

новаться иначе. а по древнему мусульманскому обряду, когда нарекают человека, то мулла трижды выкрикивает его имя, чтобы оповестить об этом весь свет.

Вот и я трижды крикнул: Каракалпак, Каракалпак, Каракалпак! Хотел крикнуть на весь свет, а насколько хватило голоса – не мне судить» [2:22-23].

Упомянутые выше символы «дерево» и «черная шапка», по нашему мнению, можно отнести к так называемым «сквозным» символам. Это явление исследовалось в русской стилистике, в частности, в работах А.Марченко, П.Ф.Юшина, А.И.Захарова, Л.Е.Малаховой. В статье А.Марченко исследуется характер построения есенинских образов: «Образы переходят из стихотворения в стихотворение, постепенно обрстая подробностями. Удачно найденный образ как бы почка: он дает росток, росток ветвится» [5:158].

Мысли о повторяемом образном арсенале в творчестве Есенина есть и в работах П.Ф.Юшина, который отмечает: «На основе традиций народного творчества он создавал собственные символы и клал их в основу собственных параллелизмов. К числу таких символов относятся многие так называемые сквозные образы есенинской лирики, окрашивающие мир его поэзий в особые неповторимые тона» [5:158]. П.Ф.Юшин называет некоторые сквозные образы Есенина: березка, клен, ива, месяц, колокольчик и бубенец, тройка; постоянные цвета: белый, синий, голубой, красный, золотой.

Л.Е.Малахова понятие сквозного образа заменяет понятием стилемы: «Береза со всеми определениями к ней стала стилемой, которая проходит через весь контекст и составляет одну из характерных черт индивидуальной стилистической системы. В этой стилеме кристаллизуются эстетические взгляды и национальный характер автора» [5:158-159].

Выделенные нами в тексте романа Т.Каипбергенова «Қарақалпақ дәстаны» символы «емен»-дерево дуб и «қара қалпақ» - черная шапка считаем целесообразным называть не стилемами, а сквозными символами в силу точности и однозначности этого термина.

В исследуемом материале находим также очень разные символы, такие, как в следующих отрывках: «Жайхун бойының пүткил халқы кек атына минеди» [1:30] – буквально: «Народ всего побережья Жайхуна сидится на коня мести»; аналогично: «Елдің айбат балтасын суғға тасладым десеш! – деди Ырыскул бийдин дауысы қарылдап» [1:57] – буквально: «Елдің айбат балтасы» - «топор гнева народа» или «орудие гнева народа». Подобные примеры ярких, метафоричных символов в каракалпакском языке позволяют сделать предположение о древности нашего народа и его языка.

Они вызывают ассоциации с такими символами в языках индейцев Америки, как «тропа войны», «закурить трубку мира» и т.п.

В заключение приведем отрывок из публицистики Т.Каипбергенова: «Мой народ древен. Свидетельство тому, старинные летописи и хроники многих царств и народов. По-разному называли в этих хрониках каракалпаков: «черные клобуки» или «кула-хтсиях», «колпак», «кара боркли». Словосочетания разные, но смысл один: и «шапки скорби», и «черные клобуки», и «траурные шапки» символизировали печальную «черную» долю каракалпаков-этих вечных изгоев.

История других народов запомнила тех, кто называл себя «черные шапки», но сами каракалпаки почти не помнили собственной истории. Говорю «почти», потому что прошлое сохранилось лишь в устных преданиях, легендарных образах храбрых батыров, коварных злодеев, страшных чудищ» [3:48].

Таким образом, символы в романе Т.Каипбергенова, как и все его богатое наследие, носят глубоко национальный характер.

### Литература

1. Каипбергенов Т. Қарақалпақ дәстаны. I. Маман бий эпсанасы. – Нөкис: «Қарақалпақстан», 1973.
2. Каипбергенов Т. Интервью, данное самому себе. – Из юрты в мир. – Нукус: «Қарақалпақстан», 1987.
3. Каипбергенов Т. Измерение времени. – Из юрты в мир. – Нукус: «Қарақалпақстан», 1987.
4. Новиков Л.А. Искусство слова. – М.: «Наука». 1991.
5. Некрасова Е.А., Бакина М.А. Языковые процессы в современной русской поэзии. – М.: «Наука», 1982.

