.Михайлов⁵⁷, В.А.Мишин⁵⁸, В.В.Морозов⁵⁹ ишларида ўқишимиз мумкин. А.В.Соколов «Информационно-поисковые системы» номли мақоласида АҚТ тузиш тамойиллари, АҚТ типлари, уларнинг вазифаси тавсифланган; индексациялаш турлари, уларнинг мавзувий кидирувдаги аҳамияти кўрсатиб берилган; автоматлашган АҚТнинг анъанавий қидирув тизимидан фарқлари очиқ берилган. А.Б.Антопольский, Г.Т.Артамоновнинг «Проблемы разработки лингвистического обеспечения

⁵⁰Соколов А.В. Автоматизация библиографического поиска / А.В. Соколов. – М.: Книга, 1981.– 167 с. – С.91.

⁵¹Пирумова Л.Н. Лингвистическое обеспечение информационно-поисковых систем библиотек АПК: Методика формирования и пути совершенствования: автор. дис.... канд. пед. наук. – Москва, 2003. (Электрон ресурс: <https://www.dissercat.com/content/lingvisticheskoe-obespechenie-informatsionno-poiskovykh-sistem-bibliotek-apk-metodika-formir>)

⁵² Антопольский А.Б. Лингвистическое обеспечение АСНТИ: структура и проблема совместимости // НТИ. Сер.2, 1983. – №7. – С.17-25.

⁵³ Арский Ю.М. Информационные ресурсы для устойчивого развития общества / Ю.М.Арский, А.И.Черный // Междунар. форум по информ. и док. –2000. –Т. 25, № 1. –С. 3-9.

⁵⁴ Белоногов Г.Г. Автоматизированная обработка научно-технической информации. Лингвистические аспекты / Г.Г. Белоногов, Б.А. Кузнецов, А.П. Новоселов // Итоги науки и техники. Сер. Информатика / ВИНТИ. –1984. –Т.8.–316.

⁵⁵ Белоозеров В.Н. Ведение и совершенствование Государственного рубрикатора НТИ (ГРНТИ) / В.Н. Белоозеров, В.И. Федосимов, Г.А. Кривенцова // Информ. продукты, процессы и технологии: НТИ-96: Материалы конф. –М., 1996.–С. 60-69.

⁵⁶ Мдивани Р.Р. Лингвистическое обеспечение автоматизированной информационной системы по общественным наукам: состав и структура / Р.Р.Мдивани, В.А.Глинский // Вопр. информ. теории и практики. 1985. –№52.–С. 37-55.

⁵⁷ Михайлов А.И. Проблемы лингвистического обеспечения АСНТИ / А.И. Михайлов // Вопр. информ. теории и практики. 1984.–№ 51. – С. 5-11.

⁵⁸ Мишин В.А. Лингвистическое обеспечение библиотечных автоматизированных систем / В.А. Мишин, Л. Санчес-Мехидо // Сов. библиотековедение. 1981. – № 2. – С. 84-93.

⁵⁹ Морозов В.В. Лингвистическое обеспечение автоматизированных систем научно-технической информации / В.В. Морозов, Б.Р. Певзнер. –М.: ИПКИР, 1987. – 76 с.

информационных сетей” номли мақоласида АҚТга мос лингвистик таъминот танлаш омиллари, глобал тармоқ учун информатсион-қидирув тизими бирликлари, шунингдек, маълум бир тармоқнинг лингвистик таъминотини яратиш зарурати изоҳланган. Б.Р.Певзнернинг “Лингвистическое обеспечение АСНТИ: расширенное толкование” номли мақоласи диалогик режимда ишловчи АҚТда лингвистик таъминот структураси тавсифига бағишланган. Н.И.Гендинанинг “Лингвистическое обеспечение автоматизированных библиотечных систем” номли монографиясида кутубхоначилик ишида ахборот-қидирув тизими лингвистик таъминоти принциплари умумлаштирилган. Н.И.Гендинанинг “Лингвистические средства автоматизации документального поиска” номли китобида маълумотлар базаси структураси, таркиби ўрганилар экан, ахборот тўплаш ва информатсион маҳсулот яратишда лингвистик воситанинг вазифаси очиб берилган; муаммони лингвистик таъминот воситасида ечиш йўллари таҳлилга тортилган. Н.И.Гендинанинг тадқиқотлари⁶⁰да лингвистик таъминотга тегишли кенг спектрдаги бир қадар саволлар (хусусан, индекслаш, ахборот қидируви) ўз жавобини топган.

Ўзбек тилида ахборотни қайта ишлаш тизимлари учун лингвистик таъминот яратишнинг назарий асосларини ишлаб чиқиш масаласи кўтарилган тадқиқотлар санокли. К.У.Шадманова “Табиий тиллар учун лексик-семантик луғат маълумотлар базасини яратиш тамойиллари” номли мақоласида *лингвистик база* атамасига таъриф беради. Мақолада ҳаракат феълларининг семантикаси ва валентлигини текшириш асосида компьютер дастурлари учун лингвистик таъмин яратиш намунасини келтиради. Тадқиқотчи лингвистик таъмин яратишда ҳаракат феълларининг олдиндан амалга оширилган семантик таснифидан фойдаланади⁶¹. Олима ўзбек тилидаги феъллардан

⁶⁰Гендина Н.И. Лингвистические средства автоматизации документального поиска / Н.И. Гендина; Под ред. В.П.Леонова. СПб.: Б-ка РАН, 1992.–188с.

⁶¹Шадманова К.У. Табиий тиллар учун лексик-семантик луғат маълумотлар базасини яратиш тамойиллари // <http://buxdu.uz/index.php/uz/>

иборат маълумотлар базаси компьютер лингвистик дастурлари, таржима, таҳрир каби дастурларга мўлжалланганини таъкидлайди, шунингдек, ундан фойдаланиш географиясини қуйидагича белгилайди:

- 1) ўзбек, инглиз тилидаги ҳаракат феълларининг тусланиши ўрганувчилар;
- 2) ўзбек, рус, инглиз тиллари бўйича тилшунос мутахассислар;
- 3) лингвистика соҳасидаги компьютер дастурчилари.

Лингвистик таъминотнинг ахборотни қайта ишлаш тизимларидаги аҳамияти. Табиий тил процессорини тузиш инсон фаолиятида мавжуд тил ифодаларини шарҳлаш тамойилларини мужассамлаштиришдан иборат. Табиий тил процессорини қуриш замонавий математик лингвистика ва сунъий интеллектни яратишнинг долзарб муаммоларидан. Бу жараён инфор­мацион портлаш оқибатида долзарблик касб эта бошлади: бирор далилни аниқлаш (теоремани исботлаш, тенгламани ечиш, маълум адабиётдан керакли маълумотни ажратиб олиш) ёки инсон бажариши катта меҳнат талаб қиладиган ишларни амалга ошириш эҳтиёжидан келиб чиқди. Айрим маълумотлар қоғоз нашр билан бир қаторда электрон шаклга эга бўлади. Агар улар автоматлаштирилган қидирувга эга бўлмаса, ахборотни қайта ишлаш дастури юкланмаган бўлса, керакли ахборотни ажратиб олиш имкони бўлмайди. Масалан, бугунги кунда (айтайлик, рус тили учун) маълум бир матннинг тўлиқ морфологик таҳлилини бера оладиган, луғат таркиби/лексикасини акс эттирувчи тизим (анча катта луғат таркибига эга бўлган, қисқа эмас) мавжуд эмас⁶². Аслида лемматизация ёки морфологик “тушуниш”нинг вазифаси бошқа инфор­мацион топшириқларни бажаришдан иборат.

Сўнгги 20 йил ичида инсониятнинг нутқи билан боғлиқ барча ҳодисаларни моделлаштириш ишига эътибор қаратилди. Дастлаб бундай

⁶²Шадманова К.У. Табиий тиллар учун лексик-семантик луғат маълумотлар базасини яратиш тамойиллари // <http://buxdu.uz/index.php/uz/>

тадқиқотларнинг бошида сўзловчи нуқтаи назаридан келиб чиқилган бўлса, 1970-йиллардан бугунги кунгача ҳаракат тушунча асосида моделлаштиришга қаратилган. Бундай ўзгариш “сўзловчи лингвистикасидан” “тушуниладиган лингвистика”га ўтиш тенденциясига кўчиш билан боғлиқ.

Математик лингвистиканинг сунъий интеллектдаги ўрни қандай? Р.Шенк ушбу саволга қуйидагича жавоб беради: “математик лингвистика – компьютернинг инсон билан табиий тилда мулоқот қилишини йўлга қўядиган муаммо ечими. Услубимиз, кишилар ўртасидаги мулоқот усуллари ўрганиш каби жараёнларни тадқиқ қилишдан иборат”⁶³. Н.Черконе ёзади: “математик лингвистика табиий тилни “тушунадиган” компьютер дастурлари учун асос бўладиган умумий назарияни яратиши керак”⁶⁴. Томпсоннинг фикрига кўра, математик лингвистика назарияси жонли тил табиатини характерлай олиши керак, жонли тилга “лингвистик ҳисоблаш”ни қўллаши – нутқ устида амал бажариши керак.

Бугунги кунда Америка математик лингвистикасида математик лингвистика назариясининг *функционал грамматика, когнитив грамматика, лексик-функционал грамматика* каби концепциялар таклиф қилинмоқда. Мукамаллаштирилган АҚТларнинг лингвистик таъминоти икки ўзаро чамбарчас боғлиқ қисмга – базавий ва бошқарув таъминотига бўлинади. Улар орасидаги алоқа махсус тил компилятор-процессори ҳамда интерпретатор орқали амалга оширилади. Базавий лингвистик таъминот АҚТ дастурий таъминотининг лисоний асоси бўлиб хизмат қилади, амал қилувчи дастурлаш тилини ташкил қилади. У АҚТ воситалари мажмуида дастурлашнинг ҳисоблаш ва моделлаштириш амалиёти умумлашган алгоритми, шунингдек, сервис вазифалари ечимини таъминловчи таъминот ҳисобланади. Бошқарув лингвистик таъминоти ихтисослашган лойиҳалашнинг умумлаштирилган

⁶³Место и назначение лингвистического обеспечения в информационных системах. Понятие информационной системы // <https://gigabaza.ru/doc/116551-pall.html>

⁶⁴Место и назначение лингвистического обеспечения в информационных системах. Понятие информационной системы // <https://gigabaza.ru/doc/116551-pall.html>

алгоритмини лойиха амаллари атамалари, бажариладиган иш ва топшириқнинг муаммоли-мўлжалли тилидан ташкил топади. Ушбу тилларда лойиҳалаш доирасига аниқ мўлжалланган луғат, синтаксис, семантика шакллантирилади. “Махсус-мўлжалланган тил”ни яратиш, уни амалиётда қўллаш автоматик лойиҳалаш жараёнининг юқори натижали бўлишига замин яратади. Хусусан, лойиҳачи ва АҚТнинг техник воситалари комплекси орасида икки томонлама мулоқот қилиш имконини беради, лойиҳалашнинг табиий нутқий савол-жавоб режимига яқинлашади⁶⁵. Таъкидлаш жоизки, пассив ва диалог тиллари нафақат матнга оид ахборотни, балки график ахборотни ҳам ифодалайди, шунинг учун улар *график тил* деб ҳам юритилади. График тилларсхема, чизма, график, расм-объектни лойиҳалаш учун ишлатилади. Зеро, график тиллар конструкциясида тасвирни генерациялаш, уларнинг турли воқеланишида ўз ифодасини топади. АҚТнинг базавий лингвистик таъминоти дастурий таъминотнинг лисоний асоси бўлгани сабабли дастурлаш тили иерархиясини ўзида мужассам этади. Унинг ҳар бир босқичи лойиҳалашнинг умумлашган алгоритмида турли амалларининг ҳар хил эффектда юзага чиқишга замин ҳозирлайди. Муваффақиятли танланган тил – сифатли дастурий маҳсулот тайёрлаш гарови⁶⁶. Ахборот тизимининг бошқарувчи лингвистик таъминоти умумий алгоритмни тавсифлайди, тил луғат таркибини тўплайди, кўрсатади, тил синтаксиси ва семантикасига тегишли у ёки бу намунали амалиётни бажариш воситаси бўлиб хизмат қилади⁶⁷. Ахборот тизимининг базавий лингвистик таъминоти – дастурий таъминотнинг лисоний асоси; у дастурлаш тилнинг ҳар бир босқичи автоматик амал ошувчи умумлаштирувчи лойиҳалаш алгоритмининг иерархиясини ҳосил қилади. Мукамал лингвистик таъминот муваффақиятли дастурий маҳсулот гарови саналади. Бунда дастур ва дастурий мажмуанинг асосий

⁶⁵www.ispu.ru/library/lessons/koposov2/14.html

⁶⁶<https://works.doklad.ru/view/IP7CrszcUzY.html>

⁶⁷Место и назначение лингвистического обеспечения в информационных системах. Понятие информационной системы // <https://gigabaza.ru/doc/116551-pall.html>

хусусияти уларнинг ишончлилиги, эффектив, етарлича компакт ва тезкор машина дастурлари, модификация ва мобилликка эгаллиги билан характерланади.

Дастурлаш тили ва табиий тил. Дастурчилик ёндашуви тил сўзининг кўпмаънолилигидан келиб чиқади; *ахборот адабиёти* атамаси остида нафақат маълумотларни акс эттириш, балки дастурлаш воситалари ёки бошқа формал тизимлар ҳам тушунилади. Кейинги пайтларда маълумотни акс эттириш формал тил вазифасига айланиб қолди; дастурий воситаларнинг оддий иш куруллари (SGML ёки XML каби разметка тиллари)га формал тил деб қараш анъана тусига кирди. Бундай ёндашув лингвистик таъминот таркибига маълумотни тавсифлаш, манипуляция қилиш, у ҳақидаги тасаввур характериға эга/эга эмаслигидан қатъи назар, тилға оид фойдаланиладиган воситалар мажмуи деб қаралади. Бундан ташқари, табиий тил лексикасининг лексик-семантик бирликлари сифатида ишлатилиш ҳолати саналади: вербал тил; дескриптор (атаманинг синоними); посткоординатор тиллар деб юритилади.

Вербал тилға электрон кутубхона, умуман олганда, ахборот қидирув тизими лингвистик таъминотининг марказий компоненти деб қараш ўринли. Амалда ҳам бутун ахборот қидируви назарияси вербал информацион дастур системаси фойдаланиши асосида яратилади⁶⁸.

Лингвистик таъминот воситаларининг иккинчи гуруҳи, том маънода, тил эмас; улар, одатда, лингвистик процессорлар деб юритилади. Бунга икки тушунча: матннинг автоматик ишлови (1) ва лингвистик маълумотлар банки (2) киради. Матнни автоматик қайта ишлаш деганда матн(файл)нинг бир ёки бир неча информацион тилда автоматик тавсифланиши, шунингдек, автоматик

⁶⁸https://studopedia.ru/6_11435_ponyatie-i-sostav-lingvisticheskogo-obespecheniya.html

индекслаш, аннотациялаш тушунилади. Бу жараёнлар, хусусан, морфологик ва синтактик таҳлили асосида лингвистик алгоритмлар ётади⁶⁹.

Лингвистик маълумотлар базаси деганда электрон кутубхонанинг энг муҳим ташкил қилувчи компоненти тушунилади. Шу ўринда лингвистик таъминот компьютер лексикографияси билан тўқнашади. Демак, (электрон кутубхонага мўлжалланган) инфорацион тизимларнинг лингвистик таъминоти қуйидаги воситалардан иборат:

1) ахборот-қидирув тили: (1.1) метамаълумотлар тизими, (1.2) таснифловчи тил, (1.3) вербал тил, (1.4) фактографик тил;

2) лингвистик процессор: (2.1) матннинг автоматик қайта ишлаш тизими, (2.2) лингвистик маълумотлар банки.

Вербал тил (лот. *verbalis*) лингвистик таъминотда сўзни лексик бирлик сифатида кўрсатадиган ҳамда табиий тил ифодаларини орфографик шаклда юзага чиқарадиган тил кўриниши саналади⁷⁰.

Автоматик лойиҳалаш системаси (рус. САПР)да дастур тили ва инсоннинг компьютер билан алоқа қилиш тили сифатида қўлланади. Лингвистик таъминот деганда инсоннинг компьютер билан алоқа қилиш тили назарда тутилади. Автоматик лойиҳалаш системасининг лингвистик таъминоти ахборотни тавсифлашнинг формал тили тўлиқ мажмуи ва автоматик лойиҳалаш жараёнида уни қайта ишлаш алгоритмини ташкил этади.

Формал тил компьютер, инсон ва турли автоматик қурилмалар орасидаги ахборот алмашуви ва ахборотни тавсифлаш учун белгилар тизимини ифодалайди⁷¹. Шу нуқтаи назардан ёндашилганда лингвистик таъминот қуйидаги тиллар билан шаклланади: дастурлаш тили, бошқарув тили, лойиҳалаш тили. *Дастурлаш тили* автоматик лойиҳалаш системасининг дастурий таъминотини ишлаб чиқишда керак бўлади. Бундай тилларга,

⁶⁹https://studopedia.ru/6_11435_ponyatie-i-sostav-lingvisticheskogo-obespecheniya.html

⁷⁰http://tsput.ru/res/informat/sist_seti_fmo/lekcii/lekcii-4.html#name3

⁷¹http://tsput.ru/res/informat/sist_seti_fmo/lekcii/lekcii-4.html#name3

масалан, Фортран, Си, Паскаль, AutoLISP, FoxPro ва бошқа тиллар киради. *Бошқарув тили* компьютер ва унинг периферик қурилмаларини бошқариш учун ишлатилади. *Лойиҳалаш тили* лойиҳаловчи-фойдаланувчига ва автоматик лойиҳалаш системасини эксплуатация қилиш (бошқариш)га мўлжалланган. Бу гуруҳга кирадиган тиллар шартли равишда *кириш*, *ички* ва *чиқиш* каби гуруҳларга ажратилади. Кириш тили фойдаланувчининг автоматик лойиҳалаш системаси (масалан, муаммони шакллантириш ва бошланғич маълумотни тайёрлаш жараёни) билан ўзаро муносабати воситаси саналади. Ички тил, одатда, оддий фойдаланувчидан яширин, автоматик лойиҳалаш системасининг турли ост тизимлари орасида узатиладиган маълумотни кўрсатишга хизмат қилади. Чиқиш тили лойиҳалаш натижасини матн ёки график шаклда расмийлаштиришни таъминлайди⁷².

Исталган бир тилнинг дастлабки элементи белги ва рамз; уларнинг тартиб ва система асосидаги тўплами – алифбо. Алифбодаги белгилар ҳарф ва махсус белги (лотин ёзувида тутуқ белгиси, кирилл алифбосида юмшатиш ва айириш белгиси)ларга ажратилади. Ҳарфларнинг маълум тартиби эса тилнинг луғат бойлигини ташкил этувчи сўзни келтириб чиқаради. Бундан ташқари, ҳар бир тилда санок системаси мавжуд, улар ҳам махсус белгилар билан ифодаланади. Сўз ва сонлар гапни, гаплар матнни ташкил этади. Сўз ясаш, сўзшакллари тил морфологиясини ташкил қилади. Морфологик қоидалар сўзни асос/ўзак, қўшимча, олд қўшимча бирикиши каби қоидалари мажмуидан иборат бўлади. Барча тилларда морфологик қоидалар мажмуи мавжуд. Улар асосида миллионлаб сўзшакллар матнда юзага чиқади, санокли сўз ясаш қолиплари асосида минглаб сўз ясаш мумкин. Сўз бирикмаси, гап ва матн тузиш принциплари тил синтаксисини ташкил этади. Сўз ясалиши, морфология, синтаксис эса грамматикани ҳосил қилади. Сўз, гап ва матннинг маъно томони *тил семантикаси* деб аталади. Алифбо, морфология, синтаксис,

⁷²Берхеев М.М., Заляев И.А., Кожевников Ю.В. и др. Основы систем автоматизированного проектирования. Учеб. пособие. – Казан, 1988. – 253с.

семантика формал тилнинг тўлиқ, мукамал характеристикасини бера олади⁷³. Амал қиладиган автоматик лойиҳалаш системасида лойиҳалаш ва моделлашда турли мураккаблик ва мақсадни мўлжалловчи, ҳар хил амални бажаришга ихтисослашган формал тилдан фойдаланилади⁷⁴.

Хулоса шуки, лингвистик таъминот – автоматлашган қидирув тизимида кўлланувчи ахборотни автоматик таҳлил қилишни унумли даражага кўтариш мақсадида табиий тилнинг формал қоидаларини яратишга хизмат қилувчи тил воситалари мажмуи. У ахборот объекти тузилиши, унинг ишлаши учун фойдаланиладиган, формаллашган, табиий тил ифодаляйдиган тушунча устида бажариладиган амаллар, унинг воситалари таснифи ҳамда тўпламидан ташкил топади.

1.2. Сўз шаклларининг морфологик анализ ва синтези

Матн/ахборотнинг мукамал лингвистик таҳлилини амалга ошириш учун лексик, морфологик, синтактик, семантик каби таҳлил турлари амала оширилиши керак. Маълумки, тартиблаштирилган таҳлил матн(ахборот)ни мантиқий кетма-кетликда лингвистик анализ қилишдан иборат бўлиб, мазкур таҳлилда ҳар бир топшириқ олдингисига нисбатан қийинроқ кўринади. Лингвистик анализаторнинг вазифаси таҳлил қилинаётган (матн ифодалаётган) тилнинг хусусиятидан келиб чиқади. Шунинг учун тузиладиган ахборот-қидирув тизимлари ўз таркибига матн ахбороти лингвистик таҳлилининг тўлиқ топшириқлар рўйхатини амалга оширувчи дастурий-техник мажмуани қамраб олиши керак. Мавжуд энг яхши ахборот-қидирув тизимлари матн ахборотини морфологик таҳлил қилиш (матнни индекслаш, фойдаланувчи сўрови) вазифасини бажаради, шунингдек, гапнинг айрим синтактик таҳлилини ҳам амалга оширади⁷⁵.

⁷³<https://works.doklad.ru/view/IP7CrszcUzY.html>

⁷⁴Берхеев М.М., Заляев И.А., Кожевников Ю. В. и др. Основы систем автоматизированного проектирования. Учеб. пособие. – Казан, 1988. – 253с.

⁷⁵Место и назначение лингвистического обеспечения в информационных системах. Понятие информационной системы // <https://gigabaza.ru/doc/116551-pall.html>

Ахборот қидирув тизимларининг ишлаш алгоритми олдиндан индексланган матнларнинг калит сўзларига асосланади ва фойдаланувчи сўрови калит сўз асосида амалга ошади. Сўров тили асосидаги қидирув нисбатан аниқроқ натижаларни топа оладики, бунда сўров тили ҳар бир ахборот қидирув тизими учун ўзига хос хусусиятга эга бўлади. Шунинг учун фойдаланувчи, кўпинча, унга мурожаат этмайди; бу эса қидирув натижаси сифатини пасайтиради, чунки қидирув натижасида катта материал топилади, фойдаланувчи уларни саралаш лозим бўлади. Бу материал қилинадиган таҳлил учун ноадекват бўлиб қолади.

Морфологик анализ (Part of Speech tagging). Морфологик анализнинг вазифаси матндаги ҳар бир сўзнинг қайси туркумга мансублигини автоматик тарзда муайянлаштириш; сўзларнинг қайси лексик-грамматик синфга тегишлилигини аниқлашдан иборат. Рус тилида формал морфология ривожланганлиги сабабли буни амалга оширишнинг кенг имкониятлари мавжуд. Инглиз тилида ҳам алгоритм нисбатан оддий: лексик кўпмаънолилик мавжуд бўлса-да, морфоанализ дастурлари (теггер) матндаги сўзнинг туркумини 90 фоизда тўғри аниқлай олади. Рус тили матнларининг морфологик анализини амалга ошириш учун А.Зализняк грамматик луғатининг компьютер версиядан фойдаланилган, инглиз тилининг морфоанализи Мюллер грамматик луғати асосида амалга оширилган. Рус ва инглиз морфоанализини яратиш тажрибаси шуни кўрсатадики, морфологик анализини амалга ошириш учун формал грамматика яратилган бўлиши керак.

Кўпмаъноли сўзларнинг туркумини аниқлаш жараёнида морфологик анализ сифатини яхшилаш учун икки хилдаги алгоритмдан: статистик-эҳтимолий ва сўз/кодларни бошқарувчи қоидалардан фойдаланилади. Статистик-эҳтимолий алгоритмнинг кўпчилиги 2та ахборот манбаига асосланади:

1) маълум тилнинг ҳар бир сўз шаклида мавжуд лексик-грамматик мансубликни билдирувчи изоҳларга эга сўзшакллар луғати. У лексик-

грамматик шаклларнинг барча учровчанлик ҳолати ҳақидаги ахборот ҳам дейилади. Бу ахборот қандай акс эттирилганига қараб биграмма, триграмма ва квадриграмма моделлари фарқ қилади;

2) қоидаларга асосланган алгоритм автоматик тарзда тил корпусидан аниқланади ёки мутахассислар томонидан тайёрланган материалга асосланади.

Икки ёндашув ҳам бир хил натижа кўрсатади. Уларни алоҳида ёки турли комбинацияларда қўллаш натижасида лексик-грамматик таҳлил 96-98 фоиз сифатли бўлади. Қўлда амалга ошириладиган лексик-грамматик таҳлилда ҳам 0,5-2 фоизгача хато ҳолатлар учраши кузатилган. Айтиш мумкинки, автоматик лексик-грамматик анализ шахс томонидан мустақил бажариладиган морфологик таҳлил даражаси ва сифатига тенглашади⁷⁶.

Матн ахборотини синтактик таҳлил қилиш гапнинг семантик элементларини – гуруҳга мансублик, предикатив асосни ажратишга имкон яратади. Бу матнда унга хос бўлган ахборотга ишлов бериш жараёнида умумлаштирилган семантик элементлар билан ишлаш интеллектуаллигини оширади. Матн ахборотини қайта ишлаш тизимлари, албатта, эксперт тизимлардан, семантик ахборотни ажратиш учун сунъий интеллект тизимларидан фойдаланишни талаб қилади. Бугунги кунда мукамал семантик таҳлил тизимининг мавжуд эмаслиги, шундай тизимни яратишдаги муаммолар сунъий интеллект тизимларини тузиш илмий йўналишининг тўлиқ йўлга қўйилмагани билан боғлиқ.

Лексик таҳлил деганда матн ахборотини миллий тил ифодаларини кўрсатувчи алоҳида сўз боши, гап, сўз шаклида таҳлиллаш назарда тутилади ва у ҳар бир тилнинг ўзига хослигини кўрсатувчи гап шакллари, лексик ифода(ўзлашма, жаргон сўз ва ҳ.)ларни аниқлаб беради. Амалиётда уни амалга ошириш қийинчилик туғдирмайди.

⁷⁶Место и назначение лингвистического обеспечения в информационных системах. Понятие информационной системы. Лекция 4 // <https://gigabaza.ru/doc/116551-pall.html>

1.2.1. Сунъий тилни яратиш: тил синтаксиси ва семантикаси. Ҳар бир тил минглаб махсус белгининг маълум тартибдаги жойлашувидан ҳосил бўлади. Шунингдек, алифбодан ташқари боғланишлар қондаси ҳам мавжуд, чунки маълум алифбодаги символлар занжирининг барчаси ҳам тилга тегишли эмас. Символлар тилнинг элементар конструкцияси – сўз ёки лексемада бирлашиши мумкин. Улар асосида эса нисбатан мураккаброқ конструкция – гаплар ҳосил қилинади. Ҳар икки ҳолат ҳам белгилар занжири саналади ҳамда ясаиш қоидаларига амал қилади. Шундай қилиб, бу қоидаларни кўрсатиш, аниқроқ айтганда, тилни белгилаш зарур. Умумий ҳолатда тилни уч хил усул билан аниқлаш мумкин⁷⁷:

- 1) тилда мавжуд барча боғланиш / занжирларни санаш;
- 2) занжир / валентликни давом эттириш усулларини кўрсатиш; тилдаги боғланишнинг ҳосил бўлиш қонуниятларини аниқлаш (грамматика тузиш билан);
- 3) занжирларни аниқлаш усулларини топиш.

Ушбу мантиқий қурилма(аниқлагич – распознаватель (русча)нинг ўзига ҳослиги шундаки, у саволни киришда белгилар занжири сифатида ўқийди, чиқишда эса жавоб беради: маълум занжир шу тилга мансубми ёки йўқлигини аниқлаб беради. Масалан, матнни ўқир эканмиз, ўзимиз ҳам морфологик анализатор (распознаватель) вазифасини бажарамиз: ўқилган матн ўзбек тилида эканлигини тасдиқлаймиз.

Исталган бир тил ҳақида фикр юритилар экан, албатта, унинг синтаксиси ва семантикаси ажратилади. Анализатор/транслятор тилнинг лексик конструкцияси (лексема) билан ҳам иш кўради. Қуйида шундай морфологик анализаторни ташкил қилувчи воситаларга изох берамиз.

Тил синтаксиси – тилда мавжудлик эҳтимоли бўлган қурилма қоидалари тўплами. Синтаксис “тил шакли”ни белгилайди, тилга мансуб бўлган белгилар

⁷⁷Место и назначение лингвистического обеспечения в информационных системах. Понятие информационной системы. Лекция 4 // <https://gigabaza.ru/doc/116551-pall.html>

занжири йиғиндисини кўрсатади. Одатда, формал тил учун қатъий синтактик қоидалар ишлаб чиқилади. Кўпинча, дастурлаш тиллари формал тил қоидаларига қўшимча киритишни, тўлдирилишни тақозо қилади. Табиий тил (нутқий мулоқот) синтаксиси эса “истиснолар қоидани тасдиқлайди” тамойилида ишлайди⁷⁸.

Маълумки, анъанавий тилшуносликда лексика тилнинг луғат таркиби мажмуи ҳисобланади. Тилнинг сўз ёки лексик бирлиги (лексема) – алифбо унсурларидан иборат тузилма; улар ўзида бошқа қурилмани сақламайди. Бошқача айтганда, лексик birlik фақат элементар белгилардан иборат, бошқа лексик birlik унга мансуб эмас. Масалан, ўзбек тили лексикасини ўзбек тилидаги сўзлар ташкил этади, тиниш белги ва ажратма (пробел)лар эса уларни ажратувчи белги саналади. Алгебранинг лексик birlikлари сон, функция ва ўлчов birlikини билдирувчи математик амал белгилари бўлгани каби дастурлаш тилида калит сўз, идентификатор, константа, математик амаллар белгиси лексик birlik сифатида қаралади. Бу лексик birlik таркибига бўлувчилар (вергул, қавс, нуқта, нуқтали вергул ва ҳ.) ҳам киради.

Грамматика ва анализатор. Грамматика – маълум тилнинг гап қурилиш усуллари тавсифи; тилни белгиловчи математик система. Грамматикани тавсифлар эканмиз, биз шу тилга мансуб белгилар занжирининг пайдо бўлиш қоидаларини кўрсатамиз. Шу сабабли грамматикани тил занжири генератори дейиш мумкин. У тилни ажратиш, танишнинг иккинчи типи – белгилар занжирининг пайдо бўлишини кўрсатади. Тил грамматикасини турли усуллар билан тавсифлаш мумкин. Масалан, рус тили грамматикасини мактабда ўқитиладиган анча қийин қоидалар мажмуи билан ифодалаш мумкин. Айрим тиллар учун (масалан, дастурлаш тили синтактик конструкцияси) қоидалар системасига асосланилган формал грамматика тавсифи асосида яратиш мумкин.

⁷⁸Место и назначение лингвистического обеспечения в информационных системах. Понятие информационной системы. Лекция 4 // <https://gigabaza.ru/doc/116551-pall.html>

Табиий тилни “танийдиган” таҳлил дастурининг умумий схемаси.

Таҳлил дастури маълум бир белгилар занжирининг қайси тилга тегишли эканлигини аниқлайдиган махсус автоматлаштирилган дастур ҳисобланади. Таҳлил дастурининг асосий вазифаси маълум бир белгилар занжирининг танланган тилга мансуб/мансуб эмаслигини аниқлашдан иборат.

Таҳлил дастури (распознаватель) турлари. Бундай дастурни уни ташкил қилувчи санаш қурилмаси, бошқариш қурилмаси ҳамда ташқи хотира каби компонентларга қараб тавсифлаш мумкин. Санаш қурилмаси турларига қараб бир томонлама ва икки томонлама таҳлил дастури мавжуд. Бу дастур белгилар занжирини аниқлаш учун фақат бир томондан (чапдан ўнгга қараб) ўқийди ва ортга қайтмайди. Икки томонлама таҳлил дастури белгилар занжирини ўнгдан чапга ва чапдан ўнгга қараб “ўқийди”⁷⁹.

Бошқариш қурилмасига кўра аниқлаштирувчи ва аниқлаштирмайдиган (*детерминированные* ва *недетерминированные* сўзларининг таржимаси сифатида олиняпти) дастурлар ажратилади. Дастурнинг ташқи хотиранинг мавжудлигига кўра қуйидаги типлари ажратилади:

- 1) ташқи хотирага эга анализатор;
- 2) ташқи хотираси чегараланаган анализатор;
- 3) чегараланмаган ташқи хотирага эга анализатор.

Ташқи хотирасиз анализаторда хотира умуман бўлмайди. Унинг иш жараёнида фақат бошқарув қурилмасининг хотираси ишлайди. Чегараланган ташқи хотирали анализаторда эса хотиранинг катталиги белгилар занжири узунлиги билан боғлиқ бўлади. Бундай дастурнинг ташқи хотираси рўйхат, модел намунаси асосида ишлайди. Учинчи турдаги дастурларда эса ташқи хотира чегараланмаган, исталган узунликдаги белгилар занжирини “ўқий олади”. Бу тасниф асосида ўзида таснифнинг уч белгисини жамлаган дастур ҳосил бўлади. Масалан, *икки томонлама аниқлаштиргичга эга чегараланган ташқи хотирали таҳлил дастури.*

⁷⁹Ўша манба, 11-маъруза.

Таҳлил дастури қанчалик мураккаб бўлса, унинг ишини амалга оширадиган алгоритм ҳам шунча мукамал тузилишга эга бўлади. Бир томонлама анализаторга нисбатан икки томонлама анализаторни ишлаб чиқиш қийинроқ.

Таҳлил мақсади. Ҳар бир дастурлаш тили учун нафақат шу тилдаги дастур матнини тузиш, балки мавжуд матннинг ушбу тилга мансублигини аниқлаш ҳам зарурий вазифа. Турли вазифалар қаторида компилятор шу функцияни ҳам бажаради. Компилятор тилни аниқлаш, таниш вазифасини бажарса, маълум дастур тилини яратган инсон шу тил белгилар занжири генератори вазифасини бажаради. Грамматикава анализатор (распознаватель) қайсидир тилни “таниш”нинг икки мустақил усули саналади.

Умуман олганда, таҳлил дастурининг вазифаси маълум тилнинг мавжуд грамматикасига асосланиб ушбу тил учун таҳлил дастури яратишдан иборат бўлади. Бу амалнинг иккаласи ҳам муқобил бўлиши лозим: иккаласи ҳам битта тилни “таний олиши”, ажрата олиши муҳим. Компилятор фақат белгилар занжирининг маълум тилга мансублигини аниқлаб қолмай, унинг маъно юкини ҳам ойдинлаштириб бериши лозим. Агар сўровга берилган сўз ушбу компилятор белгилар занжиридан бирига мос келмаса, у фойдаланувчига манфий, яъни *бундай сўз ушбу тилда мавжуд эмас* шаклида жавоб беради. Хўш, дастурнинг бундай хулосага келишига нима сабаб бўлади? Албатта, мазкур муаммонинг юзага келишини лингвистик таъминотдан излаш зарур. Демак, маълумотлар базасида шундай белгилар занжири мавжуд эмас. Тил қанча мураккаб бўлса, компиляторни ишлаб чиқиш шунча қийин бўлади. Айрим тиллар учун тилдаги матнларни математик ресурслар асосида таҳлил қилувчи компилятор ишлаб чиқишнинг умуман иложи йўқ.

Ахборот-қидирув тизимидаги айрим ёндашувлар хусусида. В.В.Пономаревнинг ахборот тизимларини фаоллаштириш муаммолари, унинг социолингвистик спецификаси, лингвистик таъминотни ишлаб чиқиш

жараёни таҳлилига бағишланган “Лингвистическое обеспечение и социолингвистическая специфика проблемы автоиндексационной актуализации информационных систем”⁸⁰ номли тадқиқотида тавсифий-таҳлилий, дистрибутив-контекстологик, қиёсий таҳлил, компонент таҳлил, контент таҳлил, мантикий таҳлил, алгоритмик моделлаштириш усулларидан фойдаланиш тамойиллари тавсифланган. Тадқиқотда ахборот тизимининг қуйидаги системалари фарқланади: маълумотлар базаси (МБ), информацион ресурс (ИР), автоматлаштирилган информацион тизим (АИТ), ахборот-қидирув тизими (АҚТ), эксперт тизими (ЭТ), автоматик лойиҳалаш тизимлари (АЛТ), автоматлаштирилган илмий-техник ахборотлар тизими (АИТАТ), информацион-қидирув тезауруси (ИҚТ) ва интернет.

Т.Б.Болтаев ва С.И.Ибрагимовнинг “О проекте программной системы морфологического анализа узбекского языка” номли мақоласи тадқиқ объекти ўзбек тили, хусусан, от ва сифат сўз туркуми морфологик таҳлилини автоматлаштириш масаласига қаратилган. Туркий тилларда сўз ясалиши, хусусан, от ясалиши жуда сермахсул бўлганлиги сабабли бу туркум морфологик анализи таҳлил усулларининг бир неча турини билишни талаб этади⁸¹. Мутахассис бу муаммо ечимларидан бири сифатида формал тиллар таҳлил тизимини лойиҳалаш жараёнида математик аппаратни қўллашни таклиф қилади. Бундай аппаратлар шундай тиллар, хусусан, дастурлаш тили ёки тилга ишлов берувчи дастурлар (компилятор, анализатор, спецификатор ва ҳ.)⁸², юқори даражадаги дастурлаш тилларида кўп қўлланилади. Ўтган

⁸⁰Пономарев В.В. Лингвистическое обеспечение и социолингвистическая специфика проблемы автоиндексационной актуализации информационных систем: автореф. диссер. канд. филол. наук. – Москва, 2005.

⁸¹Болтаев Т.Б., Ибрагимов С.И. О проекте программной системы морфологического анализа узбекского языка // <http://ziyonet.uz/uploads/books/473012/5afbccb8e9f55.pdf> (муружаат санаси: 24.06.21)

⁸²Ахо Альфред В., Лам Моника С, Сети Рави, Ульман Джеффри Д. Компиляторы: принципы, технологии и инструментарий, 2-е изд. : Пер. с англ. – М. : ООО "И.Д. Вильямс", 2008. – 1184 с.; Касьянов В.Н., Поттосин И.В. Методы построения трансляторов. – Новосибирск: Наука, 1986. – С. 344.; Болтаев Т.Б., Кузьминов Т.В., Поттосин И.В. О структурном конструирование программ и инструментах его поддержки // //Среда программирования: методы и инструменты. – Новосибирск, 1992. – С.22-37.; The Structured Constructing as a Discipline of Safe Programming and Instruments Supporting It /Aniskov M.I., Boltaev T.B., Kochetov D.V. at al//Instrumental Congress on Computer Systems and Applied Mathematics CSAM'93. St-Petersburg.July 19-23.

асрнинг 60-йилларига келиб тилга ишлов берувчи дастурий воситалардан фойдаланиш анъанага кирди. ЭТАП⁸³ тизими ҳақидаги ишларда формал тилларнинг барча таҳлил воситалари флектив тиллар (мас., рус тили) таҳлили учун қўлланган. Шунингдек, Д.С.Юравский ва Ж.Г.Мартин формал инглиз тили анализ дастури масалаларини таҳлил қилган⁸⁴.

Ушбу бўлимда ўзбек тили морфологик анализида формал тиллар таҳлилининг анъанавий воситасини қўллаш масаласи кўрилади. У аналитик тиллар сўз туркумларининг автоматик (компьютер) таҳлил жараёнини моделлаштириш борасидаги тадқиқот натижаси саналади. Иш ўз олдига ҳажман катта – ўзбек тилидаги матнлар таҳлилида формал тиллар назарияси методларини қўллаш вазифасини қўйган. Мазкур методни қўллаш бўйича тажриба ҳам мавжуд: инглиз⁸⁵ ва рус⁸⁶ тилида бу усул самарали натижа берган. Мақоланинг кириш қисмида матн берилса, чиқишида шу матн лексемалар занжири, лексемаларнинг тўлиқ морфологик характеристикаси билан чиқишини таъминловчи механизмни яратиш жараёни тавсифланади. Бу ерда формал тилларни қайта ишлашда морфологик таҳлил лексик таҳлил билан қиёсланган. Ўзбек тили сўз ясалишининг кўп схемаси ясовчи қўшимча ёрдамида ясалиш лексик таҳлил аппаратини қўллаш имконини берадики, бунда сўз ясалиш схемаси формаллаштиришда доимий ифодалар аппаратидан ҳам фойдаланиш мумкин.

Сўз ясалишининг нисбатан мураккаб схемаси: қўшма сўз, тублашиш, сўз бирикмасининг лексемалашуви автоматик ҳамда одатий ифодалар таҳлили доирасидан чиқади; улар эркин-матний грамматика асосида таҳлил қилинади. Компьютер морфологик таҳлилинини (morphological parser– MP) ўрганар

⁸³АпресянЮ.Д., БогуславскийИ.М., ИомдинЛ.Л.идр. Лингвистическое обеспечение системы ЭТАП-2., –М.: Наука 1989. – С. 296.

⁸⁴Speech and Language Processing. Daniel S. Jurafsky and James H. Martin. Contributing writers: Andrew Kehler, Keith Vander Linden, Nigel Ward 2000y. Prentice Hall, Englewood Cliffs, New Jersey 07632. pages: 950

⁸⁵The Lexical Semantics of a Machine Translation Interlingua. RickMorneau // http://www.eskimo.com/~ram/lexical_semantics.html 2006.

⁸⁶Ножов И.М. Морфологическая и синтаксическая обработка текста (модели и программы): диссертация канд. филол. наук. – Москва, 2003. // <https://docplayer.ru/26110069-I-m-nozhov-morfologicheskaya-i-sintaksicheskaya-obrabotka-teksta-modeli-i-programmy-1.html>

эканмиз, нафақат керакли математик моделлар, балки бу жараёнда кўлланиладиган информатсион структураларни ҳам кўриб чиқамиз.

МРни қуйидагича тавсифлаш мумкин: МР кириш жараёнида сўзшакл (wordform) берилади; чиқиш жараёнида сўзшаклнинг барча морфологик характеристикаси (morphological features) билан сўзнинг ўзаги (stem) олинади. Ушбу маълумот таҳлилнинг кейинги босқичи (синтактик, семантик, прагматик анализ, сўз ясаиши, тарижима) учун асос бўлади. Сўзшаклнинг морфологик тавсифида унинг қайси морфологик категорияга мансублигини кўрсатувчи айрим маъноси кўрсатилади, сўзшаклнинг қайси сўз туркумига мансублигини аниқлайди. Масалан, от туркумига (Noun-N) бирликни кўрсатувчи (Singular-SG), кўпликни билдирувчи (Plural-PL) сон категорияси шакли, от сўзшаклнинг келишик категорияси (Case) ва ҳ.лар бириктирилади. Феъл учун (Verb-V) замон (Tense-TS), шахс (Person-PR) категориялари кўрсаткичи бўлади.

Морфологик анализаторни лойиҳалашда қуйидаги воситалардан фойдаланилади:

1. Лексикон (lexicon): морфологик анализаторни шакллантириш учун керак бўладиган асос ва қўшимчанинг асос, қўшимча, информатсион структура ҳақидаги маълумот. Лексикон МРнинг сўзлари репозитори саналади. Лексиконга асос, қўшимча ва морфотактлар киритилади, аммо лексиконга ҳамма бирликларни ҳам киритишнинг имкони йўқ. Бу усул янги сўз ясаишини кузатиш учун жуда қулай.

2. Морфотактиклар (morphotactics) – сўзшаклнинг қайси морфем категорияга тегишлилигини билдирувчи информатсион модел тартиб-қоидалари. Улар маълум сўзшаклга хос бўлган морфемалар орасидаги муносабатни тартибга солади. Масалан, от туркумида кўплик шакли доим асосдан кейин жойлашади, бошқа шакллар қатъий тартибда ундан кейин қўшилади.

3. Орфографик қоида (orphographic rules): икки категорияга тегишли морфемада содир бўладиган ўзгаришларни тавсифловчи қоидалар мажмуи. Масалан, “к” ёки “қ” билан тугайдиган отга жўналиш келишиги CS(dative) кўшимчаси қўшилганда бу ўзакларда ўзгариш бўлади: элакка=элак+CS(dative).

4. Асос семантикасини шакллантирувчи бирликларнинг ахборот структураси (Stem Semantics Information Structure-SSUS). Бундай бирликлар, одатда, семантик атрибут деб аталади. Бу атрибутлар асосида таҳлил қилинаётган сўзнинг семантик жиҳатдан тўғрилигини текширувчи семантик-морфологик (Semantic-morphologic) қоидалар шаклланади. Масалан, (Proper(N)&Singular(N)) | (Abstract(N)&Uncount(N)) ʘ SG(N) теглар мажмуи бирликдаги, кўпликда қўлланмайдиган отни (Қуёш, Ер, Ой, Тошкент, Навоий) билдиради. Бундан ташқари, SSUS синтактик ва семантик таҳлилда морфологик белгиларни ажратишда ҳам қўлланилади.

5. Яқуний автомат аппарати (Finite State Automata- FSA)⁸⁷ морфотактикларни моделлаштиришда ишлатилади. Орфографик қоидаларни моделлаштиришда эса FSAнинг бошқа махсус варианты – трансдюсерлар⁸⁸ қўлланилади.

Хулоса сифатида айтиш мумкинки, жаҳон тилшунослигида грамматика ва лексикографияга бағишланган тадқиқотларнинг барчаси иллюстратив тил корпусидаги матнга таянади. Табиий тилдаги матнни қайта ишлашга мўлжалланган замонавий интеллектуал дастурий тизимлар ривожланиши катта экспериментал лингвистик базага эҳтиёж сезади. Тил корпуси фойдаланувчи томонидан такрор-такрор ишлатишга мўлжаллаб тузилади, шунинг учун унда қўлланадиган теглар тизими ва лингвистик таъминот бирхилликка келтирилиши лозим. Корпус разметкаси стандартга асосланади,

⁸⁷ Ахо А., Ульман Дж. Теория синтаксического анализа, перевода и компиляции. В двух томах. – М.: Мир, 1978, пер. сангл.

⁸⁸ Mohri M.A. Finite-state transducers in language and speech processing. Computational Linguistics, 23(3), 269-312.

бу разметкалар мажмуи “кодлаш стандарти” деб аталади. Ишнинг кейинги қисмларида мана шу стандартлар ҳақида фикр юритилади.

1.2.3. Лингвистик ва дастурий таъминотда халқаро стандартлардан фойдаланишнинг аҳамияти. Компьютер лингвистикаси ахборот-қидирув тизимларида материални қайта ишлаш масаласи долзарблигича қолмоқда. Қидирув тезлиги, натижа аниқлигига эришиш ахборот-қидирув тизимидаги табиий тил материалига берилган ишлов даражасига боғлиқ бўлади. Масалан, тил корпусларида корпус материалининг лингвистик ва экстралингвистик тегланиши фойдаланувчи томонидан исталган ахборотни ажратиб олиш имкониятини пайдо қилади. В.П.Захаров тил корпусининг теглашнинг халқаро стандартлари, ушбу стандартни тузиш, ундан фойдаланиш методологияси борасида фикр юритади⁸⁹. Жаҳон корпус лингвистикаси тажрибаси асосида рус ва туркий тиллар корпуси учун метамәълумотларни акс эттирувчи стандарт ишлаб чиқилган. Корпус ва матнлар доирасида амал қилувчи Text Encoding Initiative (TEI), ISLEProject (International Standards for Language Engineering) лойиҳаси ҳамда EAGLES (Expert Advisory Group on Language Engineering Standards) тавсияларига асосланувчи лингвистик, экстралингвистик теглашни амалга оширувчи анализаторлар шулар сирасига киради. Шунингдек, бу сирага CDIF (Corpus Document Interchange Format)⁹⁰, CES (Corpus Encoding Standard)⁹¹, XCES (Corpus Encoding Standard for XML)⁹²ларни киритиш ҳам ўринли. “Лингвистик ресурсларни бошқариш” умумий номи остидаги ушбу стандартлар қуйидагиларни тавсифлайди:

- 1) терминологияни стандартлаштириш метод ва тамойиллари;
- 2) терминологик стандартларни ишлаб чиқиш;
- 3) терминологик луғатлар ва тил ресурсларини яратиш;
- 4) компьютер лексикографияси;

⁸⁹Захаров В.П. Международные стандарты в области корпусной лингвистики //Структурная и прикладная лингвистика, 2012. – № 9. – С. 201-221. – ISSN 0202-2400.

⁹⁰www.natcorp.ox.ac.uk/archive/vault/tgcw30.pdf

⁹¹<http://www.cs.vassar.edu/CES/CES1.html#Contents>

⁹²<http://www.xces.org/>

5) терминология соҳаси ва лингвистик ресурсларни кодлаштириш.

TEI стандартлари. Нисбатан кенг ишланган воситалардан бири Text Encoding Initiative (TEI) лойиҳаси. TEI системаси маттни идентификациялаш, акс эттириш, анализ ва интерпритация қилиш, метатил тавсифи ва уни кодлаштиришга мўлжалланган. У, асосан, матнли ҳужжатларга мўлжалланган, шунингдек, график ва овозли материалларни идентификациялаш имконияти ҳам мавжуд⁹³. Ушбу лойиҳа Association for Computers and the Humanities (Ассоциация по компьютерам и гуманитарным наукам), Association for Computational Linguistics (Ассоциация по вычислительной лингвистике), Association for Literary and Linguistic Computing (Ассоциация по компьютерным технологиям в литературе и лингвистике) халқаро ташкилотлар томонидан қўллаб қувватланади.

Тавсифлаш амалиёти *теглаш* ёки *кодлаштириш* деб аталади. Кодлаш схемасини аниқлаш учун SGML ёки XML тиллари ишлатилади; бу тиллар орқали кодлаш схемасидаги атама ва элементларни шакл жиҳатдан ажрата олади.

Морфологик разметка форматлари. Тузилиши жиҳатдан матнни лингвистик ахборот билан таъминлашнинг уч тури фарқланади⁹⁴:

- 1) оддий белги қўшиш: ҳар бир сўздан кейин унга қисқа изоҳ қўшиш. Масалан, *gives_VVZ*. VVZ коди 3-шахс бирликни билдиради, яъни (Z) белгиси феълнинг (VV) 3-шахсда эканлигини ифодалайди (1-чизма).
- 2) жадвал. Ҳар бир устунда маълум морфосинтактик белги ёзилади (2-чизма).
- 3) теглаш тили. Лингвистик ахборот мажмуи ёзиш воситалари (3- чизма).

Матн разметкасининг семантик тиллари сирасига TeX, XML, SGML, XATML, HTML+CSS тилларини киритиш мумкин. Интернетда барча Web-саҳифаларнинг муштарак томони мавжуд: уларнинг барчаси HTML (Hyper

⁹³Захаров В.П. Международные стандарты в области корпусной лингвистики // Структурная и прикладная лингвистика, 2012. – № 9. – С. 201-221. – ISSN 0202-2400.

⁹⁴Кўрсатилган манба. – Б. 201-221.

Text Markup Language – гиперматнни разметкаш тили) тилида яратилади. HTML – ноодатий дастурлаш тили; Web-саҳифа яратиш дастурлашга якин жараён бўлса-да, улар алоҳида-алоҳида ҳодисалар ҳисобланади. HTML оддий матнни Web-саҳифа кўринишига келтирадиган қоидаларни белгилайди. HTML тили World Wide Wet (WWW) хизмати билан бир вақтда пайдо бўлиб, баравар ривожланиб борапти, шунингдек, яна ҳам оригинал Web-саҳифалар яратадиган, ўзига хос янги белгилар билан бойиб бормоқда⁹⁵. HTML компьютер тили; у баъзан дастурлаш тилларига қариндош сифатида қаралади. Аниқ натижаларга эришиш учун ундаги қатъий қоидаларга амал қилиш лозим. HTML тили кодлари *тег* деб аталади. Тег – учбурчак қавсга олинган махсус белгиларнинг тартибли жойлашуви: (<...>). Web-маҳсулот ўзида теглар папкасини сақлайди; шу усулда Web-саҳифа оддий матнни ўзида сақлайди; одатий матн муҳаррири билан тузилиши ва таҳрирланиши ҳам мумкин⁹⁶.

Web-ҳужжат яратишнинг ўзига хослиги шундаки, муаллифга маҳсулотнинг қанақа моделдаги компьютерда, қайси операцион система ва қайси дастур орқали акс этиши номаълум бўлади. Бу “номаълумлик” Web-ҳужжат форматланишига қатъий талаб қўйишдан ушлаб туради. Web-ҳужжат форматланиши оддий матн муҳарриридаги “документ” (компьютердаги ҳужжат, файл, дастурий маҳсулот) форматланишидан кўп функциялилиги билан фарқ қилади. Масалан, HTML кодлари шрифт, абзац, сатрнинг аниқ параметрларини белгиламайди (гарча айрим ҳолларда буни қилиш мумкин). Бунинг ўрнига сарлавҳа ва абзацнинг вазифаси белгиланади; фойдаланувчининг компьютерида Web-маҳсулотни акс эттирувчи дастур эса “ўзи истаган” форматда кўрсатади. Бу – HTML тилининг энг афзал томони. HTML кодларидан фойдаланган ҳолда қуйидаги амалларни бажариш мумкин⁹⁷:

- 1) турли сарлавҳаларни тўғрилаш, уларни экранда бошқариш;

⁹⁵Кўрсатилган манба. – Б. 201-221.

⁹⁶Кўрсатилган манба. – Б. 201-221.

⁹⁷Кўрсатилган манба. – Б. 201-221.

- 2) матнларни абзацларга ажратиш; ҳужжатнинг абзац ва бошқа қасмларини ажратишда безашнинг оддий элементларидан фойдаланиш;
- 3) Web-ҳужжатни бошқа гиперматнга ёки матннинг бошқа қисмига ҳаволалар бириктириш;
- 4) матнга расм бириктириш; матннинг расм билан биргаликда кўринишини таъминлаш; расмларнинг ҳажмини бошқариш; расмни кўрсатиш имкони бўлмаганда браузер билан ишга тушадиган альтернатив матн киритиш;
- 5) шрифтнинг ҳажми, гарнитураси ва рангини бошқариш;
- 6) дастлабки матнда асосий маъноли қисмларни ажратиш;
- 7) турли рўйхатлар ҳосил қилиш;
- 8) жадвал тузиш;
- 9) бир неча ҳужжат (маҳсулот)ни битта Web-саҳифага жойлаштириш;
- 10) матн кўринишидаги маҳсулотга мультимедиали объектларни жойлаштириш.

Хулоса шуки, ўзбек тили корпуслари ва морфологик анализаторларини тузишда разметкаланининг халқаро стандартларидан фойдаланиш мумкин. Бунда, албатта, ўзбек тилидаги матннинг ўзига хос хусусиятлари ҳам инобатга олинади.

1.3. Лингвистик таъминотнинг тузилиши, таркиби ва вазифалари

Ахборот-қидирув тилининг лингвистик таъминоти табиий тил (энг кўп тарқалган инглиз тили)га, умумқабул қилинган рамз ва график образларга асосланади⁹⁸. Лингвистик таъминот таркибига инфорацион базанинг структур бирликларини тавсифлаш учун ахборот тиллари, бошқарув ва ахборот базаси маълумотлари манипуляцияси тили, ахборот-қидирув тизими тил бирликлари, автоматик лойиҳалаш тизими тили, махсус мулоқот тили, атама ҳамда аниқловчининг бошқа тизимлари, автоматлаштирилган бошқарув

⁹⁸Митрофанов В.Г., Калачев О.Н., Схирладзе А.Г. и др. САПР в технологии машиностроения. Учеб. пособие. – Ярославль: Ярослав. гос. техн. ун-т, 1995. – 298 с.

тизими функциясини ишга тушириш, уларни ишлаб чиқиш жараёнида керак бўлувчи бошқа тиллар ҳам киритилади. “Тил воситалари”га дастурлаш тили, сўров тили ва маълумотни тавсифлаш тили киради. База устида амал бажариш учун фойдаланувчига элементар тил воситасига эгалик қилиш, тилни тушуниш, унинг барча имкониятларидан фойдаланишнинг ўзи етарли. Тил синтаксиси оддий, шу билан бирга кенг спектрдаги амалларни камраб олади⁹⁹.

Ахборот тили формал тил вазифасига кўра индексация тили; ахборот-кидирув тили; ахборот воситачи тили каби гуруҳларга бўлинади¹⁰⁰.

Яхши тузилган автоматлаштирилган ахборот тизимининг лингвистик таъминоти (русча *лингвистическое обеспечение АИС, ЛО, англ. лингвистическое обеспечение АИС, ЛО, англ. linguistic support*) таркибига қуйидаги тил воситалари киради¹⁰¹:

1) алифбо ва микросинтаксис (*alphabet and microsyntax*) – маълумотни ифодалашнинг график воситалари;

2) библиографик маълумотлар тили (*bibliographical data language*) – библиографик ахборотни етказиш воситаси;

3) таснифланган тиллар (*classification languages*) – ахборот етказишнинг мавзулаштирилган кўриниши;

4) дескриптор (*descriptor languages*) ва бошқа посткоординат тиллар (*postcoordinate languages*) – ахборот етказишнинг табақалаштирилган кўринишдаги воситаси;

5) объектив-белги тиллари (*object-sign languages*)– ахборот етказишнинг факторгафик кўриниши;

6) сўров ва маълумотни манипуляциялаш тили (*inquiry and data manipulations languages*) – дастурий амалларни бажариш воситаси.

Лингвистик таъминотга асос бўлувчи воситалар сирасига қуйидагиларни киритиш мумкин¹⁰²:

⁹⁹<https://economy.ru/info/info/152021/>

¹⁰⁰<http://www.libweb.ru/archive/resource/docs/rckk/doc26.html>

¹⁰¹<http://www.gpntb.ru/win/book/2/Doc8.HTML>

¹⁰²<http://www.gpntb.ru/win/book/2/Doc8.HTML>

1. Лингвистик процессор (*linguistic processors*)- матн ахборотини автоматик қайта ишловчи дастур ва алгоритм, шунингдек, тил воситаларини яратиш ва қўллаш жараёни.

2. Лингвистик маълумотлар базаси (*linguistic data bank*) электрон луғатлар учун маълумотлар базаси, муҳим ёзув ҳамда уларни бошқарувчи дастурий бошқарув воситаларини сақлайди. Яратиш жараёни, тил воситаларини киритиш ва қўллаш, шунингдек, таъминот воситаларини, норматив, конструкторлик ва технологик ҳужжатни мувофиқлаштириш вазифасини бажаради. Тизимнинг ташкилий, безаш воситалари сифатида улардан фойдаланишни таъминлайди.

Ҳар қандай тил воситасининг лингвистик таъминот тизими, алифбодан ташқари, қуйидагиларни қамраб олади¹⁰³:

1) луғат, лексика (*lexicon*) ва парадигматика (*paradigmatic, context-free relations*);

2) синтаксис, хусусан, форматлар мажмуи кўринишида.

Автоном ишловчи тизимлардан фарқли равишда тармоқ лингвистик таъминотининг ўзига хослиги уларнинг қўлланиш соҳаси. Шу жиҳатдан унинг қуйидаги турлари мавжуд:

1. Умумтизим тил воситалари (*general language resources*) – тизимнинг барча звенолари томонидан ёки улар орасидаги алоқани ушлаб турувчи тил.

2. Локал тил воситалари (*local language resources*) – фақат алоҳида олинган звенода қўлланувчи тил.

3. Ташқи тил воситалари (*external language resources*) – тизимнинг ички муаммоларини ечишда ишлатилмайдиган, тизимнинг фақат ташқи аъзолари билан муносабатга киришувчи тил.

4. Маълумотларни етказишнинг график воситалари – табиий тил воситалари (ёзма матннинг ифода плани) бўлиб, алифбо ва микросинтаксисдан иборат бўлади:

¹⁰³ <http://www.gpntb.ru/win/book/2/Doc8.HTML>

а) алифбо – тартиблаштирилган ҳарф, рақам (араб, рим), махсус белги (тиниш белги, математик символ) ҳамда матний ўзига хосликни ифодаловчи воситалар тизими. Автоматик ахборот тизимида қўлланувчи алифбо таркибига иккилик коди жадвалларида ифодаланадиган маълум символ (белги)лар тўплами (масалан, ASCII, KOI8-R, UNICODE, Win1251, ISO 8859-5 ва б.) киради;

б) микросинтаксис – ўзида турли жадвал ҳамда тегишли белгиларни сақловчи кирилл, лотин, грек ва бошқа алифболарнинг кенгайтирилган воситаси. Табиий тилдаги матнларда кирилл (қисқартирилган ёки тўлиқ, ўзида диакрит символларни сақловчи); лотин (қисқартирилган ёки тўлиқ, ўзида диакрит символларни сақловчи); грек; готик алифболари ишлатилади.

Демак, лингвистик таъминот таркибида таҳлил дастури яратиш мўлжалланган тил алифбоси ҳам муҳим аҳамият касб этади.

Ўзбек тилшунослигида лингвистик таъминот борасидаги ишлар. Компьютер лингвистикасида маълумотларни ахборот-қидирув тизимида қайта ишлаш масаласи асосий ўринни эгаллайди. Бу машина таржимаси, морфологик анализатор, нутқ синтезатори, тил корпуси каби барча йўналишлар учун бир хилда аҳамиятлики, ушбу йўналишлар барчаси лингвистик ва дастурий таъминотга эҳтиёж сезади.

Ўзбек компьютер лингвистикаси мутахассиси А.Пўлатов ўзбек тили грамматикаси тўлиқ ва ягона системага келтирилмаганлигини афсус билан таъкидлайди. “Компьютер лингвистикаси” дарслигида формал грамматикани яратиш ҳолати ҳақида шундай ёзади¹⁰⁴: “Грамматика асосларини баён қиларканмиз, юқоридаги камчиликларни четлаб ўтишга интилдик. Бунда ўзбек тили грамматикасини иложи борича содда тушунтириш, инглиз тили билан қиёслаш, камчиликларга эътибор қаратиш, унга компьютер имконият ва талабларини ҳисобга олган ҳолда ёндашиш мақсад қилиб олинди. Ўзбек тили грамматикасини тасвирлашдаги асосий бўлимлар:

¹⁰⁴Пўлатов А. Компьютер лингвистикаси. – Тошкент, 2011.

- 1) алфавит, товуш, тиниш белгилари;
- 2) сўз яшаш йўллари;
- 3) сўз туркумлари (от, сифат, сон, олмош, феъл, равиш, кўмакчи, боғловчи, юклама, модал, ундов, тақлид);
- 4) сўз бирикмалари;
- 5) гап бўлаклари (эга, кесим, аниқловчи, тўлдирувчи, хол);
- 6) содда гап, содда гап турлари: дарак, сўроқ, буйруқ, ундов;
- 7) қўшма гап: боғланган қўшма гап, эргашган қўшма гап, мураккаб гаплар;
- 8) услубият: бадий, илмий, расмий, публицистик, сўзлашув, компьютер (ахборот) услублари;
- 9) матн, нутқ ва дискурс (муаллиф, реципиент, мухит, матндан четланиш, сўзларнингталқини, *anaphoraresolution*);
- 10) матнтаркибий қисми (ядро, тафсилот, мисол);
- 11) рефератлаштириш, сарлавҳа қўйиш¹⁰⁵.

Ўзбек тили грамматикаси (юқорида кўрсатилган тизимда) тўла ўрганиб чиқилмаган, яъни системалаштирилмаган, формалаштирилмаган. Масалан, феъл, от ва бошқа туркумларга оид сўз шакллар, гап конструкциялари тўлиқ рўйхатга олинмаган (масалан, биргина *ишламоқ* феълинингўзбек тилида тахминан 100 минг, инглиз тилида тахминан 150та шакли мавжуд). Булар ҳақидаги маълумотлар А.Пўлатов, Т.Мўминова, И.Пўлатовалар қаламига мансуб “Дунёвий ўзбек тили. Ўзбек тилида феъл шакллари ва уларнинг рус, инглиз тилидаги кўринишлари” (Тошкент: Университет, 2003) асарида келтирилган”¹⁰⁶.Бундан келиб чиқадики, ахборот-қидирув тизимлари учун ўзбек тилининг формал грамматикаси, шу асосда дастурий воситаларнинг лингвистик таъминот ишлаб чиқиш компьютер лингвистикаси олдида турган долзарб вазифалардан бири саналади.

¹⁰⁵Пўлатов А. Компьютер лингвистикаси. – Тошкент, 2011.

¹⁰⁶Пўлатов А. Компьютер лингвистикаси. – Тошкент, 2011.

Маълумки, лингвистик таъминот лингвистик моделлар яратишга асос бўлади. Ўзбек компьютер лингвистикасида лингвистик модел ва моделлаштириш, унинг умумий тамойиллари ҳақида фикр юритилган ишлар кўп эмас. С.Мухамедов ва Р.Р.Пиотровский томонидан ўзбек тилидаги матнларнинг квантатив моделлари тузилган¹⁰⁷. Шунингдек, инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти¹⁰⁸ борасида қилинган илк тадқиқотлар ахборот-қидирув тизимида ўзбек тили сўзшакллари қайта ишлашда катта аҳамият касб этади.

Машина таржимасида лингвистик таъминот. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминотини тадқиқ этган Н.Абдурахмонова 12 мингдан ортиқ ўзбек тили феълли фраземаларни Access маълумотлар базасига жамлаган; 3 мингдан ортиқ инглиз тилидаги асосий феъллар маъноси ва ўзбек тилидаги муқобилларини берган¹⁰⁹. Унинг фикрича, инглиз ва ўзбек тилларидаги дериватив имкониятнинг ўрганилиши морфотактик ҳолатни аниқлашда, яъни ҳар бир морфеманинг семантик ва мантиқий чегарасини топишга ёрдам беради. Таржиманинг уч асосий компоненти мавжуд: асос матн (1), унга ёндашувнинг тўғрилиги (2) ва терминология (3) масалалари жуда муҳим. Матнни таржима қилишда таҳлил жараёни, сўзнинг матндаги маъносини тушуниш ҳамда матнни тиклаш жараёнига алоҳида эътибор қаратиш талаб қилинади¹¹⁰. Демак, ахборот-қидирув тизимида сифатли қидирувини амалга оширишда лингвистик таъминот муҳим омил саналади. Тадқиқотчи машина таржимасининг таҳлил бошқичлари сифатида токенизация (сўз шакллари аниқлаш), лемматизация (сўзни аниқлаш) ва стеммизация (етакчи/кўмакчи морфемаларни аниқлаш)

¹⁰⁷Мухамедов С.А., Пиотровский Г.Г. Инженерная лингвистика и опыт системно – статистического исследования узбекских текстов. –Тошкент: Фан, 1986.; Махмудов М.А., Пиотровская А.А., Садыков Т. Система машинного анализа и синтеза тюркской словоформы // Переработка текста методами инженерной лингвистики. –Минск, 1982.

¹⁰⁸Абдурахмонова Н.З. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти (сода гаплар мисолида): филол. фан. бўйича фалсафа докт. дисс. автореф. – Тошкент, 2018. – 49 б. – Б.13.

¹⁰⁹Кўрсатилган манба. – Б.16.

¹¹⁰Кўрсатилган манба. – Б.18.

кабиларни ажратади¹¹¹. Бу жараёнларни ишга тушириш учун дастурий тизимнинг лингвистик таъминоти сифатида асос/лексемалар луғати талаб қилинади.

Юқоридаги фикрдан хулоса қилиш мумкинки, луғатнинг махсус объектли модели, луғат, морфологик, фонологик ва морфофонологик коидалар морфологик таҳлил тизими учун жуда муҳим воситалар саналади.

Родолфе Делмонтенинг фикрича, морфологик анализ у ёки бу категорияларни таҳлил этибгина қолмай, лексикон учун қуйидаги лингвистик категориялар устида ҳам таҳлил олиб боради¹¹²:

- 1) грамматик категориялар;
- 2) семантик категориялар;
- 3) дискурс даражасидаги категориялар;
- 4) синтактик категориялар;
- 5) тобели категориялар;
- 6) семантик концептуал категориялар;
- 7) чегараланган лексик бирликлар;
- 8) грамматик чекловлар.

Исталган тизим (ахборот-қидирув, машина таржимаси, автоматик таҳрир)да сўзнинг грамматик (морфологик ва синтактик) таҳлиliga мурожаат этилади. Шу боис кўплаб тадқиқотларда морфологик таҳлилни оптималлаштиришга оид қатор фикрлар билдирилган¹¹³. Инглиз ва ўзбек

¹¹¹Кўрсатилган манба.

¹¹²Rodolfe D. Computational Linguistic Text Processing: Lexicon, Grammar, Parsing and Anaphora Resolution. Nova Science Publishers, Inc. – New York, 2008. – P. 4-5.

¹¹³Belonogov G. G. On the use of a similarity method for automatic processing of textual information [Об использовании метода аналогии и при автоматической обработке текстовой информации] Problems of cybernetics [Problemy kibernetiki]. Issue 28. – Moscow: Nauka Publ., 1974.; Bolshakov I. A. CrossLexica: A large electronic dictionary of collocations and semantic links between Russian words [КроссЛексика – большой электронный словарь сочетаний и смысловых связей и русских слов]. Компьютерная лингвистика и Интеллектуальный Технологии: Труды Международной Конференции «Диалог 2009» [Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference «Dialogue 2009»]. – Moscow, 2009. – P. 45-50.; Gelbukh A. F. Effectively realizable morphologic model of inflective language [Эффективно реализуемая модель морфологии и активного языка]. Научно-техническая Информация [Scientific and Technical Information], series 2, #1, 1992. – P. 24-31.; Sokirko A. V. Bystroslovar': morphological prediction of new Russian words using very large corpora [Быстрословарь: предсказание морфологии русских слов с использованием больших лингвистических ресурсов] Компьютерная лингвистика и Интеллектуальный

тилларидаги ясовчи қўшимчаларни статистик жиҳатдан ҳисоблаш, уларнинг семантик базасини яратиш дастурий таъминот тузишда ёрдам беради. Зеро, компьютер учун қўшимча (морфема)ларни алоҳида база тарзида ажратиш, моделларини аниқлаш таҳлил жараёнида муҳим омилдир. Моделлаштириш – машина таржимасида энг самарали метод; бунда лисоний ҳодисаларни муайян даражада таҳлил қилиш, текшириш ва солиштириш имконияти мавжудки, унутқий ҳодисаларнинг барчасини ўзида тўлиқ акс эттирмаса-да, улар ҳақида муайян хулоса беришда аҳамиятли. Тилда универсал қолипни яратиш мумкин эмас, чунки тилнутқий ҳодиса сифатида ҳар бир индивидда ўз аксини топади. Масаланинг оптимал ечимини топишда ҳар бир тилнинг таҳлили учун ўнлаб алгоритмлар яратилганини ҳисобга олиб, таъкидлаш лозимки, тилни формал даражада моделлаштириш самарали усул саналади. Демак, машина таржимасида лингвистик таъминот мукамал даражада бўлсагина, унинг дастурий таъминоти яратилади. Лингвистик таъминотнинг семантик қатлами, тил ва нутқ бирликлари базаси машина таржимаси омонимия ва синонимияни аниқлаш, фарқлаш, тўғри ечим бериш масаласини ҳал қилишда ёрдам беради.

Лингвофольклористика ва компьютер лексикографиясида лингвистик таъминотнинг аҳамияти. Ҳозирги замон тилшунослигида тил қонуниятларини яхлит тизим сифатида тадқиқ этиш учун лингвостатистик кузатишларга катта эҳтиёж сезилмоқда. Мазкур усулни амалиётга татбиқ этиш билан, биринчидан, лингвистик кузатишларда аниқ фактларга таянилади, иккинчидан, метатилни яратишда долзарб муаммо сари бир қадам қўйилади. Чунки матн ўз-ўзини ташкил этувчи, очиқ система сифатида ҳозирги даврда мавжуд бўлган лексик ва грамматик ахборотларни статистик сақлаш тизими ҳисобланади. Матн услубига хос жиҳат изланиши жараёнида айрим лексемалар бажарган коннотатив вазифанинг намоён бўлиши матн лингвистикасида лингвостатистик, лингвостилистик, лингвопоэтик

таҳлилларининг биргаликда олиб борилиши зарурлиги ҳақидаги назарий қарашни амалий жиҳатдан исботлади. Демак, ўзбек тилидаги ҳар бир лексик-семантик гуруҳга оид бирлик, ҳар бир сўз туркумига оид элемент шу таҳлил усулига таянилган ҳолда олиб борилиши мумкин¹¹⁴.

