

## МОРФОЛОГИК МАЪЛУМОТЛАР БАЗАСИНИ ЯРАТИШ ХУСУСИДА

**Хамроева Шахло Мирджоновна,**  
филология фанлари бўйича фалсафа доктори,  
Алишер Навоий номидаги Тошкент давлат  
ўзбек тили ва адабиёт университети  
[hamroyeva81@mail.ru](mailto:hamroyeva81@mail.ru)

**Аннотация.** Мақолада морфологик маълумотлар базасининг табиий тилга автоматик ишлов беришдаги ўрни ҳамда морфологик базани яратиш тамойиллари хусусида фикр билдирилган. Морфологик базанинг морфологик луғат, морфолексикон ҳамда морфологик анализатор ишининг сифатини оширишдаги аҳамияти тавсифланган.

**Калит сўзлар:** морфологик база, морфолексикон, морфологик луғат, грамматик таҳлил, анализ, синтез.

**Аннотация.** В статье рассматривается роль морфологических баз данных в автоматической обработке естественного языка и принципы создания морфологических баз данных. Описана роль морфологической базы в повышении качества работы морфологического словаря, морфолексикона и морфологического анализатора.

**Ключевые слова:** морфологическая база, морфолексикон, морфологический словарь, грамматический анализ, анализ, синтез.

**Resume.** The article discusses the role of morphological databases in automatic processing of natural language and the principles of creating morphological databases. The role of the morphological base in improving the quality of work of the morphological dictionary, morpholexicon and morphological analyzer is described.

**Keywords:** morphological base, morpholexicon, morphological dictionary, grammatical analysis, analysis, synthesis.

**Кириш.** 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 Access маълумотлар базаси билан ишлаш тизимидан фойдаланилган. Мазкур анализатор маълумотлар базасининг уч жадвали (асос, қўшимча ва сўз туркуми) ҳамда ушбу жадваллар ўртасидаги муносабатдан иборат [Цветков, 198: 67].

**Асосий қисм.** *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 [Гатауллин, 2019 : 175].

Мукаммал 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 қўшимчалар 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 керак [Гатауллин, 2019 : 175]:

- 1) сўзлар алифбо талабига мос ёзилган бўлиши (морфология, лексика);
- 2) гап қурилишида грамматик хатога йўл қўймаслик (грамматик, синтактик);

гап мазмуний қурилиши (семантика)нинг тўғрилиги [Хамроева, 2020: 146].

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 ҳолат 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.** Тадқиқотчи К.Шадманова мавзуга оид мақоласида ҳаракат феълларининг семантикаси ҳамда валентлигини текшириш асосида компьютер дастурлари учун лингвистик таъминот яратиш намунасини келтиради. Лингвистик таъмин яратишда тадқиқотчи ҳаракат феълларининг тайёр семантик таснифдан фойдаланади. Обьектив борликдаги нарсалар-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 келади [Кузнецов, 1999: 260]. 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 бўшлиғидаги ҳаракат* каби бирликлар эса **бирлаштирувчи** сема ҳисобланади. Ўзбек тилидаги ҳаракат феълларида намоён бўлувчи мантикий субъект ҳаракати юқоридаги кўринишда содир бўлади. Демак, объектив борлиқдаги турли ҳаракатни акс эттирувчи феъл семаси мураккаб характери мантикий равишда ҳаракат феълларининг қайд этилган кўринишдан иборат семантик гуруҳланишга олиб келди [Муҳамедова, 2006: 58].

Ўзбек тилидаги ҳаракат семантик феъллар асосидаги мазкур маълумотлар базаси компьютер дастури, таржима, таҳрир учун 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 [Муҳамедова, 2006: 62].

**Хулоса** сифатида айтиш мумкинки, табиий тилга автоматик ишлов бериш тизимлари морфологик маълумотлар базасига эҳтиёж сезади. Мукамал ишлов берилган морфологик маълумотлар базаси морфологик луғат, морфолексикон ва морфологик анализатор ишининг сифатини яхшилайдди.

### Фойдаланилган адабиётлар:

1. Кузнецов А.М. О применении метода компонентного анализа в лексике. – М., 1999. – С.260.
2. Цветков Н.В. К методологии компонентного анализа // Вопросы языкознания. – 1984, № 2, – с.67.
3. Муҳамедова С. Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. –Тошкент, 2006.
4. Шадманова К. Табиий тиллар учун лексико-семантик луғат маълумотлар базаси яратиш тамойиллари // <http://buxdu.uz/index.php/uz/>
5. Хамроева Ш. Ўзбек тили морфологик анализаторининг лингвистик таъминоти. Монография. – GlobeEdit, 2020. – 245 б.
6. 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.