### РЕЗЮМЕ

Ушбу мақолада символнинг Т.Каипбергеновнинг "Маман бий афсонаси" романи матнида рамзининг образли нутқ воситаси сифатида қўлланилиши ўрганилган.

### РЕЗЮМЕ

В этой статье рассматривается использование символа как образного средства художественной речи в тексте романа Т.Каипбергенова "Маман бий эпсанасы".

### SUMMARY

This article examines the use of symbol as a figurative means of artistic speech in the text of the novel "Maman biy epic" by T. Kaipbergenov.

### ЛИНГВИСТИК ДАСТУРНИНГ МОРФОЛОГИК ТАҲЛИЛ МОДУЛИ (МОРФОАНАЛИЗАТОР)

**М.А.Абжалова** – филология фанлари бўйича фалсафа доктори

Тошкент давлат ўзбек тили ва адабиёти университети

**О.И.Искандаров** – мустақил изланувчи

Тошкент давлат политехника университети

**Таянч сўзлар:** морфологик таҳлил, сўзшакл, морфоанализ, морфоанализатор, грамматик категория, табиий тилни қайта ишлаш.

**Ключевые слова:** морфологический анализ, словообразование, морфоанализ, морфоанализатор, грамматическая категория, автоматическая обработка естественного языка.

**Key words:** morphological analysis, word formation, morphoanalysis, morphoanalyzer, grammatical category, Natural Language Processing (NLP).

Матннинг морфологик таҳлили (МТ) тарихи қадимги ҳинд тилшуноси Панини номи билан боғлиқ. У Aṣṭādhyāyī матнида фойдаланиб, санскрит тили морфологиясининг 3959 қоидабини яратди. Юнон-рим грамматикаси анъанаси ҳам МТ билан шуғулланиш бўлган [8]. 1859 йилда эса А.Шлейхер лингвистик "морфология" терминини ўйлаб топади [7:35]. Ўтган асрнинг 60-70-йилларига келиб, табиий тилни қайта ишлаш (NLP)да машинали морфология қамровидаги

барча тадқиқотлар машина луғатини яратиш билан бошланган [1; 2; 3; 4; 5].

МТда, асосан, сўзнинг грамматик шаклланиш имкониятини таҳлил қилиш ва сўзшаклга грамматик тавсиф бериш кўзда тутилади. Ушбу мақсадга эришиш учун қуйидаги ишлар амалга оширилиши керак:

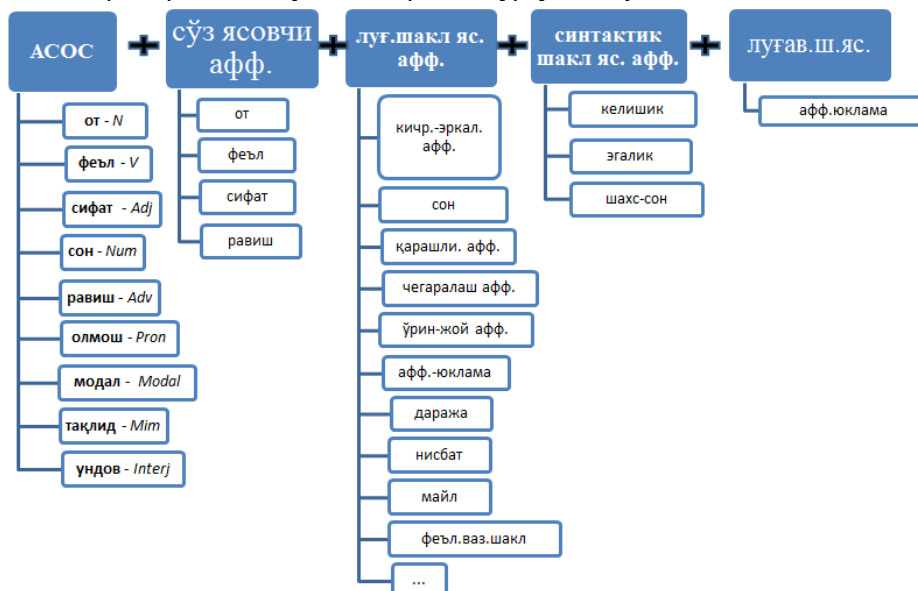
• Тилнинг имло луғатидаги глоссемалардан сўз туркумлари базаси яратилади. Бунда сўзлар туркумини топишда изоҳли луғат ўта зарур манба ҳисобланади.