Тилшунос Д.Ўринбаева монографик ишининг “Ўзбек халқ оғзаки ижоди асарлари матнининг лингвоахборот базасини яратиш тажрибасидан” деб номланган бўлимида тилнинг математик моделини яратиш, лингвистик муаммоларни ҳал қилувчи турли хил компьютер дастурлари, жумладан, лингвостатистик эҳтиёжлар ечимини топишга йўналтирилган дастурларни тузишнинг илмий-амалий аҳамияти хусусида сўз юритилган¹¹⁵. Маълумки, компьютер хотирасига киритилган тил материали лингвистик-ахборот база ҳисобланади ва ишлаб чиқилган лингвистик алгоритм, модуллар бир тизимга жамлангач, лингвистик-ахборот, математик-ахборот таъминоти ва маълумотлар базаси (банк) маълум бир матнни автоматик қайта ишлаш механизми яратилади. Маълумотлар базаси кенгайган сари лингвистик-ахборот таъминоти муаммоси камайиб боради. Д.Ўринбаева маълумотлар базасини вужудга келтиришда тилнинг лексик, грамматик, фонологик ва морфонологик блокларни атрофлича ва батафсил тавсифлайди¹¹⁶. Диссертациянинг “Ўзбек халқ оғзаки ижоди матнлари морфологик бирликларининг лингвостатистик тадқиқи” деб номланган тўртинчи бобида амалга оширилган тадқиқотлар якуни ўлароқ, ўзбек халқ оғзаки ижоди матнлари морфологиясининг лингвостатистик тадқиқида масалага грамматик умумлаштирувчилик позициясидан ёндашилган; терс луғат тузишга мўлжалланган компьютер дастури яратиш тамойиллари тавсифлаган. Тузилган луғат муайян туркумга кирувчи сўзшаклнинг миқдори, услубий чегарасини белгилаш; семантик-стилистик, лингвопоэтик таҳлил учун

¹¹⁴Ўринбаева Д. Халқ оғзаки ижоди: жанрий-лисоний ва лингвостатистик тадқиқ муаммолари: фил. фан. док. дисс. автореф. – Самарқанд, 2019.– 74 б.

¹¹⁵Кўрсатилган манба. – Б.14.

¹¹⁶Кўрсатилган манба. – Б.16.

лингвистик таъминот бўлиб хизмат қилган. Ишнинг “Мустақил сўз туркумларининг лингвостатистик тадқиқи” деб номланган бўлимида матнларнинг квантитатив таҳлили сўзшаклнинг сўз туркумига тақсимланиш тамойилларига таянилган ҳолда амалга оширилган. Бунда луғатдаги сўзшакллар қисмларга тақсимланиб, жанрлараро сўз туркумлари қиёсланган; танланма матннинг тилости услуб, тагуслубдаги ўзига хос типологик хусусиятлари аниқланган. Натижада, луғат таркибида сўз туркумларининг тақсимоти луғат ҳажми билан боғлиқ эмаслиги, луғат таркибининг миқдорий ўсиши билан сўз туркумларининг матндаги улуши ўзгариб қолмаслиги маълум бўлган¹¹⁷.

Матнга лексикографик ишлов беришнинг компьютер таъминотини такомиллаштириш, максимал лингвостатистик аниқликка эришиш, шунингдек, сўзшаклларни ягона ўзак атрофида жамлаш, уларни фонема, морфемага ажратиш, сўз бирикмалари луғатини тузиш дастурини яратиш нафақат лингвофольклористик, балки фонологик, лексик-семантик, лингвостилистик ва лингвопоэтик тадқиқотлар учун ҳам долзарблигича қолаверади. Ҳозирга қадар жамланган фольклор материалининг кенг қўламли лингвостатистик ва лексикографик тадқиқи келажакда ўзбек лингвистик корпусининг яратилишига муҳим ҳисса бўлиб қўшилади.

Кўринадики, лингвистик таъминот машина таржимаси, компьютер лексикографияси, корпус лингвистикаси, қолаверса, автоматик морфологик таҳлил жараёни учун дастурий таъминотнинг зарурий таркибий қисмидир.

Боб бўйича хулоса

1. Лингвистик таъминот ижтимоий мулоқот доирасида тил воситаларининг адекват ишлашини таъминловчи белгилар мажмуи, ахборот технологиялари дастурий таъминотида ишлатилувчи тиллар, уларни яратиш усул/воситалари, қўллаш ва фойдаланиш комплекси саналади. Ушбу атамага

¹¹⁷Кўрсатилган манба. – Б.16.

амалий ва компьютер лингвистикаси нуқтаи назаридан ёндашувда лингвистик таъминотга табиий тилда матнларни автоматик қайта ишлашда қўлланадиган воситалар мажмуи, тил процессори деб қаралади. Зеро, лингвистик таъминот автоматлашган қидирув тизимида қўлланувчи ахборотни автоматик таҳлил қилишни унумли даражага кўтариш мақсадида табиий тилнинг формал қоидаларини яратишга хизмат қилувчи тил воситалари мажмуидан иборат.

2. Автоматлашган тизимнинг лингвистик таъминоти автоматлашган тизимда қўлланувчи, ахборотни автоматик таҳлил қилишни унумли даражага кўтариш мақсадида табиий тилнинг формал қоидаларини яратишга хизмат қилувчи тил воситаларидан иборат. Ахборот тизимларининг лингвистик таъминоти эса ахборот объекти тузилиши, унинг ишлаши учун фойдаланиладиган, формаллашган, табиий тил ифодалядиган тушунча устида бажариладиган амаллар, унинг воситалари таснифи ва тўплами ҳисобланади. Автоматик лойиҳалаш воситалари комплексини турли таъминотлар мажмуи (математик, лингвистик, техник, информацион, дастурий, методик, ташкилий таъминот) ташкил этади. Жумладан, дастурий таъминот деганда ҳисоблаш техникаси воситасида маълумотни қайта ишлаш тизимини яратиш ва фойдаланиш дастурий воситалари мажмуи назарда тутилса, техник таъминот маълумотни қайта ишлаш тизимини фаоллаштиришда қўлланувчи техник воситалар комплекси бўлиб, у маълумотни қайта ишловчи намунавий амалларни амалга оширувчи қурилмаларни ўз ичига олади.

3. Лингвистик дастурий таъминот табиий тилдаги аудиомаълумот, расм, матнни таҳлил қилиш, қайта ишлаш, сақлашга оид дастур ва маълумотлар мажмуи бўлиб, унинг электрон луғат, орфоқорректор, ыидирув тизими, машина таржимаси тизими, таржимон хотирасини бошқариш дастури, нутқ анализи ва синтези тизими, овозли таржима тизими каби турлари ажратилади. Ахборот-қидирув тизимининг лингвистик таъминоти таркибига ахборот-қидирув тили (алифбо, АҚТ микросинтаксиси, луғат, тезаурус, ёрдамчи

луғатлар, классификатор, АҚТнинг граматик воситалари, табиий тил маълумотини ифодалаш воситалари); турли тилларни мослаштиришни таъминловчи воситалар (мослик жадвали, конвенрлаш жадвали, протоколлар); амалиёт воситалари; матнларга табиий тилда ишлов беришнинг ёрдамчи воситалари киради.

4. Табиий тилда ишлов беришнинг ёрдамчи воситалари морфологик ва синтактик таҳлилнинг дастурий-алгоритми; матн маълумотларини семантик қайта ишлаш; материални автоматик таснифлаш; автоматик қидирув буйруқларини тузиш, хатони автоматик аниқлаш алгоритмлари; лингвистик маълумотлар базаси, табиий тилдаги лексик ва морфологик birlikлар луғатини қамраб олади.

5. Лингвистик анализаторнинг вазифаси таҳлил қилинаётган тил хусусиятидан келиб чиққанлиги сабабли бўлғуси ахборот-қидирув тизимлари ўз таркибига матн лингвистик таҳлилининг тўлиқ топшириқлар рўйхатини амалга оширувчи дастурий-техник мажмуани қамраб олиши талаб этилади. Чунки мукамал ахборот-қидирув тизимлари матн ахборотини морфологик таҳлил қилиш вазифасини бажаради, гапнинг содда синтактик таҳлилинини амалга оширади. Морфологик таҳлилнинг вазифаси матндаги ҳар бир сўзнинг қайси туркумга мансублигини автоматик тарзда муайянлаштириш, сўзларнинг қайси лексик-грамматик синфга мансублигини аниқлашдан иборат. Матн ахборотини синтактик таҳлил қилиш гапнинг семантик элементларини – гуруҳга мансублик, предикатив асосни ажратишга имкон яратади. Лексик таҳлил эса матн ахборотини миллий тил ифодаларни кўрсатувчи алоҳида сўз боши, гап, сўз шаклида таҳлил қилиш, тилнинг ўзига хослигини кўрсатувчи гап шакллари, лексик ифода (ўзлашма, жаргон сўз ва ҳ.) ларни аниқлайди.

II БОБ. МОРФОЛОГИК АНАЛИЗАТОР ТУЗИШНИНГ НАЗАРИЙ АСОСЛАРИ

2.1. Морфологик анализатор тузишнинг умумий тамойиллари

2.1.1. Морфологик анализатор ва сўз туркуми теггери (Part-Of-Speech Tagger). Морфологик анализатор луғат(лексикон)даги сўзшакл ва алоҳида сўзларни қиёслашга мўлжалланган, сўзнинг грамматик хусусиятини ифодалайдиган алгоритмлар тўплами ҳисобланади¹¹⁸. Синтактик парсер матнни “таниш” функциясиз ҳам ишлай олса-да, таҳлил жараёнида морфологик анализатор таҳлил натижасини фаол қўллайди. Матннинг грамматик ахборот берувчи разметкаси синтактик таҳлил қоидаларини тузишга ёрдам беради. Мисол тариқасида от ва сифат орасидаги муносабатни аниқлаш қоидасини олиш мумкин. Рус грамматикаси қоидаларидан келиб чиқиб айтиш мумкинки, отдан чап томонда турувчи сифат от билан бир хил келишик, сон ва родда бўлса, у билан алоқага киришади: *На крутом (предложный, ед., муж.) берегу (локатив, ед., муж.).* Ушбу қоида от ва сифат боғланишининг ягона асоси ҳисобланмайди, аммо у нисбатан кўпроқ ҳолатларда шундай боғланишларни ҳосил қилади. Рус тили морфологиясини формаллаштириш жараёнида грамматик шаклларнинг бир хиллиги муаммосидан қочиб бўлмайди. Масалан, “книги” сўзига қараб унинг бош келишик, кўплик сонда (“старые книги лежат на столе”), тушум келишигида (“выкинь эти старые книги”) ёки қаратқич келишигида (“начало книги”) эканлигини аниқ айтиш қийин. Бу ерда грамматик омонимия келиб чиққан. Морфологик анализатор муаллифларининг фикрича, анализатор вазифалари сирасига сўзнинг барча грамматик шаклларини тавсифлаш, тавсифнинг айнан маълум бир матндаги истисно ҳолатлари ҳақида маълумот бериш кабилар киради. Омонимияни аниқлаш муаммоси бир неча алоҳида гуруҳга бўлинади:

¹¹⁸http://www.solarix.ru/for_developers/docs/morphology_analyzer.shtml

бу вазифалар махсус алгоритм ёрдамида ҳал этилади¹¹⁹. Анализаторнинг энг охирги, кўп вақт сарфланадиган вазифаси нотаниш сўзнинг эҳтимолий грамматик маъносини аниқлашдан иборат. Бошқача айтганда, агар таҳлил қилинаётган сўз лексикондан топилмаса, анализатор унинг грамматик роли ҳақида эҳтимолий хулосалар бериши керак. Ушбу вазифанинг бажарилиши учун бир қанча ахборот, хусусан, таҳлил қилиш қоидалари мажмуи талаб этилади (ушбу қоидалар тўғрисида қуйида маълумот келтирилади).

Морфотаҳлил амал қиладиган тиллар борасида. Мисол тариқасида олинган ушбу морфологик анализаторнинг муҳим томони шундаки, у битта тилга мослаштирилмаган, балки ундан бир неча тилни таҳлил қилишда фойдаланиш мумкин. Маълум бир тилнинг қоидалари мажмуи луғат базасига жойлаштирилиб, морфологик таҳлил алгоритми турли тилларга мослаштирилиши мумкин. Масалан, рус тили морфотаҳлили учун қоидаларнинг катта қисми алоҳида сўзлар морфологик таҳлили от ва сифатнинг грамматик род, сон ва келишикка оидлиги натижасига асосланган ҳолда амалга оширилади.

Луғат тайёрлаш жараёнида махсус дастур (луғат компилятори) матнларнинг ифодаланиш қоидаларини ўқийди; уни луғат базасида сақланадиган нисбатан қулай бинар кўринишда ифодалайди. Амалий дастур ва грамматик луғат таркибий қисмлари зарур ҳолларда хотирага қоидаларни юклайди: улардан таҳлил вариантларни ўчириш ва тузатиш учун ишлатилади. Юқоридаги жараён автоматик равишда амалга ошади; `sol_MorphologyAnalysis` функциясидан ташқари бирор махсус код ёки ҳаракат талаб қилмайди. Грамматик анализаторнинг ушбу версиясида рус ва инглиз тиллари учун морфологик модуллар мавжуд. Шунингдек, морфологик анализаторнинг рус тили варианты (Part-Of-Speech Tagging билан биргаликда) етакчи саналади.

¹¹⁹http://www.solarix.ru/for_developers/docs/morphology_analyzer.shtml

Автоматик анализ ва синтез хусусида. Матннинг автоматик таҳлили бир неча мураккаб операциядан иборат: компьютер табиий тилдаги матн таҳлилини белгиланган алгоритм асосида амалга оширадики, автоматик таҳлил давомида компьютер матн ҳақида ўз тилида лексик-морфологик, синтактик, семантик тасаввур ҳосил қилади. Матннинг автоматик анализ ва синтез жараёни компьютер лингвистикасининг асосий вазифаларидан бири саналади. Бу вазифа сунъий интеллект учун лингвистик моделлар яратиш ва инсониятнинг амалий эҳтиёжидан (масалан, автоматик машина таржимаси) келиб чиқади¹²⁰.

Матннинг автоматик таҳлили бир неча босқичдан иборат¹²¹:

- 1) графематик таҳлил: сўз, гап, абзац ва бошқа матний элемент (масалан, газета матни бўлаги) чегарасини аниқлаш;
- 2) морфологик таҳлил: матнда қўлланган сўзшаклнинг бош шаклини ҳамда бу сўзнинг морфологик белгиларини аниқлаш;
- 3) синтактик таҳлил: матндаги гапнинг грамматик структурасини аниқлаш;
- 4) семантик таҳлил: ибораларнинг маъносини фарқлаш.

Графематик таҳлил токенизация (инг. token. – матндаги сўз, ибора ёки бирор аҳамиятли элемент)га тенг бўлиши мумкин. Матний элементлар чегараси аниқловчиси сифатида шаклий сигналлар сўзлар орасида пробел (ажратма) бўлса, гап ва унинг қисмлари чегараси бош ҳарф ва тиниш белгилар, матннинг мазмуний қисмлари абзац билан фарқланади¹²². Аммо сўз чегарасини аниқлашда доим ҳам формал метод қўл келмайди. Масалан, хитой тилида сўзнинг шаклий чегараси мавжуд эмас. Ҳатто, Европа тилларининг аксариятида пробел (ажратма) билан ажратилган бир неча сўзларнинг битта лексемага тенг келиш ҳолати ҳам бу метод билан графематик таҳлил натижа бермаслигини кўрсатади: масалан, New York. Бундай ҳолларда графематик

¹²⁰Щипицина Л.Ю. Информационные технологии в лингвистике : учеб. пособие / Л.Ю.Щипицина. – М.: ФЛИНТА : Наука, 2013. – 128 с. – С. 43.

¹²¹Автоматическая обработка текстов на естественном языке и компьютерная лингвистика: учеб. пособие. М.: МИЭМ, 2011. – С. 94,106-107.

¹²²Щипицина Л.Ю. Кўрсатилган манба. – С.107.

таҳлилнинг аниқлигига истисно сўзлар луғати тузиш орқали эришиш мумкин, яъни таркибли сўзларнинг маълумотлар базаси шакллантирилади.

Матннинг морфологик таҳлилида унда қўлланган сўзшаклнинг бош шакли – лемма аниқланади ҳамда сўзнинг шу контекстдаги грамматик маънолари ажратилади: от учун туркум, шахс, сон, келишик, феъл учун шахс, сон, замон ва ҳ. Матнда ишлатилган ҳар бир сўз *сўзшакл* ёки *сўз қўллаш ҳолати* дейилади.

Флексив тилларда морфологик таҳлилдаги ўзак (асос) анъанавий грамматикадаги асосга тўғри келмаслиги мумкин. Морфологик таҳлилда асосни аниқлаш *стемминг* дейилади. Стем бу сўзшаклнинг синтактик шакл ясовчиларсиз қисми, флексия ҳодисасига учраган айрим туркий ўзакларнинг грамматик асоси билан морфоанализ жараёнидаги асоси – стем мос келмай қолиш ҳолатлари учраб туради. Масалан, *сон* ўзагидан ясалган *сана* феълнинг ўзаги *сон*, аммо бу сўзшаклнинг стеми *сан*. Бу ерда ўзакдан –а сўз ясовчи ажратиб олинади ҳамда *сан* стем бўлиб қолади. Лекин агглютинатив тилларда, жумладан, ўзбек тилида одатда, лемма стемга тенг келади ҳамда флексия ҳолати морфоанализ жараёнига кўп ҳалақит бермайди. Ишнинг кейинги бобларида флексияга учраган сўзларнинг морфоанализ масалалари алоҳида ҳал этилади.

Морфоанализда фақат лемматизация жараёнининг ўзи камлик қилади, бунда стеммингнинг ҳам ўз ўрни бор. Хусусан, интернетда қидирув сифатини яхшилашда стем асосий омил бўлади, яъни қидирув лемма асосида эмас, стем асосида амалга оширилади¹²³. Морфологик таҳлилда компьютернинг сўзшаклни тушуниши билан иш битмайди, яъни белгилар кетма-кетлигининг компьютер томонидан фарқланиши етарли бўлмайди. Сўзшаклнинг парадигмадаги барча аъзоларидан фарқини топиш жараёни билан таҳлил

¹²³Толдова С.Ю., Бонч-Осмоловская АА. Автоматический морфологический анализ// Фонд знаний «Ломоносов». М.,2011. www.lomonosov-fund.ru/enc/ru/encyclopedia:0127430

тугамайди. Кейинги босқич – сўзшаклнинг туркумга хослиги (таггинг) ҳамда морфологик белгиларини топиш.

Сўзшаклнинг грамматик маъносини билдирувчи қўшимчалар (компьютер моделлари) муайян лексеманинг шакл ясовчилари ҳамда унинг парадигма аъзоси сифатида шакллантирувчиси саналади. Таҳлил тизими маълумотлар базасида маълум лексемага бирикиши мумкин бўлган барча қўшимчалар парадигмаси тузиладигани, бундай парадигмалар тузилганда, айрим қўшимчаларнинг бир неча грамматик маънони англатиши, қўшимчалар омонимияси аниқланади.

Сўзларда лексик-морфологик омонимия ҳам кузатилади. Морфоанализатор бундай ҳолатда сўзшаклга морфологик белгиларни тўғри бириктириши учун гапнинг синтактик ҳамда семантик хусусиятига таянади. Бу жараёни икки хил: автоматик ҳамда инсон омили билан амалга оширилиши мумкин. Бу каби жараёнларнинг автоматик тегланишида қоидалар модуллари ишга туширилади. Масалан, инглиз тилида қуйидаги модуллар орқали грамматик омонимияни фарқлаш ва бартараф этиш мумкин:

- 1) агар омоним шакл ҳам от, ҳам феъл бўлиши мумкин бўлса, ундан олдин артикл турган бўлса, бундай ҳолда бу сўзшакл от;
- 2) агар сўзшакл ҳам предлог, ҳам боғловчи бўлиши мумкин бўлса, ундан кейин гап охиригача феъл мавжуд бўлмаса, бундай ҳолда бу шакл предлог¹²⁴.

Кўриниб турибдики, бундай қоидалар мажмуи, яъни гапда сўз туркумлари ва гап бўлақларининг тартибини аниқлаш орқали омонимияни фарқлаш имконияти пайдо бўлади.

Умуман олганда, морфологик таҳлил қуйидаги босқичлардан иборат¹²⁵:

¹²⁴Щипицина Л.Ю. Информационные технологии в лингвистике : учеб. пособие / Л.Ю.Щипицина. – М.: ФЛИНТА : Наука, 2013. – 128 с. – С. 43.

¹²⁵Толдова С.Ю., Бонч-Осмоловская АА. Автоматический морфологический анализ // Фонд знаний «Ломоносов». М., 2011. – www.lomonosov-fund.ru/enc/ru/encyclopedia:0127430

- 1) сўзшаклларнинг асосини топиш – лемматизация; сўзшаклнинг бошланғич шакли – лемма ёки стемга морфологик белгиларни бириктириш, турли сўзшаклларни бир ўзакка тенглаштириш;
- 2) таггинг – сўзшаклнинг контекстадаги сўз туркумига мансублик тўғрисидаги ахборотни аниқлаш;
- 3) тўлиқ морфологик таҳлил – сўзшаклга грамматик характеристика ёзиш.

Морфологик таҳлилда сўз ва гап орасидаги фарқ: POS Tagging ҳамда сўзшаклни аниқлаш. Грамматик анализаторда изоляцияланган (контекстдан ташқаридаги) ҳамда гап таркибидаги сўзнинг морфологик таҳлилини амалга ошириш мумкин. Таҳлил амал қиладиган тилларнинг контекстдан ташқаридаги сўзи таҳлилида проекциянинг бир неча варианты таклиф этилади¹²⁶. Рус тилида бир луғат мақоласи доирасида грамматик шаклларнинг омонимлик ҳолати жуда кўп учрайди. Масалан, *камень* сўзшакли бош (*камень лежит на земле*) ёки тушум (*взять камень*) келишиги шакли бўлиб келиши мумкин. Инглиз тилида грамматик омонимликнинг бундай ҳолатига грамматик синфларнинг мос келиши ҳам қўшилади. Масалан, *back* от ёки равиш бўлиши мумкин. Грамматик таҳлил дастури сўзшаклнинг грамматик маъносини аниқлашда барча шакл, маъноларни таклиф қилади, аниқ бир вариантни танлаш амалий код зиммасида қолади.

Матндаги сўз туркумини автоматик разметкаш *Part-Of-Speech Tagging* деб аталади. Бундай ҳолатда сўзнинг бирикма/гаптадаги контексти инобатга олинади: қуршовидаги икки сўз, узоқда турган грамматик алоқа фрагментлари, ҳатто матн мазмуни ҳосил қиладиган умумий контекст ҳам аҳамиятли. Ушбу қисқа тавсифдан маълумки, POS tagging функцияларининг юзага чиқариши учун нисбатан кўпроқ ресурс талаб қилади, бу маълум бир сўзнинг имконсиз вариантини аниқлаш имконини беради. POS tagger фаолиятини йўлга қўйиш учун бир неча алгоритмик ёндашувлар, хусусан, грамматик луғатда бир-бирини тўлдирувчи икки омил мавжуд:

¹²⁶http://www.solarix.ru/for_developers/docs/morphology_analyzer.shtml

- 1) бериладиган қонидани қўлда (мутахассис томонидан) киритиш – сўзни аниқлаш (“таниш”)нинг кам ёки умуман учрамайдиган вариантыни ҳисобга олган ҳолда шундай йўл тутилади. Масалан, “рус тилида предлогдан кейин равиш келолмайди” қонидаси, ўзбек тилида эса “равишдан кейин от келолмайди” қонидасини мисол қилиш мумкин;
- 2) разметкаланган матнлар корпусида ўрганиладиган морфологик модел сўз туркуми эталон разметкаси статистик қонуниятлари асосида иш қўради ҳамда уларни қўллайди.

Қуйида морфологик анализаторнинг маълум мақсад учун ишлатиш қонида ва техникасига тўхталамиз.

Матн аниқловчиси лексернинг бир қисми сифатида. Сўзнинг морфологик таҳлили лексер ёрдамида амалга оширилади; лексер сўзнинг белгилар занжирини аниқлайди, матнда сўзни ажратиб олади. Морфологик таҳлилда парсер нисбатан “ташаббускор” қисм саналади, чунки у лексердан гипотезани текшириш учун янги сўз сўрайди. Сўз чегараси, табиий равишда аниқ ажратилади.

Таҳлил (аниқлаш) натижаси. Морфологик таҳлил натижаси қуйида кўрсатилган бир қанча маълумот мажмуидан иборат:

- 1) сўзнинг туркумга мансублиги;
- 2) сўзнинг грамматик шакли;
- 3) сўзшаклга ишлов бериш учун керакли грамматик белгилар мажмуи (масалан, рус тилида от ва сифат учун келишик, инглиз тили учун феъл шакли ва ҳ.). Шунингдек, агар сўзшакл грамматик луғат орқали ишлов берилган, таҳлил қилинган бўлса, луғат мақоласига тегишли грамматик атрибутларни ҳам аниқлайди.

Морфологик луғат. Таҳлилнинг ушбу босқичида сўзнинг грамматик хусусиятини аниқлаш учун лексикондаги маълумотдан фойдаланилади. Лексикон маълумотлар омбори луғатининг таркибий қисми, у луғат мақоласи ва парадигмалардан ташкил топади. Бошқача айтганда, ҳар бир сўз учун унинг

грамматик шакл сифатида қайси луғат мақоласига тўғри келишини аниқлаш мумкин. Рус тили грамматик шакллари рўйхати 3 млн (такрорлар билан қўшиб саналганда) шаклни қамраб олган. Лексиконни сақлаш учун реляцион маълумотлар омбори талаб этилади: индексланган жадвал – маълумотлар омбори стандарт воситалари ишлатилади. Демак, грамматик луғат морфологик анализаторнинг ниҳоятда муҳим таркибий қисми ҳисобланади. Морфологик луғат бу маълум бир сўзнинг барча грамматик шакллари рўйхатини тўплайди.

Луғатсиз морфология ва таҳлил алгоритми. Морфологик луғат асосидаги таҳлил натижанинг ишончлилиги билан афзаллик касб этади: натижа аниқ, оддий, тезкор. Унинг камчилиги фақат маълум гуруҳ сўзларни таҳлил қилиш имкониятигагина эга, холос. Катта ҳажмли корпус таҳлил қилиниб, частотали гистограмма тузилса, юқори частотали сўзлар кам миқдорда намоён бўлади. Морфологик луғат асосидаги таҳлил 3 млн сўзшаклни муваффақиятли таҳлил қила олади. Одатий (тор соҳа матнларидан ташқари) матн таҳлилида натижа 97 фоизда аниқлик касб этади. Таҳлилнинг қолган 3 фоизи алгоритмлар мажмуи томонидан (масалан, сўзшаклнинг кўшимчасига қараб) амалга оширилади. Сон туркумига мансуб сўзлар кам миқдорни ташкил этади; чунки уларнинг нутқда қўлланиш частотаси ҳам нисбатан паст. Сонларнинг ҳар бири битта сўзга тенглаштирилади, чунки уларнинг барчасини луғатда акс эттириш имкони бўлмайд.

Луғатда мавжуд бўлмаган сўзлар таҳлили. Лексикон катта ҳажмли бўлишига қарамай, исталган матнда луғатда учрамайдиган сўзга дуч келиниши эҳтимоли бор. Кўп ҳолларда, синтактик анализатор ўз ишини давом эттириши учун ноаниқ сўзнинг қандай грамматик характерга эга эканлигини аниқлашнинг ўзи кифоя. Дастлаб, анализатор ноаниқ сўзнинг кўп турланадиган туркумга тегишли эканлигини текшириб кўради: уларнинг флексив ҳолати грамматик хусусиятларини аниқлашга ёрдам беради. Рус тилида бундай усулда сифат билан ишлаш қулай. Род, сон, келишик, қиёсий

ва орттирма даража маънолари бўлганлиги туфайли рус тилидаги сифатлар бошқа туркумларга нисбатан катта парадигмани ташкил этади. Рус тилида сифат тез-тез турланади, шу сабабли қўшимча сифатни автоматик белгилашга ёрдам беради; унинг морфологик характеристикасини аниқлайди. Рус тили лексикасида мавжуд сўзларнинг ярмини сифат туркумига оид лексемалар ташкил этади. Шу сабабли сифатни таҳлил қилиш қоидаси жуда сермахсул қоида саналади. Бундай қоидалар матн муҳаррири ёрдамида тузилади ва таҳлил амалга оширилади (Қаранг: 1-расм). Кўриниб турганидек, таҳлил таркиби сўзнинг қайси тилга мансублиги (рус тили), қоида ҳамда таҳлил натижасидан иборат. Катта қавс ичидан грамматик белгилар мажмуи ўрин олган.

Таҳлил қилинмаган сўз ва UnknownEntry луғат мақоласи. Агар анализатор берилган қоидалар алгоритми воситасида сўзни таҳлил қила олмаса, анализатор UnknownEntries:UnknownEntry{} луғат мақоласида сўз шаклини ҳосил қилади. Рус тили матнларида учрайдиган лотинча номни рус тили сўз туркуми сифатида таҳлилга тортиб бўлмайди, шундай ҳолда анализатор махсус сўзшакл билан чекланади. Бундай сўзшакл ўзгаришсиз қолдирилади; у синтактик таҳлил жараёнида таҳлилга тортилади.

Рус тилида сўзни ё ҳарфи билан ёзиш ҳақида. Ё ҳарфи билан ёзиладиган сўзни таҳлил қилишда ушбу ҳарфнинг икки асосий хусусияти инобатга олинади. Рус тилида бир қанча сўз ё ҳарфи билан фарқланади. Масалан, белый мел–мёл пол; полёт на Марс–дедушка полет траву. Одатда, ё ҳарфи ёзувда е билан ифодаланади, болалар адабиётидагина ё ва е ҳарфлари фарқланади ҳамда бундай адабиётларда ё ўз ифодаловчиси ўрнида қўлланилади. Ё ҳарфининг бундай қўлланиши анализаторнинг иш унумдорлигига унчалик таъсир кўрсатмайди. Анализатор шундай тузилганки, агар сўз ё билан ёзилган бўлса, луғат қидируви аввалдан ё ҳарфи йўқлиги маълум шаклларни четлаб ўтади. Бунда сўз махсус луғатдан танлаб олинади ва бундай луғатлар анализатор тузилаётган вақтда тўлдирилади.

Ассоциация ташиқи омонимликни фарқлаш механизми сифатида. Ушбу ёрдамчи восита бошқа усуллар қўл келмаганда иш беради. Ассоциация – бу сўзнинг одатий қўлланишини аниқлаш усули. Муаллиф мисол сифатида *Храните деньги в банке* гапани келтиради. Ушбу гапда *банке* сўз шакли – омоним. У “женский род”га мансуб *банка* ёхуд “мужской род”даги *банк* отининг шакли бўлиши мумкин. Иккала шакл ҳам бир хил келишида. Натижада, омонимликни бартараф этиш учун ҳеч қандай грамматик мезон қолмайди. Бундай пайтда таҳлилга аниқлик киритиш учун сўзнинг контекстда қўлланиш статистикаси ёрдамга келиши мумкин: *банк* ва *пул* сўзининг, одатда, ёнма-ён келиши кераклигини аниқлаш муҳим. Агар *Засаливайте огурцы в банке* гапани олсак, *банка* сўзини танлашнинг аналогик усули иш беради. *Банка* *огурцы* сўзи билан ёнма-ён қўлланади.

Омонимликни ассоциация орқали аниқлаш шажара дарахти асосида ишлайди. Агар шундай дастур мавжуд бўлса, ассоциацияланган сўзларни аниқлаш қуйидагича амалга ошади:



Омонимликни аниқлашнинг бундай усули доим ҳам иш бермайди, шунинг учун у ёрдамчи механизм саналади.

Бундай мисоллар ўзбек тилида кўп учрайди. Ўзбек тилида бир туркум доирасидаги омонимлик ҳодисаси грамматик омонимияни келтириб чиқаради.

*1.Карамингизни менадан аяманг. 2. Тангрининг карамига сизиниб, уларни Астробод томонга жўнатдим. 3. Карам шўрва. 4. Тузланган карам. 5. Бир бош карам*¹²⁷.Биринчи мисолда матнда *карам* сўзини қуршаб турган сўзлар

¹²⁷Раҳматуллаев Ш. Ўзбек тили омонимларининг изоҳли луғати. – Тошкент: Ўқитувчи, 1984. – 214 б. – Б. 72.

омонимликни аниқлай олмайди, аммо иккинчи мисолда *тангри, сизиниб* сўзшакллари ассоциатив қуршови *карам* сўзининг маъносини реаллаштиради. 3-,4-,5-мисолларда *карам* сўзининг қуршови юқорида изоҳланган алгоритм асосида омонимликни фарқлаши мумкин. Демак, МАда грамматик омонимликни граммема орқали фарқлаш имконияти бўлмаганда, ассоциациядан фойдаланиш самара беради.

Морфологик таҳлил алгоритмини оптималлаштириш ҳақидаги тадқиқот¹²⁸даги фикрлар ишимизда таҳлил этилаётган масалаларга ойдинлик киритади. Ушбу мақолада морфологик анализатор тузишнинг уч усули санаб ўтилади:

- 1) луғатга асосланган анализатор;
- 2) луғатсиз / грамматик қонуниятлар мажмуига асосланган анализатор;
- 3) луғат ва грамматик қоидалар базасига асосланган анализатор.

Юқорида санаб ўтилган усуллар ичидан луғатга асосланган анализатор ҳақида батафсил маълумот келтирамиз. Бунинг учун олтой оиласига кирувчи тиллар орасида сингармонизм ҳодисаси нисбатан кенг тарқалган қирғиз тили қоидалари ҳамда фактларига таянамиз. Шундай қилиб, табиий матн морфологик анализатори принципларини кўриб чиқишда қуйидаги босқичларни ҳисобга олиш керак:

- 1) кириш матни сўзларини грамматик шаклларга ажратиш;
- 2) сўзни лемматизациялаш, яъни асослар луғатидан сўзнинг лексема шаклини топиш/ аниқлаш;
- 3) синтактик шакл ясовчи қўшимчалар занжирини аниқ қисмларга ажратиш;
- 4) ҳар бир аффикснинг морфологик белгисини аниқлаш.

Масалан, *балдар (болалар)* сўзшакли таҳлил қилинаётганда, морфологик анализатор бу шакл *-лар* кўплик қўшимчаси қўшилган *бола* асосидан ҳосил

¹²⁸Садыков Т. Кочконбаева Б. Об оптимизации алгоритма морфологического анализа / Шестая Международная конференция по компьютерной обработке тюркских языков «Turklang-2018». (Труды конференции) – Ташкент: Издательско-полиграфический дом «Навоий университети», 2018. – 320 с. – С. 219-225.

бўлганлигини аниқлаши керак. Асосга *-лар* қўшимчаси қўшилиши натижасида қирғиз тилидаги сингармонизм ҳодисаси *бала* асосидаги охириги ҳарфнинг тушиб қолишига олиб келган. Бундай ҳолатда анализатор сўзни леммалаш, қўшимчадан олдинги асосни автоматик аниқлаши лозим. Бунда асослар луғати қидируви ишга туширилади. *Китобим* сўзшаклини таҳлил қилишда морфологик анализатор унинг *китоб* сўзига *-им* эгалик келишиги шакли қўшилганлигини аниқлаши лозим: *китоб* асос эканлиги аниқланади. Қирғиз тилида *китебим* эгалик шаклидаги сўз *китеп* ўзагига эгалик қўшимчаси қўшилганда *-им* қўшимчаси таъсирида жарангсиз *п* жарангли *б* товушига айланади. Бундай ҳолатда асослар луғати қўшимча қоидалар билан таъминланади. Ўзбек тилида фонетик ўзгаришларнинг катта миқдорни ташкил этмаганлиги морфологик таҳлил жараёнида алгоритмлар босқичини қисқартириши мумкин. Аммо фонетик ўзгариш учрайдиган ҳолатда худди қирғиз тили каби ўзгаришга учраган сўзнинг ўзагини аниқлаш жараёни 2-3-босқичли алгоритмда амалга оширилади.

Ўзбек тили морфологик анализатори тузишнинг лингвистик таъминотини ишлаб чиқиш тамойиллари орасида мана шу қоидалар блоки – лингвистик модулни шакллантиришнинг асосий вазифа саналади. Ишнинг кейинги бобларида ўзбек тилига хос бундай қонуниятлар тавсифланади, шу асосда ўзбек тили морфологик анализатори учун лингвистик таъминот ишлаб чиқилади.

Морфологик анализаторда сўзшакл таҳлилини амалга оширишнинг икки: “ўнгдан чапга”, “чапдан ўнгга” қоидаси ишлайди. Биринчи ёндашувда сўзшаклнинг охириги қисмидаги маъноли қисм – аффиксал морфемани; ундан кейин қўшимчадан ташқари ўнг томондаги қисмнинг луғатда мавжудлигини аниқлашга асосланилади. Иккинчи ёндашувда луғатдаги занжирнинг маълум бир бошланғич қисмини топишга уриниш амалга оширилади, сўнгра қолган ўнг қисми шу асос учун мумкин бўлган қўшимчалар мажмуаини ташкил этганлиги текширилади. Иккала ёндашувда ҳам, агар қидирув натижа бермаса,

сўзшаклни бошқа ҳолатда “қайта бўлиш” жараёни кечади. Табиийки, морфологик таҳлилда энг узок давом этадиган операция луғатдан маълумотни аниқлаш ҳисобланади. Аниқроқ айтганда, бунда узок кечадиган жараён – дискдаги хотирага мурожаат этиш босқичи. Шунинг учун иккала ёндашувда ҳам асосий вазифа – сўзшаклни таҳлил қилишнинг тўлиқ циклида диск хотирасига мурожаат сонини камайтириш. Бу вазифани муваффақиятли бажаришга морфологик анализаторни тузиш жараёнида қўшимчалар парадигмаси ҳамда қоидалар мажмуини тўлиқ шакллантириш билан эришиш мумкин.

Илмий-техник матндаги сўзнинг ўртача узунлиги тахминан 7-10 ҳарфни ташкил этганлиги боис “чапдан ўнгга” ёндашувида луғатга 10 мартагача мурожаат этиш талаб қилиниши мумкин. “Ўнгдан чапга” ёндашувида луғатга камроқ мурожаат қилинади. Шу сабабли “ўнгдан чапга” ёндашуви асосидаги морфологик таҳлил усулини танлаш мақсадга мувофиқ. Бироқ “ўнгдан чапга” ёндашувига нисбатан “чапдан ўнгга” ёндашуви бир қатор қулайлик ва табиий имконият жиҳатидан афзалликка эга. Шундай қилиб, дискка мурожаат қилиш сонини максимал даражада камайтириш вазифаси икки барабар муҳим, чунки унинг ечими бир қатор афзалликка эга бўлган морфологик таҳлил муаммосида “чапдан ўнгга” ёндашувини самарали амалга оширишга замин яратади.

2.1.2. Морфологик маълумотлар базасини яратиш хусусида

Маълумотлар базаси деганда фойдаланувчи эҳтиёжи/талабига жавоб берадиган кенг кўламли маълумотлар мажмуи назарда тутилади. Ушбу маълумотлар сараланади; маълумотлар базасини бошқариш тизими томонидан бошқариладиган жадвал шаклида сақланади. Ҳозирги вақтда маълумотлар базаси билан ишлашга мўлжалланган кўплаб тизим мавжуд: SQL, MySQL, Oracle, Access. Катта ҳажмдаги маълумотлар билан ишлаш доим ҳам қийин кечади; ваҳоланки, ҳар бир тизимнинг ўзига хос афзаллиги мавжуд. Морфологик таҳлилни амалга ошириш учун, табиий тилни синаб кўриш ва таклиф қилинган алгоритмни текшириш мақсадида Embarcadero

RAD Studio синов дастури яратилди; ушбу анализаторда Ассесс маълумотлар базаси билан ишлаш тизимидан фойдаланилган. Мазкур анализатор маълумотлар базасининг уч жадвали (асос, кўшимча ва сўз туркуми) ҳамда ушбу жадваллар ўртасидаги муносабатдан иборат.

Морфологик таҳлил алгоритмини оптималлаштириш. Қирғиз тили морфологик анализатор(МА)ида тизим иш босқичларни оптималлаштириш ёки дискка мурожаат қилиш сонини камайтириш учун алгоритмдан фойдаланилиши мақсадга мувофиқ бўлган:

- 1) маълумотлар базасидан “чапдан ўнгга” ёндашуви ёрдамида дастлабки иккита ҳарф таҳлил қилинадиган кириш сўзининг дастлабки икки ҳарфига тўғри келадиган сўзлар танланган;
- 2) маълумотларни сақлаш учун виртуал массив яратилган;
- 3) сўзнинг асоси “ўнгдан чапга” усули билан топилган;
- 4) кўмакчи морфеманинг грамматик категориясини аниқлаш учун маълумотлар базасидаги жадвалидан фойдаланилган.

Мукамал морфологик анализатор яратиш қийин. Шундай бўлса-да, морфологик анализатор тузиш учун қуйидаги алгоритмни амалга ошириш нисбатан аниқ таҳлилни амалга оширувчи анализатор яратишга асос бўла олади:

- 1) сўзнинг грамматик шакли фрейм-моделини яратиш. Бунда кўмакчи морфема категориясининг тўғри аниқланиши учун туркумга оид кўшимчалар жадвали тузиш назарда тутилади;
- 2) анализатор тузишда компьютер хотирасига мурожаат сонини камайтиришга мўлжалланган алгоритмлар ишлаб чиқиш;
- 3) синов дастури ишлаб чиқиш; натижаларни унда тажрибадан ўтказиш.

Морфоанализда формал грамматика ёки морфологик қонуниятлар мажмуи. Тадқиқотчи К.Шадманованинг таъкидлашича, тизимли дастурнинг назарий асосларидан бири – формал тил назарияси. Бу назария ўзининг математик тушунчалари билан оддий дастурчига мураккаб туюлади, лекин

тизимли дастур шундай назария асосида яратилади. Ҳар қандай дастурий интерфейс бирор-бир формализмга, формализм эса формал тил назариясига таянади. Умуман олганда, ихтиёрий тилда (форма ёки табиий) тузилган гап мантиқан тўғри бўлиши учун у қуйидаги шартларни қаноатлантириши керак¹²⁹:

- 1) сўзлар алифбо талабига мос ёзилган бўлиши (морфология, лексика);
- 2) гап қурилишида грамматик хатога йўл қўймаслик (грамматик, синтактик);
- 3) гап мазмуний қурилиши (семантика)нинг тўғрилиги.

Табиий тилда морфологик/синтактик хато гап ҳам маъно билдиради, яъни, тилни тўла ўрганмасдан бу тилда фикр билдириш мумкин. Лекин формал тилда бундай эмас. Синтактик жиҳатдан тўғри бўлмаган дастурни бажаришга ҳожат йўқ. Семантик (контекстли) таҳлил синтактик таҳлилдан кейин ишлайди, лекин семантик таҳлилнинг қандай тартибда ишлаши грамматика қоидалари, шажара дарахти, белги (символ)лар жадвали асосида аниқланади. Шажара дарахти ҳамда символлар жадвалининг семантик таҳлил учун керакли қисмни синтактик, лексик таҳлил тайёрлаб беради. Айтиш мумкинки, тил синтаксиси таҳлил жараёнини бошқаради. Синтактик бошқариладиган трансляция жараёнида семантик таҳлил, асосан, қуйидаги икки ёндашув асосида амалга оширилади:

- 1) грамматик продукцияга (атрибут) ҳисоблаш қоидаларини боғлаш;
- 2) грамматик продукцияга дастур фрагментини боғлаш.

Синтактик бошқариладиган трансляцияда семантик таҳлилни амалга ошириш “атрибут ечими” ёки “трансляция схемаси” концепциялари асосида амалга оширилиши мумкин.

Лингвистик база ҳақида. С.Муҳаммедованинг ишида¹³⁰да ҳаракат феълларининг семантикаси, валентлигини текшириш асосида компьютер

¹²⁹Шадманова К. Табиий тиллар учун лексико-семантик луғат маълумотлар базаси яратиш тамойиллари // <http://buxdu.uz/index.php/uz/>

¹³⁰Муҳамедова С. Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. – Тошкент, 2006.

дастурлари учун лингвистик таъминот яратиш намунасини келтирилган. Лингвистик таъмин яратишда тадқиқотчи ҳаракат феълларининг тайёр семантик таснифдан фойдаланади. Объектив борликдаги нарса-предмет, шахс ҳамда ҳайвонга хос турли фаол ҳаракатларни ифодаловчи феъллар нутқда жуда кўп ишлатилиши ва мураккаб семантик тузилиши билан фарқланиб туради. Ҳаракат феъллари семантикаси муштарак семаси “фаол ҳаракат”, у феълни бир макросистемага (семантик майдонга) бирлаштиради. Яъни муштарак семантик характеридаги “фаол ҳаракат” семаси асосида феълларнинг семантик майдони юзага келади¹³¹. Семантик таснифда ҳаракат феъллари семасидан фақат аташ (денотатив) семаси аниқланади. К. Шадманованинг таъкидлашича, феъл семаси таркибидаги ифода (коннотатив) ва бошқа семаларни текшириш алоҳида тадқиқотни талаб қилади, феъл асосида яратиладиган компьютер дастурлари учун лингвистик таъмин ишлаб чиқишда ҳозирча уларнинг кўчма маънолари ҳисобга олинмайди.

ЛСГларда *динамик (фаол) ҳаракат муштарак семантик, горизонтал, вертикал, айланма, тебранма, предметнинг тик ҳолатдан бошқа (ётиқ) ҳолатга ўтишини билдирувчи, нафас ва товуш чиқариш ҳаракатини ифодаловчи, натижали ижро, натижаси мавҳум ижро, оғиз бўшлигидаги ҳаракат* каби бирликлар эса **бирлаштирувчи** сема ҳисобланади. Ўзбек тилидаги ҳаракат феълларида намоён бўлувчи мантиқий субъект ҳаракати юқоридаги кўринишда содир бўлади. Демак, объектив борликдаги турли ҳаракатни акс эттирувчи феъл семаси мураккаб характери мантиқий равишда ҳаракат феълларининг қайд этилган кўринишдан иборат семантик гуруҳланишга олиб келди.

Ўзбек тилидаги ҳаракат семантик феъллар асосидаги мазкур маълумотлар базаси компьютер дастури, таржима, таҳрир учун

¹³¹Кузнецов А.М. О применении метода компонентного анализа в лексике. – С.260.; Цветков Н.В. К методологии компонентного анализа // Вопросы языкознания. –1984, №2, – с.67.

мўлжалланган. Маълумотлар базаси асосида яратилган дастурдан, шунингдек, ўзбек, инглиз тилидаги ҳаракат семантик феълларининг тусланишини ўрганмоқчи бўлган тадқиқотчи; ўзбек, рус, инглиз тиллари бўйича тилшунос мутахассис; лингвистика соҳасидаги дастурчилар фойдаланиши мумкин.

2.2. Морфологик анализаторнинг ахборот-қидирув тизимидаги аҳамияти

2.1.1. Матнга ишлов бериш ва таҳлил қилиш дастурлари шарҳи.

Жаҳон тажрибасида матнга ишлов бериш ва таҳлил қилишга мўлжалланган турли функциялари билан бир-биридан фарқ қилувчи кўп сонли дастурий маҳсулотлар мавжуд. Қуйида уларнинг морфологик таҳлилга хосланган айримлари ҳақида маълумот келтираамиз.

1. *Лингвистик технология ва тизимлар. AskNet (ишлаб чиқувчи ПОЛИГЛОТ)*¹³² – AskNet семантик қидирув тизими рус, инглиз тилидаги матнни тўлиқ лингвистик таҳлил қилувчи дастур. Лингвистик таҳлил модули морфология, синтаксис, семантика(комбинатор-изоҳли луғатлар)ни ўз ичига олган. Лингвистик таҳлил даражаси: графематик, морфологик, синтактик, семантик.

*Russian Context Optimizer (ишлаб чиқувчи RCO)*¹³³ – рус тили матнларини морфологик ва синтактик таҳлил қилувчи комплекс дастур. Лингвистик таҳлил модуллари морфологик луғат, графемалар луғати, синтактик қоидалар, семантик классификатор, синонимлар луғатидан ташкил топган. Морфологик луғат 115000 лексемадан иборат. Дастур Windows, Oracle, SDKлар учун ишлаб чиқилган; Windows ва Unix платформаларида ишлайди. Дастурни тижорий йўл билан олиш мумкин. Лингвистик таҳлил даражаси: графематик, морфологик, синтактик, семантик.

*Google Mini / Google Desktop*¹³⁴(ишлаб чиқувчи Google) – Google Mini дастурий-аппаратли комплекси ва Google Desktop дастури корпоратив локал

¹³²<http://asknet.ru/analytics/programms.htm>

¹³³ <http://asknet.ru/analytics/programms.htm>

¹³⁴<http://asknet.ru/analytics/programms.htm>

тармоқ ва шахсий фойдаланувчи ишлатилиши учун мўлжалланган ахборот қидирув дастури саналади. Асосий тиллар (рус, инглиз ва б. тиллар) учун калит сўз асосида қидирувни йўлга қўя олади. Лингвистик таҳлил даражаси: графематик, морфологик.

Кластреловчи қидирув тизимлар(ишлаб чиқувчи Vivisimo, Nigma)¹³⁵ – натижаларни кластерловчи метақидирув тизими; қидирув тизимида кўп учрайдиган сўров (сўзшакл)ларни ажратиб кўрсатувчи тизим. Қидирув сифатини ошириш учун синтактик ва семантик таҳлилни киритиш лозим. Лингвистик таҳлил даражаси: графематик, морфологик.

Қидирув натижасини визуализацияловчи кластерловчи тизим (ишлаб чиқувчи Quintura Searchcrystal)¹³⁶ – натижаларни график кўринишда тақдим қилувчи метақидирув тизими. Қидирув бир неча тизимлар асосида амалга оширилади, натижа статистик мезон асосида кластерланади, Quintura сўз визуал кластери учун морфологик таҳлилдан фойдаланилади. Унда семантик ва синтактик таҳлил мавжуд эмас. Масалан, *юувчи* сўзи сўралганда кластерга *ажойиб энага* бирикмасини қўшади, синоним сифатида *менинг, юувчи* сўзшакллари келтиради.

Galaktika-ZOOM (ишлаб чиқувчи Галактика), Арион (ишлаб чиқувчи SyTech) дастурлари ҳам шундай тизимлар сирасига киради, уларнинг лингвистик таҳлил даражаси ҳам юқоридаги дастурларга ўхшаш: графематик, морфологик.

Юқорида тавсифланган лингвистик таҳлил воситаларининг умумий белгиси сифатида графематик, морфологик таҳлил имкониятини таъкидлаш жоиз. Ушбу анализаторлар таҳлил даражаси, тилга мослашганлик белгиси билан фарқ қилади. Буларнинг айримида семантик/синтактик таҳлил функцияси ҳам мавжуд. Айримида ушбу функциянинг етишмаслиги (имконияти чекланганлиги) берилган маълумотда акс этган. Демак, битта

¹³⁵Ўша манба.

¹³⁶<http://asknet.ru/analytics/programms.htm>

таҳлил дастурига бир неча функцияни киритиш дастур афзаллигини оширади, шу сабабли ўзбек тили морфоанализаторини ишлаб чиқишда таҳлилнинг бир неча босқичини баравар камраб олишга алоҳида эътибор қаратиш лозим. Шунда яратиладиган ўзбек тили морфоанализатори турли лингвистик таҳлилларни бажарадиган полифункционал анализаторга айланади.

2. *Матнни лингвистик (морфологик, синтактик) таҳлил қилиш дастурлари.* Матнни лингвистик таҳлил қилиш дастурлари таҳлил даражасига кўра графематик, морфологик, синтактик, семантик бўлиши юқорида айтиб ўтилди. Қуйида синтактик таҳлил дастурлари ҳақида умумий маълумот келтирамиз¹³⁷:

1. *Link Grammar Parser* (муаллифлар: John Lafferty, Daniel Sleator, Davy Temperley, Carnegie Mellon University – USA) – инглиз тили синтактик таҳлил дастури.

2. *Link Grammar Parser for Russian* (муаллиф Сергей Протасов) – рус тили синтактик таҳлил дастури.

3. *Синтактик анализатор* (муаллиф Д.С.Ермолаев) – рус тили синтактик таҳлил дастури.

4. *Forson, a sentence generation tool* (муаллиф Alfonso Tarantini) – гапларнинг синтактик тўғри қурилишини генерациялаш дастури.

Ушбу дастурларнинг муштарак томони уларнинг синтактик таҳлил воситаси эканлигида кўринади; фарқли жиҳати сифатида турли, хусусан, инглиз, рус тилларини таҳлил қилишга мўлжалланганлини таъкидлаш мумкин. Қуйида морфологик таҳлилга мўлжалланган дастурлар ҳақида нисбатан кенгроқ маълумот келтирамиз.

1. *АОТ (автоматическая обработка текста) матнга автоматик ишлов бериш дастури* (муаллиф Алексей Сокирко) – матнга автоматик ишлов бериш мажмуи қуйидаги таркибий қисмлардан ташкил топган¹³⁸:

¹³⁷<http://asknet.ru/analytics/programms.htm>

¹³⁸<http://asknet.ru/analytics/programms.htm>

- 1) рус, англиз, немис тилларининг морфологик луғати;
- 2) матнни графематик таҳлил қилиш модули;
- 3) синтактик луғат;
- 4) омонимияни аниқлаш ва бартараф этиш модули;
- 5) дастлабки семантик таҳлил модули;
- 6) лингвистик қидирув тизими (конкорданс)
- 7) турли тезаурус ва сўзликлар.

2. *Mystem* (муаллифлар Илья Сегалович, Виталий Титов (Яndex)¹³⁹ – рус тили морфологик таҳлили дастури. Луғатда мавжуд бўлмаган сўзлар учун ишлаб чиқилган гипотезалар шу асосда ишлайди.

3. *LingSoft* (муаллиф LingSoft)¹⁴⁰ – англиз, немис, фин, дат, норвег, швед, эстон тилларидаги матнни графематик, морфологик таҳлил ва лемматизациясига мўлжалланган модул; шунингдек, “нутқни тушуниш” (распознавания речи) имконияти ҳам мавжуд. Ушбу дастур бошқа лингвистик тизимлар ишлаб чиқиш учун очик (open/free).

4. *Cibola/Oleada* (муаллиф Computing Research Laboratory)¹⁴¹ – стемминг ва автоматик морфология асосида Unicode тизимида мавжуд турли тилдаги матнлар (16 та тил учун) морфологик таҳлилни амалга оширувчи дастур. Модулдан матнни статистик таҳлил қилиш, автоматик таржимада фойдаланиш, шунингдек, луғат ва тезаурус сифатида ишлатиш мумкин; синтактик, семантик таҳлил йўлга қўйилмаган.

5. *StarLing* (муаллиф С.А.Старостин)¹⁴² – англиз ва рус тиллари морфологик анализатори.

6. *Russian Morphological Dictionary* (рус тили морфологик луғати) (муаллиф С.Цикорский)¹⁴³ – рус тили матнларининг морфологик ва синтактик

¹³⁹<http://asknet.ru/analytics/programms.htm>

¹⁴⁰Ўша манба.

¹⁴¹Ўша манба.

¹⁴²Ўша манба.

¹⁴³Ўша манба.

таҳлил дастури. ASCII форматдаги матнлар ҳамда 12000 сўздан иборат морфологик луғат билан ишлайди.

7. *MCR DLL v2.0* (муаллиф Dim)¹⁴⁴ – морфологик таҳлил ва лемматизацияни амалга оширувчи dll кутубхона шаклидаги морфологик модул. А.Зализняк луғати асосида шакллантирилган морфологик луғат базасида ишлайди. Модул MS Visual C++ да ёзилган, сақлаб олиш ва фойдаланиш мумкин.

8. *RHPMorphy*¹⁴⁵ – PHP платформасидан жой олган, бепул фойдаланиш мумкин бўлган морфологик таҳлил кутубхонаси. *RHPMorphy* куйидаги амалларни бажаришга мўлжалланган:

- 1) лемматизация (сўзшаклнинг дастлабки шакли, лексема шаклини аниқлаш);
- 2) сўзнинг барча шакллари аниқлаш;
- 3) сўзга тегишли барча грамматик маълумотларни (сўз туркуми, келишик, турланиш ва ҳ.) йиғиш;
- 4) берилган грамматик хусусиятга мувофиқ сўз шаклини ўзгартириш;
- 5) сўз шаклини берилган намунага мувофиқ ўзгартириш.

Морфологик анализатордан рус, инглиз, немис, *ispell* асосида украин, эстон тиллари таҳлилида фойдаланиш мумкин. *Myspell* луғати ёрдамида бошқа тиллар таҳлили учун контент ҳосил қилиш имконияти ҳам мавжуд.

Лемматизатор (муаллиф Мультитран)¹⁴⁶ – мультитран тизими морфологик модули имкониятларини юзага чиқарувчи лемматизатор. Рус тилида 300.000 ўзак, 15.000 қўшимчани қамраб олган тизимли луғат базаси мавжуд. Барча сўзшаклнинг грамматик шакллари билан қидирувни амалга ошира олади.

*MyStem*¹⁴⁷ – рус тилининг морфологик анализатори, анализаторда морфологик кўпмаънолилиқ бартараф этилган. Дастур «Яндекс» компанияда

¹⁴⁴<http://asknet.ru/analytics/programms.htm>

¹⁴⁵<https://novainfo.ru/article/3394>

¹⁴⁶<https://www.multitran.com/c/m.exe?a=4&MessNum=8673&topic=2&l1=4&l2=2>

¹⁴⁷<https://yandex.ru/dev/mystem/>

дастлаб Илья Сегалович, В.Титов томонидан ишлаб чиқилган. Муаллифлик ҳуқуқи фақат «ЯНДЕКС»га тегишли. Дастур луғат асосида ишлайди, нотаниш сўзлар тўғрисида морфологик гипотезани шакллантира олади; ундан нотижорий мақсадда фойдаланиш мумкин¹⁴⁸.

Юқоридаги таҳлил дастурларининг умумий жиҳати морфологик таҳлилни амалга ошириш, лемматизация, сўзшаклни аниқлаш, сўзга тегишли грамматик маълумот бириктириш хусусиятининг мавжудлигида. Шу билан бирга, уларнинг кўпчилиги бир неча тилни таҳлил қилишга мўлжалланган: битта анализатордан инглиз, рус, немис тилини таҳлил қилишда фойдаланиш мумкин. Демак, анализатор бир неча тил таҳлилиги мослаб тузилиши мумкин. Шунингдек, бир анализатор базасида бошқа бир тилнинг анализаторини тузиш имкони ҳам мавжуд. Юқорида тавсифланган МАларнинг фарқли томони шундаки, уларнинг ҳар бири ўзига хос имкониятга эга: айрими статистик таҳлилни, айрими семантик таҳлилни амалга оширади. Айтиш мумкинки, морфоанализаторнинг асосий вазифаси сўзни морфологик таҳлил қилиш бўлса-да, унга қўшимча вазифа киритиш мумкин; анализаторга турли функциялар қўшиш муаллиф имконияти ҳамда мақсадидан келиб чиқади.

5. *Матннинг статистик таҳлили воситалари. TextAnalyst 2.0*(ишлаб чиқувчи МикроСистемы) – ишлов берилаётган матндан ажратиб кўрсатилган тушунчаларнинг семантик тармоқ дастури; ҳар бир бирлик контекстга ҳавола қилинади. Таҳлил қилинаётган матндаги асосий ва иккинчи даражали мавзуларнинг иерархик дараҳтини тузиш орқали таҳлилни амалга оширади; шунингдек, матнни рефератлаш функцияси ҳам мавжуд. Морфологик таҳлил нисбатан кам сонли сўзлардагина амалга оширилади. Қидирувда сўз тартиби аҳамиятсиз. Матннинг статистик, семантик таҳлили йўлга қўйилмаган. SDK рус ва инглиз тиллари учун лемматизация функциясини юзага чиқара олади:

¹⁴⁸Морфологический анализатор текста на русском языке mystem [Электронный ресурс]; // Компания Яндекс [сайт] – 2003–2013. – URL: <http://company.yandex.ru/technologies/mystem/>

тушунчанинг частотали статистикаси, контекстда сўз қидириш имконияти мавжуд.

Лингвистик компонент (муаллиф Андрей Коваленко)¹⁴⁹ – эҳтимолий морфологик анализ дастури – стеммер ҳамда рус, украин, инглиз тили морфологик анализатори демо-версияси. Ишланма Апорт!, Рамблер, Мета, системе Пропись 4.0 каби ахборот-қидирув дастурларида фойдаланилади.

WordTabulator v2.2.3 (муаллиф С.В.Логичев)¹⁵⁰– ASCII ва HTML форматдаги матндаги сўзшакл ва сўз бирикмаларининг тизимлаштирилган индексини тузишга мўлжалланган дастур. Рус тили морфологияси коидалари асосида белгилар кетма-кетлиги билан қидирувни амалга оширади. Натижани гиперматнли индекс ёрдамида контекстда қўриш, икки корпуснинг ўхшаш ва фарқли томонларини қиёслашда қўл келади. Windows 9x/NT/2000/XP кирилл кодировкасида ишлайди.

Ушбу дастурлар орасидан айримининг тажрибаси билан чуқурроқ танишдик. Шундан келиб чиқиб, уларнинг имкониятлари тўғрисида қуйида батафсил тўхталамиз.

2.2.2. Ахборот базаси ва таҳлил дастурларининг амалий аҳамияти.

Тўлиқ матнли корпуслар билан ишлаш учун “Корпусни бошқариш тизими” ишлаб чиқилган, олимлар “ДикТУМ-1” дастурий воситасини корпус лингвистикаси воситалари мажмуи сифатида баҳолашади¹⁵¹. Корпус бирликларини таҳлил қилиш ва бошқариш “ДикТУМ-1” тизими билан амалга оширилади. Унинг асосий афзаллиги матн ва луғатлардан олинган маълумотни манипуляция қилади, таҳлил ўтказади, лексикографик ишловни амалга оширади. Тизим автоматик равишда матн/бирликни материал манбаи, унинг ҳажми, жанри, нашр санасини маркер билан белгилайди. Сўзшаклни грамматик, лексик, морфем ва бошқа категориялар асосида маркерлайди.

¹⁴⁹ <http://asknet.ru/analytics/programms.htm>

¹⁵⁰ <https://www.shouldiremoveit.com/wordtabulator-for-windows-5636-program.aspx>

¹⁵¹Амиева А. М., Филимонов В. В., Сергеев А.П., Трасов Д.А. Инструменты корпусной лингвистики //

Табиий тилга ишлов беришда автоматик маркерлаш/таҳлил қилиш дастурларининг амалий аҳамияти кенг. Масалан, шундай тизим ёрдамида “Газета матнлари корпуси” материаллари асосида газета матнлари жанрий таснифи тамойиллари (жумладан, турли публицистик жанрларнинг типик характеристикаси) ишлаб чиқилади; шу асосда жанрий тасниф амалга оширилади. Бу эса маълум матн, турли жанр, манбаларнинг тақсимланишини кўрсатувчи маълум бир профилни топишга, матнда бирликни қўллашнинг ўзига хослигини кузатишга шароит яратади. Бундай амалиётнинг ўтказилиши автоматик равишда сўзшакл, лемма, асос ва морфем моделларнинг алфавитли-частотали, частотали-тақсимловчи луғатларнинг шаклланишига олиб келади¹⁵². Шунингдек, ДикТУМ-1 тизимининг ички системалари ҳам мавжуд:

1. *Морфо-Анализатор* лексемаларни морфологик таҳлил қилишга мослашган.
2. *СегмМорф* сўзшаклларни морфологик таҳлил қилади.
3. *Линтез* матн бирликларини семантик маркировка қилади; матннинг мазмуний таҳлилини амалга оширади.
4. *ПсиАн* семантик-психологик таҳлил қилади.
5. *Линтез* калит сўз ва бирикмаларни танлайди, автоматик тарзда матннинг тезаурусини шакллантиради.
6. *Жанр* квантатив-стилистик таҳлил олиб боради.

Грамматик категориялар тўлиқ тавсифининг морфоанализатор тузишдаги аҳамияти. “RCO Morphology” морфологик анализатори тузувчиси А.Ермаков морфологик анализаторнинг аҳамияти ҳақида шундай фикрларни келтиради¹⁵³: “замонавий дунё ҳаддан ташқари кўп маълумотлар билан тўлмоқда. Айниқса, интернет материалларининг доимий ўсишини ҳисобга олсак, сўзларни қайта ишлаш масаласи жуда долзарб бўлиб қолмоқда”. А.Ермаков морфологик анализаторнинг амалий аҳамияти сифатида ахборот-

¹⁵²Ўша манба.

¹⁵³Ермаков А. Морфологический анализатор - основа поисковых систем //

<https://www.kv.by/archive/index2004154301.htm>

кидирув тизимларида сўзшаклнинг аниқ анализ-синтезини амалга ошириш эканлигини таъкидлайди. “RCO Morphology” морфологик анализаторида анализатор жойлашган интернет саҳифасида¹⁵⁴ турли таҳлилларни амалга ошириш мумкин. Бунда амалий аҳамиятга эришиш мақсадида “RCO Morphology” анализаторида алгоритмларнинг яна бир тоифаси ишлаб чиқилган, кутубхонага киритилган, у сўз кидирувини шакллантириш ва киритиш қоидаларига асосланган; сўзларни жуда юқори даражада ишончлилиқ билан таҳлил қилиш, мавжуд алгоритмларни мукамаллаштиришни таъминлайди, бу эса матндаги сўз табиати ҳақида турли фаразларни яратишга имкон беради. Икки босқичда алгоритмларни такомиллаштириш билан бир қаторда, морфологик луғатни доимий равишда тузатиш ва кенгайтириш иши олиб борилган. Бугунги кунда морфологик анализаторнинг аниқ ишлаши учун морфологик луғатда тахминан 100.000 сўз (шунингдек, тахминан 3-4 миллион сўз шакли) бўлиши керак¹⁵⁵. Натижада луғат доимий равишда янгиланиб, тўлдирилиб борилади. Масалан, “RCO Morphology” анализатори маълумотлар базасини тўлдириш ва тузатиш давомида фақат феъллар орасида 3000дан ортиқ сўз шакллари ўзгартирилган.

Морфологик анализатор тузишда асосий иш лингвистик қоидалар тизимини кенгайтириш, созлаш, шунингдек, луғатдаги хатони бартараф этишда кўп меҳнат ресурси талаб қилинади. МАга тил лексикасининг юқори частотали сўзлари, тилда мавжуд бўлган ҳар қандай сўзни қайта ишлашга оид қоидалар мажмуини ишлатадиган алгоритм киритилиши керак. Бундан ташқари, морфоанализаторда турли тахминларни инобатга олган ҳолда сўз таҳлили ҳақидаги фаразларни яратишга имкон берадиган мослашувчан созлама бўлиши талаб этилади. Бу эса модулни сўз грамматик шаклининг имло билан бир-бирига тўғри келганда омонимияни бартараф этиш

¹⁵⁴www.rco.ru/product.asp.

¹⁵⁵Ермаков А. Морфологический анализатор - основа поисковых систем // <https://www.kv.by/archive/index2004154301.htm>

воситалари билан бирга ишлатишга имкон беради. RCO Morphology анализатори Windows учун динамик кутубхона (DLL) сифатида етказиб берилади. Ушбу анализатор сўзни юқори аниқликда қайта ишлашни талаб қиладиган ахборот-қидирув ҳамда таҳлил тизими ишлаб чиқувчиларига мўлжалланган. Модул қуйидаги афзалликка эга: (1) сўз шаклнинг барча грамматик хусусиятини (сўз туркуми, грамматик маънолари ва ҳ.); (2) лексик-семантик тоифани (исм, фамилия, ташкилот, географик ном) аниқлаш, (3) сўзнинг турли грамматик шакллари меъёрий ҳолатга келтириш, (4) барча грамматик шаклларни аниқлаш орқали имлони текшириш. Аммо МА 3-4 миллион сўзшаклдан иборат луғат/маълумотлар базаси каби бундай талаб фақат флектив тил анализаторларига хос. Агглютинатив тилларда аниқ таҳлилни амалга ошириш учун “ўзак ва қўшимчалар луғати”, ўзак ва қўшимчалар қўшилиши натижасида ҳосил бўладиган фонетик ўзгариш, умумий қонуниятлар “грамматик луғати” талаб этилади. Агглютинатив тилларнинг морфологик анализатори учун 3-4 миллион сўзшаклни маълумотлар базасига киритиш ёки сўзшакллар луғатини тузишга эҳтиёж йўқ.

2.2.3. Славян тиллари морфоанализаторлари: тадқиқ ва кузатувлар

1. *“Рутения” платформасининг “Starling” морфологик анализатори ҳақида.* Маълумки, жаҳон тажрибасида бир қатор тилларнинг морфологик анализатори ишлаб чиқилган¹⁵⁶. Тарту университети (Эстония) рус тили ва адабиёти кафедраси 2000 йилда “Рутения” платформасига асос солган. Ушбу лойиҳанинг асосий мақсади рус тили мутахассиси ҳамда тарихчиларга ахборот-қидирув тизими вазифасини ўтайдиган маълумотлар базасини яратиш бўлган. Кейинчалик лойиҳа кенгайтирилиб, бир неча йўналишда ривожлантирилганки, бу йўналишдан бири морфологик анализатор яратиш бўлган. Бугунги кунда морфологик анализатор¹⁵⁷ on-line форматга эга: морфологик анализатор <http://starling.rinet.ru/morph.htm> манзилида ишлайди,

¹⁵⁶<http://www.ruthenia.ru/document/192271.html>

¹⁵⁷<http://starling.rinet.ru/morph.htm>

рус тилидаги исталган сўзни морфологик таҳлил қилиб беради. Рус тилида сўз қўллашни ўрганиш учун қулайлиги бўйича барча турдаги дидактик воситалардан афзаллик касб этади. Қуйида ушбу морфологик анализаторнинг қисқача тавсифини келтириб ўтамыз.

1. Минимал дизайнга эга бўлган (сўз киритиш ва уни тасдиқлаш ойнасидангина иборат) ушбу анализатор маълумотлар базаси С.Ожегов изоҳли луғати¹⁵⁸, А.Зализняк грамматик луғати¹⁵⁹ ҳамда Мюллер луғати¹⁶⁰ни ўзида сақлайди. Дастурга бириктирилган қўллаш бўйича йўриқнома (тушунтириш хати)да маълумотлар омбори билан ишлаш тамойиллари, маълумотларнинг ифодаланиш формати изоҳланган. Муаллифнинг ёзишича, ушбу дастур орқали “исталган рус тилидаги сўзни таҳлил қилиш, унинг тўла шакллантирилган парадигмасини олиш мумкин. Аслида, ҳамма сўзни ҳам таҳлил қилиб бўлмайди. Шеършунос Ш.Семенов таъкидлашича, морфологик анализаторнинг маълумотлар базасида *престижизитатор, занудствовать, медитировать* каби сўзларни топа олмаган. Морфологик анализатор кам ишлатиладиган, янги ўзлашган сўзларнинг таҳлилида хато натижа беради. Ушбу айрим жузъий хато натижаларни ҳисобга олмаганда, анализатор нисбатан тўлиқроқ таҳлилни амалга ошира олади. Масалан, *морфологический* сўзи сўровида жавоб сифатида сўзнинг бош шакли “морфологический” сўзи чиқади: А.Зализняк луғатидаги индекси “п За!~” шаклига эга. Сўзнинг инглиз тилига таржимаси ойнасида Мюллер луғати мақоласидаги “morphologic(al)” сўзи берилса, морфологик характеристикаси – “Nsm, Asi”; кейин жадвал берилади: сўзнинг келишиқда турланиши, род бўйича парадигмаси, бирлик/кўплик сонда эканлигини ифодаловчи кўрсаткич индекс; сифатнинг кўплиқдаги қисқа шакли келтирилади: “морфологически”. Барча сўзшаклларнинг урғуси

¹⁵⁸ОжеговС. Словарь русского языка. – 1949, 22-е издание, 1990; с 1992 – "Толковый словарь русского языка", совместно с Н. Ю. Шведовой).

¹⁵⁹Грамматический словарь русского языка. Словоизменение. – А. А.Зализняк. – М., 1977.

¹⁶⁰Мюллер В.К. Большой современный англо-русский, русско-английский словарь. Около 450 000 слов, словосочетаний и идиоматических выражений. –Москва: Новая редакция, 2009.

кўрсатиб берилади (Қаранг: 1-расм). Ушбу анализаторнинг яна бир муваффақияти шундаки, у А.Зализняк, С.Ожегов, В.К.Мюллер луғатларининг оддий электрон такрори эмас. Зеро, ушбу луғатлар маълумотлар омборига айлантирилиб, луғатлардан керакли маълумотни ажратиб олишга эришилганлиги дастурнинг катта ютуғи саналади¹⁶¹. Ушбу анализаторни кузатишларимиз МА тузишда тил лексикасини камраб олган лексикон-луғат ҳамда грамматик луғатнинг талаб қилиниши; бундай луғатлар қоғоз-луғатнинг оддий электрон такрори эмаслиги; рус тилининг флектив табиатидан келиб чиққан ҳолда сўзшакллар парадигмасининг талаб қилиниши каби хулосага келишга имкон беради. Ўзбек тили морфоанализаторини тузишда анализаторнинг лингвистик таъминотидан лексикон/луғат, грамматик луғатнинг бўлиши талаб қилинса-да, сўзшакллар парадигмаси зарурий таркибий қисм эмас, чунки ўзбек тили агглютинатив табиатга эга. Лингвистик таъминотда ўзак ва қўшимчалар базасининг мавжудлиги анализаторнинг тўлиқ режимда ишлашини кафолатлайди.

Парсер – рус тилидаги матнларнинг морфологик ва синтактик анализатори. Рус тилидаги матнларни морфологик, синтактик таҳлил қилувчи ушбу парсер қуйидаги вазифаларни бажара олади¹⁶²:

- 1) лексик таҳлил – матнни гап ва сўзга ажратиш.
- 2) сўзнинг морфологик таҳлили – сўзшаклнинг контекстадаги позициясидан келиб чиқиб туркумини аниқлаш, род, келишик, сон ва бошқа грамматик хусусиятини аниқлаш.
- 3) лемматизация – сўзни бош шакл (лемма)га келтириш, яъни сўзшаклнинг қайси лексеманинг ҳосиласи эканлиги аниқлаш.
- 4) синтактик таҳлил (dependency parsing) – гаптадаги синтактик алоқаларни аниқлаш; эга, кесим, ва б.ларни қидириш.

¹⁶¹<http://www.ruthenia.ru/document/192271.html>

¹⁶²<http://www.solarix.ru/parser.shtml>

5) соддалаштирилган синтактик таҳлил (*chunking*) – гапни иккинчи даражали бўлак: тўлдирувчи, аниқловчи, ҳол гуруҳига ажратиш.