• Ажратиб олинган сўзлар услубий (илмий, бадий, публицистик, расмий-идоравий ва сўзлашув услублари) жиҳатдан гуруҳланади.

• Сўзлар туркум доирасида лексик-семантик майдон (масалан, от туркумидаги сўзларни шахс отлари, атоқли отлар, ўсимлик ном (фитоним)лари, жой ном (топоним)лари, ҳайвон номлари (зоонимлар), мавҳум отлар каби) ва семантик гуруҳлар га иерархия тарзида ажратилади. Масалан, сифат туркумидаги ранг билдирувчи ЛСМдаги сифатлар ясовчи қўшимчаларни

олиш хусусиятларига биноан яна таркибли семантик гуруҳланади. Бундай иерархия леммага грамматик шаклларнинг бирика олиш имкониятини аниқлаштириш учун зарур ҳисобланади.

• Гуруҳланган сўзларнинг аффикслар комбинацияси тузилади. Бунда аффикслар комбинациясининг амалда адабий тил доирасида фойдаланиладиган ҳажми қамраб олинади. Комбинациялар қаторини тузишда аффикслар кетма-кетлигига эътибор қаратилади. М-н, от туркумида кўплик+эғалик +келишик категорияси ва ҳ.к.



Аi ва Еi орқали аффикслар қуйидаги ҳолатда ифодаланади:

| Аi  | Белги изоҳи                              | Рамзи    | Лемма туркуми        |
|-----|------------------------------------------|----------|----------------------|
| A1  | кўплик аффикси                           | k_a      | Исм асосли шакллар   |
| A2  | эғалик аффикси                           | e_a      |                      |
| A3  | келишик аффикси                          | ke_a     |                      |
| A4  | эркалаш-кичрайтириш аффикслари           | k_e      |                      |
| A5  | ўрин-жой от аффикси                      | õ_j      |                      |
| A6  | қарашлилиқ аффикси                       | q_a      |                      |
| A7  | чегаралаш аффикси                        | ch_a     |                      |
| A8  | сифат даражаси аффикслари                | Adj_a    |                      |
| A9  | аффиксли юкламалар                       | aff.Part |                      |
| A10 | инкор шакли ҳосил қилувчи афф.           | [inkor]  |                      |
| A11 | ажратиш (-гина)                          | aj_a     |                      |
| A12 | тегишлиқ (-лиги)                         | teg_a    |                      |
| A13 | ўхшатиш, солиштириш (-чалиқ, -дай, -дек) | õx_a     |                      |
| E1  | фееълнинг вазифадош шакллари аффикслари  | V_i      | Фееъл асосли шакллар |
| E2  | нисбат аффикслари                        | x        |                      |
| E3  | шахс-сон аффикслари                      | y        |                      |
| E4  | замон аффикслари                         | z        |                      |
| E5  | майл аффикслари                          | m        |                      |

Asos – Lemma<sub>ism</sub> = L<sub>ism</sub>

C – сўз ясовчи аффикс

L<sub>ism</sub> = иш

|                        |                    |
|------------------------|--------------------|
| L <sub>ism</sub> + A1  | derivation> ишлап  |
| L <sub>ism</sub> + A2  | derivation> ишим   |
| L <sub>ism</sub> + A3  | derivation> ишни   |
| L <sub>ism</sub> + A5  | derivation> ишдаги |
| L <sub>ism</sub> + A7  | derivation> ишгача |
| L <sub>ism</sub> + A9  | derivation> ишми   |
| L <sub>ism</sub> + A10 | derivation> ишсиз  |
| L <sub>ism</sub> + A12 | derivation> ишлиги |
| L <sub>ism</sub> + A13 | derivation> ишдек  |
| L <sub>ism</sub> + C   | derivation> ишхона |