Юқорида саналган амалиётларнинг барчаси анализаторнинг маълумотлар омборидаги луғатда бўлмаган сўзлар устида ҳам бирдек бажарилади. Шунингдек, дастурга “нусхалаш” режими қўшилса, унда орфографик хатони кўрсатиш/ тузатиш функцияси ҳам пайдо бўлади.

Парсердан фойдаланиш. Парсернинг дастлабки варианты буйруқлар сатри билан ишлайди: кириш сифатида матнли файлни utf-8 га қабул қилади; натижалар XML форматида файл кўринишида сақланади ёки тўғридан-тўғри консолда намоёниш этилади.

Одатда, анализатор чакирув кетма-кетлиги қуйидагича амал қилади¹⁶³:

```
parser.exe -verbose -tagger 0 -parser 0 -lemmatizer 0 -emit_morph 0 -d dictionary.xml input.txt -o output.xml
```

Параметрлар:

- 1) *verbose* – синтактик таҳлил натижаларини консолга киритиш.
- 2) *tagger 0* – морфологик таҳлилни амалга ошириш учун (Part-Of-Speech tagging) рус ёки инглиз тили морфологиясининг асосий модели (онтологик модели¹⁶⁴) қўлланилади.
- 3) *parser 0* – синтактик таҳлилни амалга ошириш учун (шажара дарахтини тузиш), асосий shift-reduce базавий парсер ишлатилади.
- 4) *lemmatizer 0* – лемматизация рус тилининг эҳтимолий моделидан фойдаланган ҳамда сўз контекстини ҳисобга олган ҳолда амалга оширилади.
- 5) *emit_morph 0* – натижалар XML файлида морфологик тегларни чиқармаслик биринчи режим ҳажмининг сезиларли даражада ошишига олиб келади.
- 6) *d dictionary.xml* – луғат базаси конфигурация файлига ўтилади.

¹⁶³ <http://www.solarix.ru/parser.shtml>

¹⁶⁴Изоҳ бизники

7) *o output.xml* – натижалар билан яратилган файлнинг номи, одатда, чиқиш консолга ўтади.

Ishga tushirish намунасини дистрибутивдаги *run.cmd* файлидан топиш мумкин.

2. *Рус ва украин тилларининг морфологик анализатори*¹⁶⁵. Бугунги кунда нафақат интернет, балки турли қидирув тизимлари ҳам муҳокамаларга сабаб бўляпти. Турли ахборот воситаларидан тўлиқ матнли маълумотни қидириш, унинг афзаллик/камчиликлари кенг омма, алгоритм ва дастурлари мутахассисларнинг муҳокама мавзусига айланиб улгурди¹⁶⁶.

“Профи” қидирув ҳамда “ишёқмас фойдаланувчи” каби атамалар бот-бот кулоққа чалинмоқда. Биринчи – соф “механик” машина – қатъий мантикий сўровни амалга оширувчи қидирув тизими ҳамда сўзнинг ўнг томондан юзма-юз келиш бўйича қидирув тизими тарафдорлари; уларнинг фикрича, бундай қидирув тизими қидирилаётган/керакли ахборотни ҳар қандай алгоритмдан кўра аниқроқ топади. Бошқа гуруҳ вакиллари, аксинча, қидирувнинг барча талабларини кўринмас алгоритмга кўйиб беришади; ички жараёнлар улар учун қизиқ эмас. Иккала қараш ҳам илмий асосга эга: шундай “профи”лар борки, ҳақиқатан ҳам, мантикий сўровга ниҳоятда аниқ жавоб қайтара олади (ҳатто киритмани ажратиш учун қавсни ҳам тўғри жойлаштиради). Фойдаланувчиларнинг аксарияти “ишёқмас”лар гуруҳига кирадики, фикримизни аниқроқ ифодалаш учун 1995 йилда SofTool кўрғазмасида бўлиб ўтган иллюстратив диалогни келтириш кифоя¹⁶⁷.

“Агама” компанияси томонидан компьютер локал дискига жойлаштирилган ахборот қидирув тизими ишлаб чиқилган эди. Ушбу тизим ишлов берилаётган матннинг тўлиқ луғавий морфологик анализини амалга ошириш имконига эга, рус тилидаги сўзнинг барча шакли бўйича қидириш

¹⁶⁵ Системный администратор. Ежемесячный журнал www.samag.ru (Электрон манзил – <http://samag.ru/ru/articles/>)

¹⁶⁶<http://www.keva.ru/stemka/stemka.html>

¹⁶⁷<http://www.keva.ru/stemka/stemka.html>

амалиётини бажара олади. Ушбу тизим тест режимида синовдан ўтказилаётган вақтда кўргазмадаги бир мутахассис: “Бундай дастур кимга ва нега керак? Менда кесишиш оператори бўлса, истаган нарсани бундай тизимсиз ҳам аниқлашим мумкин”, мазмунидаги амалий аҳамиятга эга бўлган саволни ўртага ташлайди. Сўров тилида анъанавий мантикий сўров оператори ҳам, кесишиш оператори ҳам бор эди. Мутахассис ушбу воситалар ёрдамида бир канча вақт сўралган амалиётни бажаришга уринди, аммо анча вақтдан кейин морфологик анализатор имкониятлари оддий мантикий сўров тили операторида мавжуд эмаслигига амин бўлди. Шундай қаршиликлар туфайли “Рус ва украин тилларининг морфологик анализатори”ни лицензиялаш осон бўлмаган¹⁶⁸.

Қидирув дастурлари, асосан, икки тамойилдан бири асосида ишлайди:

- 1) имкон қадар лингвистик алгоритмларни қўлламаслик;
- 2) стеммингга асосланиш, яъни сўзнинг асосини топиш. Сўзнинг турланиш, тусланиш каби ўзгаришларидан қатъи назар ўзгармас қисмини тиклаш.

Стемминг асосида ишловчи дастур ёки стеммерлар миқдори, ишлатилиш даражаси билан қидирув тизими миқдори бир хил. Иккала тизимдан бирдек фойдаланиш кузатилади. Таъкидлаш жоизки, инглиз тили мисолида (грамматик маъно грамматик шаклларсиз ифодаланади, ортиқча флективлик хусусияти йўқ) стеммернинг ўз олдида қўйилган амалиётни аъло даражада бажаришига гувоҳ бўламиз.

Славян тилларида бу муаммо анча жиддий тус олади. Сўзнинг ўнг томонидаги қисм маъносини хатосиз ажратиш, бундай қоидалар мажмуини ишлаб чиқиш мукамал луғатли морфологик анализатор тузиш билан солиштирилганда анча мураккаб жараён саналади. Мисол тариқасида русча *кровать* сўзини келтирамиз. Шу пайтгача барча рус тили стемминг дастурлари ушбу сўзни муваффақиятли ва мантиқан мукамал равишда феъл моделлари

¹⁶⁸ <http://www.keva.ru/stemka/stemka.html>

асосида турланадиган шакл сифатида эътироф этади: *кр ёки кров* каби асосни ажратиб олади; “я круую/кровоаю, он кровает/крует” каби натижа кўрсатарди. Аммо кейинги йилларда мисли кўрилмаган ютуқларга эришилди. Мартин Портернинг стемминг соҳасидаги ишланмалари, қоидалар мажмуи, улар билан ишлаш инструментарий(восита)ларини тармоқ фойдаланувчилари учун очик (open/free) база шаклида SourceForge лойиҳаси доирасида эълон қилди. Дастур тилга махсус ишлов беришга, ахборот кидирувида стемминг алгоритмини шакллантиришга мўлжалланган. Лойиҳа жуда тезлик билан кенгая бошлади: ҳозирга келиб инглиз, француз, испан, португал, итальян, немис, дат, швед, норвег тиллари матнлари учун стеммер вазифасини бажара олади. Дастлаб рус тили четда қолди, яқинда рус тили учун ҳам база яратилиб, стемменг рус тилида ҳам ишлай бошлади.

Юқорида тавсифланган тизим кўплаб имкониятга эга бўлишига қарамасдан Snowball стеммери мукамалликдан йироқ. Бунинг асосий сабаби унинг тилшунослар томонидан ишланган умумий тилшунослик қоидалари асосида ишлашида. Маълумки, исталган умумий қоидаларда истиснолар бўлади; натижада асос(стем)ни аниқлашда хатоликка йўл қўйилади. Яна бир жиҳат: алгоритм исталган тил таҳлили учун доим битта қоида(натижа сифатида битта ўзак)ни тақдим этади. Ушбу камчиликлар туфайли ҳар бир тил учун махсус алгоритм ва стеммингнинг қатъий қоидаларини ишлаб чиқишга қарор қилинди. Муаллифнинг таъкидлашича, дастлаб рус тилидаги сўзларнинг кесишиш қоидаларини ишлаб чиқиш режалаштирилмаган, аммо ҳарфлар жойлашуви, кетма-кетлиги чегарасини аниқловчи, сўз ўзгартирувчи кўшимчалар частотасини аниқловчи алгоритм ишлаб чиқишга жуда катта қизиқиш уйғотган. Шу сабабли катта массивли рус тили матнларига ишлов бериш асносида автоматик равишда (Кўлда эмас!) стемминг қоидалари ишлаб чиқилган. “Пропись 4.0” лойиҳаси доирасида 1994 йилда ишлаб чиқилган рус тили морфологик анализаторининг мавжудлиги ушбу вазифани анча енгиллаштирди. Анализатор бир қанча ўзгаришларга учраган бўлса-да,

Апорт, Рамблер, МЕТА қидирув тизимларида ишлайди: анализатор луғати қайта-қайта синовдан ўтди: код бир секундда 20.000-30.000 сўзни қайта ишлай олади¹⁶⁹.

Кейинчалик тўлиқ матнли ахборотга ишлов бериш дастури ишлаб чиқилди. Дастур матнни сўзларга бўлиб чиқади, морфологик анализатор ёрдамида сўзшаклларни аниқлайди. Рус тилидаги 150.000 ўзак рўйхати мавжуд луғат бўлиши, 4 млн. турли хилдаги сўзшаклни ажрата олишига қарамасдан, нотаниш сўзлар таҳлилдан четда қолади. Бирликни ажрата олмаслик ҳажми катта эмас: 1 фоизни ташкил этади, холос. Бунда ҳам имло жиҳатдан хато ёзилган бирлик, қисқартма сўз, ноанъанавий (экзотик) ном ҳамда киши номлари морфологик таҳлилида хато натижа бериши мумкин.

Компьютер томонидан фарқланган, “тушунилган” сўзшакл учун унинг асоси, яъни турланиш ва тусланишда ҳеч қандай ўзгаришга учрамайдиган қисми кўрсатилади, маълумотлар текширилгач бундай қисм янги ҳосил бўлган қоидалар мажмуи маълумотлар базасига киритилади.

Таъкидлаш лозимки, рус тилининг 3та стемминг дастури, жумладан, филология фанлари докторлари А.А.Поликарпов, О.В.Кукушкина томонидан ишлаб чиқилган Snowball стемминги 2002 йилда Москва давлат университети “Умумий ва компьютер лексикгرافияси лабораторияси”да қиёсий синовдан ўтказилган. Синов натижасида аниқландики, Snowball стеммери лингвистик нуқтаи назардан нисбатан мукамал натижа берди: максимал таҳлил ва минимал ноаниқлик (яъни ўзак/асосларни нотўғри топиш) ушбу дастурнинг асосий ютуқларидан бири бўлди. Кейинчалик худди шу тажриба асосида украин тили стемминги ишлаб чиқилди; ушбу дастур (рус ва украин тили стемминги)¹⁷⁰ қидирув тизимида рус тили билан биргаликда украин тилидаги номаълум сўзларни ҳам муваффақиятли таҳлил қила олади. Шунингдек,

¹⁶⁹<http://www.keva.ru/stemka/stemka.html>

¹⁷⁰www.meta-ukraine.com

интернет тармоғида мавжуд ISpell луғати асосида бошқа тилдаги сўзнинг стемминг қиёсий жадвалини автоматик тузувчи дастур ҳам ишлаб чиқилди.

3. *Рус тилининг автоматик морфологик таҳлили ҳақида*. Д.С.Ермолаев рус тилининг компьютер автоматик таҳлили ҳақидаги мақоласида интеллектуал ахборот қидирувида матндаги сўзларнинг морфологик таҳлилини амалга ошириш борасидаги тажрибасини баён этади¹⁷¹. Тадқиқотчи интерфейсида сўровни амалга оширишни йўлга қўйиш учун талаб этиладиган асосий элемент матндаги сўзшакл морфологик таҳлилини йўлга қўйиш муҳим босқич эканлигини таъкидлайди. Агар дастурда шундай имконият пайдо бўлса, сўзшаклларнинг катта ҳажмли луғатига эҳтиёж қолмайди; дастур муваффақиятли ишлаши учун фақат сўз ўзаги ҳақидаги маълумотнинг мавжудлиги етарли бўлади. Дастур сўров берилган сўзшаклнинг леммаси(асос)ни аниқлаши талаб этилади. Масалан, фойдаланувчи маълумотлар омборига “фирма Рик. наша фирма продает тару картонную” шаклидаги гапни киритди. Морфологик таҳлил модули ушбу ахборотни куйидаги ҳолатда қайта ишлайди: *фирма. Рик. мой фирма продать тара картонный*. Маъно нуқтаи назаридан гап мазмунини йўқотди, аммо бу ҳолат компьютер учун айни муддао: бунинг фойдаси кейинги босқичда кўзга ташланади. Бошқа бир фойдаланувчи қидирув тизимида куйидаги сўровни киритади: “продает тару картонную”. Бу ахборот ҳам автоматик ишлов натижасида “продать тара картонный” шаклига келтирилади. Тизим мос келиш ҳолати бўйича қидирувни амалга оширганда, олдин эслаб қолган маълумотини чиқаради: “фирма Рик. продать...”. Бу ерда фақат биринчи фойдаланувчи киритган маълумотни сақлаб қолиш, фақат ўша маълумотни тақдим қилишдан кўра афзалроқ жараён кечгани аниқ. Рус тилидаги сўз морфологик белгилари сўзшаклнинг аффикс/суффикси ёрдамида аниқланади. Бундай қоида морфологик таҳлил қоидаси деб аталади. Аммо шундай

¹⁷¹Ермолаев Д.С.Компьютерный морфологический разбор слов русского языка // <http://citforum.ru/programming/digest/morfolog.shtml>

қўшимчалар борки, бир неча грамматик категорияни ифодалайди ҳамда маълум сўзшаклда батамом бошқа грамматик маънони англатади. Масалан, рус тилидаги сўзшакл охирида *-ать* морфемасининг мавжудлиги ушбу морфемани сақлаётган сўзшаклнинг феъл эканлигини билдиради. Аммо рус тилида *кровать* деган сўз бор; у от туркумига мансуб. Демак, морфологик таҳлил қоидасида истисно ҳолатларини шакллантириш мақсадга мувофиқ. Айрим сўзлар борки, грамматик маъноси ўзгарган тақдирда ҳам шаклини ўзгартирмайди; *не* предлоги, *столь* равиши шундай сўзлар сирасига киради. Бундай ҳолда морфологик таҳлил қоидаларига қўшимча киритиш талаб этилади; улар ҳам қоиданинг истисноси сифатида қаралади. Муаллиф шундай тажрибалари асосида сўзнинг морфологик таҳлилининг муайян мантиқий тавсифини келтириб чиқаради. Морфологик анализатор дастурини тузиш учун эса бир неча дастурлаш тилларини таклиф қилади ҳамда ҳар бирининг афзалликларига тўхталади.

Сўзшакл морфологик таҳлил дастурининг ПРОЛОГ дастурлаш тилидаги таҳлил намунаси 13 сатрдангина иборат, қолгани маълумотлар омбори саналади. Ушбу топшириқ РЕФАЛ-5 дастурида ўзига хос алгоритм билан бажарилади. РЕФАЛ- 5 дастури таҳлилни 3та жумла билан амалга оширди. Морфологик анализаторнинг ишлаш алгоритми C++ дастурида ҳам ўзига хос тарзда намоён бўлади.

Хулоса шуки, маълум бир тилнинг морфологик анализаторини ишлаб чиқиш учун дастлаб ўша тилдаги грамматик категориялар, грамматик шакллар тўлиқ, мукамал тавсифланиши, граммема, квазиграммема ва дериватемаларнинг чегараси аниқ белгилаб олиниши лозим.

2.3. Туркий тиллар морфологик анализаторларининг муштарак ва фарқли жиҳатлари

2.3.1. Туркий тилларда морфологик анализатор: тадқиқот ва тажрибалар. Табиий тилга автоматик ишлов бериш масаласи жаҳон

тилшунослигида кенг тадқиқ этилган. Масалан, рус тилшунослиги ва компьютер лингвистикасида морфологик анализатор, унинг умумий тавсифи, ишлаш принциплари, тилга оид грамматик категория, дастур учун лингвистик модулларлар тавсифи ҳақидаги ишлар сифатида О.В.Дереза, Д.А.Каютенко, А.С.Феногенова¹⁷², М.Ҳакимов¹⁷³нинг ишларини келтириш мумкин. Туркий тилшуносликда ҳам табиий тилнинг компьютер ишлови, морфологик анализаторини тузиш масалари кенг тадқиқ этилган. Бундай ишлар сирасига А.В.Дыбо, А.В.Шеймович¹⁷⁴, А.Р.Гатиатуллин, А.М.Баширов¹⁷⁵, П.В.Желтов¹⁷⁶, Н.А.Исраилова, П.С.Бакасова¹⁷⁷ларнинг турли тадқиқотларини киритиш мумкин. Ўзбек тили морфологик анализаторини тузиш махсус тадқиқот предмети бўлмаган бўлса-да, айрим изланишларда масаланинг муайян қирралари ҳақида фикр билдирилган. Жумладан, М.Орхун¹⁷⁸, Н.Абдурахмонова¹⁷⁹, М.Абжалова¹⁸⁰, У.Турсунов¹⁸¹нинг ишларини шундай тадқиқотлар сирасига киритиш жоиз. М.Орхун ўзбек тили отларини қидирув тизимида автоматик қайта ишлаш ҳамда морфологик анализаторнинг айрим

¹⁷² Дереза О.В., Каютенко Д.А., Феногенова А.С. Автоматический морфологический анализ для русского языка сравнительный анализ систем //

¹⁷³ Ҳакимов М. Логико-лингвистические модели русского языка // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

¹⁷⁴ Дыбо А.В., Шеймович А.В. Автоматический морфологический анализ для корпусов хакасского и древнетюркского языков / Научное обозрение саяно-алтая рецензируемый научный журнал Номер 2(08), 2014 (Периодичность – 2 раза в год. серия: Филология. Выпуск 2). С. 9-31.

¹⁷⁵ Гатиатуллин А.Р., Баширов А.М. Морфологический анализатор тюркских словоформ на базе структурно-функциональной модели тюркской морфемы // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.

¹⁷⁶ Желтов П.В. Разработка морфологического анализатора чувашского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.

¹⁷⁷ Исраилова Н.А., Бакасова П.С. Морфологический анализатор кыргызского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.

¹⁷⁸ Орхун М. Computational analysis of uzbek nouns /

¹⁷⁹ Abdurakhmonova N. Uzbek ontology of Uzbek language as example of adjective // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

¹⁸⁰ Абжалова М. Матнларга авто-лингвистик ишлов бериш тизимлари // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

¹⁸¹ Турсунов А. Вопросы словообразования в формальных моделях тюркских языков (на примере узбекского языка) // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

хусусиятлари ҳақида тўхталар экан, ўзбек тили морфологик анализаторида бажарилиши мумкин бўлган жараёнлар ҳақида фикр билдиради. Зеро, табиий тилни қайта ишлашнинг биринчи босқичи морфологик таҳлил ечими билан боғлиқ.

Маълумки, ўзбек тили агглютинатив хусусиятга эга, грамматик маънонинг кўшимча билан ифодаланиш имконияти кенг. Шу сабабли морфологик анализатор тузишда грамматик шаклларнинг унификация қилиниши муҳим аҳамият касб этади. Компьютер лингвистикаси тажрибасида морфологик анализатор тузишнинг турли усули мавжуд бўлса-да, туркий тилларда яратилган кўпгина анализатор икки даражали морфологик анализатор хусусиятини намоён қилади. Демак, ўзбек тили морфологик анализаторини тузишда ҳам икки босқичли морфологик анализ назариясига асосланиш мақсадга мувофиқ.

А.В.Дыбо, А.В.Шеймович қадимги туркий тиллар ва хакас тили морфологик анализаторини яратишда дастлаб грамматик категориялар тавсифини ўрганишади. Олтой тилларини тавсифлашда қабул қилинган анъанага кўра, хакас тилида қуйидаги асосий грамматик синфлар ажраталади: номлар, феъллар. Гуруҳдаги дастлабки икки синф маълум расмий кўрсаткич (аффикс)лар билан ифодаланган грамматик тоифалар тўплами билан тавсифланади. Граммемалар тўплами сўзни сўзга тобелантириб боғлаш ҳамда синтактик шакл ясовчилар парадигмасини ажратади. Масалан, феъл учун шахс/сон, майл, замон; от учун сон; исмлар учун эгалик, келишик. Маълумки, туркий тилларда грамматик синфлар орасидаги дифференциациянинг ифодаланиши кучсиз. Хусусан, исмлар орасида бу хусусият кенг тарқалган: гапдаги вазифасига қараб бир сўз ҳам от, ҳам сифат сифатида қаралиш ҳолатлари кузатилади¹⁸².

¹⁸²Дыбо А.В., Шеймович А.В. Автоматический морфологический анализ для корпусов тюркских языков / Филология и культура. Philology and culture. – 2014. №2(36). – С. 20-26.

Шунингдек, морфологик анализатор тузишда сўз асослари луғати ҳам муҳим лингвистик таъминот вазифасини ўтайди. Хакас тили морфологик анализаторини тузишда леммаларни гуруҳлашда юқорида таҳлил қилинган масала алоҳида эътиборга олинган. Айтиш жоизки, туркий тиллар морфологик таҳлил тизими ҳамда уларнинг моделлари тавсифи, таҳлилига бағишланган бир қанча тадқиқотлар¹⁸³ мавжуд. Қуйида уларнинг айримлари билан танишамиз.

А.В.Дыбо, А.В.Шеймовичнинг туркий тиллар автоматик морфологик анализи борасидаги фикрлари аҳамиятга молик¹⁸⁴. Мақолада туркий тиллар учун автоматик морфологик анализаторнинг ишлаш тамойиллари тавсифланган. Унинг асосий таркибий қисми сифатида қуйидагилар ажратилади:

- 1) тилнинг грамматик луғати;
- 2) сўз шаклининг тартибли модели (сўзшакл позициялар тўплами ва бу позициялар учун қўшимчаларнинг морфологик ифодалари);
- 3) сўзшакл чегарасида қўшимчанинг мувофиқлик қоидалари ҳамда маълум бир қўшимчанинг алломорфларини танлаш учун икки даражали фонетик қоидалар.

Парсернинг ишлаш алгоритми Ф.Крилов томонидан StarLing тизими базаси асосида ишлаб чиқилган қоидалар тизимига бўйсунди. Намуна сифатида қадимги турк тили морфологик анализатори версияси келтирилади. Туркий тиллар учун мўлжалланган анализатор рус ёки немис тиллари каби

¹⁸³Kemal Ofi azer. Two-level Description of Turkish Morphology. Literary and Linguistic Computing, – Vol. 9, No 2, – 1994.; Altintas K., Cicekli I. A morphological analyzer for Crimean Tatar //Proceedings of the 10th Turkish Symposium on Artificial Intelligence and Neural Networks (TAINN'2001). – 2001. – P. 180-189.; Çağrı Çöltekin (2014) A Set of Open Source Tools for Turkish Natural Language Processing In: Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14) Ed. by N. Calzolari et al. 1079-1086.; G. Kessikbayeva and I. Cicekli, "Rule based morphological analyzer of Kazakh language," in Proceedings of the 2014 Joint Meeting of SIGMORPHON and SIGFSM. Baltimore, Maryland: ACL, June 2014, pp. 46-54.; Tantug C., Adali E., Ofiazer K.. Computer Analysis of the Turkmen Language Morphology, 5th International Conference on NLP(FinTAL 2006), Turku, Finland, 186-193, 2006.; Orhun M., Tantug C., Adali E.. Rule Based Analysis of the Uyghur Nouns, International Journal of Asian Lang. Proc.,19(1), 33-44, 2009.

¹⁸⁴Дыбо А.В., ШеймовичА.В.Автоматический морфологический анализ для корпусов тюркских языков / Филология и культура. Philology and culture. – 2014. №2(36). – С. 20-26.

хинд-европа тилларига мўлжалланган Abbyu Finereader, Abbyu Lingvo каби йирик дастурий маҳсулотларда ишлатиладиган автоматик морфологик анализаторлар билан бир хил тамойил асосида ишлаб чиқилади. Луғатдаги ҳар бир лемма парадигма индекси шаклини олади. Флектив тилда парадигмалар катта эмас, аммо уларнинг сони кўп. Анализатор ҳар бир сўз учун тўлиқ парадигмани шакллантиради, кейин матнда учраган сўзшаклни шу парадигма билан қиёслайди. А.В.Дыбо, А.В.Шеймовичлар олтой гуруҳига кирувчи туркий тиллар морфологик анализатори хусусиятини ёритар экан, ушбу метод барча агглютинатив тиллар учун оптимал усул эмаслигига урғу беради.

А.В.Дыбо ва А.В.Шеймович агглютинатив тилларнинг хусусиятларини кўйидагича фарқлайди¹⁸⁵:

- 1) синтактик шакл ясовчи қўшимчалар тизими кенг ривожланган; бундай қўшимчаларнинг кўпчилик қисми грамматик жиҳатдан бирмаъноли, яъни қўшимчанинг ўзи грамматик омонимияни ҳосил қилмайди (бир грамматик шакл битта грамматик маънони ифодалайди);
- 2) турланиш, тусланишнинг ягона усули мавжудлиги, яъни (флектив тилларга қараганда) исм ва феъллар орасида турланиш ҳамда тусланиш тизимида фарқ катта эмас;
- 3) мураккаб морфонологик ҳодисалар мавжуд эмас: алломорфлар кўлланилишининг аниқ қоидалари мавжуд.

Агглютинатив сўзшакллар асосга стандарт грамматик шаклларнинг қатъий тартибда қўшилишидан ҳосил бўлади: морфеманинг чегараси аниқ, морфемалар қўшилишидан ҳосил бўладиган фонетик ҳодисалар қоидаси қатъий. Маълум бир сўзшаклнинг парадигмасини тузиш сўзшаклнинг фавқулодда мураккаб, кўптармоқли эканлигини кўрсатади. Мазкур мураккаблик синтактик шакл ясовчи қўшимчаларнинг кўплиги билан изоҳланадики, бу туркий тиллар морфологик анализаторини тузишда

¹⁸⁵Дыбо А.В., Шеймович А.В. Автоматический морфологический анализ для корпусов тюркских языков / Филология и культура. Philology and culture. – 2014. №2(36). – С. 20-26.

морфемалар комбинациясининг аҳамиятни катта эканлигидан далолат беради. Туркий тилларнинг мавжуд парсерлари битта сўзшаклнинг 300гача шаклини камраб олган парадигмасини тузган ҳолатларни кўриш мумкин. Табиийки, бу эса анализаторнинг иш сифатига таъсир кўрсатмай қолмайди.

Туркий сўзшаклларнинг компьютер моделини тузишда агглютинатив тилларга хос анъанавий “таснифланган грамматика”га асосланилади. Бу дастурий таъминот ахборот воситаси агглютинатив тилларни тавсифлашнинг қулай воситаси, унинг морфологияси қуйидаги талабларга жавоб бериши керак:

- а) сўз ясовчи қўшимчалар кетма-кетлигининг белгилаб қўйилиши;
- б) сўз ясовчиларнинг грамматик бирмаънолилигига эришиш;
- в) маълум бир граммеманинг бир сўзшаклда қандай шаклда учрашининг белгиланганлиги¹⁸⁶.

Морфологик анализаторнинг асосий ташкил қилувчилари сифатида қуйидагиларни кўрсатиш мумкин:

- 1) тил луғати (сўзнинг туркумга мансублик белгиси ҳақидаги изоҳ ҳамда фонологик қоидалар мавжуд луғат);
- 2) сўзшаклнинг етарли грамматик тавсифига асосланган (тилнинг автоматик таҳлиliga йўналтирилган) компьютер модели;
- 3) маълум қўшимчанинг асосга қўшилиши учун фонетик қоидаларга бўйсунувчи алломорфларини танлаш қоидалари тўплами.

Ф.Крилов томонидан таҳлил алгоритми юқорида санаб ўтилган материал билан исталган туркий тил материални тўлдириш ҳамда ундан фойдаланиш имконини беради. У тузган алгоритм бўйича сўзшакл таҳлили ўнгдан чапга йўналтирилади. Дастлаб, дастур асослар луғатидан сўзшаклнинг тўлиқ кўринишини қидиради. Агар асослар луғатида бундай шакл (белгилар кетма-кетлиги) учрамаса, парсер сўзшаклнинг ўнг томонидаги қўшимча

¹⁸⁶ Дыбо А.В., ШеймовичА.В.Автоматический морфологический анализ для корпусов тюркских языков / Филология и культура. Philology and culture. – 2014. №2(36). – С. 20-26.

формантига мос шакли излайди. Агар базада мавжуд кўшимчага мос белгилар кетма-кетлиги топилса, кесиб олиниб, текширишга юборилади; сўзшаклнинг чап томони асослар луғат билан солиштирилади. Агар натижа мусбат бўлса, парсер шу шаклнинг таҳлилини бошлайди. Натижа манфий бўлганда, парсер яна ўнг томондаги шаклга мурожаат қилади ҳамда кўшимчалар базасидан мос формантни кидиради ва бу кидирув асослар луғатидан мос асос топилгунча давом этади. Ушбу парсер хакас тили морфологик таҳлили учун ишлаб чиқилган бўлса-да, маълумотлар базасини тўлдириш орқали шор, тува, ёкут ва қадимги турк тилини морфологик таҳлил қилиш мумкин. Бу тилларнинг “асослар луғати” SUBD STARLING воситасида тузилган ва катта туркий-рус тили луғати билан текширилган.

2.3.2. Айрим туркий тиллар морфоанализаторларининг шарҳи.
Татар тили морфологик анализатори хусусида. Қуйида татар тили морфоанализаторининг умумий хусусиятлари тўғрисида тўхталамиз. Татар тили морфологик анализаторини ўрганиш жараёнида морфоанализатор тузишнинг умумий тамойиллари билан бирга фақат татар тилига хос бўлган айрим хусусиятлар ҳам кузатилди. Жумладан, Д.Ш.Сулейманов, Р.А.Гильмуллин, Р.Р.Гатауллинлар икки босқичли морфема модели асосида татар тилининг морфологик анализаторини тузиш тамойиллари ҳақидаги тадқиқоти¹⁸⁷ да бу ҳақда батафсил маълумот берилган. Шунингдек, татар тилида морфологик таҳлил тизимлари тавсифланган ишлар¹⁸⁸ ҳам мавжудлиги аниқланди.

¹⁸⁷Сулейманов Д.Ш., Гильмуллин Р.А., Гатауллин Р.Р. Морфологический анализатор татарского языка на основе двухуровневой модели морфологии / Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.

¹⁸⁸Сулейманов Д.Ш., Гильмуллин А.А., Гильмуллин Р.А. Двухуровневое описание морфологии татарского языка Тез. Межд. научн. конф., посвященной 200-летию университета: “Языковая семантика и образ мира” (Казань, 7–10 октября 1997г.). Книга 2. – Казань: Изд-во Казан. ун-та, 1997. – С. 65-67.; Gökgöz, Ercan, et al. Two-Level Qazan Tatar Morphology. Proceedings of the 1-st International Conference on Foreign Language Teaching and Applied Linguistics, May 5-7 2011 Sarajevo, P. 428-432.; Davliyeva, A.R. An investigation of Kazan Tatar morphology (Doctoral dissertation, San Diego State University): 2011.

Маълумки, исталган бир тилнинг морфоанализаторини тузишда шу тилнинг хусусиятлари алоҳида инобатга олинади. Д.Ш.Сулейманов, Р.А.Гильмуллин, Р.Р.Гатауллинлар татар тилини “семантик мураккаб, изчил тузилишли, амалий жиҳатдан автоматик лисоний ҳодиса” сифатида баҳолашади¹⁸⁹. Табиийки, морфологик анализатор тузишда табиий тилни қайта ишлаш билан боғлиқ масалаларни ечиш зарурати туғилади. Бу масала мураккаб морфологик хусусиятга эга бошқа туркий тиллар: татар, турк, қозок, қирғиз, қрим-татар, туркман, уйғур, ўзбек, чуваш, ёқут ва бошқа тилларда ҳам долзарблик касб этмоқда. Туркий тиллар морфологик таҳлил тизими ҳамда уларнинг моделлари тавсифи, таҳлиliga бағишланган тадқиқотларда ушбу тизимларнинг аксариятида (шунингдек, татар тилида ҳам) морфологик моделнинг фонологик ва морфологик қоидалар FST (Finite State Transducer) формализмига асосланганлиги айтилади. Ушбу тизим морфологик шакллар изчиллиги таъминланган тилни таҳлил қилиш учун мос келади. Аммо тизимни ишлаб чиқувчи мутахассислар туркий тилларда (бу 30-35%ни ташкил этади) мавжуд бўлган арабча, русча ўзлашмаларни таҳлил қилишда, татар графикаси бўлган кирилл алифбосининг ноадекватлиги муаммосига дуч келишади¹⁹⁰.

Татар тили морфологик моделларини тузиш XX асрнинг 90-йилларида бошланган; ушбу давр мобайнида татар тилининг уч типга асослаган (генератив модел, падиғматик модел¹⁹¹, икки даражали модел¹⁹²) морфологик анализатори ишлаб чиқилган. Таъкидлаганимиздек, морфологик анализаторнинг бу турлари маълумотни қайта ишлашнинг аниқлиги, тўлиқлиги ҳамда тезлиги билан бир-биридан фарқ қилади. Генератив модел

¹⁸⁹Кўрсатилган манба.

¹⁹⁰Сулейманов Д.Ш., Гильмуллин Р.А., Гатауллин Р.Р. Морфологический анализатор татарского языка на основе двухуровневой модели морфологии / Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.

¹⁹¹Сулейманов Д.Ш., Гатауллин А.Р. Модель татарской аффиксальной морфемы и ее реализация // Серия: Интеллект. Язык. Компьютер. – Вып.4. -Казань: Фэн, 1996. – 113-127.

¹⁹²Сулейманов Д.Ш., Гильмуллин А.А., Гильмуллин Р.А. Двухуровневое описание морфологии татарского языка Тез. Межд. научн. конф., посвященной 200-летию университета: "Языковая семантика и образ мира" (Казань, 7–10 октября 1997г.). Книга 2. – Казань: Изд-во Казан. ун-та, 1997. – С. 65-67.

асосида татар тили матнларини тахрирлашнинг дастлабки дастури¹⁹³ яратилди. Дастур узунлигидан қатъи назар, татар тилидаги исталган сўзни орфографик текширувдан ўтказди. Ҳатто бу модел АБВУҲ матнларни автоматик сканерлаш ва таниш тизимининг OCR FineReader 4.0. версиясида, парадигматик модел Microsoft матн муҳаррирининг MS Word 2007 версиясидан бошлаб фойдаланилади. Икки даражали морфоанализатор – Яндекс. Таржима дастурининг русча-татарча таржимонида, “Туган тел” татар тили корпуси¹⁹⁴нинг грамматик разметкасида ҳамда бир неча амалий лойиҳада фойдаланилмоқда. Ушбу морфоанализатор ҳам лексикон луғати, татар тили қоидалари мажмуига асосланади. Масалан, лексиконда 15984 от-ўзак, 7281 та феъл-ўзак; сифат, равиш, олмош, сон, кўмакчи, боғловчи, ундовлар билан жами 28.411 та ўзак мавжуд.

Демак, яратилган морфоанализаторлар амалий аҳамият касб этиб, ахборот-қидирув тизимида қидирув ва қайта ишлаш жараёнларини мукамаллаштиришга хизмат қилмоқда. Татар тили морфоанализатори устида олиб борган кузатишларимиз МА тузишда лексиконни сўз туркумига ажратиш принципларини ишлаб чиқиш, лингвистик таъминот луғатида тилдаги барча сўзларни туркумга ажратиш зарур, деган хулосага келишимиз учун асос бўлди.

Чуваш тилининг морфологик анализаторини ишлаб чиқиш тажрибаси ҳақида. Туркий тиллар морфоанализаторлари орасида чуваш тили морфологик анализатори тажрибасини ўрганиш алоҳида аҳамиятга эга. Бошқа тилларда бўлгани каби, чуваш тили морфоанализаторида ҳам МА тузишнинг умумий қоидалари билан бирга унинг ўзига хос жиҳатлар ҳам мавжуд. Қуйида чуваш тили морфоанализаторининг ўзига хос хусусиятларидаги асосий ўринларга қисқача тўхталамиз.

¹⁹³Сулейманов Д.Ш., Шафигуллин Р.Н. Морфологический корректор татарских текстов ТАТКОР // Татарский язык и новые информационные технологии. Серия: Интеллект. Язык. Компьютер. – Вып.2. – Казань: Изд-во Казан. ун-та. – 1995. – С.86-89.

¹⁹⁴Tatar Nacional Corpus. <http://tugantel.tatar/?lang=en>

Чуваш тили морфологик анализатори (кейинги ўринларда ЧТМА)нинг асосий вазифаси синтактик ва семантик анализатор учун чуваш тили сўзшакллариининг морфем тузилиши, морфологик белгиларни аниқлашдан иборат. ЧТМА маълумотлар базаси морфологик белгиларни тавсифловчи типлар рўйхати, (морфема маъносини изоҳловчи) луғат, (морфема ва унинг белгилари орасида муносабатни кўрсатувчи) маълумотнома, матн/сўзларларни сақлаш, сўз таҳлили намунасини ўзида сақловчи ишчи модел структураси маълумотлари, турли сўз туркумларига тегишли сўз таркибининг умумий қонуниятларини қамраб олган¹⁹⁵. Морфологик таҳлил учун ишлаб чиқилган модел айнан бир тилга мўлжалланмаган бўлиб, чуваш ва рус тиллари учун қўлланилиши мумкин.

Чуваш тилида сўз туркум, сон, келишик, шахс, субъектив баҳо шакллари, аспект (тасдиқ, инкор, эҳтимолий) ва замон каби морфологик характерга эга. ЧТМАни ишлаб чиқиш учун ҳар бир сўз туркумига тегишли бўлган морфологик характерлар жадвали/тавсифи ишлаб чиқилган. Масалан, от – шахс, сон, келишик, квазигреммема (нокатегориал грамматик маъно)лар.

1. *Морфологик анализаторнинг тузилиши.* Морфологик анализатор кутубхонаси 7 гуруҳдан ташкил топган, қуйидаги вазифаларни бажаришга мўлжалланган:

1. MorfParser.cs – МАнинг асосий гуруҳи.
2. Affi xInfoProvider.cs сўз киритиш ва кўмакчи морфемалар билан ишлаш.
3. ContextRules.cs –сўзшаклнинг контекстга асосланган грамматик шаклини тўғри аниқлаш.
4. FeaturesDeterminers.cs – сўзшакл атрибутларини топиш.
5. Helper.cs – ёрдамчи статистик функциялар тўплами.

¹⁹⁵Желтов П.В. Разработка морфологического анализатора чувашского языка /Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

6. ImprovedAffixHandler – сўзни маъноли қисмга ажратиш учун қабул қилиб олиш ҳамда луғатда қидирувни амалга ошириш.

7. MorphConstants – лойиҳада қўлланилган константалар мажмуи.

ЧТМА тизимига (1) сўзнинг туркуми ва бошқа морфологик белгиларини аниқлаш; (2) ўзак/асос ва қўшимчани фарқлаш; (3) матнни тушуниш; (4) истисно сўзларни таҳлил қилиш; (5) натижани файл кўринишида тақдим қилиш каби талаблар қўйилган. Шунингдек, МАдан фойдаланиш учун NET Microsoft платформаси ишлатилади¹⁹⁶.

2. *Асос ва қўшимчалар базаси.* Чуваш тили сўзлари таркиби асос ва қўшимчалар бирикувидан иборат. Рус тилидан фарқли ўлароқ, чуваш тилида “приставка” ва “окончание”нинг мавжуд эмаслиги МАнинг ишлашини осонлаштиради. Демак, биринчи навбатда, МА учун асос/ўзак ва қўшимчалар луғати талаб қилиниб, ушбу луғат чуваш тилининг 30.000дан ортиқ сўзини қамраб олган: ҳар бир сўзга унинг туркуми ҳақидаги маълумот бириктирилган. Кўмакчи морфемалар ҳал қилувчи аҳамиятга эга, сўзларнинг бир-бирига нисбатан ҳоким/тобелик мавқеини аниқлайди. Чуваш тилида грамматик маъно ва нокатегориал маъно ифодалайдиган қўшимчалар сони 170та. Қўшимчалар базаси ҳам худди асослар луғати сингари морфема ва унинг маъносидан ташкил топади. Маълумотлар базасининг бундай таркибга эгаллиги машина таржимаси ва лингвопроцессорлар иши учун аҳамиятли саналади. Чунки кўмакчи морфема сўз ясаш ҳамда сўз шаклини ўзгартириш вазифасини ўтайди. Бир неча агглютинатив тилларда сўз таркиби 7 тагача кўмакчи морфемадан иборат мураккаб аналитик структурали шакллар ҳам учрайди; улар рус тилига бир нечта гап шаклида таржима қилиниши мумкин.

3. *Морфологик анализ ядроси.* Морфоанализаторни ишлаб чиқиш жараёни икки этапга ажратилади. Биринчидан, сўров сифатида киритилган сўзшакл асослар луғатидан қидирилади. Грамматик белги аниқланган сўз туркумидан келиб чиқиб белгиланади. Иккинчи босқичда бевосита таҳлил

¹⁹⁶Кўрсатилган манба.

амалга оширилади: сўзшакл асос ва қўшимчага бўлинади; уларнинг хусусиятлари аниқланади. Сўзшакл луғатдан ўзига тенг эквивалентни топмаса, натижа хусусиятлари аниқланмаган сўзшакл ҳолида чиқади.

Таъкидлаш лозимки, чуваш тилида олмош сўз туркумининг морфологик белгиларини аниқлаш бошқа туркумга нисбатан фарқланади. Чунки олмошнинг бош шакли биринчи ва иккинчи шахсда бир хил бўлади¹⁹⁷ (бу муаммо ўзбек тилидаги олмошлар орасида сал бошқа тарзда юзага чиқади). Ушбу грамматик омонимияни бартараф этишда юқорида МА маълумотлар базаси сирасида саналган истиснолар луғатидан фойдаланилади.

Чуваш тили морфоанализатори кутубхонаси (маълумотлар омбори) C+ тилида Visual Studio 2013да NET.Framework платформасида яратилган. Чуваш тили морфоанализаторини ўрганиш МАнинг тузилиши, таркибидаги асос ва қўшимчалар базаси, морфологик таҳлил ядроси ҳақидаги тасаввуримизни ойдинлаштирди. Ушбу МАнинг сўзшакл туркумини; ўзак ва қўшимчани аниқлаш; контексни таҳлил қилиш, белгини тиклаш; шаклнинг морфологик характеристикасини аниқлаш каби вазифаларни бажаришга мўлжаллангани МАларнинг имконияти ҳақида хулосаларимиз шаклланишига туртки бўлди.

Туркий тилларда яратилган морфологик анализаторлар билан танишар эканмиз, *қирғиз тили морфологик анализаторининг* ўзига хос хусусиятларига ҳам алоҳида эътибор қаратдик.

1. *Қирғиз тилининг морфологик модели.* Қирғиз тили кўмакчи морфемаларнинг етакчи морфемага қўшилишида қатъий тартибга амал қилиниши билан ажралиб туради. Масалан, от туркумига оид сўзнинг морфем таркиби куйидаги тартибда бўлади: ўзак+сўз ясовчи морфема+шакл ясовчи морфема (сон+келишик+кесимлик). Ушбу шакллардан бирининг тушиб қолиши морфемалар тартибида ҳеч қандай ўзгариш пайдо қилмайди: қайсидир

¹⁹⁷Кўрсатилган манба.

грамматик маъно бўлмаслиги мумкин, ammo тартиб қатъий¹⁹⁸. Қирғиз тилининг бу хусусияти ўзбек тили сўзшакллариининг морфем тартиби билан айнан бир хил.

2. *Қирғиз тили сўзларининг морфологик структураси.* Қирғиз тилидаги сингармонизм ҳодисаси асосида ўзакка қўшиладиган кўмакчи морфеманинг ўзгариши ўзбек тилидан фарқ қилади. Масалан, қирғиз тилида кўплик сонни билдирувчи *-лар* қўшимчасининг *-лар, -дар, -тар* каби алломорфлари бўлиб, уларнинг ўзакка қўшилиши ҳар хил позицияда турлича бўлиши мумкин. Табиийки, бундай ҳолатда таҳлилни аниқлаштириш мақсадида морфоанализатор учун қоидалар мажмуи ишлаб чиқилади¹⁹⁹. Демак, морфоанализатор тузишнинг умумий қоидалари мавжуд бўлса-да, уни тузишда ҳар бир тил учун морфологик модел ҳамда сўзшаклнинг морфологик структураси ўзига хос бўлади. Таҳлил аниқлигига морфологик шакллари тўла тавсифлаш, уларнинг парадигмасини шакллантириш билан эришилади.

3. *Қирғиз морфологиясининг ўзига хослиги ва қонуниятлари.* Н.А.Исраилова ва П.С.Бакасова морфоанализатор тузишда қирғиз тилининг морфо-синтактик ҳамда фонетик хусусиятларига алоҳида тўхталишади, чунки ушбу омиллар морфоанализ тамойилларини ишлаб чиқишда муҳим аҳамият касб этади.

Маълумки, қирғиз тили агглютинатив тиллар гуруҳига мансуб. Зеро, агглютинатив тиллар морфоанализаторини тузишда тилнинг морфологик хусусияти билан бирга унинг синтактик, фонетик хусусияти ҳам инобатга олинади. Агглютинациянинг *морфологик* белгилари сифатида сўз асоси – бош келишиқдаги сўзшакл турланиш парадигмасининг марказида туриши, маъноли қисмлар орасидаги чегаранинг яққол кўриниб туриши, кўмакчи морфеманинг асосга қўшилишида қатъий тартиб амал қилинишини эътироф

¹⁹⁸Исраилова Н.А., Бакасова П.С. Морфологический анализатор кыргызского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.

¹⁹⁹Кўрсатилган манба.

этилади. Агглютинациянинг *фонетик* белгилари эса сингармонизмнинг мавжудлиги, урғунинг билиниши(урғу сўзнинг фонетик яхлитлигини таъминлаши)дан иборат. Агглютинациянинг *синтактик* белгилари сифатида гапда сўз тартибининг қатъийлиги (1), аниқловчининг аниқланмишдан олдинда туриши (2), тўлдирувчининг тўлдирилувчидан олдинда жойлашиши (3), кесимнинг гап охирида бўлиши(4)ни кўрсатиш мумкин²⁰⁰.

4. *Қирғиз тили морфологик маълумотлар базаси структураси.* Морфологик маълумотлар базаси морфологик анализ/синтез жараёни учун зарур бўлган барча маълумотни қамраб олиши талаб қилинади. Агар морфоанализатор сўзшакллар луғати асосида ишласа, маълумотлар базасида сўзшаклнинг асоси – лексема, уларнинг туркуми, бошқа морфологик белгиларига ишора қилувчи луғат маълумотлар базасидан ўрин олиши лозим. Агар морфоанализатор ўзак/асослар луғати асосига қурилган бўлса, унда маълумотлар базасида лексемалар луғати билан бирга ҳар бир парадигмага тегишли флекция ҳолатларини қамраб олган луғат ҳам киритилади. Шунингдек, маълумотлар базасида омонимлар луғати (лексик ва грамматик омонимия) ҳам бўлиши талаб қилинади.

Қирғиз тили МАни ўрганишларимиз кўрсатадики, МА тузиш учун тилнинг морфологик модели; сўзнинг морфологик структурасини аниқлаш; маълумотлар базасида лексема, унинг туркуми, бошқа морфологик белгиларига ишора қилувчи луғат маълумотлар базасидан ўрин олиши лозим.

Ёқут тили морфологик анализатори хусусида. А.Н.Леонтьев ёқут тили морфологик анализатори²⁰¹ни яратиш тамойилларини изоҳлар экан, умумий назарий асослар ҳамда фақат ёқут тилига хос бўлган жиҳатларга тўхталиб ўтади.

²⁰⁰Кўрсатилган манба.

²⁰¹Леонтьев Н.А. Морфологический анализатор якутского языка // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) –Ташкент, 2018. – 320 с.–С. 276-279.

Тилшуносликда автоматик морфологик таҳлил – матнни компьютерда қайта ишлашнинг асосий қисмларидан бири бўлиб, ушбу турдаги таҳлил матндаги сўз қисмини ойдинлаштириш, унга тегишли грамматик белгини аниқлаш – лингвистик разметкашга ихтисослашган. Морфологик таҳлил усулларининг ривожланиши кенг тарқалган тиллар учун анализатор яратишга имкон берди. Россия туркий халқларининг кўплари, масалан, чуваш²⁰², калмиқ²⁰³, татар²⁰⁴, тува²⁰⁵, хакас²⁰⁶ учун морфологик анализатор ишлаб чиқилган. Бошқа мамлакатларда, хусусан, турк, қозоқ ва бошқа туркий тиллар лингвистик корпус ва морфологик анализатор яратиш тез суръатлар билан ўсиб бормоқда.

Ёкут тили материални морфологик таҳлил қилиш учун ёкут тилини аннотациялаш (теглаш) тизими²⁰⁷ ишлаб чиқилган. Кейинчалик ушбу изох (тег)лар тизими маълумотлар базасига ўзгартирилган. Кўмакчи морфемалар базаси, қўшимчанинг мувофиқлаштирилиши (фонетик ўзгаришнинг комбинациялари) базаси ҳамда асос/негизлар базаси шакллантирилган. Ушбу маълумотлар базаси асосида ёкут тилини морфологик таҳлил қилиш учун алгоритм, дастур яратилган. Морфологик анализаторни ишга тушириш учун асосни аниқлаш алгоритми²⁰⁸, шунингдек, мутахассис-эксперт ёрдамида

²⁰²Желтов П.В. Морфологический анализатор национального корпуса чувашского языка // В сборнике: Совершенствование методологии познания в целях развития науки. сборник статей по итогам Международной научно-практической конференции: в 2 ч.. – 2017. – С. 11-13.

²⁰³Куканова В.В., Каджиев А.Ю. Алгоритм работы морфологического парсера калмыцкого языка // В сборнике: Писменото наследство и информационните технологии. E1Manuscript–2014 Материали от V международна научной конференции. Отговорни редактори В.А.Баранов, В.Желязкова, А.М.Лаврентьев. 2014. – С. 116-119.

²⁰⁴Гатиатуллин А.Р., Баширов А.М., Осипов Г.С., Смирнов И.В., Шелманов А.О. Методы лингвистического анализа текстов на татарском языке и их применение в поисковой системе EXACTUS // Труды Института системного анализа Российской академии наук. 2016. Т. 66. №1. С. 18-25.

²⁰⁵Хертек А.Б., Ооржак Б.Ч. О морфологической разметке электронного корпуса текстов тувинского языка // Филологические науки. Вопросы теории и практики. 2012. №7-2(18). С. 214-218.

²⁰⁶Дыбо А.В., Шеймович А.В. Аппарат автоматического морфологического анализа для корпуса хакасского языка // Родной язык: лингвистический журнал. 2016. № 2 (5). С. 9-39.

²⁰⁷Торогоев Г.Г., Ноговицына А.Н. Лингвистическое аннотирование наклонений глагола якутского языка // Вестник Северо-Восточного федерального университета им. М.К. Аммосова. 2017. №3 (59). С. 108-120.

²⁰⁸Леонтьев Н.А. Вопросы автоматизированной лемматизации якутского языка // Современная наука: актуальные проблемы теории и практики. Серия: Познание. 2015. №2 (41). – С. 15-16.

сўзшаклни аннотация (тег)лаш тизими²⁰⁹ ишлаб чиқилган. Ушбу морфологик анализатор қуйидаги алгоритм бўйича ишлайди: луғат сўзи қидирилади, агар бундай сўз топилса, сўз кўрсатилади; аммо қўшимчаларни қидириш давом этади. Сўз охиридаги қўшимча бўйича қидирув амалга оширилади, агар қўшимча топилса, охирги қўшимча ўчирилади, сўз охиридаги ўзгаришни ҳисобга олиб, сўзнинг асоси/леммаси қидирилади. Натижада фонетик ўзгаришлар инобатга олинган қўшимчалар занжири ҳосил бўлади. Грамматик омонимлик/ кўпмаънолилик ҳолатида қўшимчаларнинг бир неча занжири олинади.

Ушбу морфологик анализатордан онлайн фойдаланиш учун веб-сайт учун РНР тилида код ёзилган, бу веб-сервер қобиғига коднинг интеграциясини осонлаштиради. Маълумотлар базаси СУБД MySQL маълумотлар базасига асосланган. Бу масофавий фойдаланувчиларга тизимга киришга имкон беради. Морфологик анализаторни синаб кўришда ёқут тили корпусидан фойдаланилди; таҳлиллар шунини кўрсатдики, морфологик таҳлил 80% аниқ натижа кўрсатди. Таҳлилнинг хатолари(20%)га матнда ўзлашма сўзнинг мавжудлиги, янги кириб келган сўз ҳамда грамматик кўпмаънолилик сабаб бўлган²¹⁰.

Ёқут тили МАни кузатиш, ундан фойдаланишимиз маълумотлар базасининг дастурлаш тили, МАни тест режимида ишлатишда ёқут тили корпуси, теглаш тизими ҳақидаги тасаввуримизни ойдинлаштирди. Бу эса МА тузишда дастурлаш тили, тест режими учун материал танлаш, ҳар бир тилга хос теглаш принципини ишлаб чиқиш зарурати борлиги ҳақида хулоса қилишга имкон берди.

²⁰⁹Леонтьев Н.А., Торотоев Г.Г. Многопользовательская морфологическая разметка корпуса якутского языка // В сборнике: Электронная письменность народов Российской Федерации: опыт, проблемы и перспективы Сборник материалов Международной научной конференции. – 2017. – С. 101-103.

²¹⁰Леонтьев Н.А. Морфологический анализатор якутского языка // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции)–Ташкент, 2018. – 320 с.– С. 276-279

Кузатишларимиз асосида хулоса қиламизки, морфологик анализатор маълумотлар базасида ўзлашма сўз ва неологизмларни тўлдириб бориш ойнаси ишга туширилиши мақсадга мувофиқ. Грамматик кўпмаънолиликни бартараф этиш учун эса морфологик анализаторнинг қоидалар луғатини такомиллаштириш, грамматик омонимияни филтрловчи алгоритм ишлаб чиқиш талаб қилинади (Қаранг: 4-чизма).

2.3.3 Ўзбек тилининг морфологик анализи борасидаги изланишлар хусусида. Т.Б.Болтаев, С.И.Ибрагимовларнинг ўзбек тили морфологик таҳлил дастурий тизими лойиҳаси ҳақидаги мақоласида ўзбек тилининг морфологик таҳлилни автоматлаштириш, асосан, от ва сифат сўз туркуми морфологик таҳлили ҳақида фикр юритилади. Тадқиқотчиларнинг фикрича, морфологик таҳлилнинг бир неча усуллари мавжуд, автоматлаштириш (компьютер ишлови) улар орасида энг оптимал усул саналади²¹¹. Бундай дастурнинг муваффақиятли ишлаши аниқлик, натижавийлик, турғунлик каби талабларга жавоб беришига боғлиқ. Тадқиқотчилар автоматик таҳлил дастурининг бундай талабларга жавоб бериши тизимни лойиҳалаштириш вақтида формал тилни таҳлил қилишнинг математик воситасини ишлаб чиқиш фикрини илгари суришади. Бундай воситалар жуда машҳур, улар асосан, дастурлаш тили ишлаб чиқувчилари учун (анализатор, компилятор, спецификатор)²¹² зарурий восита ҳисобланади. Мақола муаллифлари мазкур тизим тарихи ҳақида тўхталар экан, бу турдаги тизимнинг қўлланилиши ўтган асрнинг 60-йилларига тўғри келишини айтишади. Жумладан, флектив тиллар, масалан, рус тили учун таҳлил воситасининг формал тилдан ЭТАП²¹³ тизимида

²¹¹Болтаев Т.Б., Ибрагимов С.И. О проекте программной системы морфологического анализа узбекского языка // <http://buxdu.uz/index.php/uz/>

²¹²Ахо Альфред В., Лам Моника С, Сети Рави, Ульман Джеффри Д. Компиляторы: принципы, технологии и инструментарий, 2-е изд.: Пер. с англ. – М.: ООО “И.Д. Вильямс”, 2008. – 1184 с.; Касьянов В.Н., Поттосин И.В. Методы построения трансляторов. – Новосибирск: Наука, 1986. – с. 344.; The Structured Constructing as a Discipline of Safe Programming and Instruments Supporting It /Aniskov M.I., Boltaev T.B., Kochetov D.V. at al//Instrumental Congress on Computer Systems and Applied Mathematics CSAM'93. St-Petersburg. July, 19-23.

²¹³ Апресян Ю.Д, Богуславский И.М., Иомдин Л.Л. и др. Лингвистическое обеспечение системы ЭТАП-2., – М.: Наука 1989. – с. 296.

фойдаланилган, шунингдек, Д.С.Юравский, Ж.Г.Мартинлар²¹⁴ инглиз тили учун формал тиллар таҳлил воситасининг қўлланилиши борасида иш олиб боришган. Т.Б.Болтаев, С.И.Ибрагимов тадқиқотда дастур устида олиб борилган ишни тавсифлар экан, тадқиқотнинг асосий мақсади сифатида формал тиллар назарияси усулини ўзбек тилидаги матн таҳлиliga жорий этишни кўрсатишади. Чунки ушбу усулни рус²¹⁵, инглиз²¹⁶ тилида қўллаш бўйича анча тажриба мавжуд. Ушбу дастур олдида шундай асосий вазифа қўйилган: дастурга матн киритилади, сўров натижаси/чиқиш морфологик таҳлилни қамраб олган лексемалар занжиридан иборат бўлади; лексемалар морфологик характери ҳақида барча маълумот бўлади. Бу ерда морфологик анализатор формал тилларни таҳлил қилиш имконига кўра лексик анализатор билан солиштирилади. Тузилаётган таҳлил механизми сканер (лексик анализатор) хусусиятини ўзида мужассамлаштиради. Туб сўзда бўлганидек, сўз ясалишининг (сўз ясовчи ва аффиксоид билан ясалган) кўпгина қолиплари лексик анализаторни қўллаш имконини беради. Сўз ясалишининг бундай қолипни тавсифлашда кенгайтмага эга бўлган доимий ифодалар дастурий воситасидан фойдаланиш мумкин. Сўз ясалишининг композиция, лексемалашув каби нисбатан мураккаб қолиплари автоматик таҳлилдан четда қолади; истиснолар грамматикаси асосида таҳлил қилинади²¹⁷. Муаллифлар морфологик таҳлил дастури зарурий таркибий қисми ҳақида тўхталар экан, бу жараёнда математик воситалар, ахборот тизимлари зарурлигини таъкидлашади.

Муаллифлар анализаторнинг иш алгоритминини қуйидагича кўрсатишади:

1) сўров киритишда сўзшакл (wordform) берилади;

²¹⁴Speech and Language Processing. D.S.Jurafsky, J.H.Martin. Contributing writers: Andrew Kehler, Keith Vander Linden, Nigel Ward 2000y. – Prentice Hall, Englewood Cliffs, New Jersey 07632. – pages: 950.

²¹⁵ Ножов И.М. Морфологическая и синтаксическая обработка текста (модели и программы): диссертация канд. наук. –М., 1996. – с.190.

²¹⁶Morneau R. The Lexical Semantics of a Machine Translation Interlingua. http://www.eskimo.com/~ram/lexical_semantics.html 2006.

²¹⁷Болтаев Т.Б., Ибрагимов С.И. О проекте программной системы морфологического анализа узбекского языка // <http://buxdu.uz/index.php/uz/>

2) натижа (чиқиш)да сўзшакл асоси (stem) матннинг кейинги (синтактик, семантик, прагматик таҳлил, таржима) ишлови учун зарур бўладиган барча морфологик характеристикаси (morphological features) билан берилади.

Сўзшаклнинг морфологик характеристикаси асос мансуб бўлган, сўзшаклни шакллантирувчи, сўз туркумига мансубликни билдирувчи айрим маънолари саналади. Масалан, от (Noun-N) сўз туркуми учун бирлик (Singular-SG) ёки кўплик (Plural-PL)ни билдирувчи сон категорияси шакли, мазкур отнинг келишик категорияси (Case) ва ҳ. Феъл (Verb-V) сўз туркуми учун замон (Tense-TS), шахс/сон (Person-PR) категорияси кўрсатилади. Мисол тариқасида “*китобларни*” сўзшаклини оламиз. Морфологик анализатор “*kitob*”+N+PL+CS(dative) шаклини, “*ўқидим*” сўзшакли учун “*O’qi*”+V+PR(1)+TS(past) тарзидаги натижани беради. Бу ерда CS(dative) тушум келишигини, PR(1) I шахсни, TS(past) ўтган замонни билдиради.

Морфологик анализаторни конструкциялаш учун қуйидаги воситалардан фойдаланилади:

1. Лексикон (lexicon) – морфологик анализ(МА)нинг ишлаши учун керак бўладиган ўзак, қўшимча, улар ҳақидаги ахборот структураси. Лексикон МАдаги сўзларнинг репозиторийси саналади. Барча сўзларни репозиторийга киритиб бўлмаслиги сабабли лексиконда ўзак (асос), аффикс ва морфотактиклар сақланади. Ушбу услуб тилда янги пайдо бўлган сўзларни киритишда қулайлиги билан афзаллик касб этади.

2. Морфотактиклар (morphotactics) – сўзшаклдаги морфемалар тартибини аниқлашга мўлжалланган қоидаларнинг ахборот модели. У сўзшакл морфемалари орасидаги кетма-кетликни аниқлайди; маълум сўз туркумига хос сўзшаклнинг морфемалари орасидаги муносабатни кўрсатади. Масалан, от туркумида сон кетегорияси ўзакдан кейин, бошқа категориялардан олдин жойлашадики, бундай қатъий кетма-кетлик аглютинатив тилларга хос.

3. Орфографик қоидалар (orthographic rules) – икки категория комбинацияси натижасида ўзакда содир бўладиган фонетик ўзгаришларга оид қоидалар мажмуи. Масалан, “q” ёки “k” билан тугаган жўналиш келишигидаги CS(dative) отлар охиридаги ушбу ҳарфларда иккиланиш ҳодисаси учрайди, яъни қўшимча -га шаклида эмас, балки -ga, -ka шаклида ишлатилади ҳамда қўйидаги қоидага амал қилади: elakka = “elak” + CS(dative).

4. Ўзак семантикасини шакллантирувчи ахборот структураси бирликлари (Stem Semantics Information Structure – SSUS). Бундай бирликлар, одатда, семантик атрибутлар деб юритилади. Бу атрибутлар асосида таҳлил қилинаётган сўзни семантик тўғри таҳлил қилинишини таъминловчи семантик-морфологик (Semantic-morphologic) қоидалар ҳосил қилинади. Масалан, (Proper(N)&Singular(N)) | (Abstract(N)&Uncount(N)) P SG(N) формуласи таҳлил қилинаётган отнинг борлиқдаги ягона нарсани атаб келаётган атоқли от (“Quyosh”, “Yer”, “Oy”, “Toshkent”, “Navoiy”) ёки саналмайдиган мавҳум от (“mehr”, “zulmat”, “oriyat”) эканлигини кўрсатади. Бундай сўзшаклга кўплик морфемасини қўшиш мумкин эмас. Шунингдек, SSUS элементи синтактик ва семантик (контекстуал) таҳлилда, морфологик хусусиятларни аниқлашда қўлланади.

5. Якуний автоматик восита (Finite State Automata – FSA)²¹⁸ морфотактларни моделлашда фойдаланиладиган элемент саналади. Орфографик қоидаларни моделлаштириш учун FSAнинг махсус варианты – трансдюсерлар²¹⁹ қўлланилади.

Хулоса ўрнида таъкидлаш жоизки, шу пайтга қадар ўзбек тили морфоанализини тузиш борасида қилинган ишлар монографик характер касб этмаса-да, мазкур мақола бу борада қимматли материал вазифасини ўтай олади ва у ўзбек тили морфоанализи учун лексикон, морфотактика,

²¹⁸Ахо А., Ульман Дж. Теория синтаксического анализа, перевода и компиляции. Вдвухтомах. (Пер. сангл.) – М.:Мир, 1978.

²¹⁹MohriM. Finite-state transducers in language and speech processing. ComputationalLinguistics. – P.269-312.

орфографик қоидалар, яқуний автоматик восита ҳақидаги маълумот ўзбек тили морфоанализаторини тузишда муҳим назарий асос сифатида хизмат қилади.

Боб юзасидан хулоса

1. Морфологик анализатор тузиш унинг умумий тамойилларини аниқлашдан бошланади. Зеро, матннинг автоматик таҳлилида бир неча мураккаб операция белгиланган алгоритм асосида амалга ошади; автоматик таҳлил давомида компьютер матн ҳақида ўз тилида лексик-морфологик, синтактик, семантик тасаввур ҳосил қилади. Шу боис матннинг автоматик анализ ва синтез жараёни автоматлашган компьютер лингвистикасининг асосий вазифаларидан саналади. Бу вазифа ахборотни қайта ишлаш воситалари учун лингвистик моделлар яратиш ва инсониятнинг амалий эҳтиёжларидан келиб чиқади.

2. Ахборот базаси ва таҳлил дастурларининг амалий аҳамияти ахборот-қидирув тизимлари ҳамда машина таржимаси сифатини яхшилаш, тил корпуси бирликларини лингвистик разметкалаш ҳамда махсус амалларни бажаришга мўлжалланган мустақил анализатор сифатида ишлаши билан белгиланади. XXI асрга келиб замонавий дунё маълумотларининг ҳаддан ташқари кўплиги, интернет материалларининг доимий ўсишини ҳисобга олсак, табиий тилга компьютерда ишлов бериш, ахборотни қайта ишлаш масаласининг долзарблиги яққол сезилади. Бу борада морфологик анализаторнинг муҳим аҳамияти ахборот-қидирув тизимларида сўзшаклнинг аниқ анализ-синтезини амалга ошириши билан характерланади.

3. Морфологик анализатор деганда луғатдаги сўзшакл ва алоҳида сўзларни қиёслашга мўлжалланган, сўзнинг грамматик хусусиятини ифодалайдиган алгоритмлар тўплами назарда тутилади. Морфологик таҳлилда компьютернинг сўзшаклни тушуниши билан иш битмайди, яъни белгилар кетма-кетлигининг компьютер томонидан фарқланиши етарли эмас: сўзшаклнинг парадигма аъзоларидан фарқини топиш билан таҳлил тугамайди.

Шунга кўра сўзшаклнинг туркумга хослиги ва бошқа морфологик белгиларини топиш анализаторнинг асосий вазифаси ҳисобланади.

4. Морфологик таҳлил натижаси сўзнинг муайян туркумга мансублиги, сўзнинг грамматик шакли, сўзшаклга ишлов бериш учун керакли грамматик белгилар мажмуи каби маълумотлар тўпламидан иборат. Ваҳоланки, сўзшаклга грамматик луғат орқали ишлов берилса, таҳлил қилинса, луғат мақоласига тегишли грамматик атрибутлар ҳам аниқланади.

4. Грамматик луғат морфологик анализатор учун ниҳоятда муҳим таркибий қисми саналади; морфологик луғат маълум бир сўзнинг барча грамматик шакллари рўйхатини тўплайди. Морфологик маълумотлар базаси фойдаланувчи эҳтиёжи/талабига жавоб берадиган кенг кўламли маълумотлар мажмуидан иборат бўлиб, унда маълумотлар сараланади ҳамда маълумотлар базасини ҳаракатлантирувчи тизим томонидан бошқариладиган жадвал шаклида сақланади.

5. Муайян тилнинг морфологик анализаторини ишлаб чиқиш учун унинг грамматик категориялари, грамматик шакллари тўлиқ, мукамал тавсифланиши, граммема, квазиграммема ва дериватемаларнинг чегараси аниқ белгилаб олинishi лозим. Шу ўринда морфоанализатор тизимига сўзнинг туркуми ва бошқа морфологик белгиларини аниқлаш; ўзак/асос ва кўшимчани фарқлаш; матнни тушуниш; истисно сўзларни таҳлил қилиш; натижани файл кўринишида тақдим қилиш каби талаблар қўйилади.

6. Морфологик маълумотлар базаси морфологик анализ/синтез жараёни учун зарур бўлган барча маълумотни қамраб олиши талаб қилинади. Морфоанализатор сўзшакллар луғати асосида ишласа, маълумотлар базасида сўзшаклнинг асоси – лексема, уларнинг туркуми, бошқа морфологик белгиларига ишора қилувчи луғат маълумотлар базасидан ўрин олиши лозим. Морфоанализатор ўзак/асослар луғати таркибига таянса, маълумотлар базаси лексемалар луғати билан бирга ҳар бир парадигмага тегишли флекция ҳолатларини қамраб олган луғат ҳам бўлишини талаб қилади. Шунингдек,

лингвистик таъминотда омонимлар луғати, ўзлашма сўз ва неологизмларни тўлдириб бориш ойнаси ишга туширилиши мақсадга мувофиқ.

7. Татар, ёқут, чуваш, қирғиз тили морфологик анализаторлари шундай тузилма тузиш умумий тамойиллари асосида яратилган бўлса-да, ҳар бир тилга оид морфоанализаторнинг ўша тил хусусиятидан келиб чиқувчи ўзига хос томонлари ҳам мавжудлигини назардан қочирмаслик зарур. Ўзбек тилининг морфоанализаторини ишлаб чиқишда туркий тилларда яратилган морфоанализаторлар тажрибасига таяниш ўринли.

III БОБ. ЎЗБЕК ТИЛИ МОРФОАНАЛИЗАТОРИНИ ТУЗИШНИНГ ЛИНГВИСТИК АСОСЛАРИ

3.1. Ўзбек тили морфоанализатори бирликлари, тузилиши ва таркиби

3.1.1. Морфоанализатор бирликлари. Табиий тилга компьютерда қайта ишлов беришнинг дастлабки босқичи деганда морфологик таҳлилни амалга ошириш; морфоанализ модули эса матнни таҳлил қилиш дастурлари компонентлари назарда тутилади. Жаҳон компьютер лингвистикасида автоматик морфоанализ 1950-йиллардан бошлаб ўрганилган, табиий тилга ишлов бериш, математик лингвистика ва машина таржимаси амалиётида қўллаш ҳаракатлари бошланган²²⁰. Шундан буён тадқиқот ва амалий ишланмалар натижасида бир қанча иловалар ишлаб чиқилган: ахборот-қидирув тизимида фаол қўлланувчи Портер стеммери, орфографик корректор (имло муҳаррири) шулар жумласидан²²¹.

Туркий тилларнинг морфологик анализаторларини ишлаб чиқиш ўтган асрнинг 60-йилларида бошланган²²². Дастлабки анализаторларнинг асосий хусусияти шунда эдики, улар айнан битта тилга мўлжалланмаган: бир морфоанализаторни бошқа тилга мослаштириш мумкин. Чунки дастурга морфотактик қоидалар бириктирилган, қоидалар тизими тилнинг лексикон

²²⁰Дрейзин Ф.А. Об алгоритмизации составление алгоритма анализа языка (на примере морфологии агглютинативного узбекского языка)//Научные труды Ташкентского университета. – 1961, вып. 189.Матем. Науки. – С. 121.

²²¹Kemal Altıntaşh Turkish to Crimean Tatar machine translation system(a thesis submitted to the department of computer engineering and the institute of engineering and science of Bilkent university in partial fulfillment of the requirements for the degree of master of science) 2001, P. 30.; Brian Roark, Richard Sproat. Computational Approaches to Morphology and Syntax. Oxford University Press Inc. – New York, 2007. – P. 63.; Rodolfe D. Computational Linguistic Text Processing: Lexicon, Grammar, Parsing and Anaphora Resolution. NovaSciencePublishers, Inc. – New York, 2008. –P. 4-5.

²²²Халитова Н.А., Закирова Р.А., Гимадудтинова Р.У. Морфологический анализ при машинном переводе с татарского языка на русский / Вероятностные методы и кибернетика I, Сборник работ НИИММим. Н. Г. Чеботарева при Казанском университете, Учен. зап. Казан.ун-та, 122, № 4, Изд-во Казанского ун-та, Казань, 1962. – С. 98-105.; Махмудов, Масуд Ахмед оглы. Разработка системы формального морфологического анализа тюркской словоформы: на материале азербайджанского языка: диссертация ... кандидата филологических наук:10.02.06. – Баку, 1982.; Проблемы моделирования тюркской морфологии : (Аспект порождения кирг. имен. словоформы) / Т. Садыков; АН КиргССР, Ин-тяз. и лит. – Фрунзе: Илим, 1987. – 120 с.

хамда аффиксал морфемалари билан тўлдирилса, бошқа тилнинг анализатори сифатида ишлатиш имкони пайдо бўлади; бошқа тилнинг анализаторини яратиш учун алоҳида кодлар ёзишга зарурат мавжуд эмас. Ўшандан ҳозирги даврга қадар технологиялар ўзгарди, туркий тилларнинг универсал морфологик анализаторлари яратилди, луғатлар ҳажми, маълумотни қайта ишлаш тезлиги ошди. Туркий тиллар морфоанализаторини ишлаб чиқиш ҳаракати бошланганига 50 йилдан ошганига қарамасдан, бу соҳа ўсиш ҳалигача туркий тилларнинг барчасида турли даражада. Татар²²³, бошқирд²²⁴, қозоқ²²⁵, чуваш²²⁶, турк²²⁷, хакас²²⁸ тиллари ҳамда универсал²²⁹ морфологик анализатор ҳақидаги ишлар фикримизни далиллайди.

Адабиётларда автоматик таҳлилнинг стемминг (1), сўзшаклнинг луғат асосидаги таҳлили (2), мантикий ёндашув асосидаги таҳлил (3), жадвал асосида, луғатсиз таҳлил каби турлари кўрсатилади²³⁰. Мутахассислар

²²³Сулейманов Д.Ш., Гильмуллин А.А., Гильмуллин Р.А. Базаморфотактических правил для татарского глагола как основа двухуровневого морфологического анализатора // Сборник трудов Международного семинара «Диалог». – Казань, 1998. – С. 597-609.; Сулейманов Д.Ш., Гатиатуллин А.Р. К разработке Лемматорататарских словоформ // В сб. Трудов Международного семинара Диалог-99 по компьютерной лингвистике и ее приложениям в двух томах (Таруса, 31 мая – 4 июня 1999 г.). Т. 2. – Таруса, 1999. – С. 306-314.; Сулейманов Д.Ш. Реализация татарской морфологии (наангл.яз.) // In Iss.: Computational linguistics and intelligent text processing: third international conference; proceedings / CICLing 2002, Mexico, February 17–23, 2002. Alexandr Gelbukh (ed.). – Berlin: Springer, 2002 (Lecture notes in computer science; Vol. 2276). – P. 327-329.; Сулейманов Д.Ш., Гатиатуллин А.Р. Структурно-функциональная компьютерная модель татарских морфем. – Казань: Фэн, 2003. – 220 с.

²²⁴Сиразитдинов З. А. Алгоритмическая грамматика словоизменения башкирского языка // [Электронный ресурс]. URL: <http://mfbl.ru/bashdb/algram/algram.htm>; Орехов Б.В., Слободян Е.А. Проблемы автоматической морфологии агглютинативных языков и парсер башкирского языка // Информационные технологии и письменное наследие: материалы международной научной конференции (Уфа, 28-31 октября 2010 г.) / отв. ред. В. А. Баранов. – Уфа; Ижевск: Вагант, 2010. – С. 167-171.

²²⁵Шарипбаев А.А., Бекманова Г. Тошкент, Ергеш Б.Ж., Бурибаева А.К., Карабалаева М.Х. Интеллектуальный морфологический анализатор, основанный на семантических сетях // Материалы международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем» (OSTIS-2012). – Минск, БГУИР, 16–18 февраля 2012 г. С. 397–400.

²²⁶Желтов П.В. Морфологический анализатор чувашского языка. Материалы Международной конференции студентов и аспирантов по фундаментальным наукам «Ломоносов 2002». – М., 2002.

²²⁷Kemal Ofizer. Two-level Description of Turkish Morphology. Literary and Linguistic Computing, – Vol. 9, No 2, – 1994.

²²⁸Дыбо А.В., Шеймович А.В. Автоматический морфологический анализ для корпусов тюркских языков // Филология и культура. – 2014. – №2.

²²⁹Архангельский Т.А. Принципы построения морфологического парсера для разноструктурных языков: диссертация кандидата филологических наук: 10.02.21. – Москва, 2012.

²³⁰Марчук Ю. Компьютерная лингвистика. – Москва, 2006. – С. 65.

автоматик морфологик таҳлил стемматизация (1), лемматизация (2), граммедалаш (3) каби асосий блоклардан ташкил топишини уқтиришади²³¹.

Стемматизация (ёки *стемминг* инг. *stemming* сўзидан) – қидирилетган сўзнинг асосини топиш жараёни бўлиб, бунда ҳар қандай стем (сўзшакл асоси) қидирилетган сўзшаклнинг грамматик асосига тенг бўлиши шарт эмас: таҳлил жараёнида бир парадигмага мансуб сўзшаклларнинг битта стем сифатида кўрсатилиши маълум стем(асос)ни аниқлаш учун етарли. М.Абжалова ушбу жараённи шундай тавсифлайди: “Стемминг, асосан, фойдаланувчининг сўрови бўйича изланишни кенгайтириш мақсадида қидирув тизимлари учун қўлланади, матнни меъёрлаштириш жараёни қисми ҳисобланади. Сўз асосини топишнинг муайян усули стемминг алгоритми, унинг амалга оширилиши, яъни дастурнинг ўзи стеммер дейилади”²³².

Демак, стем – бир парадигмага мансуб сўзшаклларнинг максимал умумий қисми. Масалан, татар тилидаги *китаплар* ва *китабим* сўзшаклларининг умумий асоси (стем) – *кита*. Ўзбек тилидаги *билагим*, *билакка* ва *билакни* сўзшаклларининг умумий асоси (стем) – *билак*; *нонни*, *нонлар*, *ноним*, *нондан* сўзшаклларининг умумий асоси (стем) – *нон*. Кўриниб турибдики, стем грамматик асосга тенг бўлиши ҳам, тенг бўлмаслиги ҳам мумкин.

Лемматизация сўзшаклни леммага келтириш жараёни бўлса, лемма сўзнинг асосий (луғат бериладиган) шакли, лексема ҳисобланади. Лемматизация жараёни кўпроқ флектив тиллар учун (масалан, рус тили) аҳамиятли, чунки агглютинатив тилларда (жумладан, ўзбек тили), одатда, сўзнинг “стем”и лексемага тенг бўлади. Фақат флекцияга учраган сўзшаклда лемма ва стемнинг мос келмаслик ҳолатлари кузатилади. Рус тилида от сўз

²³¹Формальные модели и программные инструменты компьютерной обработки татарского языка / Р.Р.Гатауллин, А.Р.Гатиатуллин, О.А.Неврозова, Д.Р.Мухамедшин, Д.Ш.Сулейманов, Б.Э.Хакимов, А.Ф.Хусаинов. – Академия наук РТ, Институт прикладной семиотики АН РТ. – Казань: Изд-во Академии наук РТ, 2019. – 260 с. – С.39-40.

²³²Абжалова М.А. Ўзбек тилидаги матнларни таҳрир ва таҳлил қилувчи дастурнинг лингвистик модуллари (Расмий ва илмий услубдаги матнлар таҳрири дастури учун): Филол. фан. бўйича фалсафа доктори (PhD) диссер. – Фарғона, 2019. – 164 б. – Б.29.

туркумига мансуб лемма – сўзшаклнинг бирликдаги именительный падеж шакли. Масалан, *столов* сўзшаклининг лемма кўриниши – *стол*. Сифатнинг лемма кўриниши – бирлик, “мужской” роддаги шакли: *зеленых* сўзшаклининг леммаси – *зеленый*.

Ўзбек тилида *от*, *олмош* учун лемма бирлик сон, бош келишидаги шакл: *бола, мен*; сифатнинг оддий даража кўриниши – *катта, ёруғ*; равиш, сон ҳамда ёрдамчи сўзларнинг леммаси уларнинг луғатдаги шаклига тенг. Туркий тилларда, жумладан, ўзбек тилида ҳам феъл сўз туркумининг лемма шакли – луғатда қўлланадиган “ноаниқ шакл”нинг *-моқ* қўшимчасисиз шакли: *ўтирмоқ, яратмоқ, айтмоқ, югурмоқ* лексемаларининг леммаси *ўтир, ярат, айт, югур* шаклига эга бўлади²³³.

Граммемалаш (теглаш инг. *tagging* сўзидан) сўзшаклга грамматик характеристика (грамматик белги) ёзиш демакдир. Граммема (грамматик характеристика) – сўзшаклнинг маълум морфологик синфга мансублигини кўрсатувчи содда морфологик кўрсаткич²³⁴.

Рус тилининг ўзлашма сўзлар луғатида *граммема* атамасига шундай таъриф берилади: “граммема (инг. *grammete*) – синтактик шакл ясовчи морфема ёки ёрдамчи сўз воситасида (мас., кўмакчи) ифодаланадиган грамматик маъно; икки (ҳатто ундан ҳам ортиқ) граммема бир морфема орқали ифодаланиши ҳам мумкин. Масалан, замон кўрсаткичи хабар майлини ҳам билдиради: *келди* (ўтган з., хабар м., аниқ н.). Фақат битта белгиси билан фарқ қилувчи граммема (масалан, бирлик ва кўплик) битта грамматик категорияни шакллантиради”²³⁵.

В.А.Плунгяннинг граммемани қуйидагича таърифлайди: “Граммема (инг. *grammete*) грамматик категория элементларидан бири бўлган грамматик

²³³Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: филология фанлари бўйича фалсафа доктори диссертацияси. – Бухоро, 2018. – 250 б. – Б. 13-137.

²³⁴Формальные модели и программные инструменты компьютерной обработки татарского языка / Р.Р.Гатауллин, А.Р.Гатиатуллин, О.А.Неврозова, Д.Р.Мухамедшин, Д.Ш.Сулейманов, Б.Э.Хакимов, А.Ф.Хусаинов. – Академия наук РТ, Институт прикладной семиотики АН РТ. – Казань: Изд-во Академии наук РТ, 2019. – 260 с. – С.39-40.

²³⁵Словарь иностранных слов русского языка/ *dic.academic.ru* > *dic.nsf* > *dic_fwords* > граммема

маъно бўлиб бир грамматик категориянинг турли граммемалари бир-биридан фарқланади ҳамда бир пайтда ифодалана олмайди. Рус тилида бирлик ва кўплик сон категориясининг граммемаси, албатта, сўзда ёки униси, ёки буниси воқеланади, лекин иккаласи эмас. Шунингдек, грамматик кўрсаткич ҳам граммема дейиладики, граммема морфологик категорияни ташкил этувчи морфологик шакллар билан ифодаланади, шунингдек, синтактик шакл билан ифодаланувчи граммема ҳам учрайди. К.Пайк томонидан таклиф қилинган *граммема* атамаси, А.А.Зализняк томонидан анъанавий лингвистик терминга айлантирилди²³⁶. Демак, граммема грамматик маъно, у грамматик категориянинг элементи, грамматик шакл билан ифодаланади. Масалан, *СТОЛ* леммасининг *стол* сўзшаклига қуйидаги граммемалар йиғиндиси бириктирилиши мумкин: (*мр, ед, им, неод*). Ахборотни бундай усул билан кўрсатиш мисол келтирганимиз рус тили мансуб бўлган флексив тилларга хос. Туркий тилларда граммема бириктиришнинг бу тартиби амал қилмайди, чунки битта морфема, кўпинча, бир граммемага тўғри келади; айнан бир морфема бир сўзшаклда қайта-қайта келиши мумкин. Битта сўзшаклда икки марта кўплик шаклининг қўлланишини кузатамиз. Татар тилида: *урманнардагыларга* – урман+лар+ДагЫ+ЛАР+ГА. Ўзбек тилида: *ўрмонлардагиларга* – ўрмон+лар+даги+лар+га. Татар тилида сингармонизм кучли бўлганлиги сабабли сўзшаклда учраган 2та кўплик шакли ўрнидан келиб чиқиб икки шаклда учрайди, яъни бошқа-бошқа алломорфлар қўлланади; ўзбек тилида бундай эмас: *-лар* қўшимчасининг алломорфлари мавжуд эмас. Бундай ҳолат тилимизда кўп учрашини инобатга олсак, ўзбек тилининг морфологик анализатори моделларини ишлаб чиқишда қатор энгилликлар бўлишини тасаввур қилиш қийин эмас. Шу сабабли туркий

²³⁶Плунгян В.А. Классификация морфологических значений //Общая морфология: Введение в проблематику: Учебное пособие. – Изд. 2-е, исправленное. – М.: Едиториал УРСС, 2003. – С. 107. – 384 с. // <https://ru.wikipedia.org/wiki/%D0%93%D1%80%D0%B0%D0%BC%D0%BC%D0%B5%D0%BC%D0%B0>; Бондарко А.В. Граммема // Лингвистический энциклопедический словарь /Главный редактор В.Н.Ярцева. – М.: Советская энциклопедия, 1990. – 685 с. – ISBN 5-85270-031-2.

²³⁶Ўша манба. – Б. 40.

тилларда граммема бириктиришда (грамматик шаклни тавсифлашда) кўшимчалар кетма-кетлигидан фойдаланиш яхшироқ натижа беради²³⁷.

Морфоанализаторда ишлатиладиган асосий тил бирликлари сифатида қуйидагилар ажратилади:

- 1) *морф* – элементар сегмент белги: шаклан тасодифан ўхшаш бўлган айнан бир морфологик хусусиятга эга бўлган минимал ҳодиса;
- 2) *алломорф*– бир морфеманинг бир хил фонетик таркибга эга бўлган морфлари йиғиндиси. Алломорфнинг асосий хусусиятларидан бири алломорф контекстини тавсифловчи хусусияти, у иккига: алломорфнинг сўзшакл чегарасидаги контексти (1) ҳамда алломорфнинг сўзшакл чегарасидан ташқаридаги контексти (2)га бўлинади;
- 3) *сўзшакл* – таркибий қисмлари орасида морфотактика орқали аниқланадиган алоқа мавжуд бўлган алломорфлар кетма-кетлиги²³⁸.
- 4) *морфема* – тилнинг сўзни ташкил этувчи энг кичик маъноли бирлиги; у битта вазифани бажарувчи, турли умумий, ўхшаш хусусиятларга эга бўлган морф (алломорф)лар йиғиндисидан иборат бўлади;
- 5) *бирикиш қоидалари* – морфоанализатор ишини ташкил этувчи асосий элементлардан бири. Бу қоидалароддий элементдан нисбатан мураккаб бирликларни келтириб чиқаради.

Туркий тилларнинг кўптилли полифункционал интернет-сервис дастури морфоанализатори бирикишлар қоидасининг икки типи: бир сўзшакл ичидаги бирикиш қоидалари; аналитик шаклларни келтириб чиқарувчи бирикиш қоидалари билан иш кўради²³⁹. Бундай қоидаларнинг мавжудлиги морфоанализатор иш жараёнида алгоритмлар кетма-кетлигини камайтиради, жараёни соддалаштиради. Демак, ўзбек тили морфоанализаторини ишга

²³⁸Формальные модели и программные инструменты компьютерной обработки татарского языка / Р.Р.Гатауллин, А.Р.Гатиатуллин, О.А.Неврозова, Д.Р.Мухамедшин, Д.Ш.Сулейманов, Б.Э.Хакимов, А.Ф.Хусаинов. – Академия наук РТ, Институт прикладной семиотики АН РТ. – Казань: Изд-во Академии наук РТ, 2019. – 260 с. – С.39-40.

²³⁹Кўрсатилган манба.

тушириш учун морф, алломорф, сўзшакл, морфема ва бирикиш қоидалари каби муҳим бирликлар ажратилиши талаб этилади.

Туркий тиллар морфологик таҳлини амалга оширишда жаҳонда мавжуд уч асосий ёндашув (парадигматик, автоматик, генератив)нинг парадигматик²⁴⁰, автоматик²⁴¹ каби турларидан фойдаланиш мумкин²⁴². Парадигматик ёндашувда икки типдаги луғатдан: асослар луғати ҳамда парадигмалар луғатидан фойдаланилади. Бундай усулда ишлайдиган морфологик анализаторнинг асосий хусусияти шундаки, луғатдаги ҳар бир леммага парадигмага ҳавола қилувчи индекс ёзилади. Парадигматик ёндашувдан, одатда, флектив тиллар (масалан, рус тили) морфологик таҳлилда фойдаланилади. Флектив тилларда парадигмалар ҳажми катта эмас, ammo парадигмалар сони кўп. Анализаторнинг маълумотлар омборида ҳар бир типдаги асос учун парадигманинг тўлиқ шакли сақланади. Бундай усул UniParser анализатори²⁴³да қўлланган. UniParserнинг маълумотлар омбори кўйидаги файлларни ўз ичига олади²⁴⁴:

1) асослар ажратиб кўрсатилган лексемалар рўйхати, синтактик шакл ясовчилар синфи ҳамда мазкур лексеманинг сўзшаклларига бириктирилиши талаб қилинадиган мавжуд лексик ахборотлар;

2) турли синтактик шакл ясовчи кўрсаткичлар рўйхати.

²⁴⁰Тузов В.А. Морфологический анализатор русского языка //Вестник СПбГУ, сер. 1. 1996. Вып. 1 (N15). –С. 41-45.; Сегалович И., Маслов М. Русский морфологический анализ исинтез с генерацией моделей словоизменения для не описанных всло варе слов // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог'99». Т. 2. – Казань, 1998.–С.547-552.

²⁴¹Antworth, E.L. PC-KIMMO: a two-level processor for morphological analysis. Occasional Publications in Academic Computing No. 16. Dallas:Summer Institute of Linguistics, 1990. – 273 p.; Kemal Ofi azer. Two-level Description of Turkish Morphology.Literary and Linguistic Computing, – Vol. 9, No 2, – 1994.

²⁴²Гатиатуллин А.Р., Баширов А.М. Морфологический анализатор тюркских словоформ на базе структурно-функциональной модели тюркской морфемы // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с. – С. 50-71.

²⁴³Архангельский Т.А. Принципы построения морфологического парсера для разноструктурных языков: диссертация ... кандидата филологических наук : 10.02.21. – Москва, 2012.

²⁴⁴Формальные модели и программные инструменты компьютерной обработки татарского языка / Р.Р.Гатауллин, А.Р.Гатиатуллин, О.А.Неврозова, Д.Р.Мухамедшин, Д.Ш.Сулейманов, Б.Э.Хакимов, А.Ф.Хусаинов. – Академия наук РТ, Институт прикладной семиотики АН РТ. – Казань: Изд-во Академии наук РТ, 2019. – 260 с. – С.39-40.

Аммо туркий тилларда ҳинд-европа тилларидан фарқланиб турадиган бир қатор структур хусусиятлар борки, парадигматик ёндашув асосида ишлаш қатор ноқулайликлар келтириб чиқаради. Бундай хусусиятлар сифатида автомат морфологиянинг ўнг томонлама амал қилиши (1); агглютинация (2); парадигматик синфлар орасида қатъий чегаранинг мавжуд эмаслиги (3); парадигма ҳажмининг потенциал чегараланмаганлиги (4); лексик қатламнинг грамматик синф ва сўз туркумлари бўйича аниқ таснифланмаганлиги(5)ни кўрсатиш мумкин²⁴⁵.

Агглютинатив, шунингдек, туркий тиллар учун морфологик анализатор тузишнинг автоматик ёндашуви кўпроқ тўғри келади. Автоматик ёндашувга асосланган морфоанализатор FST (finite state transducer) ва WFST(англ. *weighted finite state transducer*) қайта ишлаш тизимига эга, кириш ва чиқиш (анализ/синтез) таҳлилни амалга ошира олади. Бундай анализаторларнинг моҳияти шундаки, улар “грамматик кетма-кетлик” қоидасига амал қилади: жараёнда морфологик бирликнинг сўзшаклдаги кетма-кетлиги қоидаси асос қилинади. Улар қандай бирликлар ишлатилиши билан фарқ қилади:

- 1) морфемалар кетма-кетлиги ҳамда зарурий алломорф изчиллигини кўрсатиш;
- 2) алломорфлар изчиллиги қоидалари.

Демак, ўзбек тилининг морфологик анализаторини яратишда ҳам парадигматик ҳамда автоматик усуллардан фойдаланиш мумкин.

3.1.2. Морфоанализатор тузилиши ва таркиби. Морфоанализаторни ишга тушириш учун, одатда, маълум таркибий қисмлар шакллантирилади. Манинг таркибий қисми аффиксал морфемалар базаси; асос морфемалар базаси; таснифлаш қоидалари; алломорфлар мослиги қоидаларидан иборат

²⁴⁵Кўрсатилган манба.

бўлади (Қаранг: 5-чизма)²⁴⁶. Н.А.Исраилова, П.С.Бакасовалар морфологик маълумотлар омбори тузилишини қуйидагича тавсифлашади²⁴⁷:

- 1) морфологик маълумотлар базаси морфологик анализ ва синтез жараёни учун талаб қилинадиган барча ахборотларни қамраб олиши зарур;
- 2) тилда мавжуд флекция ҳолатлари, шундан келиб чиқадиган ўзак ва қўшимча чегарасидаги фонетик ўзгаришлар ҳам маълумотлар омборида акс этиши керак;
- 3) морфологик маълумотлар омбори одатий синтактик шакл ясовчилар билан бирга супплетив, ўзгармас лексемалар ҳақидаги маълумотларни ҳам қамраб олиши лозим;
- 4) маълумотлар омбори омоним лексемалар, шунингдек, тўлиқ ва грамматик шакл таъсирида ҳосил бўлувчи омонимлар базасига ҳам эга бўлиши лозим.

Юқоридаги фикрларга асосланган ҳолда айтиш мумкинки, морфологик маълумотлар омбори қуйидаги тузилишга эга бўлади (Қаранг: 6-чизма):

Таснифлаш (классификация) типлари. Туркий тил бирликлари орасида мураккаб сўзшакл структурасига эга бирликлар борлиги сабабли грамматик белгиларини тавсифлашда нотўғри (автоматик) изоҳлаш (теглаш) ҳолатлари кузатилади. Мисол тариқасида иккита турли морфема сифатида изоҳланадиган *-ган* морфемасини олиш мумкин: бири ўтган замон категориясини ифодалаш учун, иккинчиси сифатдош категориясини билдириш учун қўлланилади. Морфоанализатор таҳлил натижаси сифатида икки вариантни тақдим этади. Бундай саволларга жавоб топиш, муаммони бартараф этиш учун тил бирликларини таснифлашнинг турли типларидан фойдаланиш тажрибалари ҳақида тўхталамиз.

Морфемани қўшиб ёзиш ва ажратиб ёзиш параметри асосида таснифлаш. Туркий тилларда бир грамматик категорияга тегишли аффиксал

²⁴⁶Кўрсатилган манба. – С.68-69.

²⁴⁷Исраилова Н.А., Бакасова П.С. Морфологический анализатор кыргызского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В2-х томах. Т2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с. – С.100-117.

морфемалар ҳам синтактик, ҳам аналитик бирлик билан ифодаланиши мумкин, яъни бир туркий тилда улар кўшиб ёзилса, бошқасида ажратиб ёзилиши кузатилади. Масалан, татар тилида ўзидан олдин сўзни кейинги сўзга боғлашда иштирок этадиган “инструментатив” грамматик категорияси фақат *белән* кўмакчиси билан ифодаланади ҳамда ўзидан олдинги сўздан ажратиб ёзилади²⁴⁸. Ўзбек тилида (шунингдек, қозоқ, турк тилларида ҳам) “инструментатив” грамматик категорияси, яъни восита маъноси ҳам келишик, ҳам кўмакчи билан ифодаланади. Солиштиринг:

- қозоқ тилида: *-бен, -мен, -пен* алломорфларига эга *-бен* аффикси ҳамда *бенен, менен, пенен* кўмакчилари воситасида;
- турк тилида: *-ла* аффикси ҳамда *иле* кўмакчиси воситасида;
- ўзбек тилида: *-да* аффикси ҳамда *билан, ила* (*-ла* алломорфи билан бирга) кўмакчилари воситасида.

Ўзбек тилининг изоҳли луғатида билан кўмакчисининг маънолари ва берилган мисолларга назар ташлаймиз.

БИЛАН кўм. 1 Бош келишикдаги турли сўзлар билан ишлатилиб: 1) иш-ҳаракатнинг амалга ошишидаги биргалик, ҳамкорликни ифодалайди. *Отам билан келдим. Кўчаларни, қишлоқни, айниқса, Маҳваши билан бирга дарс тайёрлаган жойларни... жуда-жуда кўмсадим.* Т.Ашуров, Оқ от; 2) иш-ҳаракат ёки ҳодисага восита бўлган нарсани билдиради. *Ошининг таъми туз билан, одам таъми сўз билан.* Мақол. *Сут билан кирган – жон билан чиқар.* Мақол. *Косада лагмон кўтариб кирган Марғуба бир қўли билан шижани ушлади.* С.Зуннунова, Кўк чироқлар. *Инсон хатони кўпроқ ўз тили билан содир қилади.* Газетадан²⁴⁹.

²⁴⁸Формальные модели и программные инструменты компьютерной обработки татарского языка / Р.Р.Гатауллин, А.Р.Гатиатуллин, О.А.Неврозова, Д.Р.Мухамедшин, Д.Ш.Сулейманов, Б.Э.Хакимов, А.Ф.Хусаинов. – Академия наук РТ, Институт прикладной семиотики АН РТ. – Казань: Изд-во Академии наук РТ, 2019. – 260 с. – С.73.

²⁴⁹Ўзбек тилининг изоҳли луғати. Ёжилд. – Тошкент: ЎзМЭ, 2006. – 680 б. – Б.257.

Мисоллардан кўриняптики, билан кўмакчисининг иккинчи маъноси “иш-ҳаракат ёки ҳодисага восита бўлган нарсани билдиради”, демак, восита маъноси кўмакчи билан ифодаланади.

ИЛА *кт. қ. билан. Бой ила хизматчи. тт [Кумуш] ..мактубни ўқиб тамом қилди-да, жонсиз бир товуш ила «уятсиз» деб қичқирди ва ерга йиқилиб, ўзидан кетди. А. Қодирий, Ўткан кунлар*²⁵⁰.

Ушбу мисолда ҳам восита восита маъноси *ила* кўмакчиси билан ифодаланган.

Грамматик маъно билдирувчи ёрдамчи сўзлар кўшимча билан ҳам, сўз билан ҳам ифодаланади. Шу сабабли анализаторда лексик бирлик (L); аналитик шакл (AF); аналитик конструкция (AK) каби кўрсаткичлар фарқланади²⁵¹. *Лексик бирлик* – лексик маъно билдирадиган асос морфема. *Аналитик конструкция* – лексик маъно билдирмайдиган, фақат таркибли аналитик конструкция таркибида учрайдиган асос морфема. *Аналитик шакл*– аналитик шакл ясайдиган ёрдамчи асос морфема. Аналитик шакл ясовчилар аффиксал морфема каби маълум грамматик категорияга киради. *Грамматик категория* – бир умумий ёки турли (вазифаси ўхшаш) шакллар мажмуаси. Грамматик категория сўзшакл морфологик структурасининг категориал ёки комплекс форматдаги маълумот бериш функцияси танланганда морфоанализатор томонидан тақдим этилади.

3.1.3. Морфоанализатор лингвистик таъминотида асос морфемаларни таснифлаш. Ўзбек тилшунослигида сўзларни туркумга ажратишда бир неча ёндашувлар мавжуд. У.Турсунов бошчилигидаги ҳаммуаллифларнинг “Ҳозирги ўзбек адабий тили” дарслигида сўзларни туркумларга ажратиш борасида куйидаги фикрлар билдирилган: “Сўзлар ўз мавҳум грамматик маъноси, морфологик белгиси ва грамматик

²⁵⁰Ўзбек тилининг изоҳли луғати. II жилд. – Тошкент: ЎзМЭ, 2006. – 671 б. – Б.187-188.

²⁵¹Формальные модели и программные инструменты компьютерной обработки татарского языка / Р.Р.Гатауллин, А.Р.Гатиатуллин, О.А.Неврозова, Д.Р.Мухамедшин, Д.Ш.Сулейманов, Б.Э.Хакимов, А.Ф.Хусаинов. – Академия наук РТ, Институт прикладной семиотики АН РТ. – Казань: Изд-во Академии наук РТ, 2019. – 260 с. – С.74.

категорияларига қараб, сўз яшаш, шакл яшаш ва сўз ўзгартиш хусусияти ҳамда гапда бажарган вазифаларига қараб бир-биридан фарқ қилади. Сўзларнинг мавҳум грамматик маъноси, морфологик ва синтактик белгиларига қараб гуруҳларга бўлиниши сўз туркуми дейилади. Ҳар бир туркум ўзига хос мавҳум грамматик маънони англатади ва шу туркумга хос грамматик маъно ифодаловчи воситаларга эга бўлади. Жумладан, от туркуми “предметлик” маъносини (*кўча, бино, шаҳар, садоқат, дўстлик, киши* каби), феъл туркуми “ҳаракат”ни (*ўқи, ёз, ишла, кўкар, тинчи, бошла* каби), сифат туркуми эса “белги” маъносини (*яхши, катта, гўзал, баланд, ширин, қадимги, билимли* каби) англатади. От, сифат, феъл ва бошқа сўз туркумлари англатган бу мавҳум маънолар грамматик маънога киради. Бундай мавҳум грамматик маъно сўзларни туркумларга ажратишда асосий белги ҳисобланади. У, айниқса, лексик маъноси билан грамматик маъноси бир-бирига мос бўлган сўзларда катта ўрин тутаяди, сўзларни туркумлашда грамматик белгилар ҳам, албатта, ҳисобга олинади. Сўз туркумларини характерловчи грамматик белгилар хилма-хил. Сўзнинг морфологик белгилари, асосан, унинг махсус шакл яшаш ва сўз ўзгартириш шаклларига эга ёки эга эмаслигида кўринади. Масалан, от келиши, эгалик аффиксини олиб турланса, феъл шахс-сон, замон аффиксини олиб тусланади.

Сўзнинг гапда муайян гап бўлаги вазифасида келиши, бошқа сўзлар билан синтактик муносабатга кириша олиши сўз туркумининг синтактик белгисига киради. Масалан, предметликни билдирган от гапда эга, тўлдирувчи вазифасида келади. Бу отнинг муҳим белгиларидан ҳисобланади. Бошқа сўз туркумларининг синтактик жиҳатдан эга, тўлдирувчи вазифасида келиши учун уларда маъно торайиш ҳодисаси рўй бериши зарур (*Ёмонни танқид қил, яхшига тақлид қил*). Отлар гапда феъл, сифат, сон каби қатор сўз туркумлари билан синтактик алоқага кириша олади.

Ҳозирги ўзбек тилида сўзлар маъхум грамматик маънолари, морфологик ва синтактик белгиларига қараб беш гурӯҳга бўлинади: 1) мустақил сўзлар, 2) ёрдамчи сўзлар, 3) модал сўзлар, 4) ундовлар, 5) тақлидий сўзлар²⁵².

Юқоридаги матндан кўриниб турибдики, муаллифлар ўзбек тилидаги сўз туркумларни 5 гуруҳга ажратишади. “Замонавий ўзбек тили. Морфология (Тошкент, 2008)” дарслигида эса сўзларни туркумга ажратиш ҳақида кўйидагича фикр билдирилган²⁵³: “Лексема нутқда сўз шаклида юзага чиқар экан, бунда у грамматик жиҳатдан тугал шаклланган, яъни грамматик морфема билан бириккан ҳолда намоён бўлади. Албатта, лексема ҳар қандай грамматик морфема билан бирика олмайди. Шунинг учун *грамматик категория ҳамда шакл(форма)га хилма-хил муносабатига* кўра лексеманинг яна бир қирраси аниқланади. Лексема нутқда сўз шаклида гап таркибида маълум бир синтактик боғланиш, қуршов ва гап бўлаги вазифасида келади. Лексемага “яхлитлик” (бутунлик, субстанциаллик) назарияси асосида ёндашсак, *лексеманинг гап қурилишида тута оладиган ўрни* ҳам унинг қиррасидан бири саналади. Лексема серқирралигини изчил ва қарама-қаршилиқдан холи таснифнинг назарий асоси юзасидан билдирилган фикр билан бирлаштирсак, биз унинг ҳар бир қирраси бўйича алоҳида-алоҳида таснифни беришимиз ҳамда лексема гуруҳини аниқлашимиз лозим бўлади. Шу йўл билан лексеманинг ҳар бир таснифда тутган ўрни асосида кашф этган белгининг йиғиндиси унинг (лексеманинг) нисбий моҳиятини ташкил этиши равшанроқ тушунилади. Ўзбек тили лексемасининг юқорида санаб ўтилган қирраси асосида унинг уч алоҳида изчил таснифини берамиз:

биринчи тасниф: лексеманинг маъно хусусияти асосидаги изчил таснифи;

²⁵² Турсунов У. ва бошқ. Ҳозирги ўзбек адабий тили: Олий ўқув юрт. филология фак. талабалари учун дарслик. / У.Турсунов, А.Мухторов, Ш.Раҳматуллаев; [Масъул муҳаррир Ш.Раҳматуллаев]. – Қайта ишланган ва тўлдирилган нашри. – Тошкент: Ўзбекистон, 1992. – 399 б. – Б. 247-248.

²⁵³ Замонавий ўзбек тили. Морфология. (Масъул муҳаррир: Р.Р.Сайфуллаева). – Тошкент: Мумтоз сўз, 2008. – 522 б. – Б.430-431.

иккинчи тасниф: лексеманинг морфологик категория ва шаклга муносабати жиҳатидан таснифи;

учинчи тасниф: лексеманинг гап қурилишида тута оладиган ўрни ва бажара оладиган вазифасига кўра таснифи.

Ҳар бир тасниф натижасида лексеманинг алоҳида-алоҳида *маъновий (семантик), морфологик* ҳамда *синтактик* гуруҳи ажратилади. Ажратиладиган ҳар бир гуруҳ эса нисбий мустақил”²⁵⁴.

Компьютер лингвистикасида эса сўзларни туркумларга ажратишда бошқача ёндашувни кузатамиз. Асос морфема(сўз)лар маълумотлар омборидаги “асослар луғати”да туркумга ажратилган ҳолатда берилади, шу сабабли компьютер лингвистикасида ахборот-сўров тизими нуқтаи назаридан сўзларни туркумларга ажратишда мутахассислар қуйидаги критерийларга асосланишади²⁵⁵:

- семантик (сўзнинг категориал грамматик маъноси);
- синтактик (сўзнинг маълум гап бўлаги бўлиш, маълум гуруҳдаги сўзлар билан боғланиш имконияти);
- морфологик (шакл яшаш хусусияти ҳамда грамматик категория таркиби);
- деривацион (сўз ясалиш хусусияти);
- фонологик (турли гуруҳ сўзларининг фонема ва просодик (урғуга тегишли) структураси).

Шаклан бир хил бўлган сўзлар (омонимлар) бир нечта маъноси билан бирдан ортиқ туркумга мансуб бўлиши кузатилади. Шу сабабли морфоанализатор тузиш назарияси ва амалиётида асос морфемани фарқлашнинг турли форматлари ишлаб чиқилган:

1) асос типини фарқлашда классик сўз туркумига ажратишга амал қилиш;

²⁵⁴

²⁵⁵Формальные модели и программные инструменты компьютерной обработки татарского языка / Р.Р.Гатауллин, А.Р.Гатиатуллин, О.А.Неврозова, Д.Р.Мухамедшин, Д.Ш.Сулейманов, Б.Э.Хакимов, А.Ф.Хусаинов. – Академия наук РТ, Институт прикладной семиотики АН РТ. – Казань: Изд-во Академии наук РТ, 2019. – 260 с. – С.75.

2) асос типини структур-модел морфологик тип асосида фарқлаш.

Биринчи усулдаги тасниф филологнинг иш жараёнига тегишли бўлса, иккинчи усул дастурлаш жараёнида амал қилади²⁵⁶.

Барча асос морфемалар аффиксал морфемаларнинг структур-модел морфологик таснифи асосида N, A, V, D, S каби 5 морфологик гуруҳга ажратилади:

N – ўнг томонидан исмларга бирика оладиган барча аффиксал морфемаларни бириктирдиган асос морфемалар гуруҳи. Масалан, N = {-нинг, -ни, -га, -да, -дан; -(и)м, -(и)миз, -(и)нг, -(и)нгиз, -(с)и, -(лар)и, -лар, -ман, - миз, -сан, -сиз, -дир...};

A – ўнг томонидан сифатга тегишли қўшимчаларни бириктира оладиган барча морфемалар гуруҳи. Бу қўшимчалар гуруҳига N қўшимчалар ҳамда сифатнинг қиёсий даражаси тегишли. Масалан, A= N+{-роқ};

V – ўнг томонидан феъл ўзакларга бирика оладиган қўшимчаларни қабул қилувчи асос морфемалар гуруҳи. Масалан, V={-ма, -ди, -ган, -ажак...};

D – ўнг томонидан N қўшимчалар гуруҳи ва сон шаклини ҳосил қилувчи қўшимчалар бирикадиган асос морфемалар гуруҳи. D= {-инчи, -лаб, -овлон, -нинг, -ни, -га, -да, -дан, -(и)м, -(и)миз, -(и)нг...};

S – ўнг томонидан фақат модал аффиксларни бириктира оладиган асос морфемалар синфи. S={-дир, -ми...}.

Санаб ўтилган морфологик гуруҳлар ўз ичида морфемаларни бириктира олишига кўра яна ички гуруҳларга бўлинади²⁵⁷.

Шунингдек, морфоанализаторда морфонологик қоидалар ҳам амал қилади. Морфонологик қоидалар истисно тарзида ўзакка қандайдир қўшимча қўшилишидан ҳосил бўлган фонетик ўзгариш – ўзбек тилида кам бўлса-да, учраб турадиган флексия ҳодисаси асосида ўзгаришга учраган ўзаклар ёки истиснолар учун ишлаб чиқилади. Бундай ҳодисалар, асосан, кишилик

²⁵⁶Кўрсатилган манба. – С. 76.

²⁵⁷ Кўрсатилган манба. – С. 76.

олмошларида учрайди: *менинг, мени, сенинг, сени, унга, ундан* ва ҳ. Кишилиқ олмошларига келишиқ кўшимчалари, кўрсатиш олмошларига айрим кўшимчалар (-ча, -дек, -дай, -дақа) кўшилиши типик ҳолатлардан фарқ қилгани учун тилдаги бундай ўзақлар рўйхати шакллантирилиб, улар учун морфонологик қоидалар мажмуи тузилади.

Морфемаларни синтактик-семантик белгилар асосида таснифлаш. Туркий тилларда, жумладан, ўзбек тилида ҳам сўз контекстдаги позициясига қараб турли моҳиятни англатади. Масалан: объект, объект белгиси ёки фақат объект белгиси (*олтин, темир, кумуш, тош, асфальт*); ҳаракат белгиси (*тез, чиройли, яхши*). Контекстда билдирган маъносига кўра улар турли сўз туркумига кириши мумкин. Бу эса морфоанализ жараёнида кўпмаънолилиққа олиб келади. Шундай типдаги морфемалар билан ишлаш учун сўзларни туркумга ажратишда синтактик-семантик белги асосида таснифлаш принципи амал қилади. Шундан келиб чиқиб, адабиётларда туркий тилларга ҳос асослар позиция ва нопозиция турларга ажратилади²⁵⁸.

Позиция морфемалар – гапдаги ўрнидан қатъи назар, бир семантик гуруҳга мансуб бўладиган морфемалар. Асос морфема (сўз) объект маъносини билдирувчи сўзлар гуруҳига кирса, қандай сўзлар қуршовида келишидан қатъи назар, шу гуруҳга мансуб бўлаверади.

Позиция морфемалар – гапдаги ўрнига қараб турли семантик гуруҳга мансуб бўладиган асос морфемалар гуруҳи. Бундай морфемаларга барча модда-маъдан отлари кириши мумкин. Масалан, *олтин, ёғоч, темир*. Солиштиринг: *Уста ёғочни арралади // Боғбон гулни ёғоч тувакка ўтқазди*²⁵⁹.

²⁵⁸Формальные модели и программные инструменты компьютерной обработки татарского языка / Р.Р.Гатауллин, А.Р.Гатаиатуллин, О.А.Неврозова, Д.Р.Мухамедшин, Д.Ш.Сулейманов, Б.Э.Хакимов, А.Ф.Хусаинов. – Академия наук РТ, Институт прикладной семиотики АН РТ. – Казань: Изд-во Академии наук РТ, 2019. – 260 с. – С.79.

²⁵⁹Формальные модели и программные инструменты компьютерной обработки татарского языка / Р.Р.Гатауллин, А.Р.Гатаиатуллин, О.А.Неврозова, Д.Р.Мухамедшин, Д.Ш.Сулейманов, Б.Э.Хакимов, А.Ф.Хусаинов. – Академия наук РТ, Институт прикладной семиотики АН РТ. – Казань: Изд-во Академии наук РТ, 2019. – 260 с. – С. 80.

Позицион асосларга предмет белгисини билдирадиган сўз ҳам, ҳаракат белгисини ҳам атайдиган сўзлар ҳам мансуб бўлади.

Масаланинг шу томонини ёритишда тадқиқотчи Н.Ширинованинг тил бирликларининг предметлик/белгилик хусусияти, уларнинг юзага чиқишида локализациянинг аҳамияти ҳақидаги қарашларига таянамиз. Тадқиқотчи сўз бирикмаларида предметлик ва белги-хусусият маъноларини фарқлашда релевант аҳамиятли синтактик усуллар сифатида қуйидагиларни кўрсатади²⁶⁰: “предметликни маълум даражада мўътадиллаштириб, белги-хусусиятни кучайтиришнинг барча тиллар учун кенг қўлланиладиган омили тилшуносликда локализация²⁶¹ (жойлаштириш ёки ёнма-ён қўйиш (рус. соположение) бўлиб, бунда, одатда, тобе компонент предметлик маъносини деярли йўқотиб, белги-хусусият маъносини вокелантиради. Бу усул тилимизда жуда кенг қўлланилади: *тош кўприк, олтин соат...* Бундай бирикмаларда тобе бўлаклар (*тош, олтин*) сўзларида белги-хусусият маъноси шу даражада кучлики, айрим тилшунослар туркий тилларда от ва сифатомонимияси ҳақида (*олтин* – золото; *олтин* – золотой (-ое, -ая)²⁶², айрим тилшунослар эса конверсия ҳодисаси мавжудлиги ҳақида ёзадилар²⁶³”.

Шунингдек, тадқиқотчи сўз бирикмаларида локализация (жойлаштириш) синтактик омилни оддий синтактик битишув – ҳеч қандай кўшимчаларсиз тобе сўзнинг ҳоким сўзга боғланиши (*катта кўприк*)дан фарқлаш лозимлигига эътибор қаратади²⁶⁴. Битишувда тобе бўлак ҳоким бўлакка ҳеч қандай морфологик кўшимчаларсиз боғланса-да, битишувли бирикмаларда тобе бўлак билан ҳоким бўлак бир-биридан узилиши, улар ўртасида бошқа бўлаклар жойлаштирилиши мумкин, чунончи, *катта кўприк*

²⁶⁰Ширинова Н. Ўзбек тилида белги-хусусият ва предметлик маъноларининг фарқланишида градуонимик қаторлар: филол. фан. номз. диссер. – Тошкент, 2010. – 160 б. – Б. 75-76.

²⁶¹Мещанинов И.И. Члены предложения и части речи. –Л.: Наука, 1978. – С.110-116.

²⁶²Исхаков Ф.Т. Имя прилагательное. Исследования по сравнительной грамматике тюркских языков. Ч.П. Морфология. – М.,1956. – Стр.155-156; Миртожиев М. Ўзбек тилида омонимларнинг вужудга келиши: Филол. фан. номзоди дисс. – Ташкент, 1963, – Б. 48-54.

²⁶³Ғуломов А.Ғ. Конверсия// Ўзбек тили грамматикаси, I. – Б. 46-51.

²⁶⁴Назарова С. Бирикмаларда сўзларнинг эркин боғланиш омиллари. Филол. фанлари номзоди дисс.– Тошкент, 1997.

– катта янги кўприк, Кеча келдим – Кеча Тошкентдан келдим. Локализация шундай битишувки, сўз бирикмалари компонентларини бир-биридан узиш мутлақо мумкин эмас²⁶⁵.

С.Назарованинг таъкидлашича, тилимизда локализация ва оддий битишув отларда предметлик ва белги-хусусият маъноларини воқелантиришда бошқарувли сўз бирикмалари билан биргаликда²⁶⁶ ўзига хос даражаланиш сирасини ташкил этади. Чунончи: *чўян печь* → *чўяндан печь* → *чўян печи* → *чўянни эритмоқ*. Н.Ширинова бу сирада белги-хусусият маъносининг сусайиб, предметлик маъносининг кучайиб бориши ёки *чўянни эритмоқ* → *чўян печи* → *чўяндан печь* → *чўян печь*... сирасида предметлик маъносининг сусайиб, белги-хусусият маъносининг кучайишини кўрсатади. Н.Ширинова бундай даражаланиш қаторини тузишга асосий сабаб сифатида қуйидагиларни кўрсатади: “Тилимизда модда-маъдан, материал, хомашё номлари – *олтин узук, духоби дўппи, атлас кўйлак, тош кўприк* каби бирикмаларда тўла-тўқис белги-хусусият маъносини ифодалайди ва айрим тадқиқотчилар томонидан от-сифат омонимияси ёки конверсив қўлланилиш сифатида баҳоланади. Шахсни атовчи сўзларнинг ҳам сўз туркумлари доирасидаги ўрни мунозарали; бунинг ёрқин далили шахс оти ясовчи *-увчи, -вчи* қўшимчасининг айрим тилшунослар томонидан от ясовчи, айрим тадқиқотчилар томонидан сифат ва сифатдош ясовчи қўшимча сифатида талқин этилишидир. Ҳақиқатан ҳам, бундай сўзлар тилимизда икки хил синтактик бирикмаларни беради. Чунончи, *сувчи қиз ~ сувчи қизи, тошкентлик йигит ~ тошкентлик талаффузи*. Бундай сўзларда предметлик ва белги-хусусият маъноси фақат синтактик усул билан фарқланади”²⁶⁷.

²⁶⁵Ширинова Н. Ўзбек тилида белги-хусусият ва предметлик маъноларининг фарқланишида градуонимик қаторлар: филол. фан. номз. диссер. – Тошкент, 2010. – 160 б. – Б. 75-76.

²⁶⁶Назарова С. «Модда-маъдан оти + нарса-буюм оти» қолипли сўз бирикмаси (модда-ашё отлари ва нарса-буюм отларининг бирикуви – атов бириклар ҳақида)// Тил ва адабиёт таълими, 2004 й. – №5. – Б. 52-53.

²⁶⁷Ширинова Н. Ўзбек тилида белги-хусусият ва предметлик маъноларининг фарқланишида градуонимик қаторлар: филол. фан. номз. диссер. – Тошкент, 2010. – 160 б. – Б. 85-86.

Н.Ширинова А.Потебнянинг фикрига²⁶⁸ таяниб шундай хулосага келади: “Тилимизда сифат сўз туркумида предметлик ва белги-хусусият маънолари синкретизми шу қадар ривожланганки, халигача *бетоб, қари(я), касал, ўқувчи* сингари сўзларни от ёки сифат туркумига киритиш борасида мунозаралар давом этмоқда²⁶⁹.

Тадқиқотчи синсемантик от-сифатлар (ёки сифат-отлар)²⁷⁰га кишининг руҳий, иқтисодий, маънавий, баъзан жисмоний белгиларини номловчиларни киритади: *жинни, телба, девона, бетоб, қария, касал, бой, камбагал, етим, ботир, қаҳрамон* ва ҳ. Бундай сифатлар матнда белги-хусусият маъноси мўтадиллашган/предметлик маъноси кучайган кўринишда намоён бўлади: 1. *Бир куни бой камбагалдан сўрабди... (Эртақдан)*. 2. *Ўзга юртда шоҳ бўлгунча ўз юртингда гадо бўл*. 3. *Бечора, етимларга садақа улашишни канда қилмасди*²⁷¹. Тадқиқотчи бундай сифатларнинг предметлик маъносини матнсиз, қуршовсиз ҳам бемалол ифодалай олашини, шу хусусияти билан улар мазкур градуонимик қатордаги соф сифатлардан, сифат-отлардан фарқ қилишини таъкидлайди. Морфологик ва синтактик воситаларнинг иштироки эса уларнинг бундай қобилиятини янада мустақкамлайди ҳамда ўзига хос шароит ва қуршовлар билан таъминлайди²⁷².

Позицион морфемаларнинг хусусиятини кузатиш мақсадида изоҳли луғатдан айрим сўзларнинг шарҳига диққат қаратамиз.

ЗУМРАД [ф.<юн. smaragdos – сабзаранг тош] **1** Олий сифатли, оч яшил тусли қимматбаҳо тош. *Кўрдингми? Бу нима? Мана бу – дур, мана бу – марварид. Мана бу ёнган – гавҳар. Қолганлари – ҳар хил тошлар, ёқут, зумрад [деди Мирзакаримбой].* Ойбек, Танланган асарлар. *Қозикалон сандиқларни очиб кўрсатди: баъзисида олмос, ёқут, ақиқ, зумрад ва марварид, баъзисида*

²⁶⁸Потебня А.А. Из записок по русской грамматике. Ч.І. – Харьков, 1888. – С. 32.

²⁶⁹Ширинова Н. Ўзбек тилида белги-хусусият ва предметлик маъноларининг фарқланишида градуонимик қаторлар: филол. фан. номз. диссер. – Тошкент, 2010. – 160 б. – Б. 91.

²⁷⁰Кўрсатилган манба.

²⁷¹Кўрсатилган манба. – Б.98-99.

²⁷²Кўрсатилган манба. – Б. 99-101.

олтин ва кумушдан ясалган, қимматбаҳо тошлар билан зийнатланган идиш-оёқлар, баъзиларида зеб-зийнатлар тўла эди. С.Айний, Дохунда.

2 кўчма сфт. Оч яшил, сабзаранг. Молдавия далаларининг зумрад ранги бизнинг лалмикор бугдойзорларимизни эслатади. Н.Сафаров, Оловли излар. Воҳ, юрагим ўйнарди кумуш жилғаларнинг зумрад лабида. Ш.Раҳмон, Юрак кирралари. Ариқ бўйида қалин ёш тол, ўртасида ёғоч сўри. Унинг тагидан шовуллаб оқиб турадиган зумрад сувдан доим бир хилда салқин шабада эсади. И.Раҳим, Ихлос.

3 Зумрад (хотин-қизлар исми)²⁷³.

ПЎЛАТ [ф. – жуда қаттиқ кулранг металл] **1** Темирнинг маълум микдордаги углерод билан қўшилишидан ҳосил бўлган, кулранг, чўкичланувчи қаттиқ металл. *Зангламайдиган пўлат. ит Бе-е, менда пўлат нима қилади, тақсир. Эски тақалар, темир-терсақлар билан зўрға қимирлаб турибмиз.* М. Исмоилий, Фарғона тонг отгунча.

2 Шу металлдан тайёрланган, ясалган. *Пўлат сим. Пўлат қайчи. Пўлат тиш. Пўлат пичоқ қинсиз қолмас.* Мақол.

Пўлат из *поэт.* Темир йўл. Янги машиналар келар бирин-кетин, Дўстлик қура олган пўлат излардан. Зулфия. **Пўлат от** *поэт. айн. трактор.* *Пўлат отни ўйнатиб, Тракторчи бўламан.* Қ.Муҳаммадий. **Пўлат сандиқ** *қ. сандиқ. Пўлат қуш* *поэт. айн. самолёт.* *Кўкда ранго-ранг пўлат қушлар ҳаммиша посбон.* Ҳабибий, Девон.

2 фольк. Қилич, шамшир. *Олмос пўлат боғлаб мен ҳам белимдан, Дев излаб чиққанман ўсган элимдан. “Маликаи айёр”. Пўлат билан қирди девни, кескир ханжарини урди, ялангочлаб қилич солди. “Маликаи айёр”.*

3 кўчма Метиндай мустаҳкам, бузилмас, енгилмас. *Пўлат қалъа.*

4 Кучли, забардаст. *Қиш. Деҳқон аҳлининг пўлат мускулларига ҳордиқ берадиган, кўнгил ёзадиган фасл. “Ўзбекистон кўриқлари”. Бизнинг пўлат қудратимиз – бирлигимизда.* М.Шайхзода.

²⁷³Ўзбек тилининг изоҳли луғати. 5 жилдли. 2- жилд. – Тошкент: ЎЗМЭ, 2006. – Б. 165.

5 Пўлат (эркаклар ва хотин-қизлар исми)²⁷⁴.

ҚУМ I 1 Бир-бири билан бирлашмайдиган, майда заррачалардан ташкил топган чўкинди тоғ жинси (асосан қурилишда ва қурилиш материаллари ишлаб чиқаришда хомашё сифатида фойдаланилади). *Қумдан қўрғон бўлмас. Мақол...дарвоза олдиға тўкилган қумда ўйнаб ўтирган болақайларға бирнас завқ билан қараб ўтирди. Н.Аминов, Суварак. Қурилиш материаллари оёқ остида тўкилиб, оҳақ қумға, қум тупроққа қоришиб кетади. “Муштум”.*

Қум соат Муайян миқдордаги қумнинг бир ёпиқ идишдан иккинчисига тор тешик орқали тушишига қараб вақтни ўлчовчи соат. *Ўрта асрларда бандли қум соатлар урф эди. “Фан ва турмуш”.*

2 айн. қумлик. *Карвонлар қумда адашиб, саргардон бўлмайдилар. Мирмуҳсин, Меъмор. Бир йигит чирози ўчирилган мотоциклини қўюндай учуриб, қум ичкарасига кириб кетди. С.Сиёев, Ёруғлик.*

3 сфт. *Қумға ўхшаш, майда. Қум шакар. ■ Жўрапоиша бўйнидан қум марварид маржонни чиқарди-да, ипини шарт-шурт узиб, марваридларни кигизга тўкди. М.Исмоилий, Фарғона тонг отгунча²⁷⁵.*

Юқоридаги учта мисолдан кўриниб турибдики, *зумрад, пўлат, қум* сўзларининг биринчи маъноси модда-маъда отлари, иккинчи маъноси сифат деб кўрсатилган. Бундан хулоса қилиш мумкинки, бу сўзлар позиция морфемалар, матндаги сўзлар қуршовида унинг маъноси реаллашади. Бундай морфемаларнинг маъносини фарқлаш, семантик таҳлил босқичида амалга ошади.

Позицион морфемалар билан ишлаш морфоанализаторнинг икки режими учун: кўп маъноли – фойдаланувчи ҳамда бир маъноли – тилга компьютер ишлови бериш тизими учун алоҳида аҳамиятли. Мутахассисларнинг фикрича, бу морфологик кўпмаънолилик эмас, балки семантик-синтактик кўпмаънолилик ҳодисасидир²⁷⁶.

²⁷⁴Ўзбек тилининг изоҳли луғати. 5 жилдли. 3- жилд. – Тошкент: ЎзМЭ, 2006. – Б. 328.

²⁷⁵Ўзбек тилининг изоҳли луғати. 5 жилдли. 5- жилд. – Тошкент: ЎзМЭ, 2006. – Б. 376.

²⁷⁶Кўрсатилган манба.

Юқорида морфоанализаторнинг таркибий қисми ҳақида тўхталиб, асослар луғатининг ушбу таркибга киришини эслатгандик. Қуйида шу масалани муҳокама қиламиз.

Мутахассислар морфоанализатор ишини ташкил этувчи таркибий қисмлар сирасига *асос лексемалар лексиконини* ҳам киритишади²⁷⁷. *Асос лексемалар лексикони* ҳозирги адабий тил асосида шакллантирилади. Одатда, тил луғат таркибини нисбатан тўлароқ қамраб олган луғатлар лексикон тузишга асос бўлади. Кузатишларимиз анъанавий морфоанализаторлар лексикони таркибининг HFST инструментарийси талаблари асосида тўлдирилган Nouns (от), Verbs (фeyл), Adjectives (сифат), Adverbs (равиш), Pronouns (олмош), Numerals (сон), Postpositions (кўмакчи), Conjunctions (боғловчи), Exclamations (ундов) каби 9 гуруҳдан иборат бўлишини кўрсатади. Ички лексиконлар қуйидаги уч қисмдан иборат бўлади:

- 1) лексик атом – маълум бир тилга мансуб асос сўз;
- 2) кетма-кетлик гуруҳлари – морфотактик схемага асосланган, морфема гуруҳларининг изчиллигини таъминловчи қоидалар мажмуи;
- 3) трактовка – грамматик белгилар тавсифи. Бу ўринда лексик birlikнинг мавжуд барча морфологик, грамматик, лексик, семантик хусусиятлари ўз аксини топади.

М.Абжалова графематик таҳлил ва таҳрир дастури модулини ишлаб чиқар экан, дастур лингвистик таъминоти учун асос бўладиган луғатларга алоҳида тўхталади. Бунда морфологик таҳлил босқичида морфологик луғатнинг кераклигини ўринли таъкидлайди: “Ўзбек тилидаги матнларни автоматик таҳрир қилиш дастурининг морфологик таҳлил қилиш босқичи модулларини яратишда ҳам морфологик луғат жуда муҳим. Бундай луғатнинг йўқлиги боис асосий манба вазифасини бажарувчи сифатида морфем луғатлардан фойдаланилди. Бундай луғат биринчи марта А.Ғуломов, А.Н.Тихонов, Р.Қ.Кўнғуров томонидан тузилиб, 1977 йилда “Ўқитувчи”

²⁷⁷Кўрсатилган манба.

нашриёти томонидан чоп этилган²⁷⁸. Имло луғатларидаги лексемаларнинг қайси сўз туркумига оидлигини аниқлашда “Ўзбек тилининг изоҳли луғати” муҳим ёрдамчи манба ҳисобланади, яъни лексеманинг ўзига қандай аффиксларни бириктира олиши, аффикслар комбинацияларини тузишда “Ўзбек тилининг морфем луғати”га таянилади²⁷⁹. Шунингдек, тадқиқотчи, лингвистик таъминот таркибида талаб қилинадиган асосий луғатлар ҳақида шундай хулосага келади: “Ҳар қандай дастурнинг лингвистик базасини яратишда “Ўзбек тилининг морфологик луғати” энг керакли манба саналади. Морфологик луғат тилшуноснинг вақтини тежайди; белгиланган натижага эришишни тезлаштиради. Зеро, бундай луғатда бир лексеманинг қайси туркумга оидлиги, унинг сўзшакллари келтирилади. Бу эса дастурнинг мукамал чиқишини таъминлайди. Шу боис ўзбек тилидаги матнларга ишлов берувчи дастурнинг лингвистик таъминотини яратишда қулайлик ва натижага тезкор эришиш мақсадида ушбу луғатнинг ишлаб чиқилиши мақсадга мувофиқ”²⁸⁰.

Юқоридаги фикрлардан кўриниб турибдики, ҳар қандай анализаторнинг (графематик, морфологик, семантик, синтактик) лингвистик таъминотида, албатта, асослар луғати – лексикон бўлиши талаб этилади. Ўзбек тили морфоанализатори лингвистик таъминоти лексикони учун 85 000 сўзни қамраб олган “Ўзбек тилининг имло луғати”²⁸¹ ёки беш жилдлик “Ўзбек тилининг изоҳли луғати”²⁸² асос қилиб олиниши; шу луғат сўзликлари асосида лексикон базаси қуйидаги таркибда шакллантирилиши мумкин:

1) от-лексикон (Lexicon nouns) от-ўзақдан ташкил топади;

²⁷⁸Ғуломов А. ва бошқ. Ўзбек тилининг морфем луғати. – Тошкент. 1977 – 463 б.; Самад А. “Х” ва “Ҳ”ли сўзларнинг имловий луғати. – Тошкент, 2007. – 346 б.; Менглиев Б., Баҳриддинова Б. Ўзбек тилининг сўз таркиби ўқув луғати. – Тошкент: Янги аср авлоди, 2007.

²⁷⁹Абжалова М.А. Ўзбек тилидаги матнларни таҳрир ва таҳлил қилувчи дастурнинг лингвистик модуллари (Расмий ва илмий услубдаги матнлар таҳрири учун): Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Фарғона, 2019. – 164 б. – Б.45.

²⁸⁰Абжалова М.А. Кўрсатилган манба. – Б.62.

²⁸¹Ўзбек тилининг имло луғати. (Тузувчилар: Э.А.Бегматов, А.П.Мадвалиев; Н.Маҳмудов таҳрири остида). – Тошкент: Akademnashr, 2013. – 520 б.

²⁸²Ўзбек тилининг изоҳли луғати. 5 жилдли. – Тошкент: ЎЗМЭ, 2006-2008. – 680 б.

- 2) феъл-лексикон (Lexicon Verbs) феъл-ўзакдан ташкил топади;
- 3) сифат-лексикон (Lexicon Adjectives) сифат-ўзакдан ташкил топади;
- 4) умумий лексикон сон, олмош, равиш, боғловчи, кўмакчи, юклама, ундов, модал, тақлид-ўзакдан ташкил топди.

Ўзбек тили қатъий қоидаларга бўйсунувчи агглютинатив хусусиятга эга бўлса ҳам, унинг грамматик қоидалари сирасида (хинд-европа оиласига мансуб тилларда бўлганидек) айрим истиснолар учраб туради. Масалан, сифатнинг орттирма даражаси дастлабки бўғиннинг такроридан ҳосил бўлиб, асосдан олдин дефис билан ёзилади: *қип-қизил, қоп-қора, яп-янги* ва Ҳ. Морфоанализатор лингвистик таъминотини шакллантириш учун ўзбек тилидаги бундай шаклларнинг барчаси аниқланиши, рўйхати шакллантирилиши, истиснолар луғати кўринишида сифат лексиконининг маълумотлар базасига киритилиши талаб этилади.

Б.В.Орехов бошқирд тили матнларини морфологик разметкаш муаммоларига тўхталар экан, исталган морфологик таҳлил дастури асосида икки асосий таркибий қисм: грамматик луғат ҳамда тилдаги синтактик ва луғавий шакл ясовчиларнинг формал тавсифи талаб қилинишини таъкидлайди²⁸³. Булар ичида асосий муаммо мукамал грамматик луғатнинг мавжуд/мавжуд эмаслигида. Тил луғат таркибини, максимал даражада лексемаларни қамраб олган грамматик луғатнинг мавжудлиги автоматик таҳлил учун лингвистик таъминот яратишни бир қадар осинлаштиради. Агар шундай луғат мавжуд бўлса, унда лексемалар қамров даражаси, лексик қатлам нуқтаи назаридан сўзликнинг қанча қисмини ўзлашма сўзлар ташкил этиши каби параметрларни аниқлаш талаб қилинади. Унинг фикрича, бошқирд тилида М.Х.Ахтямованинг 20000дан иборат грамматик луғатида лексемаларнинг катта қисми ўзлашмалар, шулар орасида ҳам *авиапорт, агренимия, мальтузиансылык* каби ишлатилиш частотаси паст бўлган сўзлар,

²⁸³Орехов Б.В. Проблемы морфологической разметки башкирских текстов // Труды Казанской школы по компьютерной и когнитивной лингвистике TEL-2014. – Казань: Фэн, 2014. – С. 135-140.

шунингдек, историзмлар катта фоизни ташкил этади. Албатта, бундай грамматик луғат морфологик таҳлил асосини берувчи асослар луғатини шакллантириш учун лингвистик таъминот вазифасини бажаришга ярамайди.

А.В.Чемишев компьютер таҳлили дастурлари ясама сўзлар моделини тузиш ҳақида тўхталар экан, *динамик лексикон* тузиш эҳтиёжи мавжудлигини ҳам таъкидлайди²⁸⁴. Тадқиқотчи бундай морфем луғатга асос қилиб 40.000 сўздан иборат имло луғатини (“Марийский орфографический словарь”, 2011) асос морфемалар рўйхатини тузишнинг дастлабки босқичида лингвистик таъмин сифатида танлайди. Кейинги босқичда дастурнинг “асослар луғати”ни бошқа луғатлар сўзлигидан фойдаланиб тўлдиришни таклиф этади. Бунда махсус белгилар билан белгиланадиган сўзлашув нутқидаги сўз; ишлатилиш диалекти кўрсатилган ҳолда шева сўзлари; агар имло қоидасида ўзгариш мавжуд бўлса, сўз имлосининг эскирган шакли эканлигини сақловчи ахборот/тег қўшилган ҳолда бугунги кунда имлоси ўзгарган сўзларни киритиш лозимлигини маъқул кўради. Чунки таҳлил учун бериладиган матнларда нутқнинг шундай хусусиятлари акс этадиган ҳолатлар учрайди, уларни тўғри таҳлил қилиш учун умумий қоидалардан ташқари мана шундай бирликлар умумлашмаси ҳам лингвистик таъминот базасидан ўрин олиши керак бўлади. Шунингдек, сўз ясалиши уяларини ҳам шакллантиради.

Орфографик луғат асосида шакллантирилган асослар луғатини сўз туркумлари бўйича теглайди (Қаранг: 7-чизма). Асослар туркумини белгилаш жадвали биринчи устунида сўз; иккинчи устунида сўз туркумини билдирадиган грамматик тег (**N** – от, **A** – сифат, **Num** – сон, **Pron** – олмош, **V** – феъл, **Prc** – равиш, **Ger** – сифатдош, **Adv** – равишдош, **Po** – кўмакчи, **Conj** – боғловчи, **Pcle** – юклама, **Interj** – ундов, **Imitat** – тақлид сўз); учинчи устунда “!” белгиси билан сўз урғуси акс эттирилган. Гарчи кўп учрамаса-да, айрим

²⁸⁴Чемишев А.В. Динамический лексикон: компьютерное представление моделей словообразования // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с. – С.224-231.

сўзларда урғунинг ўзгариши сўзнинг туркумида ўзгариш бўлишига олиб келади. Кейинги босқичда жадвалга асоснинг ясалиш асоси (агар ясама сўз бўлса), сўзнинг туркумига хос морфологик белгилари қўшилади (Қаранг: 8-чизма; 1-жадвал).

Юқоридаги таҳлиллардан хулоса шуки, базада бундай маълумотларнинг мавжудлиги матнларнинг морфологик, фонетик, сўз ясалиши, сўзшакллар таҳлили (шунингдек, тарихий сўзларни ҳам) автоматик анализини амалга ошириш имконини яратади. Ушбу тажрибага таянган ҳолда, ўзбек тили морфоанализатори учун динамик лексиконни шакллантиришда морфологик таҳлилни мукамаллаштириш имконияти сифатида қараш тўғри бўлади.

3.2. Ўзбек тили морфоанализатори таҳлил бирликлари, сўзшакл ва морфемалар

3.2.1. Ўзбек тили морфоанализаторида грамматик категория.

Компьютерда табиий тилга ишлов беришнинг муҳим омили морфоанализ эканлиги тадқиқотчилар томонидан такрор-такрор таъкидланади. Чунки таҳлилнинг барча шакллари: графематик, орфографик, синтактик, семантик таҳлилларга морфологик таҳлил ва унинг натижаси асос бўлади. Шунингдек, мутахассислар автоматик таржимада сўзшакллар морфоанализи грамматик категорияларини моделлаштириш машина таржимаси сифатини такомиллаштирадиган омил эканлигини таъкидлашади; дастурнинг лингвистик базаси грамматика ва луғатдан иборат бўлишини уқтиришади²⁸⁵. Морфологияда грамматик маънони изоҳлаш, деривация, сўзшакл ўзгариши қоидалари тавсифига лингвистик амалиёт жараёни сифатида қаралади. Табиий тил (ўзбек тили ҳам) компьютер қайта ишловини йўлга қўйиш учун грамматик

²⁸⁵Абдурахмонова Н. Моделирование грамматических категорий и аналитических глаголов узбекского языка как этап морфологического анализа в машинном переводе // V Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Издательство Ака де мии наук Республики Татарстан, 2017. – 327 с. – С. 155-180.

аҳборотнинг етишмаслиги лингвистик маълумотлар омбори яратиш эҳтиёжини келтириб чиқариши ҳақидаги фикрлар ҳам асосли.

Ишнинг олдинги қисмларида морфоанализаторнинг амалий аҳамияти ва функциялари баён қилинган эди. Қуйида морфоанализаторнинг таркибий қисмлари ҳамда уни ишга туширувчи энг кичик birlikлар хусусида тўхталамиз. Морфоанализатор унинг ишини ташкил этувчи, функцияларини ташкиллаштирувчи бир неча муҳим birlikлардан ташкил топади. Ишнинг олдинги бўлимида морфема, алломорф, граммема каби birlikлар тавсифланган эди. Шу birlikлар қаторида морфоанализ жараёнини автоматлаштирувчи асосий тушунчалар сифатида грамматик категория ва граммема борасида фикр юритамиз.

Грамматик категория ва грамматик шакл масаласи ҳар доим грамматиканинг долзарб ва муаммоли масаласи бўлиб келган. *Категория* асли фалсафий термин, манбаларда "...объектив борлиқ ва билишдаги моҳиятан кўпроқ қонуний алоқа ва муносабатни акс эттирувчи умумий тушунча" деб изоҳланади. Фалсафада ҳам, тилшуносликда ҳам категория бир хил нарсани атайди, яъни категория учун жуфтлик ва яккалик бўлиши шарт ва зарур. Фалсафада жуфтликни сабаб-натижа, моҳият-ходиса, бутун-қисм; яккаликни борлиқ, миқдор, макон ташкил этса, тилшуносликда категория атамаси остида қатъий зидланган ва ўзига хусусий белгига эга бўлган шакл тизими бирлаштирилади. Демак, грамматик категория деганда муайян маъно умумийлиги асосида бирлашган ва ўзаро зидланувчи шакл системаси тушунилади. Грамматик категория шаклнинг оддий арифметик йиғиндиси эмас, балки маълум турдаги шаклнинг барқарор муносабати тизмасидан иборат бўлган янги бир бутунлик²⁸⁶.

Ўзбек тили академ грамматикаси²⁸⁷ да грамматик категория қуйидагича таърифланади: "грамматик маъно билан унинг формал ифодаси биргаликда

²⁸⁶ Замоновий ўзбек тили. Морфология. (Масъул муҳаррир: Р.Р.Сайфуллаева). – Тошкент: Мумтоз сўз, 2008. – 522 б. – Б. 86.

²⁸⁷ Ўзбек тили грамматикаси. I том. Морфология. – Тошкент: Фан, 1975. – 609 б. – Б.61-62.

бир бутунлик ҳосил қилувчи, бири бошқасисиз мавжуд бўлмайдиган, ажралмайдиган, ажралмас ҳодисалар бўлиб, уларнинг бирлиги *грамматик категория*ни ташкил қилади. *Грамматик категория* термини грамматик маънога нисбатан кенг. Масалан, отларда кўплик формаси унинг бирлик формасига зиддир (кўплик маъноси *-лараффикси* орқали ифодаланади, бирлик эса бундай махсус кўрсаткичнинг йўқлиги – ноль форма билан характерланади: бундай ҳолларда бир белгининг йўқлиги ҳам, бошқаларига нисбат қилганда, ажратувчи белгидир), бу бирлик-кўплик формасидаги грамматик оппозиция сон категориясини ташкил қилади (грамматик категория шундай зид ҳолдаги ҳодисаларнинг, грамматик оппозицияларнинг, умумлаштирувчи категорияси сифатида бўлади), келишик формаларининг йиғиндиси келишик категориясини ташкил қилади; замон категорияси ўтган, ҳозирги ва келаси замон формаларини ўз ичига олади. Демак, бир хил грамматик маънони ифодалайдиган формаларнинг йиғиндиси *грамматик категория*ни ташкил қилади.

Демак, грамматик категория **кенг маънода** сўз туркумларини ўз ичига олади, **тор маънода** сон, замон каби умумлашган маънолари, форма ясашни кўрсатувчи категорияларни билдиради (бу маънолар маълум форма билан ифодаланган бўлади), сўз туркумлари (булар: лексик-грамматик категориялар) энг “йирик”, “юқори” грамматик категорияни кўрсатади. Грамматик категория сўзга хос бўлган, умумлашган характердаги умумий маънолар бўлиб, грамматик маъно бу умумий маъноларнинг хусусий кўринишидир (умумий: замон категорияси – хусусий: ҳозирги замон маъноси каби). Грамматик категориялар бевосита морфологияга хос бўлган – сўз туркумлари, сон, келишик, замон каби ҳодисаларни (морфологик категорияларни), шунингдек, синтагма, эга, кесим, тўлдирувчи каби ҳодисалар(синтактик категориялар)ни ҳам ўз ичига олади. Лекин бу ҳодисалар (масалан, сўз туркуми билан гап бўлаги, сўз формаларининг, синтактик формаларнинг, сўзларни бириктиришдаги вазифаси) бир-бири билан шу қадар боғлиқки, бу ҳол

грамматик категорияларни морфологик категориялар ва синтактик категориялар деган номлар билан ажратишнинг жуда ҳам шартли эканлигини кўрсатади”.

У.Турсунов ва бошқалар ҳаммуаллифлигидаги “Ҳозирги ўзбек адабий тили” дарслигида грамматик категорияга қуйидаги таъриф келтирилади²⁸⁸ : “ўзаро бир-бирига зид қўйилган ва қиёсан олинган грамматик маънолар системаси ва уларга хос ифода воситалари – грамматик шакллар системаси *грамматик категория* дейилади. *Грамматик категориялар*да тилнинг ўзига хос хусусияти очиқ намоён бўлади. Масалан, рус тили учун хос грамматик жинс категорияси туркий тилларда, жумладан, ўзбек тилида йўқ; рус тилида сифатларга келишик категорияси хос, ўзбек тилида эса сифат, одатда, турланмайди ва бошқалар”.

Грамматик шакл умумий грамматик маъноси муаммосига жаҳон тилшунослигида 30-, туркийшуносликда 60-йиллардан бошлаб С.Н.Иванов, В.Г.Гузев, Ҳ.Ғ.Неъматов, А.Нурмонов, Ғ.Н.Зикриллаев тадқиқотида алоҳида эътибор қаратилди²⁸⁹. Замонавий ўзбек тилшунослигида эса Ш.Шаҳобиддинова, Б.Менглиев, Б.Баҳриддинова, О.Шукуров тадқиқотларида бу масала атрофлича текширилди ва ўзбек тили барча морфологик категорияси лисон-нутқ бўлинишида тадқиқ этилди²⁹⁰.

²⁸⁸Турсунов У. ва бошқ. Ҳозирги ўзбек адабий тили: Олий ўқув юрт. филология фак. талабалари учун дарслик. / У.Турсунов, А.Мухторов, Ш.Раҳматуллаев; [Масъул муҳаррир Ш.Раҳматуллаев]. – Қайта ишланган ва тўлдирилган нашри. – Тошкент: Ўзбекистон, 1992. – 399 б. – Б.245.