|                                          |                                   |
|------------------------------------------|-----------------------------------|
| $L_{ism} + A1 + A2$                      | derivation> ишларим               |
| $L_{ism} + A1 + A2 + A3$                 | derivation> ишларимдан            |
| $L_{ism} + A1 + A2 + A9$                 | derivation> ишларимку             |
| $L_{ism} + A1 + A2 + A3 + A9$            | derivation> ишларимнику           |
| $L_{ism} + A1 + A2 + A3 + A10$           | derivation> ишларимнимас          |
| $L_{ism} + A1 + A2 + A3 + A10 + A9$      | derivation> ишларимданмасми       |
| $L_{ism} + A5 + A1 + A2 + A3 + A10 + A9$ | derivation> ишдагиларимданмасдир  |
| $L_{ism} + C + A5 + A1 + A2 + A10 + A9$  | derivation> ишхонадагиларданмаску |

Аффиксларнинг жойлашувидаги меъерий ҳолат баъзан бўзилади: *она-лар-ин*, *она-м-лар*, *айт-ди-нг-лар*, *айт-ди-лар-инг* каби. Бу сўзлашув ва бадиий услубдаги матнларда кам учрайдиган ҳодиса. Қўйида фақат шакл ҳосил қилувчи аффикслар доирасида фикр юритилди.

Ўғавий шакл ҳосил қилувчи аффиксларнинг одатдаги тартиби:

1) отда: *кичрайтириш-эркалаш* + *сон*. Кичрайтириш-эркалаш шаклининг кам қўлланадигани кўп қўлланадиганидан (ўлар базада кўрсатилади) аввал келади: *китобча*, *хатча*;

2) сифатда *-роқ* қўшимчаси *-иш*, *-имтир* аффиксидан кейин, шўнингдек, равишдош ва сифатдошга қўшилганда ҳам ўларни ҳосил қилувчи аффиксдан кейин қўшилади: *окишроқ*, *қўкимтирроқ*, *совинкураганроқ*, *тортиниброқ* ва х.к.;

3) сонда: тахмин билдирадиган *-ча* аффикси «дона» маъносини ифодаловчи *-та* (урғусиз) унсуридан кейин келади: *ўнтача* каби;

4) олмошда: гўмон билдирувчи *-ди* (урғусиз) элементи, ўмуман, юклама табиатида бўлиб, ҳамма турдаги аффикслардан кейин қўшилади: *нима-лар-ни-дир*;

5) феълда: ўзликдан бошқа нисбат кўрсаткичи орттирма нисбат аффиксидан кейин қўшилади: *тарка-тил-ди*, *югур-тир-иш-ди*, *тик-тир-иш-ди*; феълнинг тарз аффикси ва кўчайтирувчиси (*чайка*, *бура*, *тўзғи*); бўлишсизлик аффикси (*уринтирма*, *чайкатма*); замон аффикси; шарт майли аффикси ва бўйруқ майлининг кўчайтирувчиси (*борди*, *боради*, *борса*, *борай*, *боргин/боргил*), феълнинг ҳосланган шаклини ҳосил қилувчи аффикс (*бормоқ*, *борган*, *бориб* ва х.к.) охирида қўшилади.

Синтактик муносабат шакллари тартиби:

1) эгалик аффикси+келишик аффикси: *мартаба-сини*, *ўқиганимни*;

2) муносабат шакли: тасдиқ-инкор + замон / майл + шахс + сон: *ўтмаянсан*.

Ҳамма турдаги аффиксдан кейин аффикс-юклама қўшилади: *ўқимаянсанми*.

• Тузилган аффикслар комбинацияси сўзларга бириктирилади.

• Тилшунослар томонидан амалга оширилган юқоридаги ишлар дастурчи томонидан таъминотга киритилади.

Бугунги кунда морфологик таҳлил технологиялари яратилган бўлиб, шулардан самарадор технологияларга айланган стемминг ва лемматизация жараёнини қайд этиш мумкин [11:25-28; 12].

Морфологик таҳлил дастурининг амалий аҳамияти таълим жараёни ва илмий тадқиқотларда муҳим аҳамиятга эга. Муайян сўз шакллари ва унинг лингвистик изоҳларини автоматлаштириш ҳам вақт, ҳам маълумотнинг аниқлиги нуқтаи назаридан қулай имкониятларни беради. Масалан,

<китобини> [от], [турдош от], [наrsa оти], [т.к.], [бирл. с.], [III ш]

<ўқитди> [ф.], [муст. ф.], [сод. ф.], [тафаккур ф.], [ў-ли ф.], [б-ли ф.], [орт. н.], [тусл.ф.], [ў.з.], [III ш.б.] сўзшакллари китоб лемманинг китобини ҳамда ўқимоқ леммасининг ўқитди шакллари тавсифи ёрдамида таълим жараёнида ўқувчида морфологик компетенцияни шакллантириш ва илмий тадқиқотда

тадқиқот объектини атрофлича ўрганиш имкониятини беради.

Матнларни қайта ишловчи лингвистик дастурлар ҳам миллий тилнинг лингвистик қоидаларига таяниб тузилган алгоритмлар асосида ўз вазифасини бажаради. МТ алгоритмида қуйидаги белгилардан фойдаланилади:

Us – лингвистик таъминотдаги сўзлар базаси, Us="SELECT \* FROM 'Us'";

Ys – ўзбек тилидаги барча ясовчи аффикслар базаси, Ys="SELECT \* FROM 'Ys'";

Sq – ўзбек тилидаги грамматик категориялар базаси, Sq="SELECT \* FROM 'Sq'";

$S_i$  – W матндан ажратиб олинган сўзшакллар,  $1 \leq i \leq n$ , n – W матндаги сўзшакллар сони;

$S_{q_j}$  – Sq базадаги аффикслар,  $1 \leq j \leq m$ , m – Sq базадаги аффикслар сони.

Tz – сўзнинг хато ёзилганлигини визуал кўрсатувчи ва ўша сўзга мақбул сўзшакл вариантларини берувчи махсус функция.

Қуйида гапларни токенларга ажратган ҳолда, лемма бўйича ўзбек тилидаги сўзлар базасидан изланади, топилмаса ўзбек тилидаги барча ясама сўзлар базасидан изланади. Асос ёки ясалма базадан топилгач, унга бириктириш эҳтимолидаги аффикслар IDси бўйича ўзбек тилидаги барча аффикслар базасидан олинади. Демак, таҳлил алгоритми қуйидагича бўлади:

1.  $S_i$  даги ҳар бир сўз, Us базасидан излансин. Топилса, кейинги кадамга, акс ҳолда 5-кадамга ўтсин.

2.  $S_i$  сўзнинг Us базадаги ID (тартиб рақами) олинсин.

3.  $S_i$  сўзнинг ID рақамига тўғри келадиган аффикс  $S_{q_j}$  – Sq базадан излансин.

4.  $S_i + S_{q_j}$  тўғри бўлса 10-кадамга ўтилсин, акс ҳолда кейинги кадамга ўтилсин.

5.  $S_i$  даги ҳар бир сўз, Ys базасидан излансин. Топилса, кейинги кадамга, акс ҳолда 3-кадамга ўтилсин.

6.  $S_i$  сўзнинг Ys базадаги ID (тартиб рақами) олинсин.

7.  $S_i$  сўзнинг ID рақамига тўғри келадиган аффикс  $S_{q_j}$  – Sq базадан излансин.

8.  $S_i + S_{q_j}$  тўғри бўлса, 10-кадамга ўтилсин, акс ҳолда кейинги кадамга ўтилсин.

9. Tz функция амалга оширилсин ва 10-кадамга ўтилсин.

10. Сўз тўғри ёзилганлиги тасдиқлансин ва кейинги  $S_{i+1}$  сўзга ўтилсин.