²⁸⁹Гузев В.Г., Насилов Д.М. Слово-изменительные категории в тюркских языках и понятие «грамматическая категория» // Советская тюркология. 1981. – №3. – С.23-35.; Зикриллаев Ғ.Н. Фельнинг шахс, сон, хурмат категорияси системаси. – Тошкент: Фан, 1990.; Иванов С.Н. Категория залога в причастии узбекского языка. В книге «Исследования по истории культуры народов Востока»// Сборник в честь академика И.А. Орбели. – М.:Л., 1960. – С.123-134.; Неъматов Ҳ.Ғ. Тажалли тасаввуф билиш назарияси ва тилшуносликда синтаксисни ўрганиш масалалари // Ўзбек тили ва адабиёти, 1993. – №2. – Б. 27-30.; Нурмонов А., Шаҳобиддинова Ш., Исқандарова Ш., Набиева Д. Ўзбек тилининг назарий грамматикаси. Морфология. – Тошкент: Янги аср авлоди, 2001.

²⁹⁰Баҳриддинова Б.М. Фель луғавий шакли тизими. Фельнинг тур категорияси: Филол. фан. номзоди.... дис. – Самарканд, 2002.; Менглиев Б. Р. Лисоний тизим яхлитлигида сатҳлараро муносабатлар: филол. фан. док. дис. – Тошкент, 2002. Нурмонов А., Шаҳобиддинова Ш., Исқандарова Ш., Набиева Д. Ўзбек тилининг назарий грамматикаси. Морфология. – Тошкент: Янги аср авлоди, 2001.; Шаҳобиддинова Ш. Умумийлик ва хусусийлик диалектикаси ҳамда унинг ўзбек тили морфологиясида акс этиши: Филол. фан. док... дис. автореф. – Тошкент, 2001.; Шукуров О. Ҳаракат тарзи шакли парадигмаси: филол. фан. номз... дис. – Самарканд, 2005. – 126 б.

Ш.Шаҳобиддинова грамматик категория шаклининг маъновий таркибига кўра таснифини берар экан, уни содда ва мураккаб типга ажратади. Тадқиқотчи *содда категория* деганда “формаси ўзи мансуб бўлган маълум бир категория маъносини ва унга ҳамроҳ бўлган ёндош маънони ифодалашга хизмат қилиб, бутун моҳияти билан битта категорияга мансуб бўлган” категорияни тан олади ва бу қаторга сон, қиёслаш, келишиқ, нисбат, субъектив муносабат категориясини киритади²⁹¹.

Бундан маълум бўладики, олима грамматик шаклдаги “нокатегориал” маъно сифатида бошқа грамматик категорияга хос маънонигина ажратади ва ёндош маъно бундай мансубликка эга бўлмаса, уни “нокатегориал маъно” сифатида қарамайди. Масалан, келишиқ категорияси “сўз бирикмасидаги олдинги сўзни кейинги сўзга боғлаши”, нисбат категориясининг “фёъл англатган бажарувчининг тавсифини ифодалашини.”²⁹²

Мураккаб категория сифатида эса бир шаклда турли грамматик категорияга мансуб категория тушунилади. Масалан, “Эгалик категориясида уч хил маъно мужассамланган:

- 1) лисоний ва нолисоний эгалик муносабатини ифодалаш (эгалик категорияси учун нокатегориал маъно);
- 2) шахс маъноси (эгалик категорияси учун нокатегориал маъно);
- 3) сон маъноси (эгалик категорияси учун нокатегориал маъно)” дейилади ва улар сирасига эгалик, ўзгаловчи, кесимлик категорияси киритилади”²⁹³.

Туркийшуносликда грамматик категория тилшунос Ҳ.Неъматов томонидан туркий тил ёдгорликлари мисолида маълум бир гап бўлагига хос, яъни гап бўлагини, синтактик қурилмада тобе ёки ҳоким узвни шакллантирувчи (келишиқ, эгалик, кесимлик) ҳамда маълум бир туркумгагина хос (сон, нисбат, функционал шакл) категория ажратилди. Бу эса жаҳон

²⁹¹ Шаҳобиддинова Ш. Ўзбек тили морфологияси умумийлик-хусусийлик диалектикаси талқинида. I қисм. – Андижон, 1994. – Б. 23.

²⁹²Кўрсатилган манба.

²⁹³ Кўрсатилган манба. – Б. 21.

тилшунослигида грамматик шаклни лексик ва синтактик шаклга ажратиш ғоясининг ривожлантирилган кўриниши²⁹⁴.

Морфоанализ жараёнида морфологик таҳлилнинг дастлабки босқичи грамматик шаклнинг қайси грамматик категорияга мансублигини аниқлашдан иборат бўлади. Шу сабабли морфоанализаторда дастлаб тилга хос грамматик категориялар фарқланиши талаб этилади.

Туркий тилларнинг умумий морфоанализатори²⁹⁵ни ўрганишларимиз шуни кўрсатадики, ушбу анализаторнинг грамматик категорияларни белгилашда инглиз, рус тиллари грамматик категориялари асос қилиб олинган. Табиийки, бир тилнинг грамматик категориялари (хусусан, бошқа-бошқа оилага мансуб тилларда) иккинчи тил категорияларини ажратишда мукамал намуна бўлолмайди. Туркий тиллар морфоанализатори тажрибасида қуйидаги грамматик категориялар фарқланади²⁹⁶: case / падеж; definiteness / определенность; degree of comparison / степень сравнения; estimation/ оценочность; finiteness / финитность; interrogative/вопросительность; mensurative/мензуратив; mood/наклонение; number/число; part of speech/часть речи; person/лицо; personality/персональность; phase/фазовость; polarity/полярность; polite/вежливость; possessive/притяжательность; respect/респективность; tense/время; voice/залог.

Кўринадики, туркий тиллар учун тузилган умумий морфоанализаторда инглиз ҳамда рус тили грамматикаси эталон сифатида олинган, айти шу жиҳат ушбу морфоанализаторнинг ўзбек тили автоматик морфоанализини амалга оширишда ноқулайликлар келтириб чиқаришини қайд этиш мумкин. Шу сабабли умумий морфоанализаторда белгиланган грамматик категорияларни ўзбек тили морфоанализаторида қўллаб бўлмайди, чунки ўзбек тилида грамматик категориялар ўзига хос парадигмага эга.

²⁹⁴Замонавий ўзбек тили. Морфология. (Масъул муҳаррир: Р.Р.Сайфуллаева). – Тошкент: Мумтоз сўз, 2008. – 522 б. –Б.196.

²⁹⁵http://modmorph.turklang.net/uzc/gram_category

²⁹⁶http://modmorph.turklang.net/uzc/gram_category

Туркий тиллар МАни кузатишларимиз шуни кўрсатдики, ушбу анализаторда белгиланган грамматик категориялар сирасидан ўзбек тили морфоанализатори грамматик категориялари сифатида сўз туркуми; шахс; сон; келишик; эгалик; даража; замон; нисбат; майл; бўлишсизлик; ҳурмат; феълнинг вазифа шакллари; субъектив баҳо; сўроқ кабиларни ажратиш мумкин. Шундай бўлса-да, ўзбек тили морфоанализатори учун грамматик категорияларни ўзбек тили қонуниятлари асосида қатъий белгилаш лозим. Айни пайтда ўзбек тилшунослигида қуйидаги грамматик категория ажратилади²⁹⁷: (1) эгалик категорияси (2) келишик категорияси; (3) сон категорияси; (4) нисбат категорияси; (5) замон категорияси; (6) майл/модаллик категорияси; (7) шахс/сон категорияси; (8) даража категорияси; (9) бўлишли/бўлишсизлик категорияси; (10) ҳаракат тарзи категорияси; (11) ўзгалоғчи (феълнинг ҳосланган шакли) категорияси. Грамматик категорияни таснифлашда у ўз остида бирлаштирган грамматик шаклнинг семантик, морфологик, синтактик табиатига асосланилади. Айтиш мумкинки, ўзбек тили морфоанализатори учун грамматик категорияларни белгилашда юқорида кўрсатилган 11та категорияга асосланиш автоматик морфоанализ жараёнининг аниқ натижа беришини таъминлайди.

Морфоанализаторда таҳлилни амалга ошириш учун кейинги ахборот манбаини граммедалар тизими ташкил этади. Умумтуркий анализатор граммедалари орасидан ўзбек тили граммедалари сифатида қуйидагилар ажратилиши мумкин: сўз туркуми номи (1); шахс (I шахс, II шахс, III шахс) (2); келишик (бош, қаратқич, тушум, жўналиш, ўрин-пайт, чиқиш) (3); даража (4); замон ва унинг ички шакллари (5); майл+бирлик/кўплик+шахс (6); нисбат ва унинг турлари (7); феълнинг ноаниқ шакли (8); феълнинг вазифа шакллари (9); равишдошнинг маъно турлари (10); кичрайтириш/эркалалаш (11); юклама турлари (12); сўроқ (13); тасдиқ/инкор (14); кўплик (15); ҳурмат (16). Аммо

²⁹⁷Замонавий ўзбек тили. Морфология. (Масъул муҳаррир: Р.Р.Сайфуллаева). – Тошкент: Мумтоз сўз, 2008. – 522 б. – Б.186.

булар орасида ўзбек тили хусусиятидан келиб чиққан ҳолда квазиграммемага мансуб грамматик маънолар ҳам учрайди. Жумладан, кичрайтириш/эркалаш; юклама турлари; кўплик; ҳурмат каби маънолар ўзбек тилида квазиграммемалар сирасида кўрилиши лозим. Қуйида квазигреммема ва унинг шаклларига тўхталамиз.

3.2.2. Ўзбек тили морфоанализаторида грамматик маъно ва грамматик шакл: граммема, квазиграммема ва дериватемалар хусусида. Морфоанализ жараёни учун маълум тил граммемаларини аниқлашда тилдаги грамматик шаклларни тавсифлаш талаб қилинади. Грамматик шакллардан умумий граммемалар ажратиб олинади. Қуйида ўзбек тилида грамматик шакл масаласига тўхталамиз.

Ўзбек тилида грамматик маъно қуйидаги усуллар билан ифодаланади:

- 1) асосга грамматик маъно билдирувчи қўшимча қўшиш билан: *китобни, китобга, китобдан, китобнинг, ёздинг, ёзгансан, ёзасан, ёзаётирсан, ёзарсан* каби;
- 2) ёрдамчи сўзлар келтириш билан: *Биз келажакка ишонч билан қараймиз. (А.Қ.) Дўстлик биз учун ҳамиша илҳом ва куч-қувват манбаи бўлиб келган. (У.У.);*
- 3) интонация орқали. Гапни турлича интонация биланайтиш орқали ҳам турлича грамматик маъно ифодаланади: *Машина келди. – Машина келди? – Машина келди! Бу китоб (қайси?) . – Бу – китоб (нима?) каби;*
- 4) сўзларни такрорлаш ҳам грамматик маънони ифодалашга хизмат қилади: *Саида битта-битта босиб хужраси томон йуналди. (А.Қ.)*

Морфоанализатор грамматик маъно ифодалашнинг биринчи ва иккинчи усуллари таҳлил қилиш имкониятига эга. Бунинг учун маълумотлар базасида грамматик шакллар ва ёрдамчи сўзлар ҳақида маълумотлар келтириш талаб қилинади.

Грамматик маъно – мураккаб ҳодиса. Мустақил ва алоҳида олинган сўзда луғавий маъно устига қаватланадиган, ёрдамчи сўзнинг моҳиятини

ташқил этадиган, кўплаб сўз учун умумий бўлган маъно грамматик маъно дейилади. Масалан, *китоб* сўзининг луғавий маъноси устига от, турдош от, аниқ от, якка от маъноси қаватланган бўлади. “Аташ” маъноси ва “сўзни эргаш сўз сифатида бош сўзга боғлаш” учун сўзининг моҳиятини ташқил этади. Грамматик маъно лексик маънодан бир гуруҳ сўз учун **умумийлиги** билан фарқланади²⁹⁸. Грамматик маънолик ушбу хусусияти автоматик морфоанализни амалга ошириш имкониятини беради. Шу умумий грамматик маъно асосида граммедалар тизими шакллантирилади.

Луғавий маъно ҳар бир сўз учун алоҳида бўлади. Масалан, ҳатто бир маънодошлик қаторидаги ҳар бир сўзнинг луғавий маъноси фарқ қилади: башара сўзининг луғавий маъноси “одам боши олд томони пешанадан ияккача бўлган қисмининг салбий ифодаси” бўлса, юз сўзининг луғавий маъноси “одам боши олд томони пешанадан ияккача бўлган қисмининг бетараф ифодаси” дир. Фарқ: “салбий ифода” ва “бетараф ифода”. Лекин нафақат маънодош, балки нарса-предмет ва шахсни билдирувчи сўзнинг барчаси учун от грамматик маъноси умумий. Грамматик маъно барча туркум сўзида бор. Луғавий маъно эса гап бўлаги ва гап бўла оладиган сўзда мавжуд.

Мавжудлигига кўра грамматик маъно учга бўлинади:

- а) сўзнинг табиатида аввалдан мавжуд грамматик маъно;
- б) грамматик восита ёрдамида ҳосил қилинадиган грамматик маъно;
- в) сўз бирикмаси ва гап таркибида вужудга келадиган грамматик маъно.

Сўзнинг табиатида аввалдан мавжуд грамматик маъно. От туркумига мансуб от, аниқ от, мавҳум от, якка от, жамловчи от, атоқли от, турдош от; сифат туркумида сифат, аслий сифат, нисбий сифат; сон туркумида сон, миқдор; феъл туркумида феъл, ўтимли феъл, ўтимсиз феъл; равиш туркумида равиш, миқдор, пайт, сабаб; кўмакчида кўмакчи, соф кўмакчи, кўмакчи вазифасидаги сўз кабилар уларда аввалдан мавжуд грамматик маъно. Ҳар бир

²⁹⁸Замонавий ўзбек тили. Морфология. (Масъул муҳаррир: Р.Р.Сайфуллаева). – Тошкент: Мумтоз сўз, 2008. – 522 б. – Б.202.

сўз туркумида бундай маъно бор. Сўз бирор туркумга хосми, бу туркум атамаси унинг бир грамматик маъносининг номидир²⁹⁹. Демак, ҳар бир сўзга бириктириладиган биринчи граммема унинг табиатида аввалдан мавжуд грамматик маънодан келиб чиқади.

Грамматик восита ёрдамида ҳосил қилинадиган грамматик маъно. Бундай маъно мустақил сўзда мавжуд бўлиб, грамматик қўшимча ва ёрдамчи сўз воситасида ҳосил қилинади. Масалан, от туркумида бирлик ва кўплик, эгалик, келишик, кичрайтириш эркалаш ва кўмакчи ёрдамида вужудга келади. Сифат ва равишдаги даража, соннинг санок сондан бошқа барча маъно тури, феълда нисбат, бўлишли/бўлишсизлик, феълнинг хосланган шакли, майл кабилар шундай маъно сирасига киради³⁰⁰. Бундан келиб чиқадики, морфоанализда иккинчи граммема сўзга қўшилган *грамматик восита* асосида белгиланади.

Сўз бирикмаси ва гап таркибида вужудга келадиган грамматик маъно. Мустақил сўз эркин бирикма таркибида бош сўз ёки эргаш сўз, гап таркибида бирор гап бўлаги, ундалма, кириш сўз, ажратилган бўлак, сўз-гап бўлиб келади. Ёрдамчи сўз сўз ёки гапни ўзаро боғлайди (кўмакчи ва боғловчи), қўшимча маъно беради (юклама). Бу маъно фақат гап таркибида ҳосил бўлади. Гапдан ташқаридаги сўзда бундай грамматик маъно бўлмайди³⁰¹. Бундай грамматик маъно морфоанализ жараёнида таҳлилга тортилмайди, балки бу маъно таҳлилнинг кейинги босқичи – синтактик анализатор ёки парсернинг вазифасига киради. Шу сабабли бундай грамматик маъно граммема сифатида белгиланмайди.

Грамматик маъно, одатда, тилнинг турли воситалари орқали ифодаланади. Бу воситалар *грамматик шакл* дейилади³⁰². Грамматик шакл – грамматик маънони ифодалашнинг энг зарурий элементи, тилда улар бири

²⁹⁹ Кўрсатилган манба.

³⁰⁰ Кўрсатилган манба.

³⁰¹ Кўрсатилган манба.

³⁰² Замонавий ўзбек тили. Морфология. (Масъул муҳаррир: Р.Р.Сайфуллаева). – Тошкент: Мумтоз сўз, 2008. – 522 б. – Б. 241

иккинчисидан ажралган ҳолда мавжуд бўлмайди. Масалан, эгалик маъносининг *-(и)маффикси* билан ифодаланиши каби. Эгалик аффикси уч шахсни, айти вақтда бирлик ва кўпликдаги шахсларни билдиради.

Грамматик шакл сўзнинг у ёки бу грамматик восита олиб ўзгаришидир. У.Турсунов ва б. сўзнинг турли грамматикшакллар олишинисўз *ўзгариши*деб аташади. Сўз ўзгариши лексик маъно билан категориал грамматик маънолар бирлигига асосланади.

Сўзнинг грамматик шакллари системаси *парадигма* дейилади ҳамда сўзнингграмматик шакллари системаси морфологиянинг парадигматика қисмида ўрганилади. Грамматик категорияга муносабати, яъни парадигма ҳосил қилиш ёки қилмаслиги жиҳатидан сўз шакллари икки турга: категориал ва нокатегориал шаклларга бўлинади.

Категориал шакллар маълум бир сўз туркумига хос грамматик шакллар системасига бирлашувчи, демак, грамматик категория ҳосил қилувчи шакллардир: отлардаги грамматик сон, эгалик ва келишик шакллари, феъллардаги замон, шахс-сон шакллари каби. Морфоанализаторда улар граммема сифатида белгиланади. *Граммема* атамаси ҳақида юқорида тўхталган эдик, бу ўринда И.А.Мельчукнинг граммема борасидаги фикрини келтиришни ўринли деб билдик: “...сўз ўзгартирувчи шаклни *граммема* атамаси билан номлаш тўғри бўлади: атама қисқа, тушунарли эканлиги сабабли жуда қулай. Тилшуносликда грамматик маъно деганда, дастлаб, сўз ўзгартирувчилар категорияси тушунилгани сабабли шу атама (граммема) келиб чиққан”³⁰³.

И.А.Мельчук *грамматик маъно* атамасининг алоқа-муносабат шаклига ҳам, нокатегориал шаклларга ҳам, сўз ясовчи шаклларга нисбатан бирдек қўлланилиши тарафдори. Шу сабабли олим грамматик маънони синтактик шакл ясовчилар ва сўз ясовчилардан иборат деб билади: биринчисини

³⁰³Мельчук И.А. Курс общей морфологии. Т.1, Москва – Вена: «Языки русской культуры», Венскийславистический альманах, Издательская группа «Прогресс», 1997. -С.240 – 287.

граммема, иккинчисини дериватема деб атайди. Ушбу икки тушунчага хос хусусиятни қуйидагича фарқлайди³⁰⁴:

Дериватеманинг тенденциялари	Граммеманинг тенденциялари
1) аниқ бўлади; 2) бирикиш имконияти чекланган; 3) ифоданинг ноодатий воситаларига эга; 4) синтактик алоқа доирасига кирмайди; 5) лексик маънога таъсир кўрсатади; 6) кўрсаткичи граммемага нисбатан ўзакка яқин жойлашади; 7) ўзи қўшилган асоснинг туркумини ўзгартиради.	1) абстрактлик хусусиятига эга; 2) бирикиш имконияти чекланмаган; 3) стандарт ифода воситаларига эга; 4) синтактик алоқани бошқаради; 5) лексик маънога таъсир кўрсатмайди; 6) ўзакдан узоқда жойлашади; 7) ўзи қўшиган ўзак туркумини ўзгартирмайди.

Нокатегориал шакллар ҳам маълум грамматик маънони англатади, аммо улар ўзаро бир системага бирлашмайди, маълум даражада мустақилликка эга бўлади. Отлардаги кичрайтиш ва эркалаш, сифатлардаги белгининг ортик/камлиги шакллари кабилар нокатегориал шаклларга киради. Бундай шакллар морфоанализаторда квазиграммема сифатида белгиланади.

И.А.Мельчук квазиграммемани “муҳим оралик ҳодиса” деб таърифлайди. Граemmeма ва дериватемалар орасидаги чегара мутлақ эмас. Муттасил қўлланувчи, аммо бирор бир алоқа-муносабат категориясига тааллуқли бўлмаган муҳим бўлмаган грамматик маънолар ҳам мавжуд: доимий қўлланилиши инобатга олинмаса уларни дериватемаларга киритиш мумкин. Бошқа барча ҳолатларда улар граммемаларга мансуб бўлади: кенг қўлланишга эга, бирикиш имконияти чекланмаган, синтактик алоқага таъсир кўрсатади, лексик маънога таъсир кўрсатмайди. Ҳатто улар орасида локализацияси граммема каби ўзакдан узоқда бўлганлари ҳам кузатилади. Шундай грамматик маъно *квазиграммема* деб аталади³⁰⁵. Ўзбек тилида грамматик категория ҳосил қилмайдиган барча грамматик шаклларни квазиграммемаларга киритиш мумкин. Отга хос луғавий шакллар:

³⁰⁴Кўрсатилган манба.

³⁰⁵Мельчук И.А. Курс общей морфологии. Т.1, Москва – Вена: «Языки русской культуры», Венский славистический альманах, Издательская группа «Прогресс», 1997, с.240 – 287.

қарашлилик шакли *-ники*, ўрин-жой қўшимчаси *-даги*, чегара шакл *-гача*, кичрайтириш-эркалаш шакллари; сифат даражаси кўрсаткичи *-роқ*, ҳаракатнинг кучли/кучсизлигини билдирувчи барча шакллар шулар жумласидан.

Сўзнинг грамматик шакллари тузилиш жиҳатидан икки хил бўлади: синтетик шакл, аналитик шакл. *Синтетик шакл* сўзга аффикс қўшиш билан ҳосил бўлади. *Аналитик шаклда* эса грамматик маъно ёрдамчи сўз орқали ифодаланади. Ушбу шаклларнинг ҳар бири морфоанализаторда алоҳида функция билан ажратилиши талаб қилинади.

Бошқа манбага таяниб ушбу ҳодисага яна бошқа хил таъриф бериш ҳам мумкин. Грамматик воситанинг тузилишига кўра тури бир қарашда унинг синтетик ва аналитик турларининг янгича номланишига ўхшайди. Грамматик шакл тузилишига кўра содда ёки мураккаб бўлиши мумкин³⁰⁶. Содда шакл бир грамматик кўрсаткичдан иборат бўлади: келишик, эгалик, кичрайтириш, нисбат шакли. Мураккаб шакл бир неча кўрсаткичдан иборат бўлади: *уйдаги*, *айтиб қўйди*. Бундай шакллар морфоанализ жараёнида жуда муҳим бўлганлиги сабабли ишнинг кейинги бобида мазкур масалага батафсил тўхталамиз.

Туркий тилнинг миллий табиатидан келиб чиққан ҳолда ўзбек тили грамматик шакллари мутахассислар томонидан учга бўлиб ўрганилади³⁰⁷:

1) луғавий шакл ҳосил қилувчи – сўзнинг луғавий маъносини нутқ талабига мувофиқ равишда модификациялайдиган, сўзнинг аташ маъносига муайянлаштириш, торайтириш, аниқлаштириш даражасида таъсир этадиган қўшимча тури отдаги кичрайтириш, сон, шахсий муносабат шакли; феълнинг нисбат, бўлишсизлик, ҳаракатнинг кучли/кучсиз даражасини кўрсатувчи синтетик ҳаракат тарзи шакли; соннинг маъно турини ҳосил қилувчи,

³⁰⁶Mengliyev B., Xoliyurov O., Abdurahmonova N. O'zbek tilidan universal qo'llanma. (Qayta ishlangan 3-nashri). – Toshkent: Akademiya nashri, 2014. – 389 b. – B. 117-119.

³⁰⁷Замонавий ўзбек тили. Морфология. (Масъул муҳаррир: Р.Р.Сайфуллаева). – Тошкент: Мумтоз сўз, 2008. – 522 б. – Б. 221-223.

сифатнинг даража шаклини ясовчилар бу гуруҳни ташкил этади. Луғавий шакл ҳосил қилувчи сўз туркумининг таснифловчи, яъни бир сўз туркумини бошқа сўз туркумидан ажратиб турувчи грамматик категория ҳисобланиб, унинг ишлатилиши ўзи мансуб туркум билан чегараланади: кичрайтириш отга, даража сифатга, нисбат феълга, дона, чама, тартиб сонга хос ва ҳ. Булар орасидан даража, нисбат граммема, соннинг маъно турларини ҳосил қилувчилар квазиграммема сифатида қаралади;

2) синтактик шакл ҳосил қилувчи – лексеманинг муайян гап бўлаги бўлиб келишга хослайдиган ҳамда синтактик алоқага киритиш вазифасини бажарадиган шакл. Келишик, эгалик, кесимлик шакли бу гуруҳни ташкил этади. Бу шаклнинг қўлланилиши бир туркум билан чегараланмайди, сўзнинг таснифий белгисига кирмайди. Қандай сўзга қўшилишига кўра синтактик шаклни иккига ажратиш мумкин:

- 1) феълнинг муносабат шакли: замон, шахс-сон, майл шакли³⁰⁸.
- 2) исмнинг муносабат шакли: эгалик, келишик, исмни кесимга хословчи шакл³⁰⁹.

Кесим вазифасида келган ҳар қандай сўз кесимлик шаклини (замон, шахс-сон, майл, боғлама вазифасидаги синтактик бирлик), эга, тўлдирувчи, қаратқич аниқловчи вазифасида келувчи ҳар қандай сўз эгалик, келишик шаклини қабул қилиши мумкин. Мазкур шакл том моҳияти билан синтактик табиатга эга, эгалик ва келишик шакли сўз бирикмасини, кесимлик гап маркази – кесимни шакллантириш учун хизмат қилади.

Луғавий-синтактик шакл ҳосил қилувчи – лексеманинг ҳам луғавий маъносига таъсир қилади, ҳам унинг синтактик вазифасини белгилайди. Сифатдош, равишдош, ҳаракат номи шакли иккиёқлама табиати билан бу гуруҳдан ўрин эгаллайди ва оралиқ учинчи мақомига эга бўлади. Чунки мазкур шакл феълни феъл ёки отга тобелантириш билан бирга сўз луғавий

³⁰⁸ Маҳмудов Н., Нурмонов А., Собиров А. ва б. Она тили. 6- синф учун дарслик. –Тошкент: 2005. – 73-90-б.

³⁰⁹ Маҳмудов Н., Нурмонов А., Собиров А. ва б. Она тили. 7-синф учун дарслик. –Тошкент: 2005. – 36-54-б.

маъносига ҳам сезиларли таъсир этади. Аниқроғи, унда луғавий шаклга хос белги ҳам, синтактик шаклга хос белги ҳам моҳиятни белгилаш даражасида сезиларли бўлади. Ушбу шакл муайян тизим – категорияга бирлашади. Ушбу шакллар морфоанализаторда граммемалар сирасига киритилади.

Кўриб ўтилганидек, мазкур категория ва унинг қисми ниҳоятда кенг қиррали ва мураккаб табиатли, унинг ҳар бир қирраси алоҳида таснифни талаб этади. Буни қуйидагича гуруҳлаш мумкин³¹⁰:

- 1) тузилишига кўра: ички парадигмага эга мураккаб категория (кесимлик) ва содда категория (нисбат, сон);
- 2) семантик ва синтактик моҳиятига кўра: функционал морфологик (сон, даража) ва функционал-синтактик (кесимлик), морфологик-синтактик (ўзгаловчи) категория;
- 3) лексеманинг қайси семасига таъсир этишига кўра: аташ семасига таъсир этувчи, вазифа семасига таъсир этувчи, ифода семасига таъсир этувчи, ҳам аташ ҳам вазифа семасига таъсир этувчи категория.
- 4) фақат бир туркумга хослиги ёки бир неча сўз туркумига мансублигига кўра ва ҳ.

Саналганлар ичида энг муҳими категориянинг умумий моҳияти билан бевосита боғлиғи морфологик ва синтактик категорияга ажратиш ҳисобланади.

Морфологик категория сўзнинг луғавий маъносига алоқадор, сўз туркумининг таснифий белгиси, бошқа туркумдан ажратиб туради ва қўлланилиши, асосан, шу туркум билан чегараланади: отда сон, феълда нисбат, бўлишли-бўлишсизлик, ҳаракат тарзи, сифатда даража.

Синтактик категория моҳияти синтактик шакл моҳияти билан боғлиқ, сўзнинг луғавий маъноси билан алоқадор бўлмайди, гап бўлаги вазифасида кела оладиган исталган сўз туркумига қўшилиб, гап қурилишида иштирок

³¹⁰Замонавий ўзбек тили. Морфология. (Масъул муҳаррир: Р.Р.Сайфуллаева). – Тошкент: Мумтоз сўз, 2008. – 522 б. – Б.199.

этади: эгалик, келишик, кесимлик (ички парадигмасини замон, майл, шахс/сон ташкил қилади).

Морфологик-синтактик категория ҳар икки гуруҳга кирувчи категория моҳиятини ўзида кучсиз даражада сақлайди. Бу – ўзгаловчи категория³¹¹.

Б.Баҳриддинова морфологик шаклнинг янги таснифини беради. Олима луғавий шакл ҳосил қилувчи сирасига отдаги сон, феълдаги нисбат, бўлишли-бўлишсизлик, ҳаракат тарзи, сифатдаги даража, синтактик шакл ҳосил қилувчига келишик, эгалик, кесимлик, луғавий синтактик шакл ҳосил қилувчига ўзгаловчи категорияни киритади: “1. Луғавий шакл ҳосил қилувчи – лексеманинг луғавий маъносини нутқ талабига мувофиқ равишда модификациялайди. 2. Синтактик шакл – лексемани муайян гап бўлиб келишига ҳослайди ва синтактик алоқага киритиш вазифасини бажаради. Луғавий синтактик шакл лексеманинг ҳам луғавий маъносига таъсир қилади, ҳам унинг синтактик вазифасини белгилайди. 3. Луғавий-синтактик шакл том моҳияти билан луғавий ва синтактик шакл оралиқ учинчи мақомига эга бўлади”³¹².

3.3. Ўзбек тили морфоанализаторида морфотактик қоидалар

3.3.1. Морфотактика ва валентлик. Морфоанализатор тузишда яна бир муҳим омил – *морфотактик қоидалар* тизимини тузиш. Морфотактик қоидалар, асосан, феъл ҳамда от асосларнинг морфотактик имконияти асосида ишлаб чиқилади: у асос ва аффиксал гуруҳлар орасидаги ўзаро алоқани аниқлашга хизмат қилади³¹³.

Морфотактика – сўз ясаиш ёки синтактик алоқа-муносабат жараёнида морфемаларнинг валентлик (бирикиш имконияти) хусусияти, уларнинг

³¹¹Кўрсатилган манба. – Б. 201.

³¹²Баҳриддинова Б. М. Феъл луғавий шакли тизими. Феълнинг тур категорияси: филол. фан. номзоди... дис. автореф. – Самарқанд, 2002. – Б. 21.

³¹³Формальные модели и программные инструменты компьютерной обработки татарского языка. / Р.Р.Гатауллин, А.Р.Гатиатуллин, О.А.Неврозова, Д.Р.Мухамедшин, Д.Ш.Сулейманов, Б.Э.Хакимов, А.Ф.Хусаинов. – Академия наук РТ, Институт прикладной семиотики АН РТ. – Казань: Изд-во Академии наук РТ, 2019. – 260 с. – С.53.

морфематик алоқасини ўрганувчи бўлим³¹⁴. Бошқача айтганча, морфотактика деганда сўзда морфеманинг бирикиш қонуниятларини ўрганувчи морфемика бўлими назарда тутилади. Шунингдек, морфотактика морфеманинг бирикувчанлик қобилияти, валентлик хусусияти ҳамда уларнинг морфематик муносабати ўрганилувчи сўз ясаши бўлимининг таркибий қисми саналади³¹⁵.

Кўшимчаларнинг морфематик муносабати. Ҳар бир аффиксал морфема ўз қуршовига эга, маълум доирадаги сўз ёки асосларгагина қўшилади, у ёки бу сўзшаклни ҳосил қилади. Сўз ясаши ва сўз шакли ҳосил қилишдаги бундай махсус жараён *морфематик муносабат* дейилади (Н.М.Шанский, А.Н.Тихонов)³¹⁶.

Валентлик (лот. *valentia* – куч) – ясалмада ясашиш асоси ва ясовчи қисмининг бирикиш имконияти; сўз ўзгартирувчи қўшимчанинг асосга ёки бошқа турдаги қўшимча билан бирикиш имконияти. Синтаксисда валентлик атамаси сўзнинг бирикиш имкониятини ифодалайди: бундай имконият *ташқи валентлик* дейилади. Сўз ичида морфемаларнинг бирикиши *ички валентлик*, қўшимчаларнинг морфематик алоқаси ёки *морфотактика* деб юритилади³¹⁷.

М.Абузалова ўзбек тилида субстанциал морфология, валентлик ва синтактик қурилма ҳақидаги тадқиқотида валентлик масаласини монографик планда тадқиқ этган³¹⁸. Шунингдек, грамматик шакл валентлиги атамаси С.Ҳ.Муҳамедова³¹⁹, грамматик шаклнинг валентлиги муаммоси Ш.Шаҳобиддинова, Б.Менглиев тадқиқотларида ҳам кўтарилган³²⁰. М.Абузалова валентлик тушунчасини шундай таърифлайди:

³¹⁴<http://www.slovorod.ru/russian-wordform.html>

³¹⁵http://window.edu.ru/catalog/pdf2txt/516/25516/8411?p_page=10

³¹⁶http://window.edu.ru/catalog/pdf2txt/516/25516/8411?p_page=10

³¹⁷<http://www.slovorod.ru/russian-wordform.html>

³¹⁸Абузалова М.К. Субстанциал морфология, валентлик ва синтактик қурилма: Филол. фан. доктори (DSc) диссертацияси. – Самарқанд, 2018. – 224 б.

³¹⁹Муҳамедова С.Ҳ. Ўзбек тилида ҳаракат феълларининг семантик ва валентлик хусусиятлари. Докт.дисс. ... Тошкент, 2007.

³²⁰Менглиев Б. Морфологик воситаларнинг маъновий хусусиятлари ва семантик имкониятлари. Филол. фан. номз. дисс. автореф. – Тошкент, 1996.; Шаҳобиддинова Ш. Грамматик маъно талқини хусусида. Номз. дисс. – Самарқанд, 1993.

“лексик/семантик, семантик/синтактик валентлик тушунчаси дастлаб лексик ходисага нисбатан қўлланилади ва ҳозирги кунга қадар валентлик деганда, асосан, лексик/семантик валентлик тушунилади, лекин субстанциал ёндашишда лексик валентлик билан бирга грамматик валентлик – грамматик/морфологик шаклларнинг, шунингдек, синтактик категория (позиция)ларнинг валентлиги тушунчаси ҳам фарқланади”³²¹.

Тилшунос грамматик валентликни: “муайян бир сўзнинг маълум бир маъновий гуруҳ сўзлари билан меъёрий равишда бирика олиш қобилияти сўз валентлигидир. Сўзларнинг лисоний валентлигидаги бу имконият семантик (маъновий) валентлик дейилади. Лексемалар каби грамматик шакллар ҳам валентликка эга, улар лексема валентлигини ўзгартириши, лексема валентлигига таъсир қилиши мумкин. Бошқача қилиб айтганда, аффиксал морфемаларнинг бирикиш қобилияти грамматик шакл валентлигидир”³²² деб таърифлайди. Автоматик морфоанализатор морфотактика қоидаларини тузишда М.Абузалова таъкидлаган грамматик валентлик ёки морфотактика алоҳида ўрин тутди. Тилшунос кейинги йилларда ўзбек тилшунослигида валентлик тушунчаси остида уч жиҳат: сўз валентлиги; синтактик валентлик; грамматик шакл валентлик фарқланишини таъкидлайди. Унга кўра, сўз валентлиги масаласида икки муаммо: лисоний валентлик ва нутқий валентликни бир-биридан фарқламоқ лозим.

Грамматик шаклларда валентлик қобилияти, хусусан, туркий тилларнинг синтактик шакллари (Ҳ.Неъматов, Ғ.Зикриллаев, Ҳ.Латипов, Б.Баҳриддинова) ёки алоқа-муносабат шакллари деб тасниф этилувчи категориялари³²³ грамматик шаклларида валентлик қобилияти яққол кўзга ташланиб туради. Чунончи, қаратқич келишиги, албатта, эгалик қўшимчали қаралмишни талаб қилади. Қаратқич шаклида эгалик қўшимчалари билан

³²¹Кўрсатилган манба. – Б.64.

³²²Кўрсатилган манба. – Б.67-69.

³²³Нигматов Х.Г. Функциональная морфология тюркоязычных памятников XI-XII вв.– Тошкент: Фан, 1989.; Зикриллаев Ғ. Истиклол ва адабий тил. – Тошкент: Фан, 2004.; Баҳриддинова Б. Фейл луғавий шакллари тизими. Фейлнинг тур категорияси. Номз. дисс. диссер. – Тошкент, 2002.

бирикиш зарурий (облигатор) валентлик бўлса, эгалик қўшимчаларининг қаратқич келишликли шакл билан бирикиши имконий (факультатив – ихтиёрий) валентлик. Фақат бу ҳодиса энди лексик бирлик доирасида эмас, грамматик формантга ҳам хос³²⁴.

М.Абузалованинг таъкидлашича, нутқ жараёнида лексема маълум бир сўзшакл сифатида, маълум бир синтактик позицияда воқеланади. Лексеманинг лисоний (потенциал) лексик-семантик валентлиги бу жараёнда грамматик шакл валентлиги билан бирикади³²⁵. Сўзнинг лексик-семантик (потенциал) валентлигига грамматик валентликнинг таъсирини икки ўзаро боғлиқ, лекин нисбатан мустақил соҳа морфологик шакл ва синтактик позиция (мавқе), морфология ва синтаксис томондан ҳисобга олиш лозим бўлади. Шунинг учун грамматик валентликнинг ўзини иккига: а) морфологик шакл (категория) валентлиги; б) синтактик мавқе (позиция) валентлиги каби гуруҳга ажратиш лозим бўлади³²⁶.

Юқоридаги фикрга таяниб шундай хулоса қилиш мумкин: морфотактиканинг икки: лисоний морфотактика ва нутқий морфотактика кўринишлари мавжуд. Лисоний морфотактика – морфемаларнинг бирикиш имконияти (валентлик). Нутқий морфотактика лисоний валентликнинг рўёбга чиқиши ва уни белгиловчи омиллар (лексик, синтактик, морфологик фонетик омил). Масалан, нутқий вазиятда *маним* (менинг сўзшакли варианты) вариантининг танланишида ташқи омиллар таъсири бўлади.

Кетма-кет келган морфеманинг ўзаро таъсири, ҳоким сўздаги морфеманинг тобе сўздаги морфемага таъсири ҳам ташқи омилга киради. Масалан, *китобни ўқидим // китоб ўқилди*. Иккинчи қўлланишда ҳоким бўлакдаги мажҳул нисбат тобе бўлакнинг бош келишида келишини тақозо қиляпти. Морфоанализатор морфотактик қоидаларини ишлаб чиқишда ўзак морфемаларга сўз ясовчи аффикслар ва уларнинг комбинацияси моделларини

³²⁴Абузалова М.К. Кўрсатилган манба. – Б.80.

³²⁵Кўрсатилган манба. – Б. 77-78.

³²⁶Кўрсатилган манба. – Б. 80.

ишлаб чиқиш ҳамда уларни лингвистик базага киритишни оптимал йўл деб ҳисоблаймиз. Зеро, сўз ясовчи аффикслар салмоқли миқдорни ташкил этади. Морфотактика қоидаларини тузишда сўз ясовчи аффикслар ҳамда уларнинг аранжировкаси қамраб олинади. Албатта, сўз ясовчилар локализациясини моделлаштириш морфотактикадан кейинги босқичда бўлади. Бизнингча, сўзга ясовчи аффикс қўшилганда ўзакда рўй берадиган фонетик ҳодисаларга алоҳида эътибор берилмаса, морфологик таҳлилда иш ҳажми кенгайиб кетади, таҳлил сифатида кутилган натижани бермайди. Қуйида мисол тариқасида айрим шакл ясовчилар морфотактикасига оид мисол келтирамиз, сўз ясовчи, луғавий ва синтактик шакл ясовчи морфемаларнинг аранжировкаси масаласи эса кейинги бобнинг тадқиқ предмети бўлади.

Маълумки, ўзбек тилида аффиксал морфемалар ўзакдан кейин кетма-кет қўшилади. Морфотактика қоидалари сўзшаклда морфема ва алломорфларнинг қандай кетма-кетликда жойлашувини аниқлайди. Бу ҳақда адабиётлар³²⁷ да кўрсатмалар мавжуд. Шуларга таяниб, ўзбек тилидаги айрим морфемаларнинг боғланиш имкониятини қуйидаги жадвалда кузатамиз:

Чап томонда жойлашувчи морфемалар			
Морфемалар рўйхати	Белгиси: чап томонда жойлашиш хусусияти (+) / (-)	Мисол	
-лар	+	бола+лар+дан	
-ман	-	бор+a+ман	
-сан	-	ўқи+й+сан	
-миз	-	ёз+a+миз	
-сиз	-	айт+иб+сиз	
-инг	+	кел+инг	
-им	+	китоб+им	
-имиз	+	дафтар+имиз	
-си	+	ука+си	
-ни	-	ашула+ни	

Жадвалдан кўриняптики, *-лар*, *-им*, *-инг*, *-имиз*, *-си*, *-ни* каби қўшимчалар ўзакка тўғридан-тўғри қўшила олади, улар бошқа аффиксал морфеманинг чап

³²⁷Сулейманов Д.Ш., Гатиатуллин А.Р. Структурно-функциональная компьютерная модель татарских морфем. – Казань: Фэн, 2003. – 220 с. – С.34-39.

томонига қўшилади. *-ман, -сан, -миз, -сиз* қўшимчалари эса ўзакка тўғридан-тўғри қўшила олмайди. Чунки улардан олдин ўзакка аввал бошқа морфема (масалан, замон қўшимчаси) қўшилиб, ундан кейинги позицияни бу қўшимчалар гуруҳи эгаллайди, улар аффиксал морфемалардан ўнг томонда жойлашиш хусусиятига эга.

Морфема асосидаги айрим таҳлил дастурлари³²⁸да морфотактика қоидаларини белгилашнинг жуда қулай усули танланганлигини кўриш мумкин. Бунинг учун дастур базасига тилда мавжуд барча аффиксал морфемалар жойлаштирилгандан сўнг морфотактика қоидаларини белгилаш функцияси ишга туширилади. Аффиксал морфемалар рўйхатидан маълум бир қўшимча белгиланади. Шу қўшимчадан кейин қайси морфема келиши мумкин бўлса, ўша морфемага белги қўйилади ва сақланади (Қаранг: 2-3-расм). Шундай тартибда қўшимча морфемаларнинг ҳар бирига ўнг томонидан ёндош бўла оладиган морфемалар бириктирилади (Қаранг: 4-расм). Кузатишларимиз ушбу усулнинг морфотактика қоидаларини тузиш бўйича энг оптимал, қулай йўл эканлигини кўрсатади.

Айтиш мумкинки, ўзбек тилининг морфотактика қоидаларини тузиш учун тилдаги барча қўшимчалар рўйхати шакллантирилади ҳамда уларнинг жойлашишининг одатдаги тартиби белгиланади. Ишнинг кейинги бобида морфемалар аранжировкаси, яъни сўздаги аффиксал морфемалар қуршови, қуршовда маънонинг ихтисослашуви ҳамда морфемалар қуршовини моделлаштириш (аранжировка) масалаларига батафсил тўхталамиз.

3.3.2. Сўз танламас ва сўз танловчи морфемалар хусусида. Компьютер лингвистикаси, хусусан, морфоанализ мутахассислари сўз *танловчи* ва сўз *танламас* морфемаларни фарқлашади³²⁹. Туркий тилларда, жумладан, ўзбек тилида шундай қўшимчалар сираси борки, бир морфологик

³²⁸<http://modmorph.turklang.net/ru/>

³²⁹Формальные модели и программные инструменты компьютерной обработки татарского языка. / Р.Р.Гатауллин, А.Р.Гатиатуллин, О.А.Невродова, Д.Р.Мухамедшин, Д.Ш.Сулейманов, Б.Э.Хакимов, А.Ф.Хусаинов. – Академия наук РТ, Институт прикладной семиотики АН РТ. – Казань: Изд-во Академии наук РТ, 2019. – 260 с. – С.81-83.

гуруҳга мансуб барча ўзакка бирдек қўшилади. Масалан, -га қўшимчаси N морфологик гуруҳидаги барча ўзакларга қўшила олади. Бундай қўшимчалар *сўз танламас морфема* дейилади. Аммо бир гуруҳ қўшимчалар борки, бир морфологик гуруҳга мансуб ўзакларнинг барчасига ҳам қўшилавермайди. Бундай қўшимчалар гуруҳига ўзлик ва мажҳул нисбат қўшимчалари (-ин, -н, -ил, -л)ни киритиш мумкин. Улар потенциал жиҳатдан феъл-асос морфемага қўшилади, аммо амалда (контекстда) маълум гуруҳ асослар бу қўшимчани қабул қилолмайди. Масалан, *безамоқ* ўзаги иккала нисбат қўшимчасини ҳам қабул қилади: *безанмоқ, безалмоқ*. Бу қўшимчаларнинг ўзакка қўшилиш/қўшилмаслигида бирор қонуният мавжуд эмас. Нисбат қўшимчаларининг асосга қўшилиш/қўшилмаслик имкониятини аниқлаш эса морфоанализ жараёнини соддалаштиради, компьютер ишловини осонлаштиради, синтезни тезлаштиради.

Жойламоқ сўзининг нисбат категориясидаги сўзшаклларини таҳлилга тортамыз. Маънони аниқроқ ифодалаш учун замон ва шахс сонда қўллаймиз.

- 1) аниқ нисбат: *жойлади – Сабзавотларни қутиларга тартиб билан жойлади;*
- 2) ўзлик нисбат: *жойлашди – Ўктам янги квартирасига жойлашди;*
- 3) мажҳул нисбат: *жойланди – Маиший техника буюмлари картон қутиларга жойланди.*
- 4) биргалик нисбат: *жойлашди – Сабзавотларни қутига тартиб билан жойлашди;*
- 5) орттирма нисбат: *жойлаштирди – Маиший техника буюмларини картон қутиларга жойлаштирди.*

Кўриняптики, ўзлик нисбат маъноси *-и* қўшимчаси билан ифодаланган. Демак, ўзбек тилида айрим сўз сўзшаклларгагина қўшиладиган граммемалар борки, уларни *сўз танловчи қўшимчалар* сирасида алоҳида тартиблаштириш лозим. Акс ҳолда морфологик таҳлил жараёни тўғри натижа бермайди.

Қўшимчаларнинг яна бир ноанъанавий қўлланиш ҳолати кузатилади. Туркий тиллар мумтоз грамматикасида сўз ясовчилар *сўз танламас*

қўшимчалар гуруҳига мансуб. -дай, -ли, -сиз; -дай равиш, -ли, -сиз эса сифат ясайди. Масалан, *қуш+дай, бир+дай* (равиш), *туз+ли, туз+сиз* (сифат). Ушбу ясама сўзлар таркибидаги ясовчи қисм ясалма билан бирга маълумотлар базасидан ўрин олиши лозим. Табиийки, бундай ясалмаларнинг лексиконга қўшилиши база ҳажмининг фавқулодда катталашшига олиб келади. Икки тилли лексикографик ресурсларда лексикон ҳажмини сунъий равишда бир неча баробар катталашшининг олдини олиш мақсадида таржима тилида фақат лексема-эквивалентига эга бирликлар камраб олинади. Солиштиринг:

1) *тузли – соленый; ақлли – умный; истеъдодли – талантливый;*

2) *китобли - “*” – китобга эга; бурунли – “*” – бурни бор.*

Табиийки, иккинчи қатордаги сўзлар луғатдан жой олмайди, шу сабабли бундай типдаги ясалмаларнинг луғатларда акс эттирилиши тадқиқталаб лексикографик муаммо. Бу масала лексикографияда ҳам, компьютер лингвистикасидан ҳам ўз ечимини кутаётган масала бўлиб қолмоқда.

Юқоридаги тавсифлар асосида ўзбек тили морфоанализатори иш босқичларини куйидагича ажратамиз (Қаранг: 9-чизма):

1. Эҳтимолий ўзакни қидириш. Морфологик анализатор сўров ойнасига тушган сўзшакл асослар луғатида эҳтимолий бўлинадиган икки қисмга ажратилади. Масалан, *кишиларга* сўзшакли “к”, “ки”, “киш”, “киши”, “кишил”, “кишила”, “кишилар”, “кишиларг”, “кишиларга” каби эҳтимолий асосларга бўлиб чиқилади.
2. Шундан кейин луғатдан эҳтимолий мос келувчи асос қидирилади.
3. Маълумотлар омборида тавсифланган ахборотга таянилган ҳолда топилган асоснинг морфонологик типини аниқланади.
4. Кейинги босқичда луғатдан аниқланган асос чегараланиб, ўнг томондаги қисмлар таҳлили бошланади. Морфотактик қоидалар асосида *кишиларга* сўзшаклининг ўнг томондаги аффиксал гуруҳи “лар+га” эканлиги топилади.
5. Навбатдаги босқичда “киши+лар+га” мофемалар йиғиндиси аниқланади.

б. Сўнгра топилган ҳар бир алломорфнинг морфема ҳамда категорияси аниқланади. Маълумотлар омборида морфемалар ҳақидаги маълумот асосида таҳлил натижаси қуйидаги кўринишда бўлади:

N (киши) лар+га (алломорфни кўрсатяпти)

N (киши) ЛАР+ГА (морфемани кўрсатяпти)

N (киши) PL+DIR

N (киши) + PL(-Лар)+ DIR (-ГА).

Боб бўйича хулоса

1. Ўзбек тили морфоанализатори тузиш унинг бирликлари, тузилиши ва таркибини аниқлашдан бошланади. Морфоанализ жараёнининг асосий бирликлари сифатида морф, алломорф, сўзшакллар ажратиладики, бунда *морф* элементар сегмент белги: шакли тасодифан ўхшаш бўлган айнан бир морфологик хусусиятга эга минимал ҳодиса; *алломорф* бир морфеманинг турли фонетик таркибга эга бўлган морфлари йиғиндиси; *сўзшакл* деганда таркибий қисмлари орасида морфотактика орқали аниқланадиган алоқа мавжуд бўлган алломорфлар кетма-кетлиги назарда тутилади. Морфоанализатор ишини ташкил этувчи яна бир зарурий унсур *бирикиш қоидалари* бўлиб, у оддий элементдан нисбатан мураккаб бирликларни келтириб чиқаради.

2. Автоматик таҳлилнинг стемминг; сўзшаклнинг луғат асосидаги таҳлили; мантиқий ёндашув асосидаги таҳлил; луғатсиз (жадвал асосида) таҳлил турлари мавжуд. Автоматик морфологик таҳлил стемматизация, лемматизация, граммемалаш каби асосий блоклардан ташкил топади ва *стемматизация* қидирилаётган сўзнинг асосини топиш (стем – бир парадигмага мансуб сўзшаклларнинг максимал умумий қисми, у грамматик асосга тенг бўлиши ҳам, тенг бўлмаслиги ҳам мумкин), *лемматизация* сўзшаклни леммага келтириш (лемма – сўзнинг асосий (луғат бериладиган) шакли, лексема) жараёни ҳисобланади. Лемматизация жараёни флектив тиллар учун аҳамиятли, чунки агглютинатив тилларда сўзнинг “стем”и

лексемага тенг бўлади. Фақат флекцияга учраган сўзшаклда лемма ҳамда стемнинг мос келмаслик ҳолатлари кузатилади.

3. Морфоанализаторни ишга тушириш учун аффиксал морфемалар базаси; асос морфемалар базаси; таснифлаш қоидалари; алломорфлар мослиги қоидалари каби таркибий қисмлар шакллантирилади. Морфоанализаторда маълумотлар омбори муҳим аҳамият касб этадики, унга анализ/синтез жараёни учун талаб қилинадиган барча ахборотларни қамраб олиш; тилда мавжуд флекция ҳолатлари – ўзак ва қўшимча чегарасидаги фонетик ўзгаришларни акс эттириш; одатий синтактик шакл ясовчилар билан бирга супплетив, ўзгармас лексемалар ҳақидаги маълумотни қамраб олиш; омоним лексемалар (шунингдек, тўлиқ ва грамматик шакл таъсирида ҳосил бўлувчи шаклдошлар) базасига эга бўлиш каби талаблар қўйилади.

4. Морфоанализни ташкил этишда лексик бирлик (лексик маъно билдирадиган асос морфема); аналитик конструкция (лексик маъно билдирмайдиган, фақат таркибли аналитик конструкция таркибида учрайдиган асос морфема); аналитик шакл(аналитик шакл ясайдиган ёрдамчи асос морфема)лар таҳлилга тортилади ва таҳлилнинг дастлабки босқичи бирликнинг грамматик категориясини аниқлаш саналади. Грамматик категория – бир умумий ёки турли шакллар мажмуи; грамматик категория сўзшакл морфологик структурасининг категориал ёки комплекс форматдаги маълумот бериш функцияси танланганда морфоанализатор томонидан тақдим этилади.

5. Анализатор сифатини оширишнинг яна бир омили позицион ва нопозицион морфемаларни аниқлаш ҳисобланади. Зеро, позицион морфема гапдаги позициясидан қатъи назар, бир семантик гуруҳга мансуб морфема бўлиб, у гапдаги позициясига қараб турли семантик гуруҳга мансуб бўладиган асос морфемалар гуруҳини ташкил этади. Маълумотлар базасида позицион морфемалар рўйхатини шакллантириш ва локализациясини моделлаштириш морфологик таҳлил сифатини яхшилайдди.

6. Морфоанализ учун граммемаларни аниқлашда тилдаги грамматик шаклларни тавсифлаш талаб қилинади; грамматик шакллардан умумий граммемалар ажратиб олинади. Ифодаланишига кўра грамматик маъно учга бўлинади: сўзнинг табиатида аввалдан мавжуд грамматик маъно; грамматик восита ёрдамида ҳосил қилинадиган грамматик маъно; сўз бирикмаси ва гап таркибида вужудга келадиган грамматик маъно. Сўз муайян туркумга хос бўлар экан, ўша туркум атамаси унинг бир грамматик маъноси номидир. Демак, ҳар бир сўзга бириктириладиган биринчи граммема унинг табиатида аввалдан мавжуд грамматик маънодан келиб чиқса, иккинчи граммема сўзга қўшилган *грамматик восита* асосида белгиланади. Сўз бирикмаси ва гап таркибида вужудга келадиган грамматик маъно морфоанализ жараёнида таҳлилга тортилмайди, балки бу маъно таҳлилнинг кейинги босқичи – синтактик анализатор ёки парсернинг вазифасига киради. Шу сабабли бундай грамматик маъно граммема сифатида белгиланмайди.

7. Грамматик категорияга муносабати, яъни парадигма ҳосил қилиш ёки қилмаслиги жиҳатидан сўз шакллари икки турга бўлинади: категориал шакллар; нокатегориал шакллар. Категориал шакллар маълум бир сўз туркумига хос грамматик шакллар системасига бирлашувчи, демак, грамматик категория ҳосил қилувчи (масалан, отлардаги грамматик сон, эгалик ва келишик шакллари, феъллардаги замон, шахс/сон шакллари каби) шакллардир. Морфоанализаторда улар граммема сифатида белгиланади. Нокатегориал шакллар ҳам маълум грамматик маънони англатади, аммо улар ўзаро бир системага бирлашмайди, маълум даражада мустақилликка эга бўлади. Отлардаги кичрайтириш ва эркалалаш, сифатлардаги белгининг ортик/камлиги кабилар нокатегориал шаклларга киради ва бундай шакллар морфоанализаторда квазиграммема сифатида белгиланади.

8. Морфоанализатор тузишда *морфотактик қоидалар* тизими зарурий таркибий қисм; улар феъл ва от асосларнинг морфотактик имконияти асосида ишлаб чиқилади, у асос ва аффиксал гуруҳлар орасидаги ўзаро алоқани

аниқлашга хизмат қилади. Анализатор ишини ташкил этувчи таркибий қисмлар сирасига *асос лексемалар лексикони* ҳам киради. Асос лексемалар лексикони ҳозирги адабий тил асосида шакллантирилади. Тил луғат таркибини нисбатан тўлароқ қамраб олган луғатлар лексикон тузишга асос бўлади.

9. Ўзбек тили морфоанализи (1) эҳтимолий ўзакни қидириш (морфологик анализатор сўров ойнасига тушган сўзшакл асослар луғатида эҳтимолий бўлинадиган икки қисмга ажратилади); (2) асослар луғатидан эҳтимолий мос келувчи асосни қидириш; (3) маълумотлар омборидаги ахборот асосида топилган ўзакнинг морфонологик типини аниқлаш; (4) ўнг томондаги қисмларни таҳлил қилиш (аффикслар гуруҳи аниқланади) каби босқичлардан иборат бўлади.

IV БОБ. ЎЗБЕК ТИЛИ МОРФОАНАЛИЗАТОРИ УЧУН СЎЗЛАРДА ҚЎШИМЧАЛАР ҚУРШОВИНИ ЛИНГВИСТИК МОДЕЛЛАШТИРИШ МАСАЛАЛАРИ

4.1. Сўзда луғавий шакл ясовчи, синтактик шакл ясовчи ва қўшимча кўринишидаги юкламалар қуршовини лингвистик моделлаштириш

4.1.1. Луғавий шакл ясовчи морфемаларнинг сўздаги қуршови хусусида. Сўзда қўшимча морфемаларнинг одатдаги тартибида луғавий шакл ясовчи морфема сўз ясовчи морфемадан кейин келади. Бу ҳақида А.Ҳожиёв, С.Усмоновнинг қатор ишларида батафсил маълумот берилган³³⁰. Бу сўз ясовчи морфема сўзнинг материал томони, луғавий шакл ясовчи морфеманинг сўзнинг формал томони билан боғлиқлиги изоҳланади. Луғавий шакл ясовчилар грамматик маъно ифодалаш билан бирга лексик маънога ҳам таъсир кўрсатади: асос ифодалаган маъно билан боғлиқ равишда сўзни грамматик шакллантиради.

Грамматик воситанинг сўз таркибидаги тартиби маълум қонуниятга эга. Бу восита муайян тартибда жойлашади: *сўз ясовчи қўшимча+луғавий қўшимча+синтактик қўшимча*. Грамматик воситанинг жойлашишидаги тартиб, изчиллик унинг маъно ва грамматик хусусияти билан боғлиқ: янги луғавий маъно ҳосил қилувчи восита биринчи, луғавий маънога таъсир қилувчи восита иккинчи ва луғавий маънога таъсир қилмайдиган, лекин сўзни боғлайдиган восита учинчи бўлиб қўшилади. Қўшимчанинг жойлашувидаги меъёрий ҳолат баъзан бузилади: *она-лар-им // она-м-лар, айт-ди-нг-лар // айт-ди-лар-инг* каби. Бу кам учрайдиган ҳодиса³³¹.

³³⁰Сўз таркибида тартиб ҳақида қаранг: Усмонов С. Ҳозирги ўзбек тилида сўзнинг морфологик тузилиши // Низомий номидаги ТошДПИ илмий асарлари, 42 том, 2-китоб. – Тошкент, 1963.; Ўзбек тили грамматикаси. I том. – Б.103.; Ҳозирги ўзбек адабий тили. – Тошкент, 1980. – Б.163.; Ҳожиёв А. Форма ясовчи аффикслар// Ўзбек тили ва адабиёти, 1977.; Ҳожиёв А. Ҳозирги ўзбек тилида форма ясаши. – Ташкент, 1979.

³³¹Mengliyev B., Xoliyurov O., Abdurahmonova N. O'zbek tilidan universal qo'llanma. (Qayta ishlangan 3-nashri). – Toshkent: Akademnashr, 2014. – 389 b. – B.134.

“Ўзак+сўз ясовчи + луғавий шакл + синтактик шакл” қолипи бир қанча хусусиятга эга. Биз ушбу бўлимда фақат луғавий шакл ҳосил қилувчи морфемаларнинг сўздаги қуршови ҳақида доирасида фикр юритамиз.

Луғавий шакл ясовчи(ЛШЯ)ларни кузатиш кўрсатдики, улар ўзак ёки негизга маълум тартибда қўшилади ҳамда ўзига хос сектор ҳосил қилади. ЛШЯлар мустақил сўз туркумлари тизимининг деярли барчасида учраганлиги сабабли ҳар бир сўз туркуми доирасида ЛШЯнинг сўздаги қуршовини алоҳида кўриб чиқиш зарурати пайдо бўладики, бу, одатда, ЛШЯлар қуршовини феъл, от, сифат, сон, олмош, равиш доирасида тадқиқ этишни талаб этади³³².

ЛШЯларнинг феъл сўз туркумидаги қуршови. Феълнинг луғавий шакли сирасига нисбат, равишдош, сифатдош, ҳаракат номи шакли, ҳаракат тарзи (кўмакчи феъл), ҳаракатнинг кучсиз даражаси ва давомийлигини билдирувчи қўшимчалар киритилади³³³. Феълда ЛШЯларнинг махсус тартиби ҳақида Н.Баскаков, А.Шербак, А.Ғуломов, А.Ҳожиёв, ғарб тилшунослари Ж.Вандрис, Г.Глисон каби олимлар ўз ишларида мулоҳаза билдиришган. Ўзбек тилида шакл ясовчи морфемалар Т.У.Мирзақулов иши³³⁴да махсус тадқиқ этилган. Шу сабабли одатдаги тартиб ҳақида тўхталмай, кузатишларимиз асосидаги позицион қоидалар ҳақидаги хулосаларга тўхталамиз. Феъл (V) таркибида сўз ясовчи морфемадан кейин қуйидаги ЛШЯлар жойлашади³³⁵:

орттирма нисбат қўшимчаси баъзи ҳолларда феъл-асосга икки марта қўшилиши кузатилади: *том+из+дир, кий+дир+тир, кий+гиз+тир, ол+дир+тир, кел+тир+тир*. Бундай ҳолат қоидалар базасида $V+f_1^v+f_2^v$ модели

³³²Ўзбек тили грамматикаси. I том. – Б.104-105.

³³³Ne'matov N., G'ulomov A., Qodirov M., Abduraimova M. Ona tili: 6- sinf uchun darslik. – T.: O'qituvchi, 2004. – B.18-20.

³³⁴Мирзақулов Т.У. Формаобразующие сложные аффиксы в узбекском языке: автореф. дисс. ... канд. филол. наук. – Тошкент, 1980.

³³⁵Нисбат шакллариининг қўлланиши ҳақида қаранг: Тожиёв Ё. Феъл даражаларини ясовчи аффикслар синонимияси. Научные труды ТашГУ. Вып. 443. Вопросы узбекского языкознания. – Тошкент, 1973. – Б. 115-124.; Нариманова М.Д. Ҳозирги ўзбек тилида феъллардаги даража ясовчиларнинг баъзи бир хусусиятлари ҳақида. Илмий ишлар тўплами. Ўзбек тилшунослиги масалалари. ТошДУ, Тошкент, 1983. – Б. 52-59.; Ғуломов Ё.Г. Ўзбек шеваларида феъл даражалари аффикслари. Илмий ишлар тўплами. Ўзбек тилшунослиги масалалари. ТошДУ – Тошкент, 1983. – Б.64-68.

билан белгиланиши лозим. Чунки одатдаги модел $V+f_1^v+p_2^v$ шаклида бўлади, яъни феъл+ЛШЯ+СШЯ. Бундай шаклда от-кесимни шакллантирувчи –*дир* сўнг позицияда туриши билан грамматик омонимлик ҳосил қилиши мумкин. Бундай омонимия $V+f_1^v+p_2^v$ модели билан фарқланади.

Ўзлик ва мажҳул нисбат шакллари зоҳиран ўхшаш, уларни автоматик фарқлашни модел орқали ҳам амалга ошириб бўлмайди, чунки иккала ҳолатда ҳам уларнинг позицияси бир хил, аммо айрим ўзакдан кейин фақат ўзлик, айрим ўзакдан кейин фақат мажҳул нисбат шакли жойлашади. Солиштиринг:

1) фақат ўзлик нисбат маъноси: *туғилди, кўринди, сўкинди, ютинди, танилди* ва ҳ;

2) фақат мажҳул нисбат маъноси: *тешилди, қорилди, тикилди (тикмоқ, қарамоқ эмас)* ва ҳ;

3) ўзлик ва мажҳул нисбат маъноси (полисемантик шакл): *қайрилди, айрилди, бўғилди, отилди, йиғилди* ва ҳ.

Ишнинг олдинги қисмларида сўз танловчи ва сўз танламас қўшимчалар ҳақида фикр билдириб, ўзлик ва мажҳул нисбат шаклининг айрим ўзақларга қўшила олмайдиган ҳолатлари тўғрисида тўхталган эдик. Юқоридаги таснифимизда шу яққол кўринди. Ўзбек тили морфоанализаторининг лингвистик таъминоти (базаси)ни яратиш учун ушбу шаклларнинг феъл-ўзакка қўшилиш имконияти ўзбек тилидаги барча феълларда инвентаризация қилиниши талаб этилади. Демак, феълнинг ушбу шакли автоматик таҳлил модел асосида эмас, инвентаризация асосида амалга оширилади.

Феълга ҳос ЛШЯ бўлмиш *-(u)ш* шакли ҳаракат номи ва биргалик (баъзи ўзақларга қўшилиб, ўзлик) нисбат вазифасини бажариши мумкин. Бундай грамматик омонимиянинг фарқланиши, уларнинг морфоанализда тўғри таҳлил қилиниши учун маълумотлар базасига қуйидаги позицион моделлар киритилиши керак:

1) $V+f_1^v$ ёки $V+f_1^v+f_2^v$ позициясида -(и)ш шакли ҳаракат номи вазифасини бажаради: *чиқ+ар+иш*; *том+из+иш*, *кий+гиз+иш*, *кий+дир+иш*, *тасдиқла+т+иш*, *суя+н+иш*, *чўз+ил+иш*;

2) $V+f_1^v+p_1^v$ ёки $V+f_1^v+f_2^v+p_1^v$ позициясида -(и)ш шакли биргалик нисбат вазифасини бажаради: *кел+иш+ди*, *суриштир+иш+ди*, *ўтир+иш+ди*, *кел+тир+иш+ди*.

Яна бир грамматик омонимлик ҳолати борки, унинг модели алоҳида белгиланади. Масалан, феъл+биргалик н. + биргалик н.: *сўзла+ш+иш*. Бу позицияда -ш биргалик нисбат шакли ҳисобланиши лозим.

3) *жойлашди*, *керишди*, *қоришди* сўзшаклларида ўзлик нисбат маъносини беради, бу феълларнинг ўтимли/ўтимсизлик муаммолари билан боғланиб кетади (албатта, бу муаммо назарий аспектда ўз ечимини кутмоқда)³³⁶; шунинг учун бу ҳолатни моделлаштириш асосида эмас, истиснолар қоидаси асосида муайянлаштиришга тўғри келади.

Феълнинг равишдош, сифатдош, ҳаракат номи (ўзгаловчи) шаклларида бўлишли/бўлишсизлик мавжуд. Бўлишлилик ноль шакл билан ифодаланса, бўлишсизлик равишдошда *-май/масдан* (*ўқиб – ўқимасдан*), сифатдошда *-ма* (*ўқиган-ўқимаган*), ҳаракат номида *-маслик* (*кўриш – кўрмаслик*) шакллари ёрдамида юзага чиқади. Бу шакллар *-ма* ёрдамида ҳосил қилинган мураккаб кўшимчалар³³⁷. Ушбу мураккаб шакллар бошқа грамматик маънони ифодаловчи шакллар билан омонимия ҳосил қилади, шу сабабли уларни моделлаштириш орқали филтрлаймиз. Равишдошнинг *-масдан* бўлишсизлик (мураккаб) шакли отлашган бўлишсизлик шаклидаги сифатдош кўшимчаларидан фарқланиши лозим. Масалан, *ишёқмасдан ҳамма безор, борса келмасдан ҳеч ким қайтмаган* каби ҳолатларда *-мас+-дан* шаклининг ёнма-ён келиши (яъни чиқиш келишигидаги сифатдош бўлади) *-мас+-дан*

³³⁶Mengliyev B., Xoliyorov O., Abdurahmonova N. O'zbek tilidan universal qo'llanma. (Qayta ishlangan 3-nashri). – Toshkent: Akademyashr, 2014. – 389 b. – B.142.

³³⁷Mengliyev B., Xoliyorov O., Abdurahmonova N. O'zbek tilidan universal qo'llanma. (Qayta ishlangan 3-nashri). – Toshkent: Akademyashr, 2014. – 389 b. – B. 149.

алоҳида иккита морфема саналади. *Ишни тугатмасдан кетмайди; дарсни бажармасдан келмайди* каби ҳолатларда *-масдан* битта мураккаб морфема. Кўриниб турибдики, иккаласи ҳам феъл-асосга қўшилган, позицияси бир хил. Демак, автоматик морфоанализ таҳлили натижа беролмайди. Бундай ҳолатларда семантик ёки синтактик таҳлил натижаси ёрдамга келади: сўзшакллар гаптаги синтактик вазифасига қараб фарқланиши мумкин. Грамматик полисемия ва омонимия позицияларини семантик/синтактик филтёр орқали фарқлаш масалалари ушбу тадқиқот вазифалари сирасида камраб олинмади. Муаммо автоматик морфоанализатор яратиш муаммоси доирасида махсус тадқиқ қилинмоғи лозим.

Феълга хос ЛШЯ бўлмиш *-ма* бўлишсизлик шакли айрим сўзларда *-мас* шаклида қуйидаги тартибда жойлашади: *кел+моқчи+мас+ман, бор+моқчи+мас+ман, суя+н+моқда+мас*. Бундай ҳолларда *-мас* шакли аслида *эмас* тўлиқсиз феълнинг қисқарган варианты сифатида қаралади. Солиштиринг: *кел+моқчи+мас+ман = кел+моқчи+эмас+ман*. Бундай позициянинг модели қуйидаги тартибда берилади: $V + p_1^v + f_1^v + p_2^v$.

-вер форманти кўмакчи феълли сўз қўшилмаси таркибида келадиган *бер* феълнинг товуш ўзгаришига учраган шакли; бунда сўзнинг орфографиясида ҳам ўзгариш бўлади: *кел+а+вер+а+ди = кел+а бер+а+ди*. Буни қуйидаги модел билан белгилаш мумкин: $V + f_1^v + V + f_2^v + p_1^v$. Қоидаси: агар {феъл-асос + $f_1^v + вер + V + f_2^v + p_1^v$ } бўлса, *-вер* = кўмакчи феъл.

Феълнинг равишдош шакли позициясига қараб *-(и)б* ўтган замон шакли билан омонимия ҳосил қилади. Позицияларни таҳлил қиламиз:

1) *ўқиб, ёзиб, бориб* шаклида равишдош кўшимчаси; *ўқиб кўрмоқ, айтиб бермоқ, гуллаб қўймоқ* шаклида кўмакчи феълли сўз қўшилмасида етакчи ва кўмакчи феълни боғлашга, *бориб келмоқ, олиб бормоқ, кириб чиқмоқ* кўшма феълларида етакчи ва ёрдамчи феълни боғлашга хизмат қилган равишдош шакли бўлади. Демак, $V + f_1^v(и)б$, яъни {феъл-асос+(и)б+0} бўлса = феълнинг

равишдош шакли; агар $V_1+f_1^v(\text{и})б+V_2+f_1^v+f_1^v/p_1^v$, яъни {фeъл-асос+(и)б+ фeъл-асос+ЛШЯ/СШЯ} = етакчи ва кўмакчи фeъни боғловчи равишдош шакли;

2) *ўтирибман, бўлибди, борибсан, узоқлашибмиз* сўзларида ўтган замон аффикси вазифасини бажаради. Позициясини қуйидагича моделлаштирамиз: агар $V+f_1^v(\text{и})б+p_1^v$, яъни {фeъл-асос+(и)б+СШЯ} бўлса = ўтган замон шакли.

II шахс кўплик шакли *-(и)нгиз* буйруқ-истак майлини ҳам ифодалайди³³⁸: *чиқмангиз*. Оғзаки нутқда бу шакл *-(и)нглр, -(и)нгизлар* кўринишида кўлланади: *кел+ингиз, кел+ингиз+лар*. Ушбу шаклларни муқобил (эквивалент) сифатида қайд этиб қўйиш талаб қилинади, чунки фойдаланувчининг турли мақсади билан бадий асар парчаси, оғзаки нутқ (хатто шева) материали ҳам морфоанализ таҳлилига тортилиши мумкин. Бундай пайтда анализатор грамматик маънони тўғри фарқлаши учун адабий нутқ меъеридан четга чиқишининг типик ҳолатлари муқобил сифатида белгиланиши фойдадан холи эмас. Модели: $V+f_1^v+p_1^v-(и)нгиз = V+f_1^v+p_1^v-(и)нглр/-(и)нгизлар$.

-гани равишдош шакли *-ган* сифатдошининг эгалик қўшимчасини олгандаги ҳолати билан омонимик позицияни ҳосил қилади. Солиштиринг: *Ишлагани келишди; айтгани борганди* (мақсад равишдоши) // *Ишлагани йўқ, ишлаганинг йўқ; айтганини қилди, айтганингизда билардик* (эгалик шаклини олган сифатдош). Ушбу қўшимчалар позицияси моделлаштирилса, морфоанализатор хатосиз таҳлилни амалга оширади:

1) агар $V+f_1^v(\text{гани}) + 0$, яъни {фeъл-асос+гани+0} бўлса, $f_1^v =$ мақсад равишдоши шакли;

2) агар $V+f_1^v(\text{ган+и})+p_1^v+0$, яъни {фeъл-асос+(ган+и)} бўлса, $f_1^v =$ (ган+и) эгалик шаклидаги сифатдош;

3) агар $V+f_1^v(\text{гани})+йўқ$ ёки, яъни {фeъл-асос+(ган+и)+йўқ} бўлса, $f_1^v =$ эгалик шаклидаги сифатдош;

³³⁸Буйруқ-истак майли шаклининг турли кўринишлари ҳақида қаранг: Аъламова М. Туркий тилларда буйруқ-истак майли ҳақида // Ўзбек тили ва адабиёти, 1966.– № 6. – Б.30-34.