Алгоритм иш тартибини мисол орқали таҳлил қиламиз: *истеъмолчиларнинг* сўзшакли стемминг қилинади, яъни *wordform set coord* майдонида берилган аффикслар тартиби бўйича кесиб чиқилади: *истеъмол/чи/лар/нинг*. Шундан сўнг *истеъмол* лемма имлоси List of Words базасидан текширилади, *-чи*, *-лар*, *-нинг* аффикслари  $S_{q_j}$  – Sq базасидан қидирилади, бунда *-лар*, *-нинг* шакл ясовчи ва синтактик муносабат шакли эканлиги аниқланади, *-чи* ясовчи аффиксни топиш учун Ys базасига мурожаат қилинади. Лемма ва аффиксларнинг ЛТда мавжудлиги топилгач, сўзнинг тўғри ёзилганлиги тасдиқланади ва кейинги сўзга ўтилади.

Юқорида берилган алгоритм бўйича Ubuntu Linux операциялар тизимида Uztanalysis дастури яратилди [11].

Ўзбек тилида ёзилаётган матнлар, турли соҳаларга оид ҳужжатларни автоматик таҳрир ва таҳлил қилишга эришиш мақсадида юқорида илк қадамда эътиборга олинган керак бўлган жиҳатлар морфологик таҳлил доирасида кўриб чиқилди, таҳрирловчи дастурнинг лингвистик таъминоти яратиш борасида муайян ҳулосалар чиқарилди.

Хорижий таҳлил дастурларини яратиш йўналишлари ва босқичма-босқич иш принциплари ўрганиш натижасида қуйидаги илмий ҳулосаларга келинди:

— Модель тил ҳақидаги билимларни объектив ифодалаш воситаси бўлганлиги боис лингвистик ходисаларни моделлаштириш компьютер лингвистика-

сида энг самарали усул ҳисобланади.

— Лингвистик процессорни яратишда лингвистик модулнинг ўрни ва аҳамияти ўта муҳим. Шу боис ўзбек компьютер лингвистикасида матнларга ишлов берувчи лингвистик дастурларга мукамал модулларнинг яратилиши мўйян мақсад йўлида тайёр лингвистик маҳсулотнинг юзага келишига замин яратади.

— Ўзбек тилининг дастлабки жорийланишда расмий-идоравий ва илмий услубдаги матнларни таҳрир ва таҳлил қилишга мўлжалланган лингво-системанинг мукамал ишлаши ва натижаси самарали бўлиши учун ушбу тизим графематик, морфологик ва синтактик таҳлил босқичларига ажратилиб, модуллаштирилди.

#### Адабиётлар

1. Белогов Г.Г., Дуганова И.С., Кузнецов А.Б. Экспериментальная система автоматизированного обнаружения и исправления ошибок в тексте / НТИ. Серия 2, 1984. №3. -С. 20-25.
2. Клышинский, Э. С. Начальные этапы анализа текста / Э. С. Клышинский // Автоматическая обработка текстов на естественном языке и компьютерная лингвистика: учебное пособие / Е. И. Большакова [и др.]. – Москва: МИЭМ, 2011. – -С.106-140, 270.
3. Леонтьева Н.Н. Автоматическое понимание текста: системы, модели, ресурсы: учеб. пособие. – Москва: Академия, 2006. –С. 304.
4. Мельчук И.А. Порядок слов при автоматическом синтезе русского слова (предварительные сообщения) / Научно-техническая информация. 1985. № 12. – С.12-36.
5. Ножов И.М. Морфологическая и синтаксическая обработка текста (модели и программы) сегментации русского предложения: Дис. канд. филол.наук. – Москва: 2003.
6. Пиотровский Р.Г. и др. Методы автоматического анализа и синтеза текста. – Минск: «Высшая школа». 1985.
7. Schleicher, August. Zur Morphologie der Sprache. Mémoires de l'Académie Impériale des Sciences de St.-Petersbourg. VII. I, N.7. St. Petersburg. 1859. –P. 35.
8. [https://en.wikipedia.org/wiki/Morphology\\_\(linguistics\)](https://en.wikipedia.org/wiki/Morphology_(linguistics))
9. <http://www.aot.ru/docs/sokirko/Dialog2004.htm> – морфологик модуллар
10. <http://tech.yandex.ru/mystem> – MyStem морфологик таҳлил дастури сайти
11. Abjalova M.A. Tahrir va tahlil dasturlarining lingvistik modullari. [Matn] : monografiya. – Toshkent: «Nodirabegim», 2020. 25-28-b.
12. Абжалова М. Матнларга авто-лингвистик ишлов бериш тизимлари. VI Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) –Ташкент: Navoiy universiteti, 2018. – С. 244-246.

#### РЕЗЮМЕ

Мазкур мақолада ўзбек тилидаги матнларни таҳлил қилувчи дастурнинг морфологик босқичида таҳлил қилиниши назарда тутилган сўзшаклларнинг грамматик категориялар тизимида макбул аффиксларни қабул қилиш меъёрлари формал тарзда ёритилди ва морфоанализаторнинг алгоритми берилди.

#### РЕЗЮМЕ

В данной статье формально описаны критерии приёма аффиксов в систему грамматических категорий слов, подлежащих анализу на морфологическом этапе программы анализа текстов на узбекском языке, и дан алгоритм работы морфоанализатора.

#### SUMMARY

This article formally describes the norms of accepting admissible affixes in the system of grammatical categories of words to be analyzed at the morphological stage (morphoanalyzer) of the program for analyzing texts in the Uzbek language, and gives an algorithm for the morphoanalyzer.

### ЛУҒАТ МАҚОЛАСИ ВА СЕМАНТИК ТЕГЛАР МАЖМУИДАГИ МУШТАРАК ВА ФАРҚЛИ ЖИҲАТЛАР

Д.Б.Ахмедова – филология фанлари бўйича фалсафа доктори

И.О.Хамидова – талаба

Бухоро давлат университети

**Таянч сўзлар:** корпус лингвистикаси, лексикография, лексик бирликлар, семантик теглаш, изоҳли луғат, бош сўз, умумистеъмол лексика, чегараланган лексика, иллюстратив мисол.

**Ключевые слова:** корпусная лингвистика, лексикография, лексические единицы, семантические теги, аннотированный словарь, общий словарный запас, ограниченный словарный запас, иллюстративный пример.

**Key words:** corpus linguistics, lexicography, lexical units, semantic tagging, dictionary, general vocabulary, limited vocabulary, illustrative example.

Корпус лингвистикаси лексикография билан чамбарчас боғлиқ. чунки луғатдаги тил бирлиги. унинг изоҳи корпус бирликлари теглаш учун лингво-лексикографик таъминот вазифасини бажаради. Ўзбек тилида мавжуд луғатлар матнини электронлаштириш луғатни корпус билан боғлашнинг илк босқичи саналади. Ўзбек лексикографияси узок тараққийёт йўлини босиб ўтди. корпус учун материал вазифасини бажарадиган бой хазинага эга бўлди.

Профессор Э.Бегматовнинг фикрича, лексикада системалилик тилининг бошқа сатҳларидаги каби яқол кўзга ташланиб турмайди. Лексик бирликлар микдор жиҳатидан фонема, морфемаларга нисбатан анча кўп ва даврий бекарорлик ҳусусиятига эга. Шу боис лексикани бор қўлаи билан аниқлаш ва тадқиқ қилиш имконияти йўқ [1:116]. Бугун жаҳон тилшунослиги бу муаммони

тил корпуслари ёрдамида ечишга киришган ва аллақачон натижаларни қўлга киритган. Лексикани сифат ва микдор жиҳатдан инвентаризация қилиш, атрофлича тадқиқ қилиш имконияти кенгайган.

Тил бирликлари семантик теглаш учун энг асосий манба ўзбек тилининг изоҳли луғатидир [2]. Луғат ҳозирги ўзбек адабий тилида кенг истеъмолда бўлган 80 мингдан ортиқ сўз ва сўз бирикмани. фан. техника. санъат ва маданият соҳасига оид терминларни, бир неча шевада қўлланадиган сўзларни, баъзи тарихий ва эскирган атамаларни ўз ичига олган. Луғатда берилган сўзларнинг амалда қўлланиши XX аср ўзбек адабиёти ва матбуотида олинган мисоллар билан далилланган. Луғат ўзбек тилшунослиги ва туркийшунослик бўйича мутахассислар, таржимонлар, оммавий ахборот воситалари ходимлари, олий таълим муассасаси