4) агар $V+f_1^v(\text{гани})+p_1^n(-\text{га}/-\text{да}/-\text{дан})$, яъни $\{\text{феъл-асос}+(\text{ган+и})+(-\text{га}/-\text{да}/-\text{дан})\}$ бўлса, $f_1^v =$ эгалик шаклидаги сифатдош.

-дир гумон юкламаси одатдаги тартибда феъл-асосли сўз морфемалар занжирининг энг охирида жойлашади: *боргандир, келмагандир*. *-дир* кесимлик кўрсаткичи бўлганда ҳам сўз шундай позицияда туради. Бу икки шаклдош аффиксни фарқлаш учун қуйидаги модел етарли бўлади:

- 1) $V+f_1^v+f_2^v(\text{дир})$, яъни $\{\text{феъл-асос}+(\text{ЛШЯ})+\text{дир}\}$ бўлса, $f_2^v =$ гумон юкламаси;
- 2) $N/\text{Adj}/P/\text{Num}+p_1^n(\text{дир})$, яъни $\{\text{от/сифат/сон/олмош-асос}+\text{дир}\}$ бўлса, $p_1^n =$ кесимлик кўрсаткичи/боғлама.

Феъл-асосда ЛШЯ қуршовини умумий одатдаги тартиб деб қуйидаги позицияларни қатъий белгилаш мумкин:

- 1) асосдан кейин биринчи позицияда нисбат жойлашиб, уларнинг ўзаро тартиби қуйидагича: асос+ўзлик нисбат+орттирма нисбат+биргалик/мажхул нисбат: *безантиришди, ювинтиришди, безантирилди, ювинтирилди*, яъни $V+f_1^{v(\text{voice})}+f_2^{v(\text{voice})}+f_2^{v(\text{voice})}$;
- 2) нисбат шаклидан кейин бўлишсизлик шакли *-ма* жойлашади: $V+f_1^v+f_2^{v(\text{neg})}$;
- 3) бўлишсизлик шаклидан кейин феънинг ўзгаловчи категорияси шакллари (*бормаган, бормай, бормасдан, бормаслик*) ҳамда майл шакли (*бормасин, бормайлик, бормай*) жойлашади: $V+f_1^v+f_2^{v(\text{neg})}+f_3^{v(\text{fin})}$ ёки $f_1^{v(\text{mood})}$;
- 4) кейинги позицияда замон, шахс-сон шакли (*бормасин, бормайлик, бормади*) жойлашади: $V+f_1^v+f_2^{v(\text{neg})}+p_1^v+p_2^v$.

Ҳаракатнинг кучли/кучсизлигини билдирувчи ЛШЯлар квазиграммемалар қаторига киритилиши талаб қилинади. Таҳлилда фойдаланиш учун қуйидаги жадвалда морфемаларни келтирамыз.

1) ҳаракатнинг кучсиз даражасини билдирувчи қўшимча:

<i>-мсира/имсира:</i>	йиғламсира, кулимсира
<i>-қира/инқира:</i>	ишонқира, оқаринқира, ўчинқира
<i>-и/иш:</i>	тўлиш, қизиш, оқариш, тўхташ (юраги)
<i>-қ/иқ/к/ик:</i>	тутақ, толиқ, шишиқ, жунжик, кўник

2)ҳаракатнинг кучли даражасини билдирувчи қўшимча:

-ла/ала:	қувла/қувала, ишқала, савала, сийпала, чўқила, чайқала, силтала, опичла
-қи/ғи:	юлқи, сизғи, тўзғи, бижи–бижғи
-чи:	типирчила, тепчи, терчила
-а:	бур –бура, қувон –қувна, урин –урна
-гила/кила/қила/ғила:	югургила, титкила, тепкила, чопқилла, турткила, тортқила, эзғила, чўзғила.
-га/қа/қа:	сурга, сурка, чайқа.

Бундай қўшимчалар парадигмага эга бўлмаганлиги сабабли квазиграммемага мансуб бўлади. Жадвалдан кўриниб турибдики, феъл сўз туркумида ҳаракатнинг кучли/кучсизлигини билдирувчи ЛШЯларнинг аксарияти омонимик характерга эга. Уларнинг маъносини фарқлаш учун қайси туркумдаги ўзак(асос)дан кейин келишига қараб моделлаштирамиз:

$V+f_1$: -ка (сурка), -а (бура), -ала (қайтала), -имсира (кулимсира), -инқира (босинқира), -кила (титкила), -чай (букчай).

Агар $\{V+f_1 | -ка, -а, -ала, -имсира, -инқира, -кила, -чай\}$ бўлса, $\{f_1 | -ка, -а, -ала, -имсира, -инқира, -кила, -чай\} =$ ҳаракатнинг кучли/кучсизлигини билдирувчи ЛШЯ.

Бу моделimiz билан феъл-асосдан кейин қўшилган ушбу қўшимчалар ҳаракатнинг кучли/кучсизлигини билдирувчи ЛШЯ эканлиги ойдинлашади. Ушбу қўшимчалардан айрими сўз ясовчи аффикс (-ла, -чи, -а), ЛШЯ (-и, -иш), СШЯ (-га, -ка, -қа)лар билан омонимлик ҳосил қилади. Буларни қуйидаги модел орқали филтрлаш мумкин:

агар $\{V+f_1 (-ла, -чи, -а, -и, -иш, -га, -ка, -қа)\}$ бўлса, (яъни ушбу шакллар феъл-асосдан кейин қўшилса) $f_1 =$ ҳаракатнинг кучли/кучсизлигини билдирувчи ЛШЯ.

Маълумки, аффиксация усули билан феъл ясалишида феъл бўлмаган асосдан ясалади. Бу эса ҳодисани моделлаштириш имконини беради:

1) агар {N/A /Num /Adv/ Sim +d₁ (-ла, -чи, -а)} бўлса, (яъни ушбу шакллар феъл бўлмаган асосдан кейин қўшилса) d₁ = сўз ясовчи қўшимча (феъл-асосдан кейин қўшилганда ЛШЯ эканлигини юқорида таъкидланди);

2) агар {N/A /Num /Adv/ Sim /сифатдош, равишдош, ҳаракат номи +r₁-га, -ка, -қа} бўлса, (яъни ушбу шакллар феъл бўлмаган асосдан кейин қўшилса) r₁ = синтактик шакл ясовчи (феъл-асосдан кейин қўшилганда ЛШЯ эканлигини юқорида таъкидланди).

3) -и, -иш шакли позицияси қуйида батафсил моделлаштирилади.

ЛШЯларнинг от сўз туркумидаги қуршови. Отга хос ЛШЯлар сирасига шахсий муносабат шакли, кичрайтириш, эркалаш ва ҳурмат шакли киради (Қаранг: 2-жадвал). Шахсий муносабат категориясига оид шакл сўзловчининг объектив борлиққа ва ўз нутқида ижобий ёки салбий муносабатни ифодалашга хизмат қилувчи восита саналади.

Мутахассисларнинг фикрича, шахсий муносабат, кичрайтириш, эркалаш, ҳурматлаш, камситиш, кучайтириш маъносининг турли хил қўшимча ёки аффиксоид ёрдамида ифодаланиши ўзбек тилида кенг тарқалган. Масалан, *қизча, йигитча, келинчак, қўзичоқ* каби сўздаги *-ча, -чак, -чоқ* қўшимчаси кичрайтириш маъносини англатса, *болагина, укажон, холажон, акахон, Раҳимбой, Қўчқортой, Гулсинбиби, Раънохон, Моҳирабону* сўзларидаги *-гина, -жон, -хон, -той, -бой, -биби, -бону* каби аффиксоид эркалатиш (кўп ҳолда) ва камситиш (баъзан) маъносини келтириб чиқаради. Демак, шахсий муносабат шаклини кичрайтириш, эркалаш ва ҳурматлаш маъносини билдирувчи деб икки катта гуруҳга бўлиш лозим³³⁹.

Ўзбек тилида шахсий муносабат ифодалашнинг учинчи тури ҳам бор: у *-лар* қўшимчасини қавм-қариндошлик билдирувчи сўзда эгалик қўшимчасидан сўнг келтириш билан ҳосил қилинади: *опамлар, дадамлар, амакимлар, бобомлар* каби. Бунда ҳурматлаш маъносида ҳурмат шакли юзага

³³⁹Mengliyev B., Xoliyurov O., Abdurahmonova N. O'zbek tilidan universal qo'llanma. (Qayta ishlangan 3-nashri). – Toshkent: Akademyashr, 2014. – 389 b. – B. 241.

келади. Шахсий муносабат категорияси сўзнинг лексик маъносини ўзгартириб юбормайди (унинг сўз ясовчилик воситаси (*ўргимчак, қўғирчоқ, кўрпача, богча каби*) бундан мустасно), балки асосдан англашилган маънога қўшимча (коннататив) маъно киритади, яъни *стул* ва *стулча* иккита сўз бўлса ҳам, асосан битта нарсани (стулни) билдиради³⁴⁰. Ўзбек тилидаги шахсий муносабат шакллари морфоанализаторда квазиграммелар сифатида белгилаймиз ҳамда уларнинг барчасини маълумотлар базасига жойлаштирамиз.

Отда луғавий шакл ҳосил қилувчиларнинг одатдаги тартиби қуйидагича³⁴¹: кичрайтириш/эркалаш+сон. Кичрайтириш/эркалаш шаклининг кам қўлланадиган шакли кўп қўлланадиганидан аввал келади: *тойчоқча, тойлоқча*.

От-асос (N)га субъектив баҳо: кичрайтириш-эркалаш, камситиш, пичинг-кесатиқ қўшимчалари³⁴² – f_1^n , кейин кўплик шакли³⁴³ – f_2^n қўшилади: *той+чоқ+лар*. Субъектив баҳо шакллари сўз таркибида такрор қўлланиш ҳоллари ҳам учрайди³⁴⁴. Бунда пассив шакл актив шаклдан олдинда жойлашади – $N+f_1^n+f_2^n$: *-чоқ+ча: тойчоқча; -чиқ+ча: қопчиқча; -а+ча: уяча; -ка+ча: йўлкача/йўлакча; -чак+ча: тугунчакча*. Ушбу мисолларнинг морфем таркибида диахрон пландаги аффиксларнинг синхрон қўшимчадан олдин жойлашувини кузатилади. Аммо оғзаки нутқда қўлланувчи *ўғилчавой* ($N+f_1^n+f_2^n$) сўзида *-вой -ча* фаол қўшимчасидан кейин қўланади. Бундай тескари тартиб *-вой/-бой* шаклининг аффиксоид эканлиги билан изоҳланади.

³⁴⁰ Кўрсатилган манба. – В. 245

³⁴¹ Mengliyev B., Xoliyurov O., Abdurahmonova N. O'zbek tilidan universal qo'llanma. (Qayta ishlangan 3-nashri). – Toshkent: Akademnashr, 2014. – 389 b. – Б. 134.

³⁴² Субъектив баҳо шакллари тўғрисида қуйидаги ишларга қаранг: Гулямов А.Г. О некоторых особенностях аффиксов – с уменьшительно-ласкательным значением в узбекском языке / Научные труды ТашГУ. Вып. 268. Языкознание и литературоведение. – Ташкент, 1964.; Журабаева М.К. К вопросу о категории уменьшительно-ласкательности в узбекском языке Труды аспирантов ТашГУ. Вып. 360. Литература и языкознание. – Ташкент, 1970.; Кўнғуров Р. Субъектив баҳо формаларининг семантик ва стилистик хусусиятлари. – Ташкент, 1980. – Б.33-145.; Ўзбек тили стилистикаси. – Ташкент, 1983. – Б.102-108.

³⁴³ Ғуломов А. Ўзбек тилида кўплик категорияси. – Ташкент, 1943. – Б.3.

³⁴⁴ Махкамов Н. Аффиксальный и лексико-аффиксальный плеоназм в узбекском языке: Автореф. дис. ... канд. филол. наук. – Ташкент, 1983. – С.16.

Шолча, боғча (болалар боғчаси), кўрпача сўзларида янги лексик маъно вужудга келганлиги учун *-ча* ЛШЯ ҳисобланмайди, сўз маъноли қисмга ажралмайди³⁴⁵. Бундай истисно сўзлар истиснолар луғатига киритилади; *-ча* морфема таҳлилига тортилмайди. Ўнгдан чапга таҳлилида *-ча* шакли аффикс сифатида топилиши мумкин, аммо ўзак қисми текширилганда, истиснолар луғатида бу сўзлар асос сифатида топилади; таҳлил ушбу сўзларнинг битта леммага тенг эканлиги билан яқунланади.

Сўзнинг морфем таркиби қатор хусусиятларга эга: улар орасида аффиксал омонимия ҳам мавжуд. Шаклан ўхшаш морфемалар сўзда кўшимчалар жойлашуви билан боғлиқ айрим қонуниятларни келтириб чиқаради. *-гина* f_2 шаклининг *қўзичоққиналаримиз* ($N+f_1^n+f_2^n+f_3^n+r_1$) сўзидаги жойлашуви *ёшларгина* ($N+f_1^n+f_2^n$) сўзидаги тартибдан фарқ қилади. Биринчи мисолда *-гина* луғавий шакл ясовчи (*f*) бўлса, иккинчи мисолда айирув-чегаралов юкламаси вазифасини бажаради: *-гина* ЛШЯ кўплик ва эгалик шаклларида олдпозицияни, *-гина* юкламаси эса сўнгпозицияни эгаллайди. Солиштиринг: *дадагиналари* (*-гина* эркалиш маъносини билдирувчи ЛШЯ) / *дадаларигина* (*-гина* айирув-чегаралов юкламаси); *болагинага* /*болагагина*; *акагинаси* /*акасигина* сўзшаклларидаги ҳам аффиксал омоним. Морфоанализ қоидалар мажмуида *-гина* шаклининг ушбу позициялари қуйидаги модел кўринишини олади:

- 1) агар { N +гина+кўплик+эгалик} бўлса, *-гина* = кичрайтириш-эркалаш ш.;
- 2) агар { N +кўплик+эгалик+гина} бўлса, *-гина* = айирув-чегаралов юкламаси.

Ушбу шаклларнинг функцияси нутқда урғу ёрдамида фарқланади, аммо бу морфоанализатор томонидан “тушунилмайди”, шу сабабли юқоридаги модул шакл маъносини фарқлайди.

-гина шакли (f_2^n) *укачагинам, қизчагинам* каби сўзлардаги *-ча* кичрайтириш шаклидан кейин қўлланганда эркалаш маъносини беради, чунки

³⁴⁵Mengliyev B., Xoliyurov O., Abdurahmonova N. O'zbek tilidan universal qo'llanma. (Qayta ishlangan 3-nashri). – Toshkent: Akademyashr, 2014. – 389 b. – B. 175.

эркалаш шакли кичрайтириш шаклига кўшимча маъно юклайди: эркалаш маъноси кичрайтириш маъносига юклатилади³⁴⁶.

Отга хос ЛШЯлар орасида тарихан икки аффиксдан иборат, ҳозир мураккаб яхлит кўшимча саналадиган *-даги, -ники, -гача* каби шаклларнинг қисмлари *-да, -ни, -ги, -ча* каби шакллар билан омонимлик ҳосил қиладигандек туюлади. Аммо морфоанализаторнинг ўнгдан чапга қидирув тизими ушбу шаклларни яхлит: “даги”, “ники”, “гача” ҳолда топади ҳамда тўғри таҳлил қилади, чунки базада ушбу шаклларга квазиграммемалиги ҳақидаги изох бириктирилган. Таҳлилда хатолик бўлмаслигининг яна бир асоси шуки, *-ги, -ки, -ча* шакллари доим *-да, -ни, -га* шакллари билан олдин келади ва позиция омонимликни келтириб чиқармайди.

-лар кўшимчаси ҳурмат маъносини ифодалаганда эгалик шаклидан кейин, соф кўпликни билдирган эгалик шаклидан олдин жойлашади. Қиёсланг: *опамлар // опаларим*. Шунингдек, *опаларим* сўзшаклида *-лар* шахсга нисбатан умумийликни, *опамлар* сўзшаклида конкретликни билдиради. Морфоанализ ушбу ҳолатни фарқлашини қуйидаги модел билан таъминлаймиз. $f =$ -лар бўлганда: (1) агар $\{R+f+r\}$ бўлса, $f =$ кўплик; (2) агар $\{R+r+f\}$ бўлса, $f =$ ҳурмат маъноси.

Борларинг // боринглар, ўзларинг // ўзинглар сўзшакллари маъносида услубий фарқ сезилиб туради: *борларинг, ўзларинг* менсимасликни билдиради³⁴⁷.

Бизларники (-лар кўплик кўрсаткичи – ЛШЯ) ва *бизникилар* (-лар ЛШЯ ва СШЯ) сўзшакллари морфемалари позициясида ҳам функционал, ҳам семантик фарқ бор:

³⁴⁶Кўнғуров Р. Субъектив баҳо формаларининг семантик ва стилистик хусусиятлари. – Ташкент, 1980. – Б.145-160.; Махкамов Н. Аффиксальный и лексико-аффиксальный плеоназм в узбекском языке: Автореф. дис. канд. филол. наук. – Ташкент, 1983. – С.18.

³⁴⁷-лар морфемаси позициясининг ўзгаришининг маънога таъсирини А.М.Шербак ҳам таъкидлайди. Бу ҳақда қаранг: Шербак А.М. Последовательность морфем в словоформе как предмет специального лингвистического исследования// Вопросы языкознания, 1983. – № 3.

1) *бизларники* шакли тегишлиликни ифодалаган олмош: ноаниқликни ифодалайди, шахс ёки нарсага ноаниқ ишора қилади;

2) *бизларники* шакли тегишлиликни ифодалаган олмош, аммо шахс ва унинг кўплигини билдирувчи сўзга аниқ ишора қилади.

Демак, ушбу маъновий фарқ ҳам морфологик-услубий фильтрация қилиниши лозим.

Шунингдек, *болаларсиз* // *боласизлар* каби сўзларида *-лар* қуршовининг ўзгариши маъно ва вазифасига таъсир кўрсатади. *Болаларсиз яшаи зерикарли* // *боласизлар йиғилиши* қўлланишларида семантик фарқ мавжудлиги яққол сезилиб турибди³⁴⁸. Уларнинг вазифаси қуйидагича:

1) ҳаракат, ҳолатнинг инкорини (*болаларсиз яшамоқ* – қандай яшамоқ?) билдиради: сўз ясовчи вазифасини бажаради;

2) шахснинг мавжудлиги инкорини билдиради: (*боласизлар* – *боласиз оила*) сўз ясовчи вазифасини бажаради.

Аммо маънода кескин фарқ мавжуд: ҳатто сўзнинг туркуми ўзгаради. Бу каби ҳолатлар учун ҳам семантик фильтрацияси ишлаб чиқирилиши керак.

2. ЛШЯларнинг сифат сўз туркумидаги қуршови. Сифатга хос луғавий шакл ҳосил қилувчининг одатдаги тартиби мавжуд³⁴⁹. Сифатда *-роқ* кўшимчаси *-иш*, *-имтир* кўшимчасидан, шунингдек, равишдош, сифатдошга кўшилганда ҳам уларни ҳосил қилувчи кўшимчадан кейин кўшилади: *оқшироқ*, *кўкимтирроқ*, *совинқираганроқ*, *тортиниброқ*.

Сифат асосга (А) аввал миқдорнинг озлигини, камликни билдирувчи шакл ёки субъектив баҳо шакли кўшилади $A + f_1^a$: *кўкимтир*, *оқчил*, *оқиш*; *яхшигина*. Ундан кейин қиёсий даража шакли кўшилади: $A + f_1^a + f_2^a$: *кўкимтирроқ*, *оқшироқ*, *сарғишироқ*, *кичкинагина*. Сифат-асосга кўшилган -

³⁴⁸Лингвистик бирликларнинг валентлиги ҳақида қаранг: Степанова М.Д. Теория валентности и валентный анализ. – М., 1973.; Қўчқортоев И.К. Сўз маъноси ва унинг валентлиги. – Ташкент, 1977.; Абузалова М.К. Субстанциал морфология, валентлик ва синтактик қурилма: Филол. фан. доктори (DSc) диссер. Автореферати. – Самарқанд, 2018. – 84 б.

³⁴⁹Mengliyev B., Xoliyurov O., Abdurahmonova N. O'zbek tilidan universal qo'llanma. (Qayta ishlangan 3-nashri). – Toshkent: Akademyashr, 2014. – 389 b. – B. 134.

гина шакли отдагидан фарқли ўлароқ кучайтирув-таъкид маъносини билдиради, шу сабабли анализатор уни қуйидагича фарқлаши керак: агар {А+гина} бўлса, *-гина* = кучайтирув-таъкид маъноси. Айнан шу тартиб равиш-асосда ҳам кузатилади: *астагина, кўпгина, тезгина*.

3. ЛШЯларнинг сон туркумидаги қуршови. Сонга хос луғавий шакл ҳосил қилувчи морфемаларнинг одатдаги тартиби мавжуд³⁵⁰. Тахминни ифодалайдиган *-ча* аффикси “дона” маъносини ифодаловчи *-та* (урғусиз) унсуридан кейин келади: *ўнтача* каби.

Сон туркумидаги асосда ЛШЯларнинг қуршови қуйидаги позицияларни эгаллайди:

- 1) санок сонни шакллантирувчи урғусиз *-та* морфемаси биринчи позицияда туради: f_1^{num} ҳамда ундан кейинги позициялар қуйидагича бўлади:
- 2) тақсим маъносини билдирувчи *-дан* морфемаси – Num+ $f_1^{num}+f_2^{num}$: *учтадан* (*-дан* морфемасини бундай моделлаштириш бу шаклни чиқиш келишиги шаклидан фарқлаш заруратидан келиб чиқади);
- 3) чама сон ҳосил қилувчи *-ча* шакли – Num+ $f_1^{num}+f_2^{num}$: *ўнтача, бештача* (ушбу позицияда чама-тахмин маъносини ифодовчи *-ча* шакли равиш ясовчи ва кичрайтириш шаклидан фарқланади).

4. ЛШЯларнинг олмош сўз туркумидаги қуршови. Олмошга хос луғавий шакл ҳосил қилувчининг одатдаги тартиби мавжуд³⁵¹. Гумон билдирувчи *-дир* (урғусиз) унсури, умуман, юклама табиатида, ҳамма турдаги қўшимчадан кейин қўшилади: *нима-лар-ни-дир*.

Тахмин маъносини билдирувчи *-дир* аффикси олмош-асос(Pr)дан кейин жойлашганда, одатда, юклама характериға эға бўлади ҳамда бошқа қўшимчалар қаторида сўнгпозицияни эгаллайди: Pr+ $f_1+f_2+f_3^p$ – *кимларнидир, бизларникидир, улардандир*. Бундай позициядаги *-дир* шаклининг лингвистик

³⁵⁰Кўрсатилган манба.

³⁵¹Mengliyev B., Xoliyurov O., Abdurahmonova N. O`zbek tilidan universal qo`llanma. (Qayta ishlangan 3-nashri). – Toshkent: Akademyashr, 2014. – 389 b. – B. 134.

моделли мазкур кўринишда бўлади: агар $\{Pr+f_1+f_2+dip\}$ бўлса, $-dip =$ гумон юкламаси.

Хулоса шуки, сўзларда луғавий шакл ясовчи морфемалар қуршовини таҳлил қилиш, моделлаштириш морфоанализаторда ушбу шаклларнинг сўз ясовчи ва синтактик шакл ясовчилардан фарқланишини таъминлайди. Морфоанализаторда луғавий шакл ясовчилар позициясини аниқлаш учун 5 та инвариант модел асосида феълга хос 20 та, отга хос 5 та, сифат ва сонга хос 1 тадан, олмош учун 2 та субмодел ва унинг кўринишлари (жами 29 та субмодел) ишлаб чиқилди.

4.1.2. Синтактик шакл ясовчи морфемаларнинг сўздаги қуршовини лингвистик моделлаштириш. Синтактик шакл ҳосил қилувчи морфема сўзнинг луғавий маъносига таъсир этмайди, балки сўзни бир-бири билан боғлашга ёки унга маълум синтактик вазифа беришга хизмат қилади. Мутахассисларнинг фикрича, синтактик шакл ясовчининг қўлланиши луғавий шаклдек маълум сўз туркуми билан чегараланмайди, яъни бу шакл истаган мустақил сўз туркумига мансуб сўзга қўшилиб, уни ўзидан кейин ёки олдин келган сўз билан алоқасини таъминлайди: *қорасини олмоқ, қизили ёқади, келганимни ёқтирмади, ўқиши яхши эмас, шундайидан бер, кўпини ўтказди* ва ҳ³⁵².

Сўздаги жойлашувига кўра синтактик шакл ясовчи (СШЯ) морфема – асосдан энг узоқ жойлашадиган морфема: бу шакл асосдан кейин жойлашадиган ЛШЯ морфемадан кейин келади. Одатда, бу қўшимчанинг сўздаги позицияси қуйидаги қуршовга эга бўлади: асос+(агар мавжуд бўлса, сўз ясовчи)+ЛШЯ+СШЯ ёки R+d+f+r. Синтактик шаклнинг сўздаги ўзаро тартиби ҳам мавжуд³⁵³:

а) эгалик қўшимчаси+келишиқ қўшимчаси: *боламни, ўқиганимни*;

³⁵²Замонавий ўзбек тили. Морфология. (Масъул муҳаррир: Р.Р.Сайфуллаева). – Тошкент: Мумтоз сўз, 2008. – 522 б. – Б. 217-218.

³⁵³Замонавий ўзбек тили. Морфология. (Масъул муҳаррир: Р.Р.Сайфуллаева). – Тошкент: Мумтоз сўз, 2008. – 522 б. – Б. 220.

б) муносабат шакли: тасдиқ-инкор+замон/майл+шахс+сон: *ўқимаяпсан*.

Ҳамма турдаги қўшимчадан кейин қўшимча типдаги юклама қўшилади: *ўқимаяпсанми*. Қўшимча қўшилишида истисно ҳолат ҳам бор. Тил ходисаларининг даврга кўра ўзгариши, диалектал фарқи ва ўзгариш хусусияти бадий талаб билан боғлиқ бўлади. Мисол: *Билурлар эрди (Қутб) – билар эдилар, қилмагайларсиз («Бобурнома») – қилмагайсизлар, солурлар эрди қулоқ – қулоқ солур эдилар*. Бундай тартиб ўзгариши ўзида нозик маъновий фарқни акс эттиради. Қуйида шундай шаклларнинг позициясини морфоанализатор тадқиқи учун лингвистик моделлаштириш масалалари таҳлил қилинади.

Исмларнинг СШЯ морфемаларнинг одатдаги тартибини қуйидагича моделлаштирамиз: $N / A / Num / Pr + r_1^n + r_2^n$. Бунда: N = от-асос; A = сифат-асос; Num = сон-асос; Pr = олмош-асос; r_1^n = эгалик; r_2^n = келишик.

Морфоанализатор таҳлилида исмларнинг муносабат шакллари – синтактик шакл ясовчиларнинг қуршовини аниқлаш ва моделлаштириш учун эгалик ва келишик шакллари парадигмасини келтирамиз. Ушбу шакллар феълнинг синтактик шакл ясовчи морфемаларидан фарқлашда керак бўлади.

Эгалик шакллари парадигмаси: $r_1^n = -им / -инг / -(с)и / -(й)и / -(и)миз / -(и)нгиз$.

Келишик шакллари парадигмаси: $r_2^n = -нинг / -ни / -га / -да / -дан$.

Феълларнинг СШЯ морфемалари барча қўшимчалардан кейин келади. Ҳозирги ўзбек адабий тилида мавжуд шахс-сон кўрсаткичининг қуйидаги тартиби кузатилади³⁵⁴:

- 1) феъл-асос + { -ган / -(и)б / -яп / -а / -й / -моқчи } + { -ман / -сан / -0 / -миз / -сиз / -(ди)лар }.
- 2) Феъл-асос + { -ди / -са } + { -м / -нг / -0 / -к / -нгиз / -(ди)лар }.
- 3) Феъл-асос + { -ай / -0 / -син / -айлик / -гин / -(и)нгиз / -(син)лар / +-ишсин }.

³⁵⁴Замонавий ўзбек тили. Морфология. (Масъул муҳаррир: Р.Р.Сайфуллаева). – Тошкент: Мумтоз сўз, 2008. – 522 б. –Б. 420-426.

Ушбу кўшимчалар куршовини морфоанализ учун қуйидагича моделлаштирамиз:

- 1) $V+f_1^v+r_1^f$. Бунда: $f_1^v = \{-ган /-(и)б / -яп / -а / -й /-моқчи\}; r_1^f = \{-ман /-сан / -0 / -миз / -сиз/ -(ди)лар\};$
- 2) $V+f_2^v+r_2^f$. Бунда: $f_2^v = \{-ди /-са\}; r_2^f = \{-м /-нг / -0 / -к / -нгиз/ -(ди)лар\};$
- 3) $V+ r_3^f$. Бунда: $r_3^f = \{-ай/-0 / -син / -айлик / -гин/ -(и)нгиз / -(син)лар / +-ишсин\}.$

Юқорида келтирилган маълумотлардан кўринадикки, исм ва феълларга хос СШЯ морфемалар грамматик омонимия ҳосил қилади. СШЯлар доим сўзнинг охирида келганлиги сабабли морфоанализатор иш жараёнида ўнгдан чапга қидирувда уларнинг айрими шаклдош бўлиб қолади, яъни анализатор маълумотлар омборидаги маълумотга таяниб, *укангиз* ва *келдингиз*, *бордим ва укам*, *келсанг ва уйқунг* сўзларида -нгиз шаклини СШЯ деб топади. Шу ерда савол туғилади: ушбу СШЯ отга хосми ёки феълга хосми? Морфоанализатор жавоб топиши учун иккинчи алгоритмни ишга солади, бунинг учун модел талаб қилинади. *-нгиз* шаклининг қайси СШЯ эканлигини қуйидаги модел ёрдамида аниқлаш мумкин:

- 1) агар $N / A / Num / Pr + r_1^n$, яъни $\{\text{исм-асос}+-(и)м / -инг / -(и)миз\}$ бўлса, $r^n =$ эгалик шакли;
- 2) агар $V + r^f$, яъни $\{\text{феъл-асос}+-(и)м / -инг / -(и)миз\}$ бўлса, $r^f =$ шахс-сон кўрсаткичи.

Ҳозирги ўзбек адабий тилида сўзнинг морфем таркиби тартибида одатдаги тартибдан чекиниш ҳолати – СШЯ морфеманинг ЛШЯдан олдинда келиши кузатилади: *айтинглар*, *юринглар*, *боринглар*. Бунинг сабабини қуйидагича изоҳлаш мумкин:

-(и)нг аффикси юқоридаги мисоллар таркибида II шахс кўпликни эмас, балки ҳурмат маъносини билдиради. Бу маънонинг ифодаланиш табиати билан боғлиқ. II шахс бирлик шаклидаги *ўқи*, *ёз*, *кел*, *кет* сўзлари -нг ёрдамида синтактик алоқага киришиши мумкин, аммо бу грамматик маънони

ифодаламайди. *Айтинг* сўзшакли ҳурмат маъносида *сиз* олмоши билан (сиз олмоши грамматик жихатдан кўпликни билдиради), *айт* сўзшакли *сен* олмоши билан келади. Бу ўринда *сиз* олмоши (*сен* олмоши билан зидланишда) кўплик шаклини билдирса-да, у одатда, ҳурматни билдиради. Бу сўзнинг семантик тараққиёти билан боғлиқ. Ҳурмат маъноси англашиладиган контекстда *-инг* морфемасидан кейин *-лар* шакли келади: *-инг + -лар*. Демак, синтактик шакл ясовчидан кейин ЛШЯ морфема келиши ҳам мумкин.

Адабиётларда айтилганидек, *айтинглар*, *боринглар* сўзшакллари таркибидаги *-лар* фақат луғавий шакл ясовчи эмас: у алоқа-муносабат шакли сифатида юзага чиқади: эга ҳамда кесимни шахс-сонда мослаштиради. Шу қоидадан келиб чиқиб, *-лар* шакли сўз охирида келади ҳамда СШЯ вазифасини бажаради: *бордилар*, *ёздилар*. Адабий тил нормасидан четлашиш ҳолати – шевада *борларинг*, *ёзларинг* шакли ҳам учрайди; қаҳрамоннинг индивидуал лексикаси, характерини яратиш учун бундай шакллар бадиий асарларда ҳам ишлатилади. Морфоанализатор исталган матнни таҳлил қила олиши кераклигини инобатга олсак, унинг лингвистик таъминотини тузиш жараёнида тилдаги адабий меъёр билан бирга номеъёрий, аммо кенг ишлатиладиган ҳолатларни инобатга олиш зарурати мавжудлиги маълум бўлади.

Таъкидлаш жоизки, мутахассисларнинг ҳозирги ўзбек тилида *-лар* кўшимчасининг икки турли функцияси: кўпликни ифодалаши (луғавий шакл ясовчи шакл –f) ҳамда синтактик муносабатни юзага чиқариши (синтактик шакл ясовчи –r) ҳақида асосли фикрлари мавжуд³⁵⁵. МА учун қуршов асосида синтактик шакл ясовчилар вазифасини фарқлашнинг 1 та инвариант модели асосида 6 та субмодели ишлаб чиқилди.

4.1.3. Кўшимча кўринишидаги юкламаларнинг қуршовдаги маъносини аниқлаш масаласи. Ўзбек тилида кўпгина лексоид ёрдамчи сўзларнинг аффиксоид кўринишлари мавжуд. Бу юкламага ҳам тегишли: *ҳам*

³⁵⁵Жўрабоева М. Аффиксал омонимик қаторни белгилашга доир (-ган ва -лар аффикслари мисолида). – Научные труды ТашГУ. Вып. 427. Вопросы узбекского языкознания. – Ташкент, 1972. – С. 169-177.

-ам /-ям; учун, -чун; -у, -ю; билан, -ила, -ла шулар жумласидан. Юкламалар шаклий жиҳатдан уч хил кўринишга эга:

- 1) кўшимчасимон юкламалар (-ми, -чи, -а, -я...).
- 2) соф юкламалар (*худди, фақат, ахир...*).
- 3) нисбий юкламалар (*ёлғиз, баъзан, танҳо, бир...*)³⁵⁶.

Одатдаги тартибда кўшимчасимон юкламалар сўз таркибида барча кўшимчалардан кейин келади. Бу тартиб барча туркумлар учун қарийб бир хил: *бор+иш+ма+ди-ку, сув+чи+лари+миз+ни+ми*. Юкламанинг сўзга кўшилиш тартибини қуйидагича моделлаштириш тўғри бўлади: $R+f_1+f_2+r_1+r_2$ +юклама ёки $R+d+f_1+f_2+r_1+r_2$ +юклама. Аммо айрим юкламалар сўзда қуршовини ўзгартириши ҳам кузатилади. Қуйида шундай ҳолатлар, уларнинг сабаблари ва ўзгарган тартибни аниқлашнинг лингвистик моделлари хусусида фикр юритамиз.

Сўроқ-таажжуб юкламаси -ми, -чи, -а, -я доим сўз охирида келади: *биз-чи? бизни-чи? ўқувчилар-чи? бординг-а? айтмади-я. -у, -ю, -ку* кўшимча юкламалар ҳам барча турдаги кўшимчалардан кейин келади: *бўлади-ю, келди-да, борлар-ку, борганлар-у.*

-гина айирув-чегаралов юкламаси ҳам, одатда, сўз охирида келади. Агар ушбу юклама позициясини ўзгартирса – синтактик шакл ясовчилардан олдин келса, у луғавий шакл ясовчи вазифасини бажаради. Бу масала ишнинг олдинги қисмларида кўриб ўтилганлиги сабабли бу ерда масалага такрор тўхталмаймиз.

-ми сўроқ юкламаси турли сўз туркумларида турли сабаблар билан ўз ўрнини ўзгартириш ҳоллари кузатилади:

- 1) кўпинча, *эди* тўлиқсиз феъли ёки унинг қисқарган шаклидан кейин, баъзан олдин жойлашади: *кўрмабмидинг (кўрмабми эдинг), келиб турармиди (келиб турардимми), сизмидингиз (сизми эдингиз), ётганмидинг (ётгандингми),*

³⁵⁶Замонавий ўзбек тили. Морфология. (Масъул муҳаррир: Р.Р.Сайфуллаева). – Тошкент: Мумтоз сўз, 2008. – 522 б. – Б.497.

кетармиди (кетардимми), ишлармидик (ишлардикми), бормасмиди (бормасдимми) ва ҳ. Бундай тартиб ўзгариши услубий ўзига хосликни таъминлаш, тарихий ва диалетик омиллар билан боғлиқ, демак, ўзига хос маъно ифодалайди. Одатдаги тартибдан чекланиш маънони ўзгартиришга хизмат қилса, юклама позициясининг маъновий ўзига хослигини қўйидагича моделлаштирамиз. Агар {феъл/олмош-асос + ЛШЯ + ЛШЯ + юклама + СШЯ + СШЯ} бўлса, юклама тарихий/диалектик шаклда қўлланилган. МАдаги модели: $\{V \setminus P + f_1 + f_2 + \text{юклама} + r_1 + r_2 = \text{тарихий/диалектик қўлланиш}\}$;

2) *уйдамикин, чиқармикин, ростмикин, бормасмиканман* каби сўзларда *-микин/-микан* шакли *-ми* экан шаклининг қисқарган кўриниши саналади. Анализатор базасига қўйидаги қоида киритилади: агар {асос + ЛШЯ/СШЯ + *ми* + *кин*} бўлса, *-кин* унсури экан тўлиқсиз феълнинг қисқарган шакли. МАдаги модели: $\{V \setminus P + f_1/r_1 + \text{ми} + \text{кин} = \text{сўроқ юкламаси} + \text{экан тўлиқсиз феъли}\}$;

Келарканми, борарканми, юрибдиканми сўзларида *-ми* юкламаси экан тўлиқсиз феълнинг қисқарган шаклидан олдин жойлашган. Бундай ҳолатда морфоанализ жараёни ўнгдан чапга таҳлил қоидаси орқали ҳеч қандай моделсиз тўғри таҳлил қилади, чунки, анализатор базасида экан шаклининг сўз охирида келиши, у қисқарган шаклда *-кин* морфемаси сифатида қўлланиши ҳақида ахборот мавжуд;

3) *омонмисан, бормисиз, шу ердამисиз, бормоқчимисиз, демоқчимисиз* сўзшаклларида *-ми* юкламаси синтактик шакл ясовчи морфемадан олдин келган; *борасизми, дейсанми, танияпсанми, келдингми, боргансизми, келдингизми, чўчияпсанми* сўзларида синтактик шакл ясовчи морфемадан кейин жойлашган. Юкломанинг икки хил қўлланиши унинг маъновий ўзига хослиги билан боғлиқ. Бундай жойлашув морфоанализ таҳлилига таъсир кўрсатмайди, махсус моделга эҳтиёж йўқ;

4) эски ўзбек тилида *-ми* юкламаси кесимдан ташқари барча гап бўлаклари билан қўлланади. Масалан, *сен китобними келтирдинг* гапида *-ми*

юкламаси тўлдирувчига қўшилган. Юкламанинг бундай қўлланилиши диалектик ва тарихий ўзига хослик билан изоҳланади ҳамда адабий тил меъёрига зид келади³⁵⁷. Аммо тил материалларини кузатишимиз ҳозир ҳам шундай қўлланилиш учрашини кўрсатди. Масалан, *Китобнингми, журналнингми саҳифасида; Китобними келтирди; Ойшагами қаради; Ойшадами гапи бор; Ойшаданми қарз экан* гапларида келишик шаклларида кейин келади ҳамда қисман ўзак(сўз)дан англашилган тушунча (предмет, шахс)ни инкор этади: аниқ айтилмаган *китобми, Ойшами? Китобни келтирдингми* гапида *-ми* тусланган феълдан кейин қўшилган, аниқроғи, бутун гапга сўроқ маъносини юклаган. Бу ўринда *-ми* фақат синтактик-модаллик муносабатини билдирувчи сўроқ-таажжуб юкламаси. *-ми* юкламасининг гап ва отлардан кейин қўлланилиши улар орасида нафақат семантик, балки грамматик фарқ мавжудлигини кўрсатди. Солиштиринг: от билан келган *-ми* – модаллик кўрсаткичи; феъл билан қўлланилган *-ми* эса синтактик модаллик кўрсаткичи. Ушбу ахборотни морфоанализатор базасига қуйидаги қоида ва модел орқали киритамиз:

- 1) агар {от-асос + ЛШЯ/СШЯ + ми} бўлса, *-ми* юкламаси – модаллик кўрсаткичи: {N + f₁/r₁ + ми} бўлса, {-ми = M};
- 2) агар {феъл-асос + ЛШЯ/СШЯ + ми} бўлса, *-ми* юкламаси – синтактик модаллик кўрсаткичи: {N + f₁/r₁ + ми} бўлса, {-ми = SM}.

Сўз таркибида бирдан ортиқ юклама қўлланиши мумкин. Табиийки, уларнинг ўзаро жойлашишида ҳам маълум тартиб кузатилади:

- 1) гумон юкламаси кучайтирув-таъкид юкламасидан олдин жойлашади: *келгандир-да;*
- 2) кучайтирув-таъкид юкламасидан кейин таажжуб маъносини ифодаловчи сўроқ-таажжуб юкламаси жойлашади: *майлику-я, кўринасизку-а;*

³⁵⁷*-ми* юкламасининг турли позицияда қўлланилиши хақида қаранг: Ўзбек тили грамматикаси. 1-том. – Б. 438-439.; Бозоров О. Ўзбек тилининг юклама категорияси. – Ўзбек тили ва адабиёти, 1977. – № 5. – Б.24-28.

3) агар *-а* юкламаси таажжуб, ҳайратланишни билдирса, сўроқ юкламасидан кейин келади: *нарсами-а?*, *сенмисан-а?*

Шунингдек, юкламалар қуршови ва сўздаги позицияси услубни ҳам белгилайди; у ҳақида ахборот беради: *ишлармидик // ишлардикми, шу ердасизми // шу ердамисиз* ва х. Демак, юкламанинг қуршови ҳам унинг семантикаси билан боғлиқ бўлади. Бундай қўлланиш ҳақидаги ахборот маълумотлар базасига киритилса, гапларда юкламаларнинг қўлланиши орқали матн услубини аниқлаш имкони туғилади. Юқоридаги моделлар асосида қўлланиш позицияси қайси услубга тегишли эканлиги ҳақида тег бириктириш кифоя қилади. Қўшимча кўринишида юкламаларнинг сўздаги қуршовини аниқлаш учун 2 та инвариант модел асосида 4 та субмодел ишлаб чиқилди.

Умуман олганда, морфоанализаторнинг ўзбек тили луғавий шакл ясовчи морфемалар қуршовини аниқлаши учун 9 та инвариант модел асосида 39 та субмодел кифоя эканлиги аниқланди ва улар ишлаб чиқилди.

4.2. Сўз ясовчи морфемаларнинг сўздаги қуршови масалалари

Туркий тилдаги сўзлар структураси мураккаб бўлишига қарамай, асосан, ўзак ва қўшимчадан иборат. Маълумки, агглютинатив тилларда қўшимчалар асосга кетма-кетлик асосида қўшилади. А.М.Шербак таъкидлаганидек, “исталган туркий сўзда таркибли элементларда қатъий структура мавжуд: чап томонда асос морфема, ўнг томонда аффиксал морфемалар жойлашади. Аффиксал морфемаларнинг контекстдаги ўрни қатъий, улар маълум изчилликда кетма-кет жойлашади ҳамда маълум қуршовга эгаллиги билан характерланади”³⁵⁸.

³⁵⁸Шербак А.М. Очерки по сравнительной морфологии тюркских языков. – Л., 1977. – С. 21.

Сўз ясаши, сўз асоси морфологиясининг айрим масалалари Н.Д.Арутюнов³⁵⁹, сўзнинг морфологик структураси Н.А.Баскаков³⁶⁰, Э.В.Севортян³⁶¹, турк тилида сўзнинг морфологик структураси М.С.Джикия³⁶², туркий агглютинация табиати А.Н.Кононов³⁶³ларнинг ишларида ўрганилган.

Сўз ясовчи морфеманинг асосий хусусияти – деривация; унинг ўзакка қўшилишидан янги лексик бирлик ҳосил бўлади. Ушбу бўлимда сўз ясовчи морфема қуршови – позициясининг сўз ясовчи, луғавий шакл ясовчи (ЛШЯ), синтактик шакл ясовчи (СШЯ) морфемалар билан омонимлик ҳосил қилишга таъсири ҳамда унинг натижалари таҳлил қилинади. Морфемалар сўздаги позициясининг маънога таъсири Н.К.Якубова³⁶⁴ томонидан ўрганилган. Олиманинг таъкидлашича, сўз ясовчи ва ЛШЯларнинг позицияси унинг функциясини белгилайди.

Маълумки, “ясама сўз ясовчи компонентга эга бўлади, шу компонент бирикувидан ясалма ҳосил бўлади. Сўз ясашида иштирок этувчи бундай компонентлар ясама сўз структурасини ташкил этади, яъни ҳар бир конкрет ясама сўзни юзага келтирувчи компонентлар бирликда сўз ясашиш структурасини ҳосил қилади. Ясама сўзнинг структурасида камида иккита сўз ясовчи компонент иштирок этади”³⁶⁵.

Морфоанализаторнинг ясама сўз, ясовчи қисм ва ясашиш асосини аниқлай олиши учун унинг лингвистик таъминотида ясама сўз структураси ҳақидаги мукамал маълумот талаб қилинади. Ясама сўз структурасида

³⁵⁹Арутюнов Н.Д. Некоторые вопросы образования и морфологии основ слов. – Филологические науки, 1958. – № 1.

³⁶⁰Баскаков Н.А. Морфологическая структура слова и части речи в тюркских языках. – Советское востоковедение, 1957. – № 1.; Баскаков Н.А. Порядок и иерархия аффиксов в основе слова в тюркских языках. Баскаков Н.А. В сб.: Морфологическая типология и проблема классификации языков. – М.-Л., 1965.

³⁶¹Севортян Э.В. Морфологическое строение слова в связи с другими его характеристиками. – В. Кн.: Тюркологический сборник. – М., 1972.

³⁶²Джикия М.С. Морфологическая структура слова в турецком языке: автореф. дисс. ... канд. филол. наук. – Тбилиси, 1975.

³⁶³Кононов А.Н. О природе тюркской агглютинации// Вопросы языкознания, 1976. – № 4.

³⁶⁴Якубова Н.К. Расстановка аффиксов в узбекских словаформах. – Ташкент, 1979.

³⁶⁵Ҳожиёв А. Ўзбек тили сўзясалиши. Олий ўқув юртарининг филология факультетлари студентлари учун қўлланма. – Тошкент: Ўқитувчи, 1989. – 109 б. – Б. 13.

маъноси ва вазифасига кўра ўзаро фарқланадиган икки хил компонент (икки хил ҳодиса) мавжуд бўлади, яъни ясама сўзда: 1) сўз ясаиш асоси (ёки мотивловчи асос); 2) сўз ясовчи формант бўлади. Ясама сўзнинг ҳосил қилинишида асос бўладиган, ясама сўзнинг маъносига асос бўладиган (ясама сўз маъносини мотивлайдиган) компонент сўз ясаиш асоси (мотивловчи асос) ҳисобланади. Мас., *туртки, ишчан, қорай* сўзларида *турт, иш, қора* сўз ясаиш асосидир (сўз ясаиш асоси *ясовчи асос* деб ҳам юритилади)³⁶⁶.

Анализаторнинг асослар луғатида ўзбек тилидаги барча ўзақлар туркумга тегишлилик ҳақидаги маълумот билан сақланади, яъни барча ўзақлар леммаланган, қайси сўз туркумига тегишлилиги ҳақидаги ахборот – тег бириктирилган бўлади.

А.Ҳожиёвнинг таъкидлашича, аффиксация усули билан ясалган ҳар бир сўзда сўз ясаиш асоси битта бўлади. Бу асос ўзак-лексемага, содда ясама сўзга, қўшма сўзга, сўз бирикмаси, ҳатто, гапга тенг бўлиши мумкин: *терим* (тер-им), *гавдали* (гавда-ли), *гулла* (гул-ла), *кийимлик* (кийим-лик), *бирлашма* (бирлаш-ма), *теримчи* (терим-чи), *теримчилик* (теримчи-лик); *асаларичи* (асалари-чи), *белбоғли* (белбоғ-ли); *беш йиллик* (беш йиллик), *темирйўлчи* (темирйўл-чи), *ўзибўларчилик* (ўзибўлар-чилик) каби. Демак, аффиксация усули билан ясалган сўзнинг структураси қандай мураккаб бўлмасин, шунингдек, бир сўзнинг таркибида бир неча сўз ясовчи аффикс бўлмасин, унинг сўз ясаиш структураси шу ясаишга асос бўлган қисм – сўз ясаиш асоси ва шу сўзни ясалган қисм – сўз ясовчи аффиксдан иборат бўлади³⁶⁷.

Аввалги бобда таъкидлаганимиздек, МА морфемаларни ўнгдан чапга қараб аниқлайди, таҳлил қилади. Бу ўринда МА уч нарсани: сўзнинг асоси, ясаиш асоси ва сўз ясовчи аффикс аниқлаши керак. Уларни аниқлаш қуйидаги алгоритмда бўлади:

1) ўзак асослар луғатида мавжуд бўлади, демак, асос топилади;

³⁶⁶Кўрсатилган манба.

³⁶⁷Кўрсатилган манба. Б.14-15.

2) маълумотлар базасида ясовчи қўшимчалар ҳақидаги маълумотга асосланиб, ўнгдан чапга қараб таҳлил қилинади: сўз ясовчи қўшимча аниқланади;

3) ясовчи асосдан олдинги қисм ясаиш асоси деган хулоса келиб чиқади: ясаиш асоси топилади³⁶⁸.

Морфоанализатор ясама сўзларни ушбу алгоритм асосида аниқ таҳлил қилиши маълумотлар базасидаги моделларнинг мукамал ишланиши билан боғлиқ: моделлар қанча аниқ ишланса, таҳлил шунча аниқ бўлади. Бунинг учун инвариант модел ва субмоделлар ишлаб чиқилиши лозим.

Сўз ясаишининг асосий инвариант моделини қуйидагича белгилаймиз: $R+d$. Бунда R – ҳар қандай асос; d – дереватема, яъни сўз ясовчи морфема. Асосий $R+d$ инвариант модели қуйидаги субмоделларда воқелашади: R_v+d ; R_n+d ; R_a+d ; $R_{num}+d$; $R_{pr}+d$; $R_{adv}+d$; $R_{sim}+d$.

4.2.1. R_n+d субмоделидаги сўз ясовчи морфемаларнинг сўздаги куршови. Дереватемаларнинг хусусиятидан келиб чиқиб, R_n+d субмодели R_n+d^n , R_n+d^{adj} , R_n+d^{adv} , R_n+d^v , яъни от-ўзак + от ясовчи, от-ўзак + сифат ясовчи, от-ўзак + равиш ясовчи ва от-ўзак + феъл ясовчи каби типларга бўлиш мумкин. Қуйида ушбу моделларнинг кўринишлари сўздаги куршов ва маънога таъсири нуқтаи назаридан батафсил ёритилади.

Юқорида берилган *терим, гавдали, гулла, кийимлик, бирлашма, теримчи, теримчилик; асаларичи, белбоғли; беш йиллик, темирйўлчи, ўзибўларчилик* каби ясама сўзларнинг таҳлили қуйидагича бўлади: Морфоанализатор ўнгдан чапга таҳлилда сўз ясовчи морфемаларни аниқлаши керак: *-им, -ли, -ла, -лик, -ма, -чи, -чилик, -арчилик*. Бир бўғиндан иборат қўшимчаларни аниқлашда муаммо бўлмайдими, аммо *-чилик, -арчилик* шаклларининг ўнг томондан ўқилишида дастлабки *-лик* шакли ажратиб олиниши мумкин. Шунинг учун маълумотлар омборига қоида ва модел

³⁶⁸Степанова М.Д. Вопросы моделирования в словообразовании и условия реализации моделей. – Вопросы языкознания, 1975. – № 4.

киритамиз. Агар {ўзак + лик} бўлса, *-лик* сўз ясовчи бўла олади. Модели: 1) агар $\{R + d_1(\text{лик})\}$ бўлса, $\text{лик} = d$; 2) агар $\{R + d_1 + d_2(\text{лик})\}$ бўлса, $d_1 + \text{лик} = d$.

А.Ҳожиев ясама сўз структурасида сўз ясалиш асосидан ташқари сўз ясовчи *формант* ҳам бўлишини (сўз ясовчи компонент *формант* деб ҳам юритилади) уқтиради. Бу компонент сўз ясалиш асосига нисбатан умумийроқ маънога эга бўлади, бир типдаги бир неча сўзларнинг ясалишида иштирок этади. Мотивланувчи сўз худди шу формантга хос маъно билан мотивловчи асосдан фарқланади: ясама сўзга хос сўз ясалиш маъноси мана шу сўз ясовчи формантга оид бўлади. Сўз ясовчи формант, асосан, сўз ясовчи аффикс ва аффиксоидга тенг бўлади. Масалан, *бугдойзор, бедазор, тутзор, жийдазор; ишхона, тўйхона, касалхона* каби³⁶⁹.

Ҳозирги ўзбек тилида мустақил лексик маънога эга бўлган бирлик (лексема) бир неча ясама сўзнинг ҳосил бўлишида иштирок этиши ҳамда ҳаммасида бир хил маъно билан қатнашиши мумкин. Бундай компонент ҳам сўз ясовчи формантга хос хусусиятга эга бўлади, бинобарин, сўз ясовчи аффикс ёки аффиксоидга хос функция бажаради. Масалан, *камгап, камдаромад, камқатнов, камҳосил* каби сўзлар таркибидаги *кам* компоненти шундай хусусиятга эга. Уларни ўз антонимларига қиёсласак, бу ҳодиса яна ҳам очик кўринади: *камгап – сергап, камдаромад – сердаромад, камқатнов – серқатнов, камҳосил – серҳосил*. Худди шу критерий асосида кўшма феъл ҳосил қилувчи *қил (эт, айла)*, бўл феълларини ҳам сўз ясовчи формант ҳисоблаш мумкин: *хурсанд қилмоқ, жабр қилмоқ, сабр қилмоқ, гина қилмоқ; карахт бўлмоқ, касал бўлмоқ* ва б. Сўз ясалиш формантига хос белгига эга бўлгани учун ҳам бундай феъллардаги *қил, бўл* феъллари ўрнида феъл ясовчи аффикс(формант)ни қўллаш мумкин бўлади: *тасдиқ қилмоқ – тасдиқламоқ, афсус қилмоқ – афсусланмоқ, баҳс қилмоқ – баҳслашмоқ; дўст бўлмоқ*

³⁶⁹Ҳожиев А. Ўзбек тили сўз ясалиши. Олий ўқув юр்தларининг филология факультетлари студентлари учун қўлланма. – Тошкент: Ўқитувчи, 1989. – 109 б. – Б. 15.

– дўстлашмоқ, зада бўлмоқ – задаланмоқ, лоҳас бўлмоқ – лоҳасланмоқ, мағрур бўлмоқ – мағрурланмоқ ва б³⁷⁰.

Кўринадики, сўз яшаш модели ясама сўзнинг тузилиш схемаси, демак, бу схема сўз яшаш схемасидир. У сўзнинг яшалиш структурасини текширишнинг умумлаштирилишидан туғилган: сўзнинг яшалиш структурасини аниқлаш билан умумлаштириш натижасида сўз яшаш моделлари белгиланади, бу моделлар тилнинг сўз яшалиш системасини ташкил қилади³⁷¹.

А.Ҳожиёв академграмматикадаги ушбу фикрга таяниб ўринли таъкидлаганидек, сўз яшалишига оид қуйидаги умумий ҳодиса (тушунча)лар мавжуд: 1) сўз яшаш модели; 2) сўз яшалиши структураси; 3) сўз яшалиши маъноси; 4) сўз яшаш моделининг унумли-унумсизлиги; 5) сўз яшаш моделининг ички турлари; 6) сўз яшалиши системаси³⁷².

Сўз ясовчи морфемаларнинг кўпчилиги яшалиш жараёнида категориал ўзгаришларга сабаб бўлади: бир сўз туркумидан бошқаси ясалади, бу сўз ташқи сўз яшалиши саналади (отдан сифат, сифатдан от, феълдан от яшалиши каби); ички сўз яшалиши (сўз туркумида ўзгариш бўлмайди: тунока+чи) нисбатан камроқ учрайди.

М.Д.Степанов сўз яшалиши моделлари ва уларнинг воқеланиши масалаларини атрофлича ёритаркан³⁷³, моделларнинг инварианти ва вариантини тузиш тамойилларини белгилайди.

Сўз яшалишининг одатдаги тартибида ясовчи морфема ўзак/негиздан кейин жойлашади. Кузатишлар кўрсатадики, ўзбек тилида бир асосга бештагача сўз ясовчи морфема қўшилиши мумкин. Бунда, албатта, сўз ясовчиларнинг ички кетма-кетлиги бўлади. Қуйида яшалиш асоси ва битта

³⁷⁰Кўрсатилган манба. Б. 15.

³⁷¹Ўзбек тили грамматикаси, I том, Морфология. – Тошкент: Фан, 1975. – Б. 12-13.

³⁷²Ҳожиёв А. Ўзбек тили сўз яшалиши. Олий ўқув юртлирининг филология факультетлари студентлари учун қўлланма. – Тошкент: Ўқитувчи, 1989. – 109 б. – Б. 17.

³⁷³Степанова М.Д. Вопросы моделирования в словообразовнии и условия реализации моделей. – Вопросы языкознания, 1975. – № 4.

ясовчининг позицияга қараб ўзгариши ҳамда унинг маънога таъсирини аниқлаш жараёни учун зарур бўладиган моделлар тавсифланади.

1. **Отдан от ясалиши** инвариант R_n+d^n модели билан белгиланади. Таркибий қисми бир хил, ушбу инвариантдан маъноси билан фарқ қилувчи структураларни кўриб чиқамиз. Бундай қонуният асос ва СЯ морфеманинг турли хил бирикишларига барҳам беради³⁷⁴. Масалан, *сувчи, ишчи* сўзларидаги *-чи сув, иш* ўзакларига қўшилади, аммо бу СЯ морфема сифат-асосга *оқчи, кўкчи, баландчи, юмшоқчи* тарзида қўшила олмайди³⁷⁵. Демак, *-чи* шахс оти ясовчиси R_n+d_1 модели билангина сўз ясай олади:

1) шахс оти ясалиши $R_n+d_1^n$ моделига эга бўлади, бунда $d_1^n = -чи, -шунос, -кор, -каш, -хон, -гар, -соз, -наз, -боз, -хўр, -параст, -гўй, -дор, -бон, -фуруш, -дош, -вачча, -ат, -лик, -парвар, -тарош, -манд, -и$;

2) ўрин-жой оти ясалиши $R_n+d_2^n$ моделига эга бўлади, бунда $d_2^n = -зор, -лоқ, -истон, -зоҳ, -хона, -ов, -кент, -обод$;

3) предмет маъносини билдирувчи от ясалиши $R_n+d_3^n$ моделига эга бўлади, бунда $d_3^n = -дон, -нома, -лик$ ³⁷⁶;

4) касб-кор ва фаолиятни билдирувчи от ясалиши $R_n+d_4^n$ моделига эга бўлади, бунда $d_4^n = -лик, -чилик, -гарчилик, -гарлик$ ³⁷⁷.

А.Ҳожиёв сўз ясалиши типини тушунчасини ажратади. Бир хил сўз ясалиши маъносига эга бўлган ясама сўзлар сўз ясалишининг бир типини ташкил этади. Масалан, спорт ўйини турини билдирувчи *волейбол, футбол, бокс* каби сўзлардан ясалган *волейболчи, футболчи, боксчи* сўзлари *-чи* аффикси билан ясалган шахс отларининг бир типини; *арра, ранда, эгов, дазмол* каби асбоб-қурол

³⁷⁴Улуханов И.С. Значения словообразовательных аффиксов и части речи мотивирующих слов. – Филологияческие науки, 1974. – № 4.

³⁷⁵Степанова М.Д. Вопросы моделирования в словообразовании и условия реализации моделей//Вопросы языкознания, 1975. – № 4. – С. 53-63.

³⁷⁶Ўзбек тилида –лик аффикси билан от ясалиши масаласи ҳақида қаранг: Нуъмонова Ш. Ўзбек тилида -лик аффикси билан ясалган отларнинг семантик классификацияси ҳақида. – Тилшунослик ва адабиётшуносликка оид тадқиқотлар. – Тошкент, 1965.

³⁷⁷Ушбу морфемаларнинг ҳақида тўлиқ маълумот олиш учун қаранг: Шарипова К.А. Словообразующие аффиксы -лик, -чи, -чилик в узбекском языке: автореф. дисс. ... канд. филол. наук. – Тошкент, 1972.; Ҳожиёв А. Ўзбек тилида -лик аффиксининг вазифаси ҳақида//Ўзбек тили ва адабиёти, 1985. – № 1.

билдирувчи сўзлардан ясалган *аррала, рандала, эговла, дазмолла* сўзлари -ла аффикси билан феъл ясалишининг бир типи ҳисобланади³⁷⁸.

“Сўз ясалиш типи” тушунчаси – сўз ясалиш сатҳига оид энг асосий тушунчалардан бири. А.Ҳожиев “сўз ясалиш типи” тушунчасига тўғри ёндашади. Аммо айрим адабиётларда ушбу атаманинг бошқача талқин қилиниши ҳам кузатилади: “...ўзбек тилида сўз ясашнинг воситалари ҳар хил бўлиб, уларнинг типлари шундай кўринишларга эга... Сўз ясашнинг кенг қўлланадиган энг унумли, етакчи типлари аффиксал йўл ва композиция сўз кўшиш йўлидир”³⁷⁹. Кўринадик *тип* атамаси сўз ясаш воситаларига, сўз ясаш усулига нисбатан қўлланыпти. Шу ишда “тип” маъносида *модель* сўзи ҳам қўлланган бўлиб, ушбу қўлланиш мантиқан тўғри.

Бир хил сўз ясалиш маъносига эга бўлган сўзлар гуруҳи сўз ясалишининг бир типи ҳисобланиши юқорида айтилди. Бир хил сўз ясалиш маъносига эга бўлишнинг сабаби эса бундай ясама сўзларнинг бир хил лексик-семантик гуруҳга оид сўзлардан бир сўз ясовчи ёрдамида ҳосил қилинишидир. Шунга кўра, сўз ясалиш типига “Маълум бир лексик-семантик гуруҳга оид сўзлардан муайян бир сўз ясовчи ёрдамида ҳосил қилинган, бир хил сўз ясалиш маъносига эга бўлган сўзлар”³⁸⁰ бир сўз ясалиш типини ҳосил қилади” деган таъриф берилса, ҳодисанинг моҳияти тўғри ва тўла акс этади. Сўз ясалишида бу ҳодисани қайд этиш жуда муҳим.

Сўз ясалиш маъноси, бир томондан, бир типга оид ясама сўзларнинг ҳаммасини уларнинг сўз ясалиши асосларидан фарқлайди. Иккинчи томондан, сўз ясалишнинг шу типини бошқа сўз ясалиш типларидан фарқлайди. Учинчи томондан, бир типга оид ҳар бир ясама сўзнинг маъноси сўз ясалиши маъносига тенг бўлмаслигини кўрсатади.

³⁷⁸Ҳожиев А. Ўзбек тили сўз ясалиши. Олий ўқув юрглариининг филология факультетлари студентлари учун қўлланма. – Тошкент: Ўқитувчи, 1989. – 109 б. – Б. 41.

³⁷⁹ Ўзбек тили грамматикаси. 1- том. Морфология. – Тошкент: Фан, 1975. – Б. 21-24.

³⁸⁰Акобировна З. Сўз ясаш қолиплари, уларда унумлилик ва унумсизлик//Ўзбек тили ва адабиёти, 2001. – 6-сон. – Б. 68-70.

Юқорида от асосдан от ясалиш моделларини тузишда шу қоиданинг амал қилишини мисолларда кузатдик, аммо модел тузишда истисно ҳолатларни аниқлаш учун анализаторни тест режимида ишлатиш, натижалар асосида хулоса чиқариш лозим.

Отдан сифат ясалиши инвариант R_n+d^a модели (от-асос + сифат ясовчи морфема) билан белгиланади. Унинг турларини қуйидагича моделлаштирамиз:

1) асосдан англашилган предметга эгалик маъноси $R_n+d_1^a$ моделига эга бўлади, бунда $d_1^a =$ -ли, -дор, -манд, -кор, -гай, -акам, -вор, -хон; масалан, *кунгай, ҳазилакам, доривор, саводхон*;

2) хосликни билдирувчи сифат маъноси $R_n+d_2^a$ моделига эга бўлади, бунда $d_2^a =$ -йи, -ки, -вий, -й, -шумул, -ик, -лик; масалан, *ҳавойи, филологик, йиллик*;

3) ўзак билдирган тушунчага ўхшашлик маъносини билдирувчи сифат структураси $R_n+d_3^a$ моделига эга бўлади, бунда $d_3^a =$ -намо, -симон;

4) ўзак билдирган тушунчанинг кўплиги маъносини билдирувчи сифат структураси $R_n+d_4^a$ моделига эга бўлади, бунда $d_4^a =$ -чан, -дон, -каш, -овар (*жанговар*);

5) асосдан англашилган предметга эга эмаслик маъноси $R_n+d_5^a$ моделига эга бўлади, бунда $d_5^a =$ -сиз.

Отдан равиш ясалиши инвариант R_n+d^{adv} модели (от-асос + равиш ясовчи морфема) билан белгиланади, унинг турларини қуйидагича моделлаштирамиз:

1) ҳаракат тарзи ва ҳолатни билдирувчи равиш ясалиш структураси $R_n+d_1^{adv}$ моделига эга бўлади, бунда $d_1^{adv} =$ -ча, -часига, -ларча, -она, -сиз, -ан, -чан, -акам (*ёлгондакам*), -лай, -сига;

2) пайт маъносини билдирувчи равиш ясалиш структураси $R_n+d_2^{adv}$ моделига эга бўлади, бунда $d_2^{adv} =$ -лаб, -ин: *ойлаб, қишин-ёзин*.

Отдан феъл ясашиши инвариант R_n+d^v (от-ўзак + феъл ясовчи) билан белгиланади, унинг турларини қуйидагича моделлаштирамиз:

1) аниқ отдан феъл ясашиш структураси $R_n+d_1^v$ моделига эга бўлади, бунда $d_1^v =$ -лаш, -а, -қар, -и, -ик, -ки, -ир, -лаштир, -са;

2) мавҳум отдан феъл ясашиш структураси $R_n+d_2^v$ моделига эга бўлади, бунда $d_2^v =$ -ла, -ай, -лан, -ан, -сит.

2. Сифат асосдан ясашиш инвариант R_a+d модели билан белгиланади. Таркибий қисми бир хил, ушбу инвариантдан маъноси билан фарқ қилувчи структураларни кўриб чиқамиз. R_a+d модели R_a+d^n , R_a+d^{adv} , R_n+d^v , яъни сифат-ўзак + сифат ясовчи, сифат-ўзак + равиш ясовчи ва сифат-ўзак + феъл ясовчи моделларини ўзида бирлаштиради.

Сифатдан от ясашишида R_a+d^n модели қуйидаги позицияларга эга бўлади:

1) шахс оти ясашиш структураси $R_a+d_1^n$ моделига эга бўлади, бунда $d_1^n =$ -чи, -хўр: *аълочи, порахўр*;

2) предмет маъносини билдирувчи оти ясашиш структураси $R_a+d_2^n$ моделига эга бўлади, бунда $d_2^n =$ -ат (*кўкат*);

3) фаолият-жараён маъносини билдирувчи от ясашиш структураси $R_a+d_3^n$ моделига эга бўлади, бунда $d_3^n =$ -лик³⁸¹, -чилик.

Сифатдан равиш ясашишида R_a+d^{adv} модели қуйидаги позицияларга эга бўлади:

1) белги-хусусият билдирувчи равиш структураси $R_a+d_1^{adv}$ моделига эга бўлади, бунда $d_1^{adv} =$ -ларча, -она, -часига;

2) ҳолат билдирувчи равиш структураси $R_a+d_2^{adv}$ моделига эга бўлади, бунда $d_2^{adv} =$ -лаб, -лайн.

³⁸¹-лик морфемасининг кўпмаънолилиги ва омонимияси ҳақида қаранг: Гулямов А. Об аффиксе –лик в узбекском языке. Научные труды ТашГУ. Вып. 260. Языкознание. – Ташкент, 1964. – С. 50-57.; Нуъмонова Ш. Ўзбек тилида -лик Аффикси билан ясалган отларнинг семантик классификацияси ҳақида. – Тилшунослик ва адабиётшуносликка оид тадқиқотлар. Тошкент, 1965. – Б. 44-49.; Шарипова К.А. Словообразующие аффиксы -лик, -ли, -чилик в узбекском языке: автореф. канд. филол. наук. – Ташкент, 1972. – С.9.; Ҳожиёв А. Ўзбек тилида -лик аффиксининг вазифаси ҳақида// Ўзбек тили ва адабиёти, 1985. – №1. – Б. 24-27.

Сифатдан феъл ясашида R_a+d^v модели куйидаги позицияларга эга бўлади:

1) ранг билдирувчи асосдан феъл ясашиш структураси $R_a+d_1^v$ моделига эга бўлади, бунда $d_1^v = -й, -ар$: *қорай, кўкар*;

2) предмет белгисини билдирувчи асосдан феъл ясашиш структураси $R_a+d_2^v$ моделига эга бўлади, бунда $d_2^v = -а, -ла$: *бўшамоқ, текисламоқ*;

3) шахсга тегишли белги-хусусият билдирувчи асосдан феъл ясашиш структураси $R_a+d_3^v$ моделига эга бўлади, бунда $d_3^v = -си$: *гарангсимоқ*;

4) умуман белги-хусусият билдирувчи асосдан феъл ясашиш структураси $R_a+d_4^v$ моделига эга бўлади, бунда $d_4^v = -лаш, -син, -ик$: *қийинлашмоқ, оғриқсинмоқ, зўриқмоқ*.

3. Сон-асосдан ясашиш инвариант $R_{num}+d$ модели билан белгиланади. Таркибий қисми бир хил, ушбу инвариантдан маъноси билан фарқ қилувчи структураларни кўриб чиқамиз. $R_{num}+d$ модели $R_{num}+d^n$, $R_{num}+d^a$, $R_{num}+d^{adv}$, $R_{num}+d^v$, яъни сон-ўзак + от ясовчи, сон-ўзак + сифат ясовчи, сон-ўзак + равиш ясовчи ва сон-ўзак + феъл ясовчи моделларини ўзида бирлаштиради.

1) сон-асосдан шахс ва мавҳум от ясашиш структураси $R_{num}+d^n$ моделига эга бўлади, бунда $d^n = -чи, -лик$: *иккичи, бирлик*;

2) сон-асосдан сифат ясашиш структураси $R_{num}+d^a$ моделига эга бўлади, бунда $d^a = -дай$ (*бирдай*); $-лик$ (*бирлик сон*);

3) сон-асосдан равиш ясашиш структураси $R_{num}+d^{adv}$ моделига эга бўлади, бунда $d^{adv} = -лаб, -ларча, -ин$: *битталаб, мингларча, бирин*;

4) сон-асосдан феъл ясашиш структураси $R_{num}+d^v$ моделига эга бўлади, бунда $d^v = -лн, -лаш, -ик$: *иккиланмоқ, бирлашмоқ, бирикмоқ*.

4. Олмош-асосдан ясашиш инвариант $R_{pr}+d$ модели билан белгиланади. Ушбу модел $R_{pr}+d^{adv}$, $R_{pr}+d^v$, яъни олмош-ўзак + равиш ясовчи; олмош-ўзак + феъл ясовчи каби камунум моделларга эга.

1) олмош асосдан равиш ясаиш структураси $R_{pr}+d^{adv}$ моделига эга бўлади, бунда $d^{adv} =$ -сиз, -дек, -дай, -ча, -чаки, -чалик: *усиз, бизсиз, бундай, биздай*³⁸², *бунча, шунчаки, бизчалик*;

2) олмошдан феъл ясаиш структураси $R_{pr}+d^v$ моделига эга бўлади, бунда $d^v =$ -ла, -сира: *сенла, сенсира*.

5. Равиш-асосдан ясаиш инвариант $R_{adv}+d$; модели билан белгиланади. Ушбу модел $R_{adv}+d^n$, $R_{adv}+d^a$, $R_{adv}+d^v$, яъни равиш + от ясовчи, равиш + сифат ясовчи, равиш + феъл ясовчи каби типларга бўлинади:

1) миқдор, белги, хусусият билдирувчи равиш асосдан ҳаракат белгиси ва миқдори билдирувчи от ясаиш структураси $R_{adv}+d^n$ моделига эга бўлади, бунда $d^n =$ -лик, -чилик: *тезлик, озчилик, кўпчилик*;

2) равиш асосдан ўрин-пайт белгисини билдирувчи сифат ясаиш структураси $R_{adv}+d^a$ моделига эга бўлади, бунда $d^a =$ -ги, -ки. d^a нинг воқеланиши: *пастки, дастлабки, илгариги*;

3) равиш асосдан феъл ясаиш структураси $R_{adv}+d^v$ моделига эга бўлади ҳамда қуйидаги маъноларни ифодалайди:

а) миқдор билдирувчи равиш асосдан англашилган маънонинг бор/йўқлигини феъл ясаиш структураси $R_{adv}+d_1^v$ моделига эга бўлади, бунда $d_1^v =$ -ай; d_1^v нинг воқеланиши: *кўпаймоқ, озаймоқ*;

б) ҳолат билдирувчи равишдан феъл ясаиш структураси $R_{adv}+d_2^v$ моделига эга бўлади, бунда $d_2^v =$ -ит, -ла, -лаш; d_2^v нинг воқеланиши: *беркитмоқ, пастламоқ, тезлашмоқ*;

в) шакл-ҳажм билдирувчи равишдан феъл ясаиш структураси $R_{adv}+d_3^v$ моделига эга бўлади, бунда $d_3^v =$ -ар; d_3^v нинг воқеланиши: *қисқармоқ*;

г) пайт билдирувчи равишдан феъл ясаиш структураси $R_{adv}+d_4^v$ моделига эга бўлади, бунда $d_4^v =$ -ик; d_4^v нинг воқеланиши: *кечикмоқ*.

³⁸²Айрим тадқиқотларда -дек/дай форманти ЛШЯ, айрим тадқиқотларда сўз ясовчи сифатида қаралади. Бу ҳақда қаранг: Раҳматуллаев Ш. -дек форманти ҳақида. – Ўзбек тили ва адабиёти, 1977. – № 1. – Б. 48-52.

6.Тақлид-асосдан ясалиш инвариант $R_{sim}+d$; модели билан белгиланади. Ушбу модел $R_{sim}+d^n$, $R_{sim}+d^a$, $R_{sim}+d^{adv}$, $R_{sim}+d^v$, яъни тақлид + от ясовчи, тақлид + сифат ясовчи, тақлид + равиш ясовчи, тақлид + феъл ясовчи каби типларга бўлинади.

Тақлид-асосдан от ясалиш структураси $R_{sim}+d^n$ моделига эга бўлади; куйидаги кўринишлари бор:

1) тақлид-асосдан аниқ от ясалиш структураси $R_{sim}+d_1^n$ моделига эга бўлади, бунда $d_1^n = -a, -ак, -илдоқ$; d_1^n нинг воқеланиши: *жизза, гижжак, бизилдоқ, шипилдоқ*;

2) тақлид-асосдан ҳаракат-ҳолат жараёнини атовчи от ясалиш структураси $R_{sim}+d_2^n$ моделига эга бўлади, бунда $d_2^n = -ак, -ос, -ға$; d_2^n нинг воқеланиши: *қарсак (пахта), гиргирак (юла), шаррос (узум тури), гуввос (шаррос қуядиган ёмғир), говга (тўполон)*;

3) тақлид-асосдан ўрин-жой оти ясалиш структураси $R_{sim}+d_3^n$ моделига эга бўлади, бунда $d_3^n = -а$; d_3^n нинг воқеланиши: *шаршара*.

Тақлид-асосдан сифат ясалиш структураси $R_{sim}+d^a$ камунум моделига эга бўлади; бунда $d^a = -аки, -илдоқ, -ма$; d^a нинг воқеланиши: *жиртаки, чийилдоқ, лақма*.

Тақлид-асосдан равиш ясалиш структураси $R_{sim}+d^{adv}$ камунум моделига эга бўлади; бунда $d^{adv} = -а$; d^{adv} нинг воқеланиши: *шартта, таққа*.

Тақлид-асосдан феъл ясалиш структураси $R_{sim}+d^v$ моделига эга бўлади; бунда $d^v = -а, -илла, -ир, -ла, -ши$; d^{adv} нинг воқеланиши: *чулдирамоқ, шовилламоқ, тупирмоқ, войвойламоқ, гингшимоқ*.

Ҳар бир сўз туркумида сўз ясалишидаги сўз ясалиш типлари мажмуи шу сўз туркуми сўз ясалиш тизимини ташкил этади ва “от ясалиши тизими”, “феъл ясалиши тизими” каби атамалар билан юритилади.

Муайян тилнинг ўзаро алоқа(муносабат)даги сўз ясалиш типлари мажмуи сўз ясалиши тизими ҳисобланади. А.Ҳожиев таъкидлаганидек, сўз ясалиши билан боғлиқ ҳодисалар ва уларнинг моҳияти ҳар бир ясама сўзнинг

ёки ҳар бир сўз ясалишининг ўзи мисолида эмас, сўз ясалиш типлари мисолида ёритилиши керак. Шу асосда сўз ясалиш тизими таҳлили юзага келади³⁸³.

Биз сўз ясовчилар қуршовини моделлаштиришда сал бошқача йўлни танладик. Юқорида от-асосдан ясалиш, сифат-асосдан ясалиш каби олтига мустақил сўз туркумидан сўз ясалиш ҳолатларини моделлаштирдик. Бунга асосларимиз бор, албатта. Бизнинг мақсадимиз маълум бир туркумдаги сўз ясалиш тизимини таҳлил қилиш эмас. (Агар шундай бўлганда А.Ҳожиев ёндашуви, яъни сўз ясалиш тизимига асосланиб моделлар тузиларди.) Биз морфеманинг позицияси унинг маъносини белгилаши нуқтаи назаридан морфоанализатор шаклини таниши, асос туркумини аниқлаши, асосдан кейин кўшилган морфемалар тартиби ва бу тартибнинг маънога таъсирини таҳлил қилишига мўлжаллаб сўз ясалиш моделларини таклиф қилдик.

7. Феъл-асосдан ясалиш инвариант R_v+d ; модели билан белгиланади. Ушбу модел R_v+d^n , R_v+d^a , яъни феъл+от ясовчи, феъл + сифат ясовчи каби типларга бўлинади.

Зеро, А.Ҳожиев тўғри таъкидлаганидек, бир сўз ясовчи ёрдамида бир ёки бошқа-бошқа сўз ясалиш маъносига оид сўзлар ясалиши мумкин. Бу ҳолат сўз ясовчининг бир ёки бирдан ортиқ лексик-семантик гуруҳга оид сўзлардан сўз ясай олиши билан боғлиқ. Бинобарин, бунда сўз ясалиш асоси ва унинг хусусияти муҳим роль ўйнайди³⁸⁴. Бундай ҳолатлар моделларда ўз ифодасини топади.

R_v+d^n модели қуйидагича воқеланади:

1) феъл-асосдан шахс оти ясалиш структураси $R_v+d_1^n$ моделига эга бўлади, бунда $d_1^n = -ар, -оқ, -увчи$; d_1^n нинг воқеланиши: *қочоқ, чонар, учувчи*;

2) феъл-асосдан нарса оти ясалиш структураси $R_v+d_2^n$ моделига эга бўлади, бунда $d_2^n = -ги, -им, -ма, -ин, -инди, -қич, -гич, -моқ, -имта, -кон, -моч,$

³⁸³Ҳожиев А. Ўзбек тили сўз ясалиши. Олий ўқув юр்தларининг филология факультетлари студентлари учун қўлланма. – Тошкент: Ўқитувчи, 1989. – 109 б. – Б. 42.

³⁸⁴Ҳожиев А. Ўзбек тили сўз ясалиши. Олий ўқув юр்தларининг филология факультетлари студентлари учун қўлланма. – Тошкент: Ўқитувчи, 1989. – 109 б. – Б. 39.

-чоқ, -ат, -к, -ок, -чи, чик, -чак, -из; d_2^n нинг воқеланиши: *супурги, кийим, сузма, тугун, қиринди, ювинди, қисқич, сузгич, илмоқ, қуймоқ, ўсимта, босқон, қирмоч, овунчоқ, кекирдак, қовурдоқ, ўроқ, суюнчи, ёпинчиқ, кемирчак, бўғиз, буламиқ*;

3) феъл-асосдан фаолият-жараён оти ясалиш структураси $R_v+d_3^n$ моделига эга бўлади, бунда $d_3^n =$ -ик, -ки, -ги, -им, -кин, -гин, -ин, -мачоқ, -миш, -в, -он, -килик; d_3^n *оғриқ, кулги, тепки, чидам, билим, тўлқин, ёнгин, бекинмачоқ, кечмиш, қилмиш, тергов, қўзғолон, тўзон, ичкилик* каби сўзларда воқеланади;

4) феъл-асосдан фаолият-жараён натижасини атовчи от ясалиш структураси $R_v+d_4^n$ моделига эга бўлади, бунда $d_4^n =$ -ик, -инди, -ч, -диқ, -он, -итки, -гин, -ғоқ, -гу, -илон, -имча, -ит, -инг; d_4^n *кўпик, ўлик, ювинди, севинч, ишонч, қолдиқ, қўзғолон, ачитқи, ёнгин, тўлғоқ, урғу, қўшимча, тушунча, чиқит, йиринг* каби сўзларда воқеланади;

5) феъл-асосдан ўрин-жой оти ясалиш структураси $R_v+d_5^n$ моделига эга бўлади, бунда $d_5^n =$ -ув, -оқ, -қоқ; d_5^n *кечув, ётоқ, ботқоқ* каби сўзларда воқеланади.

Бирдан ортиқ семантик гуруҳ сўзларидан бир сўз ясовчи ёрдамида ясалган сўзлар сўз ясалишининг бирдан ортиқ типига мансуб бўлади, яъни бунда айнан бир сўз ясовчи бирдан ортиқ сўз ясалиш типига оид янги сўзларни юзага келтиради. Масалан, феъл ясовчи *-ла* аффикси ёрдамида ҳосил қилинган *зираламоқ, тузламоқ, ўғитламоқ* феъллари сўз (феъл) ясалишининг бир типи; *бурдаламоқ, бўлакламоқ, тилимламоқ* феъллари бошқа типи ҳисобланади³⁸⁵.

А.Ҳожиёв ушбу ҳодисани, яъни бир сўз ясовчи ёрдамида бирдан ортиқ сўз ясалиши типига мансуб сўзларнинг ҳосил қилинишини сўз ясовчининг кўп маънолилиги (бирдан ортиқ маънога эгаллиги) билан боғлайди. Олимнинг фикрича, сўз ясовчи морфемаларни бир ва кўп маъноли каби гуруҳларга

³⁸⁵Ҳожиёв А. Ўзбек тили сўз ясалиши. Олий ўқув юр்தларининг филология факультетлари студентлари учун кўлланма. – Тошкент: Ўқитувчи, 1989. – 109 б. – Б. 44.

ажратиш, таърифлашнинг сабаби ҳам шунда. Бироқ бу ҳам биртомонлама ёндашув. Чунки *-ла* сўз ясовчиси юқорида кўрилганидан яна бошқа типдаги феъллар ясайди. Лекин бу нарса мазкур аффикснинг кўп маънога эгалигини кўрсатмайди. Чунки келтирилган икки типдаги ясама феълларнинг ўзиданок кўриш мумкинки, бу ясама сўзларнинг икки хил сўз ясалиши типига мансублиги (икки хил сўз ясалиши маъносига эгалиги) сўз ясалиш асосларининг икки хил семантик гуруҳга мансублигидан(икки хил семантик гуруҳга оид сўзлардан ясалганлигидан)дир.

Шунинг учун юқоридаги моделларда асос билдирган маънодан келиб чиқиб гуруҳлаш усулидан фойдаланилди. Моделларда феъл-асосдан бешта маъновий гуруҳ асосида беш хил модел таклиф қилинди. моделларда айрим сўз ясовчи аффиксларнинг қайталанишига сабаб ҳам, А.Ҳожиёв айтганидек, бир сўз ясовчи аффикснинг бир неча тип сўз ясаганидандир. Масалан, *-ги, -ин* сўз ясовчи морфемалари иккинчи ва учинчи, *-он* эса тўртинчи ва бешинчи моделларда учрайди; улар бошқа тип сўз ясалишини ҳосил қилади. Тўртинчи ва бешинчи моделлардаги кўпмаъноли қўшимчалар *-зоқ, -қоқ* икки типдаги сўз ясалишини ҳосил қилади, яъни ўрин-жой ва фаолият-жараён оти; *-гин* сўз ясовчи морфема эса фаолият-жараёни оти ва фаолият-жараён натижасини; *-оқ* қўшимчаси ўрин-жой ҳамда шахс оти ясайди. Юқоридаги моделлар эса морфоанализ жараёнида унинг маъносини ойдинлаштиради.

R_v+d^a модели, яъни феълдан сифат ясалиши қуйидагича воқеланади:

1) феъл-асосга тўғридан-тўғри қўшилиб, сифат ясалиш структураси $R_v+d_1^a$ моделига эга бўлади, бунда $d_1^a =$ *-қоқ, -гир, -ағон, -қи, -к, -гин, -кин, -ма, -арли, -оқ, -ри, -чак, -чан, -ма, -ч, -мсиқ, -м, -им, -қа, -а, -аси, -гич, -дан, -ин, -акса, -чиқ, -ғоқ*; d_1^a шакли *қочқоқ, урушқоқ, сезгир, ўткир, билагон, сайроқи, вайсақи, чирик, эгик, ёпиқ, озгин, сотқин, кўчма, ачинарли, қолоқ, эгри, эринчоқ, курашчан, ўтмас, жирканч, ачимсиқ, қарам, айрим, қисқа, тўла, келаси, билгич, югурдак, тўлин, ўлакса, сирпанчиқ, тойгоқ* каби сўзларда воқеланади;

2) феъл-асосга грамматик шаклдан кейин қўшилган сифат ясаиш структураси $R_v+d_2^a$ моделига эга бўлади, бунда $d_2^a=$ -чоқ, -чан, -ғич; d_2^a мақтанчоқ, яшовчан, қопонғич каби сўзларда юзага чиқади.

Биринчи ва иккинчи моделлардан кўриняптики, -чоқ, -чан, -ғич (-ғич) икки типда ҳам учрайди, демак, бу кўшимчаларни моделлаштириш морфоанализда тўғри натижани келтириб чиқаради.

Хулоса шуки, $R+d$ модели, яъни асосга битта сўз ясовчи қўшилиши билан сўз ясаишида маълум қоидаларга амал қилинади. Жумладан:

1) одатда, сўз ясовчилар ўзакка ЛШЯ морфемадан олдин қўшилади;

2) феъл-асосдан сўз ясалганда сўз ясовчи морфемадан олдин маълум грамматик шакл келиш ҳолатлари учрайдики, бу ҳолат бошқа туркумдан сўз ясаишига хос эмас: $(ачи+m)+қи$, $(иси+m)+ма$.

3) $R+d$ моделининг воқеланишидан ҳосил бўлган сўзларда лексемалашуш ҳолати ҳам кузатилади; ҳатто лексемалашуш натижасида сўз тамоман бошқа маъно ифодалайди. Бундай сўзлар таркибидаги сўз ясовчидан олдин жойлашувчи ЛШЯнинг вазифаси пассивлашади, сўз ясовчи морфема бўлиб кўринади: *ўтмаслик*, *сезувчан*, *ўқувчи*, *терговчи*. Бундай модел ҳосиласи синхрон аспектда таҳлил қилинади.

А.Ҳожиёв сўз ясовчи бирликнинг умумий маъноси ҳақида шундай дейди: “Ҳар қандай сўз ясовчи бирдан ортиқ сўз ясади. Акс ҳолда, яъни биргина сўз таркибида мавжуд бўлган ва ташқи кўринишдан ясовчи кўшимчадек кўринган нарса сўз ясовчи бирлик бўла олмайди. Чунки яққа ходисалар тизимдан ўрин олмайди. Шунга кўра, муайян бир сўзнинг ўзига хос маъноси сўз ясаиши тизимига хос маъно бўла олмайди. Ана шу фактнинг ўзиёқ кўрсатадики, айна бир сўз ясовчи ёрдамида ҳосил қилинган сўзлар маъносида ҳар бир ясама сўзга хос муайянлик билан бирга, улар учун умумий бўлган қисм ҳам бўлади. Масалан, ҳар қандай иш қуроли маълум бир ишни бажариш учун хизмат қилади. Ҳар бир қуролни билдирувчи сўзнинг семемаси (маъноси) таркибида *қурол* ва қурол бажарадиган *ишни* қайд этувчи сема

мавжуд бўлади. *Қурол (асбоб)* семаси шу типдаги сўзлар учун умумий бўлади. Қурол ёрдамида бажариладиган иш семаси эса муайян қуролни билдирувчи сўзнинг семемаси учун хос, яъни фарқловчи сема бўлади. *Бажармоқ (қилмоқ)* маъноли *-ла* сўз ясовчиси қурол (асбоб) билдирувчи сўзлардан, умуман, шу сўзлар билдирадиган қурол (асбоб)га хос ишни *бажармоқ (қилмоқ)* маъноли феъллар ҳосил қилади. Ҳар бир ясама феъл эса шу феъл ясалиши учун асос бўлган сўз (сўз ясалиш асоси) ва унга хос ишни бажариш (қилиш, амалга ошириш) маъносини билдиради. Масалан, *арра, ранда, дазмол* сўзларининг ҳаммасининг маъносида бир умумий сема, яъни *қурол* семаси бор (уларнинг ҳаммаси, умуман олганда, қуролни билдиради).

-ла аффикси буларнинг ҳаммасидан бир умумий маъноли феъллар ҳосил қилади. Бу умумий маъно “қурол ва унга оид ишни бажариш”га тенг бўлади. Бу маъно асбоб (қурол) маъноли сўзлардан *-ла* аффикси ёрдамида ясалган сўзларнинг ҳаммаси учун умумий (уларни бир гуруҳга бирлаштирувчи) маъно ҳисобланади.

Демак, ҳар бир ясама сўзнинг маъносида фарқловчи ва умумий сема бўлади. Муайян сўз ясовчи ёрдамида муайян умумий сема асосида ҳосил қилинган сўзлар маъноси *сўз ясалиш маъноси* дейилади. Ясама сўзлар таркибидаги умумий қисм, умумий сема бошқа-бошқа бўлса, уларнинг сўз ясалиш маъноси ҳам шундай бўлади. Мас., *-ла* сўз ясовчиси билан ҳосил қилинган *мойламоқ, тузламоқ, лакламоқ* сўзлари бир хил сўз ясалиш маъносига эга ва шу маъноси билан юқорида кўрилган (*арраламоқ* типдаги) феъллардан фарқланади³⁸⁶.

Морфоанализатор таҳлили учун сўз ясалиш моделларини ишлаб чиқишда ушбу қонуниятга тўла амал қилинади. Ваҳоланки, моделда таклиф қилинаётган сўз ясовчи қисм умумий маънога эга бўлади. Ўзакнинг умумий маъносини ифодалаш учун турли индекслар билан белгиланади. Аммо ушбу

³⁸⁶Ҳожиёв А. Ўзбек тили сўз ясалиши. Олий ўқув юр்தларининг филология факультетлари студентлари учун қўлланма. – Тошкент: Ўқитувчи, 1989. – 109 б. – Б. 35-36.

моделларнинг юқорида олим айтганидек, хусусий маъноларнинг барча сўзларда қай даражада амал қилиши ҳақида фақат анализатор ишга туширилиб, синов тариқасида сўзлар массиви таҳлил қилингандагина хулоса қилиш мумкин.

А.Ҳожиёвнинг фикрича, сўз ясалиш типи тушунчаси сўз ясалиши сатҳининг асосий тушунчаларидан бири, унинг моҳиятини тўғри англаш сўз ясалиши билан боғлиқ кўпгина асосий ҳодисаларнинг моҳиятини белгилашга ёрдам беради. Сўз ясалиши типи ҳақида айтилганларнинг ўзиёқ сўз ясалиш тизимини сўз ясовчиларни таҳлил этиш йўли билан эмас, сўз ясалиш типларини таҳлил этиш йўли билан ёритиш тўғри йўл эканлигини кўрсатади (Биз моделлаштиришда шу йўлдан бордик). Шундай йўл билан иш кўрилганда, биринчидан, сўз ясалишидек кўринган, лекин сўз ясалиши ҳисобланмайдиган ҳодисалар аниқ белгиланади ва сўз ясалиш объектига киритилмайди. Иккинчидан, шу вақтгача сўз ясовчига боғлаб ўрганилган кўпгина ҳодиса (масала)лар энди сўз ясалиш типлари мисолида ўрганилади (Масалан, маҳсулли-маҳсулсизлик ҳодисаси)³⁸⁷.

Бундан келиб чиқадики, морфоанализаторнинг сўз ясалиш таҳлилида фақат морфологик шаклларнинг аниқланиши, яъни ясовчи қисмни аниқлаш билан сўз ясалиш ҳодисаларини тўғри таҳлил қилишга эришиб бўлмайди. Сифатли таҳлилга эриши учун сўз ясалиш типлари ҳам таҳлил қилиниши лозим. Сўз ясовчи қўшимчалар қуршовининг автоматик таҳлили учун 1та инвариант модел асосида 7 та субмодел ҳамда 60та субмодел кўринишлари ишлаб чиқилди.

4.2.2. $R+d_1+d_2$ ва бошқа мураккаб субмоделдаги сўз ясовчи морфемалар қуршови. Юқорида таъкидланганидек, ўзбек тилида бир асосга бирдан ортиқ сўз ясовчи морфема қўшилиши мумкин: бунда сўз ясовчи морфемаларнинг ўзаро тартиби ва занжири мавжуд.

³⁸⁷Кўрсатилган манба. – Б.47.

Асосий инвариант $R+d_1+d_2$ куйидаги субмоделлар орқали воқеланади:
 $R_n+d_1+d_2$; $R_a+d_1+d_2$; $R_{num}+d_1+d_2$; $R_{pr}+d_1+d_2$; $R_{adv}+d_1+d_2$; $R_{sim}+d_1+d_2$; $R_v+d_1+d_2$.
Бунда феъл, от, сифат, сон, олмош, равиш, тақлид асосга кетма-кет икки сўз ясовчи морфема кўшиш билан сўз ясалади.

$R_n+d_1+d_2$ субмоделининг $R_n+d_1^n+d_2$; $R_n+d_1^a+d_2$; $R_n+d_1^n+d_2$; $R_n+d_1^{adv}+d_2$; $R_n+d_1^v+d_2$ каби кўринишлари мавжуд.

$R_n+d_1^n+d_2$ тип модели d_2 нинг хусусиятларидан келиб чиққан ҳолда турлича воқеланади. Бунда d_2 от, сифат, равиш, феъл ясовчи бўлиши мумкин.

1) $R_n+d_1^n+d_2^n$ типда от-асосга иккита сўз ясовчи кўшилиб, биринчи ясовчи шахс оти, иккинчиси фаолият, касб-кор отини ясайди: *беданабозлик, бинокорлик, аравакашлик, китобхонлик, хунармандлик, заргарлик*.

2) $R_n+d_1^n+d_2^n$ да от-асосга ясовчи шахс оти ва белги, хусусият отини ясовчи кўшимча кўшилади: *димогдорлик, безорилик, ватанпарварлик, насиҳатгўйлик*.

Умуман олганда бундай моделдаги сўзларни анализатор тўғри таҳлил қилади: ўнг чапга қараб кўшимчаларни “танийди”, яъни аввал энг охириги кўшимчани, кейин ундан чап томондаги кўшимчани аниқлайди. Шундай бўлса-да, $R_n+d_1^n+d_2$; $R_n+d_1^a+d_2$; $R_n+d_1^n+d_2$; $R_n+d_1^{adv}+d_2$; $R_n+d_1^v+d_2$ моделлар маълумот сифатида маълумотлар базасида ахборот сифатида мавжуд бўлиши лозим. Аммо от, сифат, баъзан равиш-асосга кўшилиб сўз ясайдиган -лик камунум морфемаси феъл-асосга кўшилувчи -лик луғавий шакл ясовчи морфема билан омоморф ҳисобланади: *боришлик, ўқимаслик*. Ушбу омоморфлик ҳолатини бартараф этиш учун маълумотлар базасига филтр-модел киритиш талаб қилинади. Ишнинг олдинги қисмида феъл-асосга хос ЛШЯ морфемалар сирасида *-(u)ш* ҳаракат номи кўшимчасининг бўлишсизлик шакли *-маслик* кўринишида бўлиши ҳақида маълумот берилган, ҳолат моделлаштирилганди. Шу асосда аввалги қоидага кўшимча киритамиз:

1) агар {феъл-асос+мас+лик} бўлса, {-маслик} = ҳаракат номининг бўлишсизлик шакли: { $R_{adv} + f_1 + \text{лик}$ } да, -лик = f_2 , яъни граммема;

2) агар {от/ сифат/ сон +лик} бўлса, {-лик} = сўз ясовчи: { R_n / R_a / R_{num} / R_{adv} + лик} да -лик = d^n , яъни дериватема;

3) $R_n+d_1^n+d_2^n$ да от-асосга ясовчи шахс оти ва ўрин-жой отини ясовчи кўшилади: *элчихона*;

4) $R_n+d_1^n+d_2^n$ да от-асосга иккита ўрин-жой оти ясовчиси кўшилади: *ўтлоқзор* (бу кўлланиш адабий тилга хос бўлмаса-да, оғзаки нутқда учраб туради);

5) $R_n+d_1^n+d_2^n$ да от-асосга ясовчи ўрин-жой ва шахс оти ясовчилари кўшилади: *чойхоначи*.

Ўзакка учта сўз ясовчи кўшимча кўшилишини ифодаловчи асосий инвариант $R+d_1+d_2+d_3$ модели ва бир қанча субмоделлар орқали воқеланади.

Ўзбек тилида сўз таркибида тўртта сўз ясовчи морфема учрайдиган ҳолатлар ҳам мавжуд. Бундай сўзлар кам учраса-да, уларнинг таҳлилини оптималлаштириш учун $R+d_1+d_2+d_3+d_4$ моделини морфоанализатор базасига киритиш ўринли бўлади. Шунингдек, ушбу инвариант модел куйидаги субмоделлар орқали реаллашади: $R_n+d_1+d_2+d_3+d_4$, $R_{sim}+d_1+d_2+d_3+d_4$, $R_v+d_1+d_2+d_3+d_4$. Кўриняптики, бундай тузилишга эга ясалмалар одатда, от, тақлид ва феъл асосдан ясалади. Масалан: *чойхоначичалик*, *газламачилик*, *ўтоқчилик*, *тилаксизлик*, *йўртмадўзчилик*, *супургифурушчалик*, *унумдорсизлик*, *билимдонамолик*. Мураккаб сўз ясалиши автоматик таҳлили учун, асосан, инвариант ва субмоделлар ишлаб чиқилди. Жумладан, ўзакка иккита сўз ясовчи морфема кўшилишини фарқлаш учун 1 та инвариант асосида 7 та субмодел, унинг 5 та кўриниши; ўзакка учта, тўртта сўз ясовчи морфема кўшилишини фарқлаш учун 1 та инвариант модел аниқланди.

4.3. Сўз ясовчи кўшимчалар қуршовини моделлаштиришнинг айрим муаммолари хусусида

А.Ҳожиёв сўз ясалиш моделлари ҳақида шундай ёзади: “Сўз ясалиш асоси” ва “сўз ясовчи”нинг таҳлилидан маълум бўлдики, ҳар қандай ясама сўз

сўз ясалиш асосига сўз ясовчи қўшиш йўли билан ҳосил қилинади. Демак, ўзбек тилида сўз ясалишининг умумий модели “сўз ясалиш асоси + сўз ясовчи”дан иборат бўлади. Сўз ясалиш асоси ва сўз ясовчи эса материал жихатдан турлича бўлиши мумкин. Лекин ҳар бир ясама сўзнинг, масалан, қандай сўздан, сўз ясовчининг қандай аффикс ёки ёрдамчи сўздан бўлишига қараб, уларнинг ҳар бирини алоҳида модель деб қараш керак эмас, бунга эҳтиёж ҳам йўқ”³⁸⁸.

Олимнинг ушбу фикри анъанавий сўз ясалиш таҳлили учун ўринли, аммо автоматик морфоанализ жараёни учун моделлар жуда муҳим, чунки сўз ясовчининг ҳар бир позицияси алоҳида маъно билдириши, бошқа-бошқа вазифа бажариши мумкин. Бундан ташқари молеллаштириш табиий тилни қайта ишлашнинг асосий қуроли саналади. Шу сабабли анализаторлар учун асос ва сўз ясовчининг турли позициялари, албатта, моделлаштирилиши лозим.

Ҳ. Неъматов, Р.Расулов эса қолиплар ҳақида қуйидагиларни қайд этади: “... қолип тил бирликларининг ўзаро бирикиш тизими ва бу бирикишнинг – нутқий ҳосиланинг моҳияти сифатида тушунилади. Бу қисмлар тенглама каби (=) белгиси билан ажратилади. Масалан, {[от] +[чи]} = (от англатган нарса-предмет билан шуғулланувчи, алоқадор шахс атамаси)”³⁸⁹. Англашиладики, сўз ясалиш қолипларини тузиш сўз ясалишини ўрганиш учун муҳим бўлганидек, морфоанализаторнинг сўз ясалиши таҳлилида ҳаракатлантирувчи қисм бўлиб хизмат қилади.

Олимлар ўзбек тилида сўз ясалишининг баҳсли масалалари³⁹⁰га ҳам алоҳида эътибор қаратишган. Морфоанализатор учун моделларни шакллантиришда ушбу ишлардан ўринли фойдаланиш мумкин.

³⁸⁸Ҳожиёв А. Ўзбек тили сўз ясалиши. Олий ўқув юрларининг филология факультетлари студентлари учун қўлланма. – Тошкент: Ўқитувчи, 1989. – 109 б. – Б. 26.

³⁸⁹Ҳ.Неъматов, Р.Расулов. Ўзбек тили систем лексикологияси асослари. – Тошкент: Ўқитувчи, 1995. – Б. 38-39.

³⁹⁰А.Ходжиёв, А.Нурманов, Н.Махмудов. Спорные вопросы в системе словообразования тюркских языков. // Советская тюркология, 1989. – №2.

А.Ҳожиёв маҳсулли ва маҳсулсиз, каммаҳсул ва сермаҳсул ясовчи морфемалар ҳақида шундай дейди: “У ёки бу сўз ясовчига нисбатан “ясовчи” терминини қўллаш хато бўлмайди. Чунки улар ичида маҳсулли ва маҳсулсизи ҳам, каммаҳсул ва сермаҳсули ҳам бор. Шунинг учун ҳар бир сўз ясовчи ҳақида гап борганида, унинг ана шундай хусусияти айтилиши мумкин. Лекин бунда ҳам у ёки бу қўшимча ёрдамида ҳозирда янги сўз ясалиши ёки ясалмаслиги, кам ёки кўп миқдорда сўз ясалиши қайд этилади, холос. Унинг сўз ясалиши тизимидаги бошқа хусусиятлари ёритилмай қолаверади. Сўз ясалиши тизимини ёритишда сўз ясалиши типларининг маҳсулли-маҳсулсизлигини (каммаҳсул-сермаҳсуллигини) белгилаш, бунинг сабабларини аниқлаш илмий ва амалий жиҳатдан муҳим аҳамият касб этади. Бундан ташқари, у ёки бу сўз ясовчи ёрдамида кам ёки кўп сўз ясалиши фақат шу ясовчининг хусусияти билангина изоҳланмайди, балки бунда сўз ясалиш асосининг хусусияти ҳам муҳим роль ўйнайди. Бу ҳол ҳам мазкур ҳодисани сўз ясалиши типига нисбатан қўллашни тақозо этади”³⁹¹.

Ҳ. Неъматов, Р. Расуловлар сўз яшаш қолипларини унумли ва унумсиз қолипларга ажратишади³⁹². Унумли қолиплар деганда, ҳозирда амал қилаётган, маҳсул бераётган қолиплар, унумсиз қолип деганда, тарихий қолиплар назарда тутилади. А.Ҳожиёв бу фикрга шундай муносабат билдиради: “Ҳозирги тил нуқтаи назаридан тарихий қолипларни “унумсиз қолип” дейиш хато эмас. Лекин тил ҳодисаларини ўрганишнинг тарихий (диахрон) ва ҳозирги (синхрон) аспектлари мавжудлиги, уларни қориштириш мумкин эмаслиги маълум ва буни мазкур иш муаллифлари ҳам қайд этадилар. Шундай экан, сўз яшаш қолипининг маҳсулли-маҳсулсизлиги ҳам шу икки аспектнинг ҳар бирида алоҳида қаралиши керак. Чунки уларнинг ҳар бирида, жумладан, ҳозирги сўз ясалишида ҳам, унумлилик-унумсизлик ҳодисаси бор ва ҳозирги

³⁹¹Ҳожиёв А. Ўзбек тили сўз ясалиши. Олий ўқув юр்தларининг филология факультетлари студентлари учун қўлланма. – Тошкент: Ўқитувчи, 1989. – 109 б. – Б. 46.

³⁹²Ҳ.Неъматов, Р.Расулов. Ўзбек тили систем лексикологияси асослари. – Тошкент: Ўқитувчи, 1995. – Б. 38-39.

сўз ясалиши таҳлил этилганда, бу ҳодиса алоҳида қайд этилиши, унинг сабаблари, қонуниятлари ва бошқа томонлари ёритилишини тақозо этади. Ҳозирги сўз ясалиши(унинг қолипи)ни унумли, тарихий сўз ясалиши қолипни унумсиз деб фарқлаш, бизнингча, мақбул эмас ва бундай ажратиш унумли-унумсизликни белгиламайди, шу нуқтаи назардан ҳеч нарса бермайди ҳам. Ҳозирги сўз ясалиши нуқтаи назаридан, “тарихий қолип” деб аталаётган қолип (қўполроқ бўлса ҳам), умуман йўқ бўлган қолип, ҳозирда ишламайдиган қолип ва унинг унумли-унумсизлиги ҳақида гапиришнинг ўрни ҳам эмас. Қолаверса, ҳозирги сўз ясалишига оид бирор қолипни унумли, серунум (маҳсулли, сермаҳсул) деб аташ умумий гапдан бошқа нарса эмас, бу гап илмий ва амалий жиҳатдан ҳеч нарса бермайди. Масалан, “сўз ясалиш асоси + -ла кўшимчаси ёрдамида феъл ясалиши модели (қолипи) маҳсулли (сермаҳсул)” дейиш билан шу модел (қолип) асосида кўп миқдорда феъл ясалиши қайд этилади, холос. Ваҳоланки, бу моделда феъл ясалишининг бир неча типлари мавжуд, улардан баъзилари маҳсулли (сермаҳсул ёки кammaҳсул), баъзилари эса маҳсулсиздир.

А.Ҳожиев Ҳ.Неъматовнинг қолипларнинг унумли ёки унумсизлиги ҳақидаги фикрига эътироз билдирар экан, бунга ҳеч қандай исбот келтирмайди. Зеро, бундай эътирозни исботлаш учун юзлаб қолипларни минглаб мисолларда синаб кўриш лозим, шундагина аниқ хулоса бериш мумкин бўлади. Морфоанализатор ишда таклиф этилаётган қолиплар асосида уларнинг унумли/унумсизлигини асосли далиллайди, мисоллар массивини тақдим этади, аниқ хулоса чиқаришга ёрдам беради.

Шунингдек, А.Ҳожиев ясалма таркибидаги ясовчиларнинг жойлашуви ҳақида шундай ёзади: “Сўз ясовчилар ўзбек тилида сўз ясалиш асосига ва сўз охирига қўшилади, яъни суффикс, суффиксоид характерида бўлади: *ишчи, ишла, ишхона* каби. Тожик тилидан ўзлашган айрим сўз ясовчиларгина сўз ясалиш асосига унинг олдидан қўшилади: *нотўғри, серҳаракат, камган* каби.

Лекин бу ҳол ясама сўзнинг маъно ёки бошқа хусусиятига таъсир этмайди, Шунинг учун уларни алоҳида таҳлил этишнинг ҳожати йўқ³⁹³”.

Сўз ясалишини ўрганишда ясовчининг жойлашуви ҳодисани ўрганишда аҳамиятсиз бўлиши мумкин, аммо морфоанализда бу жараён аҳамиятли. Агар сўз ясовчи кўшимча асосдан кейин, яъни ўнг томонда жойлашган бўлса, анализатор ўнгдан чапга таҳлил қилишда сўз ясовчи аффиксни аниқлайди. Масалан, *супурги, кийим, сузма, тугун, қиринди, ювинди, қисқич, сузгич, илмоқ, қуймоқ, ўсимта, босқон, қирмоч, овунчоқ, кекирдак, қовурдоқ, ўроқ, суюнчи, ёпинчиқ* каби ясама сўзларни ўнгдан чапга таҳлил қилар экан, ясалиш асоси ва ясовчи қисмни маълумотлар омборидаги ахборотга асосланиб *супур+ги, кий+им, суз+ма, туг+ун, қир+инди, юв+инди, қис+қич, суз+гич, ил+моқ, қуй+моқ, ўс+имта, бос+қон, қир+моч, овун+чоқ, кекир+дак, қовур+доқ, ўр+оқ, суюн+чи, ёпин+чиқ* тарзида аниқлайди ва натижани кўрсатади. Аммо анализатор маълумотлар омборида {сўз ясовчи + ясалиш асоси} қолипи бўлмаганда ясовчи асосдан олдин жойлашган сўз ясовчилар – олдқўшимчалар таҳлилини амалга оширишда натижа беролмайди. Шу сабабли анализаторда бундай ясалиш типи ҳам алоҳида модел билан қоидалаштирилиши керак бўлади. Анализатор таҳлил жараёнида *нотўғри, серҳаракат, камган* каби сўзларни таҳлил қилар экан, ўнгда чапга морфемаларни аниқлаш жараёнида *тўғри, ҳаракат, ган* каби морфемаларни аниқлайди, ундан олдин турган қисмларни ажратиб қўяди, лекин уларнинг ясовчи қисм эканлиги ҳақида хулоса беролмайди. Агар базада {сўз ясовчи + ясалиш асоси} қолипи ҳақида маълумот бўлса, анализатор ўзакдан олдин жойлашган *но-, сер-, кам-* морфемаларини сўз ясовчи қисм сифатида танийди.

С.Усмоновнинг сўз ясовчи аффиксоидларга бағишланган махсус тадқиқотида масала атрофлича таҳлил қилинганлиги сабали бу ўринда қисқа тўхталамиз³⁹⁴.

³⁹³Ҳожиев А. Ўзбек тили сўз ясалиши. Олий ўқув юр்தларининг филология факультетлари студентлари учун қўлланма. – Тошкент: Ўқитувчи, 1989. – 109 б. – Б. 18.

³⁹⁴Усмонов С. Сўз ясовчи аффиксоидлар//Ўзбек тили ва адабиёти, 1963. – № 3.

Сўз ясовчи вазифасида айрим аффиксоид ва ёрдамчи сўзлар ҳам кўлланиши кузатилди, лекин улар жуда кам ва тожик тилидан ўзлашган: *кам-*, *ҳам-*, *хуш-*, *хона-* сўз ясовчилари. Буларнинг ҳар бири ёрдамида ясалган сўзлар ҳам кўп эмас. Яна шуни ҳам айтиш керакки, булар сўз ясовчи сифатида, сўз ясовчи аффикслардан фарқли бирор белгига эга эмас. Аксинча, сўз ясовчи аффиксларга хос умумий хусусиятларга бу сўз ясовчилар ҳам эга. Ҳатто, *кам-* ва *ҳам-* сўз ясовчилари сўз ясовчи *сер-* ва *-дош* аффиксларига антоним, маънодош: *камҳосил* – *серҳосил*, *ҳамсуҳбат* – *суҳбатдош*. Шунинг учун буларни сўз ясовчи аффикслардан ажратиб (алоҳида олиб) таҳлил этишнинг ҳожати йўқ. Қандай сўз ясашига қараб, ҳар бири ўз ўрнида берилавергани маъқул³⁹⁵. А.Ҳожиёв буларнинг қўшма сўз ясамаслигини батафсил изоҳлаган.

Демак, аффиксоидлар рўйхати ҳам маълумотлар омборидан жой олиши лозим, шунда анализатор таҳлилни моделлар асосида олиб бораверади.

Ўзбек тилида сўздаги морфемалар қуршови ва тартибини моделлаштиришда қуйидаги хулосаларга асосланиш мумкин:

1) сўз таркибида морфемаларнинг маълум тартиби мавжуд бўлиб, уларда одатдаги тартиб қуйидагича: асос+ сўз ясовчи морфема + ЛШЯ морфема + СШЯ морфема;

2) сўз ясовчи морфемалар қуршови ўзбек тилининг агглютинативлик хусусияти билан боғлиқ ҳамда уларнинг кет-кет қўшилиш ҳолатлари кўп. Шунингдек, ички ва ташқи сўз ясашиш ҳам фарқланади;

3) ўзбек тилида таркибида 4 та сўз ясовчи учрайдиган сўзлар кам бўлса ҳам учрайди, шунингдек, лингвистик иқтисод қилиш принципига амал қилмасликдан келиб чиқади;

4) барча сўз ясовчилар асосдан кейин, ЛШЯдан олдин жойлашади;

5) агар сўз таркибида иккита сўз ясовчи мавжуд бўлса, аввал аниқ от ясовчи, кейин мавҳум от ясовчи қўшилади;

³⁹⁵ Ҳожиёв А. Ўзбек тили сўз ясашиши. Олий ўқув юртларининг филология факультетлари студентлари учун кўлланма. – Тошкент: Ўқитувчи, 1989. – 109 б. – Б. 23.

б) ўзлашма сўз ясовчилар ўзбек тилидаги сўз ясовчилардан кейин қўшилади;

7) кучсиз семантик характерли сўз ясовчи сўзда сўнг позицияни эгаллайди;

8) каммасул ясовчи олдин, сермахсул сўз ясовчи кейин жойлашади;

9) сўз таркибида кичрайтириш-эркалаш қўшимчалари бўлганда, аввал кичрайтириш, кейин эркалаш қўшимчаси қўшилади;

10) иккита функционал бир хил ЛШЯлар кетма-кет қўшилади, яъни плеонастик характер касб этади;

11) феъл луғавий шакл ясовчиларининг тартиби қуйидагича бўлади: {нисбат шакли + бўлишсизлик шакли + феълнинг функционал шакллари + замон шакли};

12) синтактик шакл ясовчи қўшимчалар ҳақиқий сўз шаклини ҳосил қилувчилардир, чунки улар сўзнинг синтактик вазифасини шакллантиради;

13) ўзбек тилида қўшимча шаклидаги юкламалар доим сўзнинг охирига қўшилади (уларнинг ҳам кетма-кет қўшилиш ҳолатлари учрайди);

14) сўз таркибида одатдаги тартибнинг бузилиши маълум лингвистик факторлар билан боғлиқ бўлади.

Омоним ва полисемантик сўз ясовчи морфемаларни лингвистик моделлаштириш хусусида. Морфемалар омонимияси рус тилшунослигида П.А.Соболева, ўзбек тилшунослигида М.А.Джурабаева ишларида атрофлича ўрганилган³⁹⁶. Шунингдек, ўзбек тили бўйича дарслик ва қўлланмаларда қўшимчалар омонимияси бўйича таҳлилий маълумотлар етарли³⁹⁷. Ушбу материаллар омоним қўшимчалар қуршовини ўзбек тили морфоанализатори учун моделлаштиришда бой лингвистик таъминот вазифасини ўтайди.

³⁹⁶Соболева П.А. Словообразовательная полисемия и омонимия. – М., 1980.; Джурабаева М.А. Аффиксальная омонимия в узбекском языке: автореф. дисс. ... канд. филол. наук. – Тошкент, 1975.

³⁹⁷Омоним қўшимчалар жадвали ҳақида қаранг: Mengliyev B., Xoliyorov O., Abdurahmonova N. O`zbek tilidan universal qo`llanma. (Qayta ishlangan 3-nashri). – Toshkent: Akademnashr, 2014. – 389 b. – Б.267.

Маълумки, кўшимчалар қуршовида, хусусан, омоним кўшимчалар алоҳида аҳамият касб этади. Уларнинг асос ва қуршовдаги кўшимчаларини моделлаштириш автоматик таҳлилда ниҳоятда аҳамиятли. Тилда кўшимчалар шаклдошлиги бир турдаги кўшимча орасида ҳам, турли кўшимчалар орасида ҳам бўлиши мумкин: сўз ясовчи кўшимчалар шаклдошлиги, грамматик шакллар шаклдошлиги, сўз ясовчи ва грамматик кўшимча шаклдошлиги ходисаси учрайди³⁹⁸.

Шундай шаклдош морфемалар борки, қуршовда улардаги омонимияни автоматик аниқлаш ва фарқлашни модел асосида амалга ошириб бўлмайди. Масалан, от, сифат, равиш ва феъл ясовчи, феълга хос луғавий шакл ясовчи бўлиши мумкин: *жизза* (от ясовчи), *кўтара* (савдо – сифат ясовчи), *қона* (феъл ясовчи), *шигга*, *чиппа* (равиш ясовчи); *бора*, *бура* (феълга хос ЛШЯ); *тароқ* (от ясовчи), *юмшоқ* (сифат ясовчи). Бундай кўшимчаларни на асосга нисбатан позицияси, на ундан кейин келувчи морфемага нисбатан умумий бир модел тушириб бўлмайди. Бундай сўз ясовчиларни морфоанализатор ўнгдан чапга таҳлилда кўмакчи морфема сифатида аниқлайди, аммо қайси туркум ясовчиси бўлганини моделда бериб бўлмайди.

Аммо қатор кўшимчалар борки, уларнинг омонимияси, маълум сўз/сўз шаклдаги функциясини модел орқали аниқлаш мумкин. Масалан: *қорай* ва *ўқий* сўзларидаги *-й* омоним кўшимча, у феъл ясовчи ва феълга хос СШЯ бўла олади. Биламизки, феъл ясовчи кўшимча феъл бўлмаган асосдан феъл ясайди. Шу қоидага асосланиб моделлаштирамиз:

- 1) агар {феъл бўлмаган асос + *-й*} бўлса, *й* = феъл ясовчи;
- 2) агар {феъл-асос + *-й*} бўлса, *й* = луғавий шакл ясовчи.

Демак, модел асосида филтрлаш йўли топилди. *-кин* морфемаси *кескин* сўзида сифат ясовчи, *тўккин* сўзида феълга хос луғавий шакл ясовчи

³⁹⁸ Кубрякова Е.С. Основы морфологического анализа. – М., 1974.

вазифасини бажаради. Икки ҳолатда ҳам феъл асосга қўшилган, асосга нисбатан позицияси орқали моделга тушириб бўлмайди.

Демак, ўзбек тилидаги барча омоним ва полисемантик морфемаларнинг қуршовдаги маъносини аниқлаш моделлари алоҳида тадқиқотни талаб қилади.

Боб бўйича хулоса

1. Агглютинатив тилларда грамматик воситанинг сўздаги тартиби сўз ясовчи қўшимча+луғавий қўшимча+синтактик қўшимча кўринишидаги қонуниятга эга эканлигини биламиз. Грамматик шаклларнинг сўздаги қуршови унинг маъно/вазифасини аниқлайди. Шакли ва сўздаги қуршови ҳар хил бўлган грамматик шаклларнинг мавжудлиги сўз ясовчи, луғавий шакл ясовчи ва синтактик шакл ясовчилар қуршови феъл, от, сифат, сон, олмош, равишда ўзига хос бўлиб, уларнинг сўздаги қуршови морфоанализатор учун махсус моделлаштирилиши лозим. Бироқ шакли ва сўздаги қуршови бир хил бўлган грамматик шаклларни моделлаштириш билан ҳам автоматик морфоанализга эришиб бўлмайди, зеро, улар автоматик ишловда семантик/синтактик таҳлил қилиниши лозим.

2. Сўзларда луғавий шакл ясовчи морфемалар қуршовини таҳлил қилиш, моделлаштириш унинг сўз ясовчи, синтактик шакл ясовчилардан фарқланишини таъминлайди. Жумладан, исм ҳамда феълларга хос синтактик шакл ясовчи морфемалар грамматик омонимия ҳосил қилади. Шаклдош қўшимчалар, одатда, бир гуруҳга мансуб бўлса-да, уларнинг ўзаро шаклдошлиги фақат моделлаштириш орқали фарқланади.

3. Адабий тил нормасидан четлашиш ҳолати қаҳрамоннинг индивидуал лексикаси, характерини яратишга хизмат қилади. Адабий тил учун типик бўлмаган, аммо нутқда кўп учрайдиган ҳодисалар ҳам морфоанализатор базасида эквивалент сифатида қайд этилиши лозим, чунки фойдаланувчи мақсади билан бадий асар парчаси, оғзаки нутқ (баъзан шева) материали ҳам морфоанализ таҳлилига тортилиши мумкин. Қўшимча морфемалардаги истиснолар “истиснолар луғати”дан жой олиши лозим.

Бундай шакллар лингвистик таъминотини тузиш жараёнида адабий меъёр билан бирга номеърий, аммо кенг ишлатиладиган шаклларни инобатга олиш зарурати пайдо бўлади.

4. Юкламалар қуршови ва сўздаги позицияси услубни ҳам белгилайди: бу сўзларнинг семантикаси билан боғлиқ бўлиб, услуб ҳақида ахборот беради. Шундай қўлланишлар маълумотлар базасига киритилса, гапларда юкламаларнинг ишлатилиши орқали матн услубини аниқлаш имкони туғилади. Зеро, морфоанализаторга қўлланиш позицияси қайси услубга тегишли эканлиги ҳақида тег бириктириш кифоя.

5. Морфоанализатор муайян сўзнинг ясама сўз, ясовчи қисм ва ясаиш асосини аниқлай олиши учун унинг лингвистик таъминотида ясама сўз структураси ҳақида мукамал маълумот бўлиши талаб қилинади. Анализаторнинг асослар луғатида ўзбек тилидаги барча ўзақлар маълум туркумга тегишлилиги ҳақидаги маълумот билан сақланади: уларнинг қайси сўз туркумига тегишлилиги тегланган бўлади. Сўз ясаиш моделлари асосида анализатор ўнгдан чапга қараб уч нарсани: сўзнинг асоси, ясаиш асоси ва сўз ясовчи аффиксни аниқлайди.

6. Морфоанализаторда отдан от ясаиши R_n+d^n , отдан сифат ясаиши R_n+d^a модели, отдан равиш ясаиши R_n+d^{adv} , отдан феъл ясаиши R_n+d^v , сифатдан от ясаишида R_a+d^n , сифатдан равиш ясаиши R_a+d^{adv} , сифатдан феъл ясаиши R_a+d^v , сон-асосдан ясаиш $R_{num}+d$, олмошдан ясаиш $R_{pr}+d$, равишдан ясаиш $R_{adv}+d$, тақлиддан ясаиш $R_{sim}+d$, феълдан ясаиш R_v+d инвариант моделлари ҳамда махсус субмоделлар орқали белгиланади. Моделлар морфоанализаторнинг шаклни таниши, асос туркумини аниқлаши, асосдан кейин қўшилган морфемалар тартиби ҳамда бу тартибнинг маънога таъсирини таҳлил қилиши нуқтаи назаридан ишлаб чиқилди.

7. Моделларда асос билдирган маънодан келиб чиқиб гуруҳлаш усулидан фойдаланилди. Феълдан бешта маъновий гуруҳ асосида беш хил моделда айрим сўз ясовчи аффиксларнинг такрорланиши бир сўз ясовчи

аффиксининг бир неча тип сўз ясагани билан изоҳланади. Шунингдек, сўз ясовчи ўзакка луғавий шакл ясовчи морфемадан олдин қўшилиши меъёр эканлиги маълум, бироқ феълдан сўз ясалганда сўз ясовчи морфемадан олдин маълум грамматик шакл келиш ҳолатлари ҳам учрайди, бу ҳолат бошқа туркумдан сўз ясалишига хос эмас. Буни инобатга олиш морфоанализатор учун муҳим.

8. R+d моделининг воқеланишидан ҳосил бўлган сўзларда лексемаланиш ҳолати ҳам кузатилади; ҳатто лексемаланиш натижасида тамоман бошқа маъно ифодаланади. Бундай сўзлар таркибидаги ясовчидан олдин жойлашувчи луғавий шакл ясовчининг вазифаси пассивлашади, сўз ясовчи морфема бўлиб кўринади. Бундай модел ҳосиласи синхрон аспектда таҳлил қилинади.

9. Сўз ясалиш типларини моделлаштиришда сўз ясовчиларни эмас, балки сўз ясалиш типларини таҳлил этиш йўлидан бориш тўғри. Бунда сўз ясалишидек кўринган, аммо ясалиш ҳисобланмайдиган ҳодисалар аниқ белгиланади ва сўз ясалиш объектига киритилмайди. Иккинчидан, шу вақтгача сўз ясовчига боғлаб ўрганилган ҳодисалар сўз ясалиш типлари мисолида ўрганилади. Морфоанализатор сўз ясалиш таҳлилида фақат морфологик шаклларни аниқлаш билан сўз ясалиш ҳодисасини тўғри таҳлил қилиб бўлмайди, сифатли таҳлилга сўз ясалиш типлари таҳлили билан эришилади.

10. Сўз таркибида бирдан ортиқ ясовчи морфема бўлган тип учун асосий инвариант $R+d_1+d_2$ ва унинг субмоделлари: $R_n+d_1+d_2$; $R_a+d_1+d_2$; $R_{num}+d_1+d_2$; $R_{pr}+d_1+d_2$; $R_{adv}+d_1+d_2$; $R_{sim}+d_1+d_2$; $R_v+d_1+d_2$ орқали воқеланиши кузатилади. Бунда феъл, от, сифат, сон, олмош, равиш, тақлид асосга кетма-кет икки сўз ясовчи морфема қўшиш билан сўз ясалади.

11. Ўзбек тилига форс-тожик тилидан ўзлашган баъзи ясовчилар олдпозицияни эгаллайдики, сўз ясалишини ўрганишда ясовчининг жойлашув ўрни аҳамиятсиз бўлиши мумкин, аммо морфоанализда бу жараён муҳим. Сўз

ясовчи қўшимча асосдан кейин жойлашган бўлса, анализатор ўнгдан чапга таҳлил қилишда сўз ясовчи аффиксни аниқлайди. Маълумотлар омборида {сўз ясовчи + ясалиш асоси} қолипи бўлмаганда ясовчи асосдан олдин жойлашган сўз ясовчи – олдқўшимча таҳлилини амалга оширишда натижа беролмаслиги бундай ясалиш типи ҳам алоҳида модел билан қоидалаштирилиши талаб қилинади. Шу сабабли аффиксоидлар рўйхати ҳам маълумотлар омборидан жой олиши лозим.

12. Морфоанализаторнинг ўзбек тили луғавий шакл ясовчи морфемалар қуршовини аниқлаши учун 9 та инвариант модел асосида 39 та субмодел; сўз ясовчи қўшимчалар қуршовининг автоматик таҳлили учун 1 та инвариант модел асосида 7 та субмодел, 60 та субмодел кўриниши ишлаб чиқилди. Мураккаб сўз ясалиши автоматик таҳлили учун, асосан, инвариант ва субмоделлар ишлаб чиқилди. Жумладан, ишимизда ўзакка иккита сўз ясовчи морфема қўшилишини фарқлаш учун битта инвариант асосида 7 субмодел, унинг 5 та кўриниши; ўзакка учта, тўртта сўз ясовчи морфема қўшилишини фарқлаш учун битта инвариант модел аниқланди.

ХУЛОСА

1. Лингвистик таъминот ахборот технологиялари дастурий таъминотида ишлатилувчи тиллар, уларни яратиш усул, восита, қўллаш ҳамда фойдаланиш комплекси дир. Табиий тилда матнларни автоматик қайта ишлашда қўлланиладиган воситалар мажмуи, тил процессори бўлса, ахборот тизимларининг лингвистик таъминоти ахборот объекти тузилиши, унинг ишлаши учун фойдаланиладиган, формаллашган, табиий тил ифодаладиган тушунча устида бажариладиган амаллар, унинг воситалари тўплами ҳисобланади.

2. Лингвистик дастурий таъминот деганда табиий тилдаги аудиомаълумот, расм, матнни таҳлил қилиш, қайта ишлаш, сақлашга оид дастур ва маълумотлар мажмуи назарда тутилади. Унинг электрон луғат, орфоқорректор, кидирув тизими, машина таржимаси системаси, таржимон хотирасини бошқариш дастури, нутқ анализ/синтези тизими, овозли таржима тизими каби турлари ажратиладики, бунда ахборот-кидирув тизимининг лингвистик таъминоти таркибига алифбо, микросинтаксис, луғат, тезаурус, классификатор; турли тилларни мослаштиришни таъминловчи воситалар; амалиёт воситалари киради.

3. Морфологик таҳлилнинг вазифаси матндаги сўзнинг туркумий белгисини автоматик тарзда белгилаш, лексик-грамматик синфга мансублигини аниқлашдан иборат. Матн ахборотини синтактик таҳлил қилиш гапнинг семантик элементи(гуруҳга мансублик, предикатив асос)ни ажратишга имкон яратади. Лексик таҳлил тилнинг ўзига хослигини кўрсатувчи гап шакллари, лексик ифода (ўзлашма, жаргон сўз)ни аниқлайди. Шунингдек, автоматик таҳлил давомида компьютер матн ҳақида ўз тилида лексик-морфологик, синтактик, семантик тасаввур ҳосил қилади. Ахборот базаси, таҳлил дастурларининг амалий аҳамияти АҚТ ҳамда машина таржимаси сифатини яхшилаш, тил корпуси бирликларини лингвистик разметкалаш, шунингдек, мустақил анализатор сифатида ишлаши билан

белгиланади. Зеро, морфологик анализатор луғатдаги сўзшакл, алоҳида сўзларни қиёслашга мўлжалланган, сўзнинг грамматик хусусиятини ифодалайдиган воситадир.

4. Морфоанализаторнинг лингвистик таъминоти маълумотлар базаси кўринишида сақланади: база асослар луғати ва қоидалар луғатидан иборат. Маълумотлар базаси кўламли маълумотлар йиғиндиси бўлиб, унда маълумот сараланади, база тизимида жадвал шаклида сақланади. Асослар луғати тил лексикасидан ташкил топади, грамматик луғатда тилдаги грамматик шакллар рўйхати тўпланади; аффикслар базаси; таснифлаш қоидалари; алломорфлар мослиги қоидалари маълумотлар базасидан жой олади. Морфоанализатор лингвистик таъминоти ўзбек тилидаги флекция ҳолатлари (ўзак ва кўшимча чегарасидаги фонетик ўзгаришнинг акс этиши); одатий синтактик шакл ясовчилар билан бирга супплетив, ўзгармас лексемалар ҳақидаги маълумотни камраб олиши; омоним лексемалар базасини ўзида мужассамлаштиради.

5. Морфоанализатор тузиш учун тилнинг грамматик категория, грамматик шакллари тавсифланади, граммема, квазиграммема ва дериватеманинг чегараси аниқ белгиланади. МА тизимига сўз туркуми ва бошқа морфологик белгиларини аниқлаш; ўзак/асос ва кўшимчани фарқлаш; матнни тушуниш; истисно сўзларни таҳлил қилиш; натижани файл кўринишида тақдим қилиш талаблари кўйилади. Билъакс, морфоанализаторнинг асосий бирликлари морф, алломорф, сўзшакл бўлиб, унда морф шакли тасодифан ўхшаш бўлган айнан бир морфологик хусусиятга эга бўлган минимал ҳодиса; алломорф бир морфеманинг турли фонетик таркибга эга бўлган морфлари йиғиндиси; сўзшакл эса таркибий қисмлари орасида морфотактика орқали аниқланадиган алоқа мавжуд бўлган алломорфлар кетма-кетлиги ҳисобланади. Морфотактик қоидалар эса асос ва аффиксал гуруҳлар орасидаги ўзаро алоқани аниқлашга хизмат қилади.

6. Морфологик таҳлил стемматизация, лемматизация, граммедалаш каби асосий блоклардан ташкил топади. Бунда стемматизация қидирилаётган

сўзнинг асосини топиш; лемматизация сўзшакли леммага келтириш жараёни бўлиб, у флектив тиллар учун аҳамиятли, чунки агглютинатив тилларда сўзнинг “стем”и лексемага тенг бўлади. Фақат флекцияга учраган сўзшаклда лемма ва стемнинг мос келмаслик ҳолати кузатилади. Таҳлил сифатини оширишнинг яна бир омили позицион ва нопозицион морфемаларни аниқлашдан иборат. Позицион морфема гапдаги позициясидан қатъи назар бир семантик гуруҳга мансуб бўлади. Нопозицион морфема гапдаги позициясига қараб турли семантик гуруҳга мансуб бўлади. Маълумотлар базасида позицион морфемалар рўйхатини шакллантириш, уларнинг локализациясини моделлаштириш морфологик таҳлил сифатини оширади.

7. Грамматик категорияга муносабати жиҳатидан сўз шакллари *категориал ва нокатегориал шаклларга* бўлинади. Категориал шакл маълум бир сўз туркумига хос грамматик шакллар системасига бирлашувчи, грамматик категория ҳосил қилувчи шакл, улар граммема сифатида белгиланади. Нокатегориал шакл ҳам маълум грамматик маънони англатади, аммо ўзаро бир системага бирлашмайди, маълум даражада мустақилликка эга бўлади, улар квазиграммема саналади. МАда сўз ясовчи шакл дериватема саналади. Ўзбек тили морфоанализи (1) эҳтимолий ўзакни қидириш (морфологик анализатор сўров ойнасига тушган сўзшакл асослар луғатида эҳтимолий бўлинадиган икки қисмга ажратилади); (2) асослар луғатидан эҳтимолий мос келувчи асосни қидириш; (3) маълумотлар омборидаги ахборот асосида топилган ўзакнинг морфонологик типини аниқлаш; (4) ўнг томондаги қисмларни таҳлил қилиш (аффикслар гуруҳи аниқланади) каби босқичлардан иборат бўлади.

8. Грамматик воситанинг сўздаги тартиби маълум қонуниятга эга, сўздаги қуршови эса унинг вазифасини белгилайди. Шакли бир хил, сўздаги қуршови ҳар хил бўлган грамматик шакллар морфотаҳлил учун алоҳида моделлаштирилиши лозим. Шакли, сўздаги қуршови бир хил бўлган грамматик шаклини аниқлашга морфоанализ билан ҳам эришиб бўлмайди, улар

семантик/синтактик автоматик ишловдан ўтиши лозим. Сўзда луғавий шакл ясовчи морфема қуршовини моделлаштириш морфотаҳлилда унинг сўз ясовчи, синтактик шакл ясовчидан фарқланишини таъминлайди. Морфоанализатор ясама сўз, ясовчи қисм ва ясалиш асосини аниқлаши учун лингвистик таъминотда ясама сўз структураси ҳақида маълумот бўлиши лозим. Сўз ясалиш моделлари асосида анализатор ўнгдан чапга қараб уч нарсани: сўзнинг асоси, ясалиш асоси ва сўз ясовчи аффиксни аниқлайди. Модел морфоанализаторнинг шаклини таниши, асос туркумини аниқлаши, асосга қўшилган морфемалар тартиби ҳамда бу тартибнинг маънога таъсирини таҳлил қилиши нуқтаи назаридан ишлаб чиқилади.

9. Адабий тил нормасидан четлашиш ҳолати қаҳрамоннинг индивидуал лексикаси, характерини яратишга хизмат қилади. Адабий тил учун типик бўлмаган, аммо нутқда кўп учрайдиган ҳодисалар ҳам морфоанализатор базасида эквивалент сифатида қайд этилиши лозим, чунки фойдаланувчи мақсади билан бадий асар парчаси, оғзаки нутқ (баъзан шева) материали ҳам морфологик таҳлилга тортилиши мумкин. Қўшимча морфемалардаги истисно “истиснолар луғати”дан жой олиши талаб қилинади. Бундай шакллар рўйхатини тузишда адабий меъёр билан бирга номеъёрий, аммо кенг ишлатиладиган шакллар инобатга олиниши зарур.

10. Морфоанализаторнинг ўзбек тили луғавий шакл ясовчи морфемалар қуршовини аниқлаши учун 9 та инвариант модел асосида 39 та субмодел; сўз ясовчи қўшимчалар қуршовининг автоматик таҳлили учун 1 та инвариант модел асосида 7 та субмодел, 60 та субмодел кўриниши ишлаб чиқилди. Мураккаб сўз ясалиши автоматик таҳлили учун, асосан, инвариант ва субмоделлар ишлаб чиқилди. Жумладан, ишимизда ўзакка иккита сўз ясовчи морфема қўшилишини фарқлаш учун битта инвариант асосида 7 субмодел, унинг 5 та кўриниши; ўзакка учта, тўртта сўз ясовчи морфема қўшилишини фарқлаш учун битта инвариант модел аниқланди.

Фойдаланилган адабиётлар

I. Норматив-ҳуқуқий ҳужжатлар ва ижтимоий-сиёсий адабиётлар

1. “Ўзбекистон Республикаси давлат тили ҳақида”ги Қонуни. – Тошкент: Ўзбекистон, 1989.
2. Каримов И.А. Юксак маънавият – енгилмас куч. – Тошкент: Маънавият, 2008. – 176 б.
3. Мирзиёев Ш.М. Буюк келажакимизни мард ва олижаноб халқимиз билан бирга қураимиз. – Тошкент: Ўзбекистон, 2017. – 488 б.
4. Мирзиёев Ш.М. Қонун устуворлиги ва инсон манфаатларини таъминлаш – юрт тараққиёти ва халқ фаровонлигининг гарови. – Тошкент: Ўзбекистон, 2017. – 48 б.
5. Ўзбекистон Республикаси Президенти Шавкат Мирзиёевнинг «Ўзбек тилининг давлат тили сифатидаги нуфузи ва мавқеини тубдан ошириш чора-тадбирлари тўғрисида»ги фармони // www.xabar.uz

II. Илмий рисола, монография, дарслик, ўқув ва методик қўлланмалар

1. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика: учеб. пособие. – М.: МИЭМ, 2011. – 118 с.
2. Апресян Ю.Д., Богуславский И.М., Иомдин Л.Л. и др. Лингвистическое обеспечение системы ЭТАП-2. – М.: Наука 1989. – 296 с.
3. Ахборотлар тизимлари ва технологиялари (С.С.Фуломовнинг умумий таҳрири остида). – Тошкент: Шарқ, 2000. – 196 б.
4. Ахо А., Ульман Дж. Теория синтаксического анализа, перевода и компиляции. В двух томах. пер. с англ. – М.: Мир, 1978. – 306 с.
5. Ахо Альфред В., Лам Моника С, Сети Рави, Ульман Джеффри Д. Компиляторы: принципы, технологии и инструментарий, 2-е изд. : Пер. с англ. – М. : ООО “И.Д. Вильямс”, 2008. – 1184 с.

6. Берхеев М.М., Заляев И.А., Кожевников Ю.В. и др. Основы систем автоматизированного проектирования. Учеб. пособие. – Издательство Казанского университета, 1988. – 253 с.
7. Brian Roark, Richard Sproat. Computational Approaches to Morphology and Syntax. Oxford University Press Inc. – New York, 2007. – 163 p.
8. Замонавий ўзбек тили. Морфология. (Мас. муҳаррир: Р.Р.Сайфуллаева). – Тошкент: Мумтоз сўз, 2008. – 522 б.
9. Зикриллаев Ғ. Истиқлол ва адабий тил. – Тошкент: Фан, 2004. – 119 б.
10. Зикриллаев Ғ.Н. Феллинг шахс, сон, ҳурмат категорияси системаси. – Тошкент: Фан, 1990. – 110 б.
11. Кубрякова Е.С. Основы морфологического анализа. – М., 1974. – 254 с.
12. Қўнғуров Р. Субъектив баҳо формаларининг семантик ва стилистик хусусиятлари. – Ташкент, 1980. – 180 б.
13. Қўчқортоев И.К. Сўз маъноси ва унинг валентлиги. – Ташкент, 1977. – 205 б.
14. Марчук Ю. Компьютерная лингвистика. – Москва: АСТ Восток-Запад, 2007. – 174 с.
15. Mahmudov N., Nurmonov A., Sobirov A. va b. Ona tili. 6-sinf uchun darslik. – Toshkent: 2005. – 112 b.
16. Mahmudov N., Nurmonov A., Sobirov A. va b. Ona tili. 7-sinf uchun darslik. – Toshkent: 2005. – 128 b.
17. Мельчук И.А. Курс общей морфологии. Т.1. – Москва-Вена: Прогресс, 1997. – 596 с.
18. Мещанинов И.И. Члены предложения и части речи. –Л.: Наука, 1978. – 256 с.
19. Митрофанов В.Г., Калачев О.Н., Схирладзе А.Г. и др. САПР в технологии машиностроения. Учеб. пособие. – Ярославль: Ярослав. ГТУ, 1995. – 298 с.

20. Морозов В.В. Лингвистическое обеспечение автоматизированных систем научно-технической информации / В.В. Морозов, Б.Р. Певзнер. –М.: ИПКИР, 1987.–76 с.
21. Мухамедов С.А., Пиотровский Г.Г. Инженерная лингвистика и опыт системно-статистического исследования узбекских текстов. –Тошкент: Фан, 1986. – 178 с.
22. Мухамедова С. Ҳаракат феъллари асосида компютер дастурлари учун лингвистик таъминот яратиш. –Тошкент, 2006. – 178 б.
23. Мухаммедова С. Ҳаракат феъллари асосида компютер дастурлари учун лингвистик таъминот яратиш. Методик қўлланма. – Тошкент, 2006. – 156 б.
24. Mengliyev B., Xoliyorov O., Abdurahmonova N. O`zbek tilidan universal qo`llanma. (Qayta ishlangan 3-nashri). – Toshkent: Akademnashr, 2014. – 389 b.
25. Ne`matov N., G`ulomov A., Qodirov M., Abduraimova M. Ona tili: 6-sinf uchun darslik. – T.: O`qituvchi, 2004. – 142 б.
26. Нигматов Х.Г. Функциональная морфология тюркоязычных памятников XI-XII вв. –Тошкент: Фан, 1989. – 298 с.
27. Норов А. Компютер лингвистикаси асослари. – Қарши, 2017. – 136 б.
28. Нурмонов А., Шаҳобиддинова Ш., Искандарова Ш., Набиева Д. Ўзбек тилининг назарий грамматикаси. Морфология. – Тошкент: Янги аср авлоди, 2001. – 256 б.
29. Неъматов Ҳ., Расулов Р. Ўзбек тили систем лексикологияси асослари. – Тошкент: Ўқитувчи, 1995. – 112 б.
30. Потенбня А.А. Из записок по русской грамматике. Ч.І. –Харьков, 1888. – 382 с.
31. Пўлатов А. Компютер лингвистикаси. – Тошкент: Akademnashr, 2011. – 156 б.
32. Севортян Э.В. Морфологическое строение слова в связи с другими его характеристиками. – В. Кн.: Тюркологический сборник. – М., 1972. – 268 б.

33. Соболева П.А. Словообразовательная полисемия и омонимия. – М., 1980. – 294 с.
34. Соколов А.В. Автоматизация библиографического поиска / А.В. Соколов. – М.: Книга, 1981. – 167 с.
35. Степанова М.Д. Теория валентности и валентный анализ. – М., 1973. – 110 с.
36. Сулейманов Д.Ш., Гатиатуллин А.Р. Структурно-функциональная компьютерная модель татарских морфем. – Казань: ФЭН, 2003. – 220 с.
37. Турсунов У. ва б. Ҳозирги ўзбек адабий тили. – Тошкент: Ўзбекистон, 1992. – 399 б.
38. Ўзбек тили грамматикаси, I том, Морфология. – Тошкент: Фан, 1975. – 608 б.
39. Ўзбек тили стилистикаси. – Ташкент, 1983. – 146 б.
40. Ўринбоева Д.Б. Ўзбек фольклори матнларининг лингвостатистик тадқиқи. – Тошкент: Фан, 2010. – 168 б.
41. Гуломов А. Ўзбек тилида кўплик категорияси. – Ташкент, 1943. – 78 б.
42. Ҳожиев А. Ўзбек тили сўз ясалиши. Олий ўқув юртларининг филология факультетлари студентлари учун қўлланма. – Тошкент: Ўқитувчи, 1989. – 109 б.
43. Шаҳобиддинова Ш. Ўзбек тили морфологияси умумийлик-хусусийлик диалектикаси талқинида. II қисм. – Андижон, 1994. – 34 б.
44. Щербак А.М. Очерки по сравнительной морфологии тюркских языков. – Л., 1977. – 456 с.
45. Щипицина Л.Ю. Информационные технологии в лингвистике : учеб. пособие / Л.Ю.Щипицина. – М.: ФЛИНТА : Наука, 2013. – 128 с.
46. Якубова Н.К. Расстановка аффиксов в узбекских словаформах. – Ташкент, 1979. – 158 с.

III. Матбуот материаллари ва илмий мақолалар

47. Ходжиев А., Нурманов А., Махмудов Н. Спорные вопросы в системе словообразования тюркских языков // Советская тюркология, 1989. – №2.
48. Абдурахмонова Н. Моделирование грамматических категорий аналитических глаголов узбекского языка как этап морфологического анализа в машинном переводе // V Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Ака де мии наук, 2017. – 327 с. – С. 155-180.
49. Абжалова М. Матнларга авто-лингвистик ишлов бериш тизимлари // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.
50. Акобирова З. Сўз яшаш қолиплари, уларда унумлилик ва унумсизлик// Ўзбек тили ва адабиёти, 2001 й.– 6-сон. – Б. 68-70.
51. Антопольский А.Б. Лингвистическое обеспечение АСНТИ: структура и проблема совместимости // НТИ. Сер.2, 1983. – №7. – С.17-25.
52. Арский Ю.М. Информационные ресурсы для устойчивого развития общества / Ю.М.Арский, А.И.Черный // Междунар. форум по информ. и док. –2000. –Т. 25, № 1. –С. 3-9.
53. Арутюнов Н.Д. Некоторые вопросы образования и морфологии основ слов//Филологические науки, 1958. – № 1.
54. Аъламова М. Туркий тилларда буйруқ-истак майли ҳақида //Ўзбек тили ва адабиёти, 1966. –№ 6. – Б.30-34.
55. Баскаков Н.А. Морфологическая структура слова и части речи в тюркских языках//Советское востоковедение, 1957. – № 1.
56. Баскаков Н.А. Порядок и иерархия аффиксов в основе слова в тюркских языках. Баскаков Н.А. В сб.: Морфологическая типология и проблема классификации языков. – М.-Л., 1965.
57. Белоногов Г.Г. Автоматизированная обработка научно-технической информации. Лингвистические аспекты / Г.Г. Белоногов, Б.А. Кузнецов, А.П.

- Новоселов // Итоги науки и техники. Сер. Информатика / ВИНТИ. –1984. –Т.8.–316.
58. Белоозеров В.Н. Ведение и совершенствование Государственного рубрикатора НТИ (ГРНТИ) / В.Н. Белоозеров, В.И. Федосимов, Г.А. Кривенцова // Информ. продукты, процессы и технологии: НТИ-96: Материалы конф. –М., 1996.–С. 60-69.
59. Бозоров О. Ўзбек тилининг юклама категорияси // Ўзбек тили ва адабиёти , 1977. – № 5. – Б.24-28.
60. Болтаев Т.Б., Кузьминов Т.В., Поттосин И.В. О структурном конструировании программ и инструментах его поддержки // Среда программирования: методы и инструменты. – Новосибирск, 1992, С22-37.
61. Гатиатуллин А.Р., Баширов А.М. Морфологический анализатор тюркских словоформ на базе структурно-функциональной модели тюркской морфемы // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Академиинаук, 2017. – 327 с.
62. Гатиатуллин А.Р., Баширов А.М., Осипов Г.С., Смирнов И.В., Шелманов А.О. Методы лингвистического анализа текстов на татарском языке и их применение в поисковой системе EXACTUS // Труды Института системного анализа Российской академии наук. Т. 66., 2016. –№1. –С. 18-25.
63. ГендинаН.И. Лингвистические средства автоматизации документального поиска / Н.И. Гендина;Под ред. В.П.Леонова. СПб.: Б-ка РАН, 1992.–188с.
64. Гузев В.Г., Насилов Д.М. Слово-изменительные категории в тюркских языках и понятие «грамматическая категория» // Сов. тюркология. 1981. – №3. – С.23-35;
65. Гуломов Ё.Г. Ўзбек шеваларида феъл даражалари аффикслари. Илмий ишлар тўплами. Ўзбек тилшунослиги масалалари.– Тошкент: ТошДУ, – 1983. – Б.64-68.

66. Гулямов А. Об аффиксе -лик в узбекском языке. Научные труды ТашГУ. Вып. 260. Языкознание. – Ташкент, 1964. – С. 50-57.
67. Гулямов А.Г. О некоторых особенностях аффиксов – с уменьшительно-ласкательным значением в узбекском языке / Научные труды ТашГУ. Вып. 268//Языкознание и литературоведение. – Ташкент, 1964.
68. Дереза О.В., Каютенко Д.А., Феногенова А.С. Автоматический морфологический анализ для русского языка сравнительный анализ систем // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018.
69. Дрейзин Ф.А. Об алгоритмизации составление алгоритма анализа языка (на примере морфологии агглютинативного узбекского языка)//Научные труды Ташкентского университета. – 1961, вып. 189.Матем. Науки. –С. 121.
70. Дыбо А.В., Шеймович А.В. Аппарат автоматического морфологического анализа для корпуса хакасского языка // Родной язык: лингвистический журнал. – 2016. № 2 (5). – С. 9-39.
71. Дыбо А.В., Шеймович А.В. Автоматический морфологический анализ для корпусов хакасского и древнетюркского языков / Научное обозрение саяно-алтая рецензируемый научный журнал Номер 2(08), 2014 (Периодичность – 2 раза в год. серия: Филология. Выпуск 2). С. 9-31.
72. Дыбо А.В., Шеймович А.В. Автоматический морфологический анализ для корпусов тюркских языков / Филология и культура. Philology and culture. – 2014. № 2 (36). – С. 20-26.
73. Желтов П.В. Разработка морфологического анализатора чувашского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Академии наук, 2017. – 327 с.
74. Журабаева М.К. К вопросу о категории уменьшительно-ласкательности в узбекском языке Труды аспирантов ТашГУ. Вып. 360. Литература и языкознание. – Ташкент, 1970.

75. Жўрабоева М. Аффиксал омонимик қаторни белгилашга доир (-ган ва -лар аффикслари мисолида). – Научные труды ТашГУ. Вып. 427. Вопросы узбекского языкознания. – Ташкент, 1972. – С. 169-177.
76. Захаров В.П. Международные стандарты в области корпусной лингвистики // Структурная и прикладная лингвистика, 2012. – № 9. – С. 201-221. – ISSN 0202-2400.
77. Иванов С.Н. Категория залога в причастии узбекского языка. В книге «Исследования по истории культуры народов Востока»// Сборник в честь академика И.А. Орбели. – М.:Л., 1960. – С.123-134;
78. Исраилова Н.А., Бакасова П.С. Морфологический анализатор кыргызского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Академии наук, 2017. – 327 с.
79. Исхаков Ф.Т. Имя прилагательное. Исследования по сравнительной грамматике тюркских языков. Ч.II, Морфология. – М.,1956. – С.155-156.
80. Касьянов В.Н., Поттосин И.В. Методы построения трансляторов – Новосибирск. Наука, 1986. – с. 344.
81. Кононов А.Н. О природе тюркской агглютинации. – Вопросы языкознания, 1976. – № 4.
82. Кочконбаева Б.Об оптимизации алгоритма морфологического анализа / Шестая Международная конференция по компьютерной обработке тюркских языков «Turklang-2018». (Труды конференции) – Ташкент: Издательско-полиграфический дом «Навоий университети», 2018. – 320 с.
83. Кузнецов А.М. О применении метода компонентного анализа в лексике. –С.260.; Цветков Н.В. К методологии компонентного анализа // Вопросы языкознания. –1984, №2, –с.67.
84. Куканова В.В., Каджиев А.Ю. Алгоритм работы морфологического парсера калмыцкого языка // В сборнике: Писменото наследство и информационните технологии. El'Manuscript–2014 Материали от V

- международна научной конференции. Отговорни редактори В.А.Баранов, В.Желязкова, А.М.Лаврентьев. 2014. – С. 116-119.
85. Леонтьев Н.А. Вопросы автоматизированной лемматизации якутского языка // Современная наука: актуальные проблемы теории и практики. Серия: Познание. 2015. №2 (41). – С. 15-16.
86. Леонтьев Н.А. Морфологический анализатор якутского языка // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) –Ташкент, 2018. – 320 с.–С. 276-279.
87. Леонтьев Н.А., Торотоев Г.Г.Многопользовательская морфологическая разметка корпуса якутского языка // В сборнике: Электронная письменность народов Российской Федерации: опыт, проблемы и перспективы Сборник материалов Международной научной конференции. – 2017. – С. 101-103.
88. Махмудов М.А., Пиотровская А.А., Садыков Т. Система машинного анализа и синтеза тюркской словоформы // Переработка текста методами инженерной лингвистики. –Минск, 1982.
89. Садыков Т. Проблемы моделирования тюркской морфологии : (Аспект порождения кирг. имен. словоформы) / Т. Садыков; АН КиргССР, Ин-тяз. и лит. – Фрунзе : Илим, 1987. – 120 с.
90. Мдивани Р.Р. Лингвистическое обеспечение автоматизированной информационной системы по общественным наукам: состав и структура / Р.Р.Мдивани, В.А.Глинский // Вопр. информ. теории и практики. 1985. –№52.–С. 37-55.
91. Михайлов А.И. Проблемы лингвистического обеспечения АСНТИ / А.И. Михайлов // Вопр. информ. теории и практики. 1984.–№ 51. – С. 5-11.
92. Мишин В.А. Лингвистическое обеспечение библиотечных автоматизированных систем / В.А. Мишин, Л. Санчес-Мехидо // Сов. библиотековедение. 1981. – № 2. – С. 84-93.

93. Назарова С. «Модда-маъдан оти + нарса-буюм оти» қолипли сўз бирикмаси (модда-ашё отлари ва нарса-буюм отларининг бирикуви – атов бирликлар ҳақида) //Тил ва адабиёт таълими, 2004 й. – №5. – Б. 52-53.
94. Нариманова М.Д. Ҳозирги ўзбек тилида феъллардаги даража ясовчиларнинг баъзи бир хусусиятлари ҳақида. Илмий ишлар тўплами. Ўзбек тилшунослиги масалалари. –Тошкент: ТошДУ, 1983. – Б. 52-59.
95. Неъматов Ҳ.Ғ. Тажалли тасаввуф билиш назарияси ва тилшуносликда синтаксисни ўрганиш масалалари // Ўзбек тили ва адабиёти, 1993. – №2. – Б. 27-30.
96. Нуъмонова Ш. Ўзбек тилида -лик Аффикси билан ясалган отларнинг семантик классификацияси ҳақида // Тилшунослик ва адабиётшуносликка оид тадқиқотлар. –Тошкент, 1965. – Б. 44-49.
97. Орехов Б.В. Проблемы морфологической разметки башкирских текстов // Труды Казанской школы по компьютерной и когнитивной лингвистике TEL-2014. – Казань: Фэн, 2014. – С. 135-140.
98. Орехов Б.В., Слободян Е.А. Проблемы автоматической морфологии агглютинативных языков и парсер башкирского языка //Информационные технологии и письменное наследие: материалы международной научной конференции (Уфа, 28-31 октября 2010 г.) /отв. ред. В. А. Баранов. – Уфа; Ижевск: Вагант, 2010. С. 167-171.
99. Орхун М. Computational analysis of uzbek nouns / Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.
100. Раҳматуллаев Ш. -дек форманти ҳақида//Ўзбек тили ва адабиёти, 1977. – № 1. – Б. 48-52.
101. Совершенствование методологии познания в целях развития науки. сборник статей по итогам Международной научно-практической конференции: в 2 ч. –Ташкент, 2017. – С. 11-13.

102. Степанова М.Д. Вопросы моделирования в словообразовании и условия реализации моделей//Вопросы языкознания, 1975. – № 4.
103. Сулейманов Д.Ш. Реализация татарской морфологии (на англ.яз.) // In Iss.: Computational linguistics and intelligent text processing: third international conference; proceedings / CICLing 2002, Mexico, February 17–23, 2002. Alexandr Gelbukh (ed.). – Berlin: Springer, 2002 (Lecture notes in computer science; Vol. 2276). – P. 327-329.
104. Сулейманов Д.Ш., Гатиатуллин А.Р. Модель татарской аффиксальной морфемы и ее реализация // Серия: Интеллект. Язык. Компьютер. – Вып.4. –Казань: Фэн. – 1996. – С. 113-127.
105. Сулейманов Д.Ш., Гатиатуллин А.Р. К разработке Лемматорататарских словоформ // В сб. Трудов Международного семинара Диалог-99 по компьютерной лингвистике и ее приложениям в двух томах(Таруса, 31 мая – 4 июня 1999 г.). Т. 2. – Таруса, 1999. – С. 306-314.
106. Сулейманов Д.Ш., Гильмуллин А.А., Гильмуллин Р.А. Базаморфотактических правил для татарского глагола как основа двухуровневого морфологического анализатора // Сборник трудов Международного семинара «Диалог», Казань, 1998. – С. 597-609.
107. Сулейманов Д.Ш., Гильмуллин А.А., Гильмуллин Р.А. Двухуровневое описание морфологии татарского языка Тез. Межд. научн. конф., посвященной 200-летию университета: “Языковая семантика и образ мира” (Казань, 7–10 октября 1997г.). Книга 2. – Казань, 1997. – С. 65-67.
108. Сулейманов Д.Ш., Гильмуллин А.А., Гильмуллин Р.А. Двухуровневое описание морфологии татарского языка Тез. Межд. научн. конф., посвященной 200-летию университета: “Языковая семантика и образ мира” (Казань, 7–10 октября 1997г.). Книга 2. – Казань, 1997. – С. 65-67.
109. Сулейманов Д.Ш., Гильмуллин Р.А., Гатауллин Р.Р. Морфологический анализатор татарского языка на основе двухуровневой модели морфологии / Пятая Международная конференция по компьютерной обработке тюркских

языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Академии наук, 2017. – 327 с.

110. Сулейманов Д.Ш., Шафигуллин Р.Н. Морфологический корректор татарских текстов ТАТКОР // Татарский язык и новые информационные технологии. Серия: Интеллект. Язык. Компьютер. – Вып.2. – Казань, 1995. – С.86-89.

111. Тожиев Ё. Феъл даражаларини ясовчи аффикслар синонимияси. Научные труды ТашГУ. Вып. 443. Вопросы узбекского языкознания. – Тошкент, 1973. – Б. 115-124.

112. Горотоев Г.Г., Ноговицына А.Н. Лингвистическое аннотирование наклонений глагола якутского языка // Вестник Северо-Восточного федерального университета им. М.К. Аммосова. 2017. №3 (59). С. 108-120.

113. Тузов В.А. Морфологический анализатор русского языка //Вестник СПбГУ, сер. 1. 1996. Вып. 1 (N15). –С. 41-45.

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

115. Улуханов И.С. Значения словообразовательных аффиксов и части речи мотивирующих слов//Филологияческие науки, 1974. – № 4.

116. Усмонов С. Сўз ясовчи аффиксоидлар//Ўзбек тили ва адабиёти, 1963. – № 3.

117. Усмонов С. Ҳозирги ўзбек тилида сўзнинг морфологик тузилиши // Низомий номидаги ТошДПИ илмий асарлари, 42 том, 2-китоб. – Тошкент, 1963.

118. Формальные модели и программные инструменты компьютерной обработки татарского языка / Р.Р.Гатауллин, А.Р.Гатиатуллин, О.А.Неврозова, Д.Р.Мухамедшин, Д.Ш.Сулейманов, Б.Э.Хакимов, А.Ф.Хусаинов. – Академия

наук РТ, Институт прикладной семиотики АН РТ. – Казань: Академии наук, 2019. – 260 с. – С.39-40.

119. Хакимов М. Логико-лингвистические модели русского языка // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

120. Халитова Н.А., Закирова Р.А., Гимадудтинова Р.У. Морфологический анализ при машинном переводе с татарского языка на русский / Вероятностные методы и кибернетика I, Сборник работ НИИММ им. Н. Г. Чеботарева при Казанском университете, Учен. зап. Казан. ун-та, 122, № 4, Изд-во Казанского ун-та, Казань, 1962. – С. 98-105.

121. Хертек А.Б., Оржак Б.Ч. О морфологической разметке электронного корпуса текстов тувинского языка // Филологические науки. Вопросы теории и практики. – 2012. №7-2(18). – С. 214-218.

122. Ҳожиёв А. Ўзбек тилида -лик аффиксининг вазифаси ҳақида//Ўзбек тили ва адабиёти, 1985. – №1. – Б. 24-27.

123. Ҳожиёв А. Форма ясовчи аффикслар//Ўзбек тили ва адабиёти, 1977.

124. Ҳожиёв А. Ҳозирги ўзбек тилида форма ясалиши.– Ташкент, 1979.

125. Чемышев А.В. Динамический лексикон: компьютерное представление моделей словообразования // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 2. – Казань: Академии наук, 2017. – 327 с. – С.224-231.

126. Шарипбаев А.А., Бекманова Г.Тошкент, Ергеш Б.Ж., Бурибаева А.К., Карабалаева М.Х. Интеллектуальный морфологический анализатор, основанный на семантических сетях // Материалы международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем» (OSTIS-2012). Минск, БГУИР, 16–18 февраля 2012 г. – С. 397–400.

127. Щербак А.М. Последовательность морфем в словоформе как предмет специального лингвистического исследования//Вопросы языкознания, 1983. – № 3.
128. Abdurakhmonova N. Uzbek ontology of Uzbek language as example of adjective // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.
129. Akhmedova D.B., Mengliev B.R. Semantic Tag Categories in Corpus Linguistics: Experience and Examination International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-3S, October 2019. – P. 208-212.
130. Altintas K., Cicekli I. A morphological analyzer for Crimean Tatar //Proceedings of the 10th Turkish Symposium on Artificial Intelligence and Neural Networks (TAINN'2001). – 2001. – P. 180-189.
131. Altintas K., Cicekli I. A morphological analyzer for Crimean Tatar //Proceedings of the 10th Turkish Symposium on Artificial Intelligence and Neural Networks (TAINN'2001). – 2001. – P. 180-189.
132. Aniskov M.I., Boltaev T.B., Kochetov D.V. The Structured Constructing as a Discipline of Safe Programming and Instruments Supporting It. //Instrumental Congress on Computer Systems and Applied Mathematics CSAM'93. St-Petersburg.July 19-23.
133. Antworth, E.L. PC-KIMMO: a two-level processor for morphological analysis. Occasional Publications in Academic Computing No. 16. Dallas:Summer Institute of Linguistics, 1990. – 273 p.
134. Belonogov G. G. On the use of a similarity method for automatic processing of textual information [Обиспользовании метода аналогии при автоматическо жобработкетекстовожин формации] Problems of cybernetics [Problemy kibernetiki]. Issue 28. – Moscow: Nauka Publ., 1974.

135. Bolshakov I. A. CrossLexica: A large electronic dictionary of collocations and semantic links between Russian words [КроссЛексика – большой электронный словарь сочетаний и смысловых связей русских слов]. Компьютерная лингвистика и Интеллектуальные Технологии: Труды Международной Конференции «Диалог 2009» [Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference «Dialogue 2009»]. – Moscow, 2009. – P. 45-50.
136. Çağrı Çöltekin (2014) A Set of Open Source Tools for Turkish Natural Language Processing In: Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14) Ed. by N. Calzolari et al. 1079-1086.
137. Structured Constructing as a Discipline of Safe Programming and Instruments Supporting It /Aniskov M.I., Boltaev T.B., Kochetov D.V. at al//Instrumental Congress on Computer Systems and Applied Mathematics CSAM'93. St-Petersburg. July 19-23.
138. Kessikbayeva G. and Cicekli I. “Rule based morphological analyzer of Kazakh language,” in Proceedings of the 2014 Joint Meeting of SIGMORPHON and SIGFSM. Baltimore, Maryland: ACL, June 2014. – p. 46-54.
139. Gelbukh A. F. Effectively realizable morphologic model of inflective language [Эффективно реализуемая модель морфологии и эктивного языка]. Научно-техническая Информация [Scientific and Technical Information], series 2, #1, 1992. – P. 24-31.
140. Gökgöz, Ercan, et al. Two-Level Qazan Tatar Morphology. Proceedings of the 1-st International Conference on Foreign Language Teaching and Applied Linguistics, May 5-7 2011 Sarajevo. – P. 428-432.
141. Jurafsky D.S., Martin J.H. Speech and Language Processing. (Contributing writers: Andrew Kehler, Keith Vander Linden) Nigel Ward 2000y. – Prentice Hall, Englewood Cliffs, New Jersey 07632. – pages: 950.

142. Karimov R., Mengliev B. Theoretical fundamentals of uzbek-english parallel corpus / Journal of critical reviews. ISSN- 2394-5125. – VOL 7, ISSUE 17, 2020. – P. 73-76.
143. Karimov R.A., Mengliev B.R. The Role of the Parallel Corpus in Linguistics, the Importance and the Possibilities of Interpretation International Journal of engineering and Advanced Technology (IJEAT). ISSN: 2249 – 8958, Volume-8, Issue-5S3 July 2019. – P. 388-391.
144. Kemal Altıntaş Turkish to Crimean Tatar machine translation system(a thesis submitted to the department of computer engineering and the institute of engineering and science of Bilkent university in partial fulfillment of the requirements for the degree of master of science) 2001, P. 30.;
145. Kemal Oflazer. Two-level Description of Turkish Morphology. Literary and Linguistic Computing, – Vol. 9, No 2, – 1994.;
146. Kemal Oflazer. Two-level Description of Turkish Morphology. Literary and Linguistic Computing, – Vol. 9, No 2, – 1994.
147. Kessikbayeva G., Cicekli I. “Rule based morphological analyzer of Kazakh language,” in Proceedings of the 2014 Joint Meeting of SIGMORPHON and SIGFSM. Baltimore, Maryland: ACL, June 2014, pp. 46-54.;
148. Mohri M. Finite-state transducers in language and speech processing / Computational Linguistics. – P.269-312.
149. Morneau R. The Lexical Semantics of a Machine Translation Interlingua. http://www.eskimo.com/~ram/lexical_semantics.html 2006.
150. Orhun M., Tantug C., Adali E. Rule Based Analysis of the Uyghur Nouns / International Journal of Asian Lang. Proc.,19(1), 33-44, 2009.
151. Rodolfe D. Computational Linguistic Text Processing: Lexicon, Grammar, Parsing and Anaphora Resolution / Nova Science Publishers, Inc. – New York, 2008. – P. 4-5.
152. Sokirko A. V. Bystroslovar’: morphological prediction of new Russian words using very large corpora [Быстроговорарь: предсказание морфологии русских слов] / Proceedings of the 2014 Joint Meeting of SIGMORPHON and SIGFSM. Baltimore, Maryland: ACL, June 2014, pp. 46-54.

слов с использованием больших лингвистических ресурсов]. Компьютерная лингвистика и Интеллектуальные Технологии: Труды Международной Конференции «Dialog 2010»– Moscow, 2010. – P. 450.

153. Tantug S., Adali E., Oflazer K.. Computer Analysis of the Turkmen Language Morphology // 5th International Conference on NLP(FinTAL 2006), Turku, Finland, 186-193, 2006.

IV. Диссертациялар ва диссертация авторефератлари

154. 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. – 52 б.

155. Абжалова М.А. Ўзбек тилидаги матнларни таҳрир ва таҳлил қилувчи дастурнинг лингвистик модуллари (Расмий ва илмий услубдаги матнлар таҳрири дастури учун): Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Фарғона, 2019. – 164 б.

156. Абузалова М.К. Субстанциал морфология, валентлик ва синтактик қурилма: Филол. фан. доктори (DSc) диссертацияси. – Самарқанд, 2018. – 224 б.

157. Абузалова М.К. Субстанциал морфология, валентлик ва синтактик қурилма: Филол. фан. доктори (DSc) диссер. автореферати. – Самарқанд, 2018. – 84 б.

158. Айымбетов М.К. Опыт лингвостатистического анализа лексики и морфологии каракалпакского публицистического текста: Автореф. дисс... канд. фил. наук.–Ташкент, 1987.

159. Архангельский Т.А. Принципы построения морфологического парсера для разноструктурных языков: дисс... канд. филол. наук. – Москва, 2012.

160. Бабанаров А. Разработка принципов построения словарного обеспечения турецко-русского машинного перевода: Автореф. дисс... канд. фил. наук. - Л., 1981.

161. Баҳриддинова Б. М. Феъл луғавий шакли тизими. Феълнинг тур категорияси: филол. фан. номзоди... дис. автореф. – Самарқанд, 2002. – 25 б.
162. Баҳриддинова Б. Феъл луғавий шакллари тизими. Феълнинг тур категорияси. Филол. фан номз. дисс... – Тошкент, 2002. – 146 б.
163. Davliyeva A.R. An investigation of Kazan Tatar morphology (Doctoral dissertation, San Diego State University).– Казань, 2011.
164. Джикия М.С. Морфологическая структура слова в турецком языке: автореф. дисс. ... канд. филол. наук. – Тбилиси, 1975. – 25 с.
165. Джурабаева М.А. Аффиксальная омонимия в узбекском языке: автореф. дисс. ... канд. филол. наук. – Тошкент, 1975. – 25 с.
166. Жуманазарова Г.У. Фозил Йўлдош ўғли дostonлари тилининг лингвопоэтикаси: Фил. фан. док. дис...автореф. – Тошкент, 2017. – 65 б.
167. Махкамов Н. Аффиксальный и лексико-аффиксальный плеоназм в узбекском языке: Автореф. дис. ...канд. филол. наук. – Ташкент, 1983. – 25 с.
168. Махмудов, Масуд Ахмед оглы. Разработка системы формально-морфологического анализа тюркской словоформы: на материале азербайджанского языка: дисс... канд. филол. наук.– Баку, 1982.– 145 с.
169. Менглиев Б. Морфологик воситаларнинг маъновий хусусиятлари ва семантик имкониятлари. Филол. фан. номз. дисс автореф. – Тошкент, 1996. – 156 б.
170. Менглиев Б. Р. Лисоний тизим яхлитлигида сатҳлараро муносабатлар: филол. фан. док.... дис. – Тошкент, 2002. – 212 б.
171. Мирзакулов Т.У. Формаобразующие сложные аффиксы в узбекском языке: автореф. дисс. ... канд. филол. наук. – Тошкент, 1980. –25 с.
172. Миртожиев М. Ўзбек тилида омонимларнинг вужудга келиши: Филол. фан. номзоди дисс. – Ташкент, 1963, – 25 б.
173. Муҳамедова С.Х Ўзбек тилида ҳаракат феълларининг семантик ва валентлик хусусиятлари. Филол. фан докт.дисс. ... Тошкент, 2007. – 225 б.

174. Мухаммедов С.А. Статистический анализ лексико-морфологической структуры узбекских газетных текстов: Автореф. дисс... канд. фил. наук.- Ташкент, 1980. – 25 б.
175. Назарова С. Бирикмаларда сўзларнинг эркин боғланиш омиллари. Филол... фанлари номзоди... дисс. – Тошкент, 1997. – 26 б.
176. Ножов И.М. Морфологическая и синтаксическая обработка текста (модели и программы): диссертация канд. филол. наук. – Москва, 2003. – 26 с.
177. Ножов И.М. Морфологическая и синтаксическая обработка текста (модели и программы): диссертация канд. наук. – М., 1996. – 190 с.
178. Пирумова Л.Н. Лингвистическое обеспечение информационно-поисковых систем библиотек АПК: Методика формирования и пути совершенствования: автор. дис.... канд. пед. наук. – Москва, 2003. – 26 с.
179. Пономарев В.В. Лингвистическое обеспечение и социолингвистическая специфика проблемы автоиндексационной актуализации информационных систем: автореф. диссер. канд. филол. наук. – Москва, 2005. – 26 с.
180. Ризаев С. Ўзбек тилининг лингвостатистик тадқиқи: Фил. фан. док. дис... автореф. – Тошкент, 2008. – 26 б.
181. Ўринбаева Д. Халқ оғзаки ижоди: жанрий-лисоний ва лингвостатистик тадқиқ муаммолари: фил. фан. док. дисс. автореф. – Самарқанд, 2019. – 74 б.
182. Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол. фан. бўйича фалсафа доктори (PhD)... диссер. – Қарши, 2018. – 250 б.
183. Шарипова К.А. Словообразующие аффиксы -лик, -чи, -чилик в узбекском языке: автореф. дисс. ... канд. филол. наук. – Тошкент, 1972. – 27 б.
184. Шаҳобиддинова Ш. Умумийлик ва хусусийлик диалектикаси ҳамда унинг ўзбек тили морфологиясида акс этиши: Филол. фан. док... дис. автореф. – Тошкент, 2001 – 25 б.

185. Шаҳобиддинова Ш. Грамматик маъно талқини хусусида. Номз. дисс.... – Самарқанд, 1993. – 166 б.
186. Ширинова Н. Ўзбек тилида белги-хусусият ва предметлик маъноларининг фарқланишида градуонимик каторлар: Филол. фан. номз. диссер. – Тошкент, 2010. – 160 б.
187. Шукуров О. Ҳаракат тарзи шакли парадигмаси: филол. фан. номз... дис. – Самарқанд, 2005. – 126 б.
188. Эшмўминов А.А. Ўзбек тили миллий корпусининг синоним сўзлар базаси: Филол. фан. бўйича фалсафа доктори (PhD)... диссер. – Қарши, 2019. – 140 б.

VI. Луғатлар

189. Большой современный англо-русский, русско-английский словарь. – Новая редакция, Около 450 000 слов, словосочетаний и идиоматических выражений, Мюллер В.К., 2009.
190. Грамматический словарь русского языка. Словоизменение. – А.А.Зализняк. – М., 1977.
191. Гуломов А. ва б. Ўзбек тилининг морфем луғати. – Тошкент, 1977. – 463 б.
192. Каримов С., Қаршиев А., Исроилова Г. Абдулла Қаҳҳор асарлари тилининг луғати. Алфавитли луғат. Частотали луғат. Терс луғат. – Тошкент, 2007.
193. Лингвистический энциклопедический словарь. – Гл. Ред. В.Н.Ярцева. – М.: Советская энциклопедия, 1990. – 685 с.
194. Менглиев Б., Баҳриддинова Б. Ўзбек тилининг сўз таркиби ўқув луғати. – Тошкент: Янги аср авлоди, 2007.
195. Ожегов С. Словарь русского языка. – 1949, 22-е издание, 1990; с 1992 - "Толковый словарь русского языка", совместно с Н. Ю. Шведовой).
196. Раҳматуллаев Ш. Ўзбек тили омонимларининг изоҳли луғати. – Тошкент: Ўқитувчи, 1984. – 214 б.

197. Самад А. “Х” ва “Х”ли сўзларнинг имловий луғати. – Тошкент, 2007. – 346 б.
198. Ўзбек тилининг изоҳли луғати. 5 жилдли. I жилд. – Тошкент: ЎЗМЭ, 2006. – 680 б.
199. Ўзбек тилининг изоҳли луғати. 5 жилдли. 2-жилд. – Тошкент: ЎЗМЭ, 2006.
200. Ўзбек тилининг изоҳли луғати. 5 жилдли. 3-жилд. – Тошкент: “ЎЗМЭ, 2006.
201. Ўзбек тилининг изоҳли луғати. 5 жилдли. 4-жилд. – Тошкент: ЎЗМЭ, 2000-2006.
202. Ўзбек тилининг изоҳли луғати. 5 жилдли. 5-жилд. – Тошкент: ЎЗМЭ, 2006.
203. Ўзбек тилининг имло луғати. (Тузувчилар: Э.А.Бегматов, А.П.Мадвалиев; Н.Маҳмудов таҳрири остида). – Тошкент: Akademnashr, 2013. – 520 б.

VII. Интернет материаллари, электрон ресурслар

204. Болтаев Т.Б., Ибрагимов С.И. О проекте программной системы морфологического анализа узбекского языка // <http://buxdu.uz/index.php/uz/>
205. Ермаков А. Морфологический анализатор - основа поисковых систем // <https://www.kv.by/archive/index2004154301.htm>
206. Ермолаев Д.С. Компьютерный морфологический разбор слов русского языка // <http://citforum.ru/programming/digest/morfolog.shtml>
207. Место и назначение лингвистического обеспечения в информационных системах. Понятие информационной системы. Лекция 4 // <https://gigabaza.ru/doc/116551-pall.html>
208. Морфологический анализатор текста на русском языке mystem [Электронный ресурс]; URL: <http://company.yandex.ru/technologies/mystem/>
209. Плунгян В.А. Классификация морфологических значений // Общая морфология: Введение в проблематику: Учебное пособие. – Изд. 2-е,

- исправленное. – М.: Едиториал УРСС, 2003. –С. 107. –384 с. // <https://ru.wikipedia.org/wiki/%D0%93%D1%80%D0%B0%D0%BC%D0%BC%D0%B5%D0%BC%D0%B0>;
210. Сиразитдинов З.А. Алгоритмическая грамматика словоизменения башкирского языка // <http://mfbl.ru/bashdb/algram/algram.htm>;
211. Системный администратор. Ежемесячный журнал www.samag.ru (Электрон манзил – <http://samag.ru/ru/articles/>)
212. Словарь иностранных слов русского языка / dic.academic.ru › [dic.nsf](http://dic.academic.ru/dic.nsf) › [dic_fwords](http://dic.academic.ru/dic.nsf/dic_fwords) › граммема
213. Tatar Nacional Corpus. <http://tugantel.tatar/?lang=en>
214. Толдова С.Ю., Бонч-Осмоловская АА. Автоматический морфологический анализ // Фонд знаний «Ломоносов». М., 2011. www.lomonosov-fund.ru/enc/ru/encyclopedia:0127430
215. Шадманова К. Табий тиллар учун лексико-семантик лугат маълумотлар базаси яратиш тамойиллари // <http://buxdu.uz/index.php/uz/>

VIII. Интернет манбалари

216. <http://asknet.ru/analytics/programms.htm>
217. <http://asknet.ru/analytics/programms.htm>
218. <http://asknet.ru/analytics/programms.htm>
219. <http://asknet.ru/analytics/programms.htm>
220. <http://modmorph.turklang.net/ru/>
221. http://modmorph.turklang.net/uzc/gram_category
222. <http://starling.rinet.ru/morph.htm>
223. <http://tamgasoft.kg/morfo/>;
224. <http://tatmorphan.pythonanywhere.com/>
225. http://tsput.ru/res/informat/sist_seti_fmo/lekcii/lekciiy-4.html#name3
226. http://tsput.ru/res/informat/sist_seti_fmo/lekcii/lekciiy-4.html#name3
227. http://tsput.ru/res/informat/sist_seti_fmo/lekcii/lekciiy-4.html#name3
228. http://window.edu.ru/catalog/pdf2txt/516/25516/8411?p_page=10

229. <http://www.cs.vassar.edu/CES/CES1.html#Contents>
230. http://www.eskimo.com/~ram/lexical_semantics.html 2006.
231. <http://www.gpntb.ru/win/book/2/Doc8.HTML>
232. <http://www.keva.ru/stemka/stemka.html>
233. <http://www.keva.ru/stemka/stemka.html>
234. <http://www.libweb.ru/archive/resource/docs/rckk/doc26.html>
235. <http://www.ruthenia.ru/document/192271.html>
236. <http://www.slovorod.ru/russian-wordform.html>
237. http://www.solarix.ru/for_developers/docs/morphology_analyzer.shtml
238. <http://www.solarix.ru/parser.shtml>
239. <http://www.solarix.ru/parser.shtml>
240. <http://www.xces.org/>
- 241. <https://economy-ru.info/info/152021/>**
242. <https://novainfo.ru/article/3394>
243. https://ru.wikipedia.org/wiki/%D0%9B%D0%B8%D0%BD%D0%B3%D0%B2%D0%B8%D1%81%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%BE%D0%B5_%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%BC%D0%BD%D0%BE%D0%B5_%D0%BE%D0%B1%D0%B5%D1%81%D0%BF%D0%B5%D1%87%D0%B5%D0%BD%D0%B8%D0%B5
244. https://sociolinguistics.academic.ru/915/%D0%9B%D0%B8%D0%BD%D0%B3%D0%B2%D0%B8%D1%81%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%BE%D0%B5_%D0%BE%D0%B1%D0%B5%D1%81%D0%BF%D0%B5%D1%87%D0%B5%D0%BD%D0%B8%D0%B5
245. https://studopedia.ru/6_11435_ponyatie-i-sostav-lingvisticheskogo-obespecheniya.html
246. <https://works.doklad.ru/view/IP7CrszcUzY.html>
247. <https://www.multitran.com/c/m.exe?a=4&MessNum=8673&topic=2&l1=4&l2=2>

248. <https://www.shouldiremoveit.com/wordtabulator-for-windows-5636-program.aspx>
249. <https://yandex.ru/dev/mystem/>
250. <tp://www.gpntb.ru/win/book/2/Doc8.HTML>
251. www.ispu.ru/library/lessons/koposov2/14.html
252. www.meta-ukraine.com
253. www.natcorp.ox.ac.uk/archive/vault/tgcw30.pdf
254. www.rco.ru/product.asp

ИЛОВА

(1- чизма)

Associated Press Corpus (горизонтальный формат в кодировке университета
Ланкастер) корпуси разметка намунаси

<i>The door, which was equipped with neither bell nor knocker, was blistered and distained.</i>
[N The_AT door_NN1 ,_, [Fr [Nwhich_DDQ N] [V was _VBDZ
equipped_VVN [P with_IW [N neither_LE [bell_NN1 nor_CC
knocker_NN1] N] P] V] Fr] N] ,_,
[V was_VBDZ [blistered_VVN and_CC distained_VVN] V] ._.

(2- чизма)

Associated Press Corpus корпусининг вертикал морфосинтактик кодировка
намунаси

The	AT	[N
door	NN1	
,	,	
which	DDQ	[Fr[N]
was	VBDZ	[V
equipped	VVN	
with	IW	[P
neither	LE	[N
bell	NN1	[
nor	CC	
knocker	NN1]N]P]V]Fr]N]
,	,	
was	VBDZ	[V
blistered	VVN	
and	CC	
distained	VVN	V]

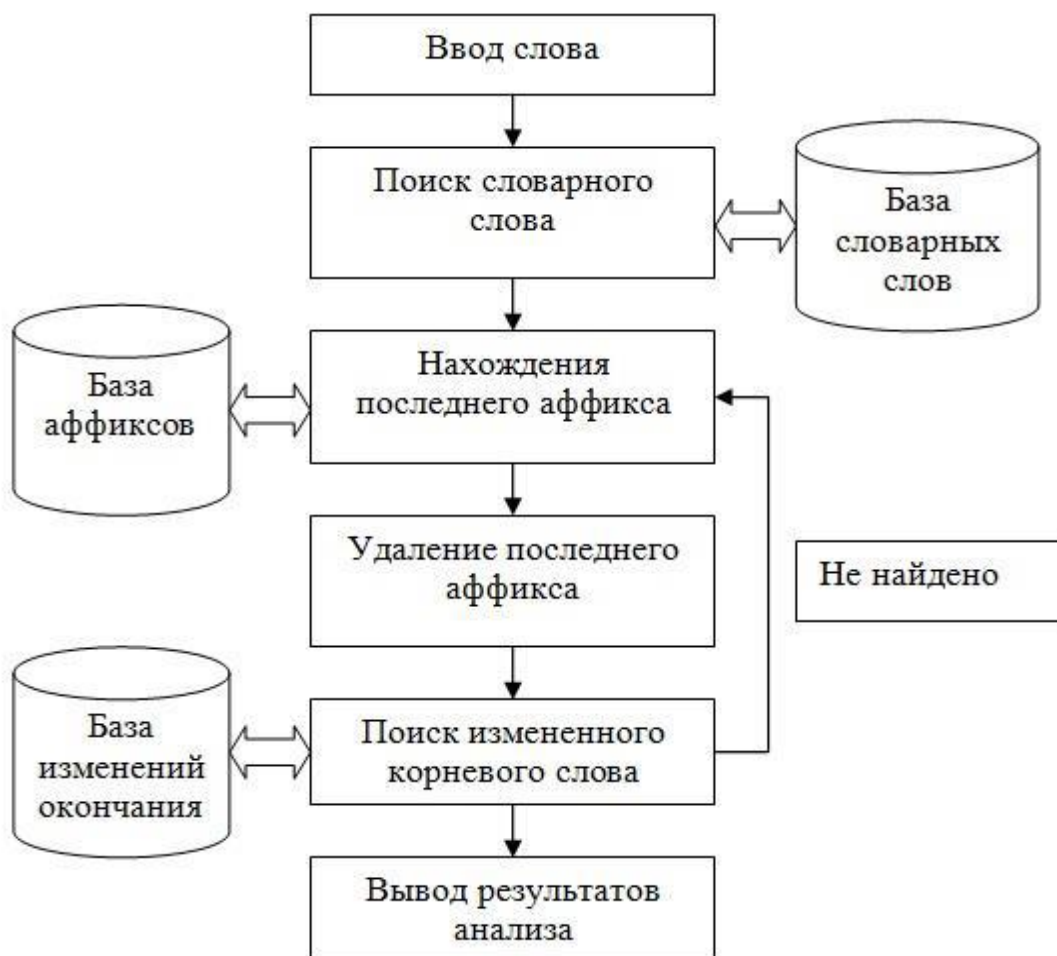
(3- чизма)

Рус матнлари корпуси разметка намунаси (морфологик кодировкада разметка тили)

```
<?xml version="1.0" encoding="windows-1251" ?><text><p>
<s><w>Звонили<ana lemma="ЗВОНИТЬ" pos="Г" gram="мн,нс,нп,дст,прш," /></w><w>к<ana
lemma="К" pos="ПРЕДЛ" gram="" /></w><w>вечерне<ana lemma="ВЕЧЕРНЯ" pos="С"
gram="жр,ед,дт,пр,но," /><ana lemma="ВЕЧЕРНИЙ" pos="П" gram="ср,ед,кр,"
/></w><pun>.</pun></s>

<s><w>Торжественный<ana lemma="ТОРЖЕСТВЕННЫЙ" pos="П" gram="мр,ед,им,вн,"
/></w><w>гул<ana lemma="ГУЛ" pos="С" gram="мр,ед,им,вн,но," /></w><w>колоколов<ana
lemma="КОЛОКОЛ" pos="С" gram="мр,мн,рд,но," /><ana lemma="КОЛОКОЛОВ" pos="С"
gram="мр,фам,ед,им,од," /></w>.....<pun>.</pun></s></p></text> .....
```

Морфологик анализаторнинг таркибий қисми ва тузилиши



1-расм

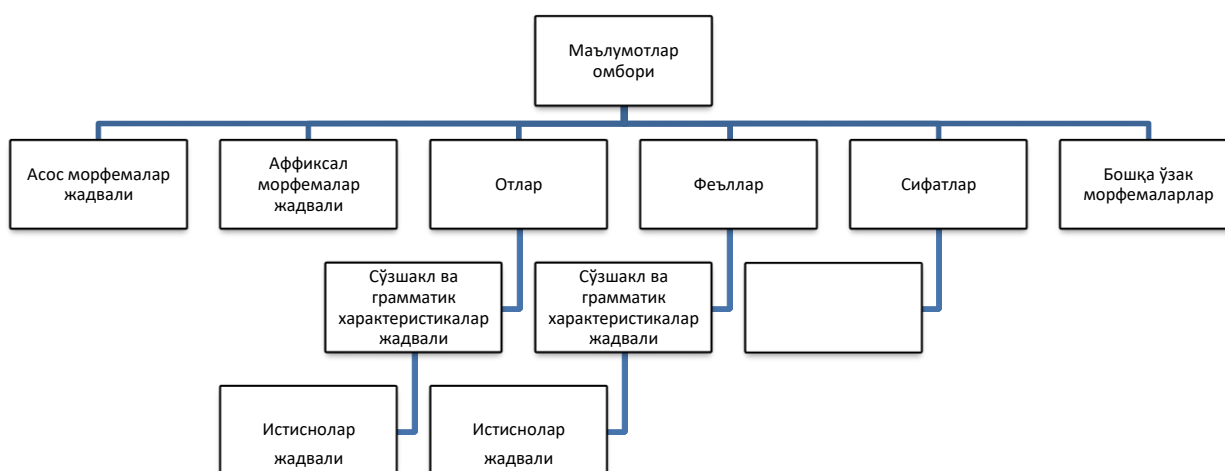
```

recognition ПервыйИм1 language=Russian
{if ".+ЫЙ" thenПРИЛАГАТЕЛЬНОЕ:"???" {СТЕПЕНЬ:АТРИБ
~КРАТКИЙ ЧИСЛО:ЕД ПАДЕЖ:ИМ РОД:МУЖ }
  
```

Ўзбек тили морфоанализаторининг маълумотлар омбори таркибий қисмлари



Морфологик маълумотлар омбори тузилиши



(2-расм)

Портал "Тюркская Морфема"

Вы вошли как: Әлфия Галиева

Вики Форум Профиль RU

Выбранный язык базы данных: Татарский

Общая часть

Грамматические категории

Грамматические значения

Граммемы

Квазиграммемы

Дериватемы

Морфотактика: Для аффиксальных морфем

Левая аффиксальная морфема:

-Лаp : Personality 3-st person plural : Персональность 3-е лицо множественное число

Связанные правые аффиксальные морфемы:

Выберите аффиксальную морфему

Не связанные правые аффиксальные морфемы:

Выберите аффиксальную морфему

(3-расм)

Портал "Тюркская Морфема"

Вы вошли как: Әлфия Галиева

Выбранный язык базы данных: Татарский

Общая часть

Грамматические категории

-сыз : Abessive : Лишительный падеж

-Дан : Ablative : Исходный падеж

-н[ы] : Accusative : Винительный падеж

-[Г]А : Dative/Directive : Дательно-направительный падеж

-ча : Equative : Экватив

-ның : Genitive : Притяжательный падеж

-Дыр : Indefinite : Неопределенный

-мы : Interrogative : Вопросительный

-мыни : Interrogative-Mirative : Вопросительно-удивительная частица

-ДА : Locative : Местно-временной падеж

-быз : Personality 1-st person plural : Персональность 1-е лицо множественное число

-сыз : Personality 2-st person plural : Персональность 2-е лицо множественное число

-[ы]быз : Possessive 1-st person plural : Притяжательность 1-го лица множественного числа

-[ы]м : Possessive 1-st person singular : Притяжательность 1-го лица единственного числа

-[ы]гыз : Possessive 2-st person plural : Притяжательность 2-го лица множественного числа

-[ы]ң : Possessive 2-st person singular : Притяжательность 2-го лица единственного числа

-[с]ы[н] : Possessive 3-st person : Притяжательность 3-го лица

-ДАЙ : Similative : Уподобительно-сравнительный падеж

(4-расм)

Вы вошли как: Элфия Галиева

Вики Форум Профиль RU

Выбранный язык базы данных: Татарский

Общая часть

Грамматические категории

Грамматические значения

Графемы

Квазиграммы

Дериваты

Языковая часть

Морфемы

Морфотактика: Для аффиксальных морфем

Левая аффиксальная морфема:
-ЛАр : Plural : Множественное число

Связанные правые аффиксальные морфемы:
Выберите аффиксальную морфему

Не связанные правые аффиксальные морфемы:
-кА[й] : Hipocoristic : Ласкательный аффикс

1. (гы)-лар-(га)	<input type="checkbox"/> 1. кай	<input type="checkbox"/> 2. кай	<input type="checkbox"/> 3. ка	<input type="checkbox"/> 4. ка	✖
2. (ге)-лар-(га)	<input type="checkbox"/> 1. кай	<input type="checkbox"/> 2. кай	<input type="checkbox"/> 3. ка	<input type="checkbox"/> 4. ка	✖
3. (сыман)-лар-(га)	<input type="checkbox"/> 1. кай	<input type="checkbox"/> 2. кай	<input type="checkbox"/> 3. ка	<input type="checkbox"/> 4. ка	✖
4. (чэн)-лар-(га)	<input type="checkbox"/> 1. кай	<input type="checkbox"/> 2. кай	<input type="checkbox"/> 3. ка	<input type="checkbox"/> 4. ка	✖

7-чизма

а-а	Interj	{а-а}
абажур	N	{абажу!р}
абажуран	A	{абажура!н}
аббревиатур	N	{аббревиату!р}
абзац	N	{абза!ц}
абиссин	N	{абисси!н}
абитуриент	N	{абитурие!нт}
абонемент	N	{абонеме!нт}
абонент	N	{абоне!нт}
aborиген	N	{aborиге!н}
абрикос	N	{абрико!с}
абсолютизм	N	{абсолюти!зм}
абстрактнын	Adv	{абстра!ктнын}
абстракций	N	{абстра!кций}
абстракт	N	{абстра!кт}

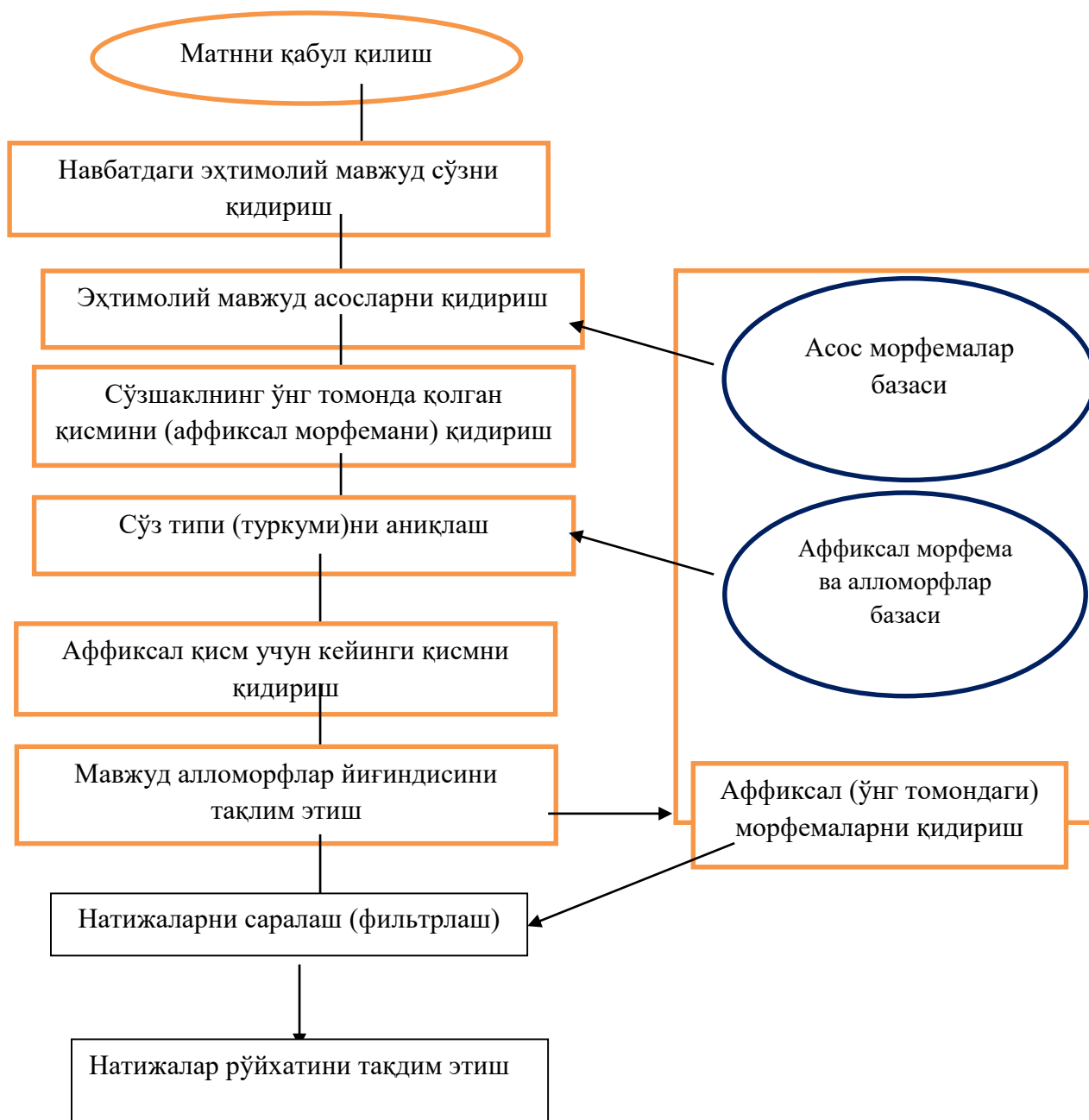
8-чизма

Слово	Часть речи	Разметка слова	Опорная морфема	Словообразовательный аффикс	Неодушевл./ Одушевл./ Состояние/ Процесс
а-а	Interj	{а-а}	а	-	Signal
абажур	N	{аба жу!р}	абажур	-	неодушевл.
абажуран	A	{аба жу ра!н}	абажур	ан	ATTR
аббревиатур	N	{аб бре ви ат у!р}	аббревиатур	-	неодушевл.
абзац	N	{аб за!ц}	абзац	-	неодушевл.
абиссин	N	{аб ис си!н}	абиссин	-	одушевл.
абитурант	N	{аби ту ри е!нт}	абитурант	-	одушевл.
абонемент	N	{або не ме!нт}	абонемент	-	неодушевл.
абонент	N	{або не!нт}	абонент	-	неодушевл.
абориген	N	{або ри ге!н}	абориген	-	одушевл.
абрикос	N	{аб ри ко!с}	абрикос	-	неодушевл.
абсолютизм	N	{аб со лю ти!зм}	абсолютизм	-	неодушевл.
абстрактнын	Adv	{аб стра!кт нын}	абстрактн	ын	состояние
абстракций	N	{аб стра!к ций}	абстракций	-	неодушевл.
абсурд	N	{аб су!рд}	абсурд	-	неодушевл.
абхаз	N	{аб ха!з}	абхаз	-	одушевл.

S	substantive	Существительное
V	verb	Глагол
ADJ	adjective	Адъектив
NUM	numeral	Числительное
SPRO	pronoun	Местоимение
PART	particle	Частица
POST	postposition	Послелог
CONJ	conjunction	Союз
INTJ	interjunction	Междометие
PL	plural	множественное число
SG	singular	единственное число
1,2,3SG	first (second, third) person, singular	первое (второе, третье) лицо единственного числа
1,2,3PL	first (second, third) person, plural	первое (второе, третье) лицо множественного числа
UNCERT	uncertainty	клитика со значением неуверенности
UNDEF	undefiniteness	клитика со значением предположительности
INTERROG	interrogative	клитика со значением вопросительности
REQUEST	request	клитика со значением просьбы
PRED	predicative	Сказуемость
POSS	possessive	Принадлежность
IPOSS	impersonal possessive	неличная принадлежность
NOM	nominative case	основной падеж
GEN	genitive case	родительный падеж
DAT	dative case	дательный падеж
ACC	accusative case	винительный падеж
ABL	ablative case	исходный падеж
LOC	locative case	местно-временной падеж
ABE	abessive case	Абессив
DERIV.ABSTR	abstract noun	абстрактное существительное
DERIV.AGEN	agens noun	со значением деятеля
S		
ASSIM	assimilation	Уподобление
DIST	distinctive	выделительное (числительное)
COL	collective	Собирательное
ORD	ordinary	Порядковое
APRX	approximative	Приблизительное
DIVIS	divisive	Разделительное
COMP	comparative	сравнительная степень
ACYCL	acyclic	непериодическое действие
GER	gerund	Деепричастие
IMP	imperative mood	повелительное наклонение
REFL	reflexive genus	возвратный залог
PASS	passive genus	страдательный залог

RECP	reciprocal genus	взаимный залог
CAUS	causative genus	понудительный залог
NEG	negation	Отрицание
PRES	present tense	настоящее время
PST.DEF	past definite tense	прошедшее определённое
PST.INDF	past indefinite tense	прошедшее неопределённое
FUT.INDF	future indefinite tense	будущее неопределённое
FUT.DEF	future definite tense	будущее определённое
DESI	desiderative mood	наклонение намерения
COND	conditional mood	условное наклонение
OPT	optative mood	желательное наклонение
INF	infinitive	Инфинитив
SUP	supin	имя действия

Ўзбек тили морфоанализаторининг иш алгоритми



Ўзбек тилидаги шахсий муносабат шакллари

Шахсий муносабат шакли	Кўрсаткичи	Мисол
Кичрайтириш шакли	-ча -чоқ -чак	қизча қўзичоқ келинчак
Эркалаш шакли	гина (-кина, -қина) -лоқ -жон -хон -бой -бек -қул -той -ой -бека -биби -бону -нисо -пари -гул -пошиша	қизгина қизалоқ Баҳодиржон Раънохон Азимбой Суръатбек Асрорқул Омонтой Чўлпоной Раҳимбека Ойшабиби Хуришдабону Лутфинисо Дилнозапари Санамгул Каримапошиша
Хурмат шакли	-лар	бобомлар

- Грамматик шаклларни автоматик фарқлашнинг лингвистик моделлари иловаси
1. $V+f_1^v$ ёки $V+f_1^v+f_2^v$ позициясида $-(и)ш$ шакли ҳаракат номи вазифасини бажаради: *чиқ+ар+иш; том+из+иш, кий+гиз+иш, кий+дир+иш, тасдиқла+т+иш, суя+н+иш, чўз+ил+иш;*
 2. $V+f_1^v+p_1^v$ ёки $V+f_1^v+f_2^v+p_1^v$ позициясида $-(и)ш$ шакли биргалик нисбат вазифасини бажаради: *кел+иш+ди, суриштир+иш+ди, ўтир+иш+ди, кел+тир+иш+ди.*
 3. агар {феъл-асос+ f_1^v +вер+ $V+f_2^v+p_1^v$ } бўлса, -вер = кўмакчи феъл.
 4. $V+f_1^v(и)б$, яъни {феъл-асос+(и)б+0} бўлса = феълнинг равишдош шакли; агар $V_1+f_1^v(и)б+V_2+f_1^v+f_1^v/p_1^v$, яъни {феъл-асос+(и)б+ феъл-асос+ЛШЯ/СШЯ} = етакчи ва кўмакчи феълни боғловчи равишдош шакли;
 5. *ўтирибман, бўлибди, борибсан, узоқлашибмиз* сўзларида ўтган замон аффикси вазифасини бажаради. Позициясини қуйидагича моделлаштирамиз: агар $V+f_1^v(и)б+p_1^v$, яъни {феъл-асос+(и)б+СШЯ} бўлса = ўтган замон шакли.
 6. агар $V+f_1^v$ (гани) + 0, яъни {феъл-асос+гани+0} бўлса, f_1^v = мақсад равишдоши шакли;
 7. агар $V+f_1^v(ган+и)+p_1^v+0$, яъни {феъл-асос+(ган+и)} бўлса, f_1^v = (ган+и) эгалик шаклидаги сифатдош;
 8. агар $V+f_1^v(гани)+йўқ$ ёки, яъни {феъл-асос+(ган+и)+йўқ} бўлса, f_1^v = эгалик шаклидаги сифатдош;
 9. агар $V+f_1^v(гани)+ p_1^n(-га/ -да/ -дан)$, яъни {феъл-асос+(ган+и)+ (-га/ -да/ -дан)} бўлса, f_1^v = эгалик шаклидаги сифатдош.
 10. $V+f_1^v+f_2^v(дир)$, яъни {феъл-асос+(ЛШЯ)+дир} бўлса, f_2^v = гумон юкламаси;
 11. $N/Adj/P/Num+p_1^n(дир)$, яъни {от/сифат/сон/олмош-асос+дир} бўлса, p_1^n = кесимлик кўрсаткичи/боғлама.
 12. $V+f_1^{v(voice)}+f_2^{v(voice)}+f_2^{v(voice)}$;
 13. нисбат шаклидан кейин бўлишсизлик шакли *-ма* жойлашади: $V+f_1^v+f_2^{v(neg)}$;

14. бўлишсизлик шаклидан кейин феънинг ўзгаловчи категорияси шакллари (*бормаган, бормай, бормасдан, бормаслик*) ҳамда майл шакли (*бормасин, бормайлик, бормай*) жойлашади: $V+f_1^v+f_2^{v(neg)}+f_3^{v(fin)}$ ёки $f_1^{v(mood)}$;
15. кейинги позицияда замон, шахс-сон шакли (*бормасин, бормайлик, бормади*) жойлашади: $V+f_1^v+f_2^{v(neg)}+p_1^v+p_2^v$.
16. агар $\{V+f_1$ (*-ла, -чи, -а, -и, -иш, -га, -ка, -қа*) $\}$ бўлса, (яъни ушбу шакллар феъл-асосдан кейин қўшилса) f_1 = ҳаракатнинг кучли/кучсизлигини билдирувчи ЛШЯ.
17. агар $\{N/A /Num /Adv/ Sim +d_1$ (*-ла, -чи, -а*) $\}$ бўлса, (яъни ушбу шакллар феъл бўлмаган асосдан кейин қўшилса) d_1 = сўз ясовчи қўшимча (феъл-асосдан кейин қўшилганда ЛШЯ эканлигини юқорида таъкидланди);
18. агар $\{N/A /Num /Adv/ Sim /сифатдош, равишдош, ҳаракат номи +r_1$ -*га, -ка, -қа* $\}$ бўлса, (яъни ушбу шакллар феъл бўлмаган асосдан кейин қўшилса) r_1 = синтактик шакл ясовчи (феъл-асосдан кейин қўшилганда ЛШЯ эканлигини юқорида таъкидланди).
19. агар $\{N+гина+кўплик+эгалик\}$ бўлса, *-гина* = кичрайтириш-эркалаш ш.;
20. агар $\{N+кўплик+эгалик+гина\}$ бўлса, *-гина* = айирув-чегаралов юкламаси.
21. f = -лар бўлганда: (1) агар $\{R+f+r\}$ бўлса, f = кўплик; (2) агар $\{R+r+f\}$ бўлса, f = ҳурмат маъноси.
22. $V+f_1^v+r_1^f$. Бунда: $f_1^v = \{-ган /-(и)б / -яп / -а / -й /-моқчи\}$; $r_1^f = \{-ман /-сан / -0 / -миз / -сиз/ -(ди)лар\}$;
23. $V+f_2^v+r_2^f$. Бунда: $f_2^v = \{-ди /-са\}$; $r_2^f = \{-м /-нг / -0 / -к / -нгиз/ -(ди)лар\}$;
24. $V+ r_3^f$. Бунда: $r_3^f = \{-ай/-0 / -син / -айлик / -гин/ -(и)нгиз / -(син)лар / +-ишсин\}$.
25. агар $N / A / Num / Pr + r_1^n$, яъни $\{исм-асос+-(и)м / -инг / -(и)миз\}$ бўлса, r_1^n = эгалик шакли;
26. агар $V + r^f$, яъни $\{феъл-асос+-(и)м / -инг / -(и)миз\}$ бўлса, r^f = шахс-сон кўрсаткичи.
27. $\{V\Pr + f_1/r_1 + ми+ кин = сўроқ юкламаси + экан тўлиқсиз феъли\}$;

28. агар {от-асос + ЛШЯ/СШЯ + ми} бўлса, *-ми* юкламаси – модаллик кўрсаткичи: {N + f₁/r₁ + ми} бўлса, {-ми = M};
29. агар {феъл-асос + ЛШЯ/СШЯ + ми} бўлса, *-ми* юкламаси – синтактик модаллик кўрсаткичи: {N + f₁/r₁ + ми} бўлса, {-ми = SM}.
30. шахс оти ясалиши R_n+d₁ⁿ моделига эга бўлади, бунда d₁ⁿ = *-чи, -шунос, -кор, -каш, -хон, -гар, -соз, -паз, -боз, -хўр, -параст, -зўй, -дор, -бон, -фуруш, -дош, -вачча, -ат, -лик, -парвар, -тарош, -манд, -и*;
31. ўрин-жой оти ясалиши R_n+d₂ⁿ моделига эга бўлади, бунда d₂ⁿ = *-зор, -лоқ, -истон, -гоҳ, -хона, -ов, -кент, -обод*;
32. предмет маъносини билдирувчи от ясалиши R_n+d₃ⁿ моделига эга бўлади, бунда d₃ⁿ = *-дон, -нома, -лик*;
33. касб-кор ва фаолиятни билдирувчи от ясалиши R_n+d₄ⁿ моделига эга бўлади, бунда d₄ⁿ = *-лик, -чилик, -гарчилик, -гарлик*.
34. асосдан англашилган предметга эгалик маъноси R_n+d₁^a моделига эга бўлади, бунда d₁^a = *-ли, -дор, -манд, -кор, -гай, -акам, -вор, -хон; масалан, кунгай, ҳазилакам, доривор, саводхон*;
35. хосликни билдирувчи сифат маъноси R_n+d₂^a моделига эга бўлади, бунда d₂^a = *-йи, -ки, -вий, -й, -шумул, -ик, -лик; масалан, ҳавойи, филологик, йиллик*;
36. ўзак билдирган тушунчага ўхшашлик маъносини билдирувчи сифат структураси R_n+d₃^a моделига эга бўлади, бунда d₃^a = *-намо, -симон*;
37. ўзак билдирган тушунчанинг кўплиги маъносини билдирувчи сифат структураси R_n+d₄^a моделига эга бўлади, бунда d₄^a = *-чан, -дон, -каш, -овар (жанговар)*;
38. асосдан англашилган предметга эга эмаслик маъноси R_n+d₅^a моделига эга бўлади, бунда d₅^a = *-сиз*.
39. ҳаракат тарзи ва ҳолатни билдирувчи равиш ясалиш структураси R_n+d₁^{adv} моделига эга бўлади, бунда d₁^{adv} = *-ча, -часига, -ларча, -она, -сиз, -ан, -чан, -акам (ёлгондакам), -лай, -сига*;

40. пайт маъносини билдирувчи равиш ясашиш структураси $R_n+d_2^{adv}$ моделига эга бўлади, бунда $d_2^{adv} =$ -лаб, -ин: *ойлаб, қишин-ёзин*.
41. аниқ отдан феъл ясашиш структураси $R_n+d_1^v$ моделига эга бўлади, бунда $d_1^v =$ -лаш, -а, -қар, -и, -ик, -ки, -ир, -лаштир, -са;
42. мавҳум отдан феъл ясашиш структураси $R_n+d_2^v$ моделига эга бўлади, бунда $d_2^v =$ -ла, -ай, -лан, -ан, -сит.
43. шахс оти ясашиш структураси $R_a+d_1^n$ моделига эга бўлади, бунда $d_1^n =$ -чи, -хўр: *аълочи, порахўр*;
44. предмет маъносини билдирувчи оти ясашиш структураси $R_a+d_2^n$ моделига эга бўлади, бунда $d_2^n =$ -ат (*кўкат*);
45. фаолият-жараён маъносини билдирувчи от ясашиш структураси $R_a+d_3^n$ моделига эга бўлади, бунда $d_3^n =$ -лик, -чилик.
46. белги-хусусият билдирувчи равиш структураси $R_a+d_1^{adv}$ моделига эга бўлади, бунда $d_1^{adv} =$ -ларча, -она, -часига;
47. ҳолат билдирувчи равиш структураси $R_a+d_2^{adv}$ моделига эга бўлади, бунда $d_2^{adv} =$ -лаб, -лайн.
48. ранг билдирувчи асосдан феъл ясашиш структураси $R_a+d_1^v$ моделига эга бўлади, бунда $d_1^v =$ -й, -ар: *қорай, кўкар*;
49. предмет белгисини билдирувчи асосдан феъл ясашиш структураси $R_a+d_2^v$ моделига эга бўлади, бунда $d_2^v =$ -а, -ла: *бўшмоқ, текисламоқ*;
50. шахсга тегишли белги-хусусият билдирувчи асосдан феъл ясашиш структураси $R_a+d_3^v$ моделига эга бўлади, бунда $d_3^v =$ -си: *гарангсимоқ*;
51. умуман белги-хусусият билдирувчи асосдан феъл ясашиш структураси $R_a+d_4^v$ моделига эга бўлади, бунда $d_4^v =$ -лаш, -син, -ик: *қийинлашмоқ, оғриқсинмоқ, зўриқмоқ*.
52. сон-асосдан шахс ва мавҳум от ясашиш структураси $R_{num}+d^n$ моделига эга бўлади, бунда $d^n =$ -чи, -лик: *иккичи, бирлик*;
53. сон-асосдан сифат ясашиш структураси $R_{num}+d^a$ моделига эга бўлади, бунда $d^a =$ -дай (*бирдай*); -лик (*бирлик сон*);

54. сон-асосдан равиш ясаиш структураси $R_{num}+d^{adv}$ моделига эга бўлади, бунда $d^{adv} =$ -лаб, -ларча, -ин: *битталаб, мингларча, бирин*;
55. сон-асосдан феъл ясаиш структураси $R_{num}+d^v$ моделига эга бўлади, бунда $d^v =$ -лин, -лаш, -ик: *иккиланмоқ, бирлашмоқ, бирикмоқ*.
56. олмош асосдан равиш ясаиш структураси $R_{pr}+d^{adv}$ моделига эга бўлади, бунда $d^{adv} =$ -сиз, -дек, -дай, -ча, -чаки, -чалик: *усиз, бизсиз, бундай, биздай, бунча, шунчаки, бизчалик*;
57. олмошдан феъл ясаиш структураси $R_{pr}+d^v$ моделига эга бўлади, бунда $d^v =$ -ла, -сира: *сенла, сенсира*.
58. миқдор, белги, хусусият билдирувчи равиш асосдан ҳаракат белгиси ва миқдори билдирувчи от ясаиш структураси $R_{adv}+d^n$ моделига эга бўлади, бунда $d^n =$ -лик, -чилик: *тезлик, озчилик, кўпчилик*;
59. равиш асосдан ўрин-пайт белгисини билдирувчи сифат ясаиш структураси $R_{adv}+d^a$ моделига эга бўлади, бунда $d^a =$ -ги, -ки. d^a нинг воқеланиши: *пастки, дастлабки, илгариги*;
60. миқдор билдирувчи равиш асосдан англашилган маънонинг бор/йўқлигини феъл ясаиш структураси $R_{adv}+d_1^v$ моделига эга бўлади, бунда $d_1^v =$ -ай; d_1^v нинг воқеланиши: *кўпаймоқ, озаймоқ*;
61. ҳолат билдирувчи равишдан феъл ясаиш структураси $R_{adv}+d_2^v$ моделига эга бўлади, бунда $d_2^v =$ -ит, -ла, -лаш; d_2^v нинг воқеланиши: *беркитмоқ, пастламоқ, тезлашмоқ*;
62. шакл-ҳажм билдирувчи равишдан феъл ясаиш структураси $R_{adv}+d_3^v$ моделига эга бўлади, бунда $d_3^v =$ -ар; d_3^v нинг воқеланиши: *қисқармоқ*;
63. пайт билдирувчи равишдан феъл ясаиш структураси $R_{adv}+d_4^v$ моделига эга бўлади, бунда $d_4^v =$ -ик; d_4^v нинг воқеланиши: *кечикмоқ*.
64. тақлид-асосдан аниқ от ясаиш структураси $R_{sim}+d_1^n$ моделига эга бўлади, бунда $d_1^n =$ -а, -ак, -илдоқ; d_1^n нинг воқеланиши: *жизза, гижжак, бизилдоқ, шипилдоқ*;

65. тақлид-асосдан ҳаракат-ҳолат жараёнини атовчи от ясаиш структураси $R_{sim}+d_2^n$ моделига эга бўлади, бунда $d_2^n = -ак, -ос, -ға$; d_2^n нинг воқеланиши: *қарсақ (пахта), гиргаирак (юла), шаррос (узум тури), гуввос (шаррос қуядиган ёмгир), говга (тўполон)*;

66. тақлид-асосдан ўрин-жой оти ясаиш структураси $R_{sim}+d_3^n$ моделига эга бўлади, бунда $d_3^n = -а$; d_3^n нинг воқеланиши: *шаршара*.

67. феъл-асосдан шахс оти ясаиш структураси $R_v+d_1^n$ моделига эга бўлади, бунда $d_1^n = -ар, -оқ, -увчи$; d_1^n нинг воқеланиши: *қочоқ, чопар, учувчи*;

68. феъл-асосдан нарса оти ясаиш структураси $R_v+d_2^n$ моделига эга бўлади, бунда $d_2^n = -ги, -им, -ма, -ин, -инди, -қич, -гич, -моқ, -имта, -кон, -моч, -чок, -ат, -к, -оқ, -чи, чик, -чак, -из$; d_2^n нинг воқеланиши: *супурги, кийим, сузма, тугун, қиринди, ювинди, қисқич, сузгич, илмоқ, қўймоқ, ўсимта, босқон, қирмоч, овунчоқ, кекирдак, қовурдоқ, ўроқ, суюнчи, ёпинчиқ, кемирчак, бўғиз, буламиқ*;

69. феъл-асосдан фаолият-жараён оти ясаиш структураси $R_v+d_3^n$ моделига эга бўлади, бунда $d_3^n = -иқ, -ки, -ги, -им, -кин, -гин, -ин, -мачоқ, -миш, -в, -он, -килик$; d_3^n *огриқ, кулги, тепки, чидам, билим, тўлқин, ёнгин, бекинмачоқ, кечмиш, қилмиш, тергов, қўзғолон, тўзон, ичкилик* каби сўзларда воқеланади;

70. феъл-асосдан фаолият-жараён натижасини атовчи от ясаиш структураси $R_v+d_4^n$ моделига эга бўлади, бунда $d_4^n = -ик, -инди, -ч, -дик, -он, -итки, -гин, -ғоқ, -гу, -илон, -имча, -ит, -инг$; d_4^n *кўпик, ўлик, ювинди, севинч, ишонч, қолдиқ, қўзғолон, ачитқи, ёнгин, тўлғоқ, урғу, қўшимча, тушунча, чиқит, йиринг* каби сўзларда воқеланади;

71. феъл-асосдан ўрин-жой оти ясаиш структураси $R_v+d_5^n$ моделига эга бўлади, бунда $d_5^n = -ув, -оқ, -қоқ$; d_5^n *кечув, ётоқ, ботқоқ* каби сўзларда воқеланади.

72. феъл-асосга тўғридан-тўғри қўшилиб, сифат ясаиш структураси $R_v+d_1^a$ моделига эга бўлади, бунда $d_1^a = -қоқ, -гир, -ағон, -қи, -к, -гин, -кин, -ма, -арли, -оқ, -ри, -чак, -чан, -ма, -ч, -мсиқ, -м, -им, -қа, -а, -аси, -гич, -дан, -ин, -акса, -чиқ, -ғоқ$; d_1^a шакли *қочқоқ, урушқоқ, сезгир, ўтқир, билагон, сайроқи*,

вайсақи, чирик, эгик, ётиқ, озгин, сотқин, кўчма, ачинарли, қолоқ, эгри, эринчоқ, курашчан, ўтмас, жирканч, ачимсиқ, қарам, айрим, қисқа, тўла, келаси, билгич, югурдак, тўлин, ўлакса, сирпанчиқ, тойгоқ каби сўзларда воқеланади;

73. феъл-асосга грамматик шаклдан кейин қўшилган сифат ясалиш структураси $R_v+d_2^a$ моделига эга бўлади, бунда $d_2^a =$ -чоқ, -чан, -гич; d_2^a *мақтанчоқ, яшовчан, қопонгич* каби сўзларда юзага чиқади.

74. агар {феъл-асос+мас+лик} бўлса, {-маслик} = ҳаракат номининг бўлишсизлик шакли: $\{R_{adv} + f_1 + \text{лик}\}$ да, -лик = f_2 , яъни граммема;

75. агар {от/ сифат/ сон +лик} бўлса, {-лик} = сўз ясовчи: $\{R_n// R_a// R_{num}// R_{adv} + \text{лик}\}$ да -лик = d^n , яъни дериватема;

76. $R_n+d_1^n+d_2^n$ да от-асосга ясовчи шахс оти ва ўрин-жой отини ясовчи қўшилади: *элчихона*;

77. $R_n+d_1^n+d_2^n$ да от-асосга иккита ўрин-жой оти ясовчиси қўшилади: *ўтлоқзор* (бу қўлланиш адабий тилга хос бўлмаса-да, оғзаки нутқда учраб туради);

78. $R_n+d_1^n+d_2^n$ да от-асосга ясовчи ўрин-жой ва шахс оти ясовчилари қўшилади: *чойхоначи*.