

**АЛИШЕР НАВОИЙ НОМИДАГИ ТОШКЕНТ ДАВЛАТ ЎЗБЕК
ТИЛИ ВА АДАБИЁТИ УНИВЕРСИТЕТИ ҲУЗУРИДАГИ
ПЕДАГОГ КАДРЛАРИНИ ҚАЙТА ТАЙЁРЛАШ ВА
УЛАРНИНГ МАЛАКАСИНИ ОШИРИШ ТАРМОҚ МАРКАЗИ**



2019



КОМПЮТЕР ЛИНГВИСТИКАСИ

ВА АСОСЛАРИ

Тингловчилар учун ўқув-услубий мажмуа

ЎЗБЕКИСТОН РЕСПУБЛИКАСИ
ОЛИЙ ВА ЎРТА МАХСУС ТАЪЛИМ ВАЗИРЛИГИ

**ОЛИЙ ТАЪЛИМ ТИЗИМИ ПЕДАГОГ ВА РАЎБАР КАДРЛАРИНИ
ҚАЙТА ТАЙЁРЛАШ ВА УЛАРНИНГ МАЛАКАСИНИ ОШИРИШНИ
ТАШКИЛ ЭТИШ БОШ ИЛМИЙ - МЕТОДИК МАРКАЗИ**

**АЛИШЕР НАВОИЙ НОМИДАГИ ТОШКЕНТ ДАВЛАТ ЎЗБЕК ТИЛИ
ВА АДАБИЁТИ УНИВЕРСИТЕТИ ҲУЗУРИДАГИ ПЕДАГОГ
КАДРЛАРИНИ ҚАЙТА ТАЙЁРЛАШ ВА УЛАРНИНГ МАЛАКАСИНИ
ОШИРИШ ТАРМОҚ МАРКАЗИ**

“Тасдиқлайман”

Тармоқ маркази директори

_____ **Н.М.Эгамбердиева**

“ _____ ” _____ **2019 йил**

“КОМПЮТЕР ЛИНГВИСТИКАСИ ВА АСОСЛАРИ”

модули бўйича

ЎҚУВ –У СЛУБИЙ МАЖМУА

Малака ошириш курси йўналиши: Барча йўналишлар

**Тингловчилар контингенти: олий таълим муассасаларининг
профессор-ўқитувчилари**

Тошкент – 2019

Мазкур ўқув-услугий мажмуа Олий ва ўрта махсус таълим вазирлигининг 2019 йил 18 октябрдаги 5-сонли буйруғи билан тасдиқланган ўқув режа ва дастур асосида тайёрланди.

Тузувчи: Абдурахмонова Н.З.

**ТошДўТАУ “Ахборот ва замонавий педагогик технологиялар”
кафедраси ўқитувчиси, ф.ф.д (PhD).**

Такризчи: Менглиев Б.,

ТошДўТАУ, “Ўзбек тилшунослиги” кафедраси профессори, ф.ф.д.

*Ўқув-услугий мажмуа Алишер Навоий номидаги ТошДўТАУ
Кенгашининг 2019 йил 28-августдаги 1-сонли қарори билан нашрга
тавсия қилинган.*

МУНДАРИЖА

I. ИШЧИ ДАСТУР.....	3
II. МОДУЛНИ ЎҚИТИШДА ФОЙДАЛАНИЛАДИГАН ИНТРЕФАОЛ ТАЪЛИМ МЕТОДЛАРИ.....	6
III. МАЪРУЗА МАТНЛАРИ.....	14
IV. АМАЛИЙ МАНБУЛАТ УЧУН МАТЕРИАЛЛАРИ, ТОПШИРИҚЛАР ВА УЛАРНИ БАЖАРИШ БЎЙИЧА ТАВСИЯЛАР..ОШИБКА! ЗАКЛАДКА НЕ ОПРЕДЕЛЕНА.	
V.БИТИРУВ ИШЛАРИ УЧУН МАВЗУЛАР.....	37
VI. КЕЙСЛАР БАНКИ.....	38
VII. ГЛОССАРИЙ.....	
4ОШИБКА! ЗАКЛАДКА НЕ ОПРЕДЕЛЕНА.	
VIII. АДАБИЁТЛАР РЎЙХАТИ.....	46

I. ИШЧИ ДАСТУР

Кириш

Дастур Ўзбекистон Республикаси Вазирлар Маҳкамасининг 2012 йил 26 сентябрдаги “Олий таълим муассасалари педагог кадрларини қайта тайёрлаш ва уларнинг малакасини ошириш тизимини янада такомиллаштириш чора-тадбирлари тўғрисида”ги 278-сон қарори, Ўзбекистон Республикаси Президентининг 2015-йил 12-июндаги “Олий таълим муассасаларининг раҳбар ва педагог кадрларини қайта тайёрлаш ва малакасини ошириш тизимини янада такомиллаштириш чора-тадбирлари тўғрисида”ги ПФ-4732-сон, 2016-йил 13-майдаги “Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университетини ташкил этиш тўғрисида”ги ПФ-4797-сон, 2017-йил 7-февралдаги “Ўзбекистон Республикасини янада ривожлантиришга бўйича Ҳаракатлар стратегияси тўғрисида”ги ПФ-4947-сон Фармонлари, 2017 йил 20 апрелдаги “Олий таълим тизимини янада ривожлантириш чора-тадбирлари тўғрисида”ги ПҚ 2909-сон Қарори йўналишлари мазмунидан келиб чиққан ҳолда тузилган бўлиб, у замонавий талаблар асосида қайта тайёрлаш ва малака ошириш жараёнларининг мазмунини такомиллаштириш ҳамда олий таълим муассасалари педагог кадрларининг касбий компетентлигини мунтазам ошириб боришни мақсад қилади.

Модулнинг мақсади ва вазифалари

Мазкур курс тингловчиларга тилшунослик ва адабиётшунослик соҳаларида компьютер технологияларидан унумли фойдаланиш ҳамда лингвистикага доир масалалар (тилларни ўқитиш, билимларни баҳолаш, матнларни таҳрир қилиш, бир тилдан иккинчи тилга таржима қилиш кабилар)ни компьютер воситасида ҳал қилиш йўллари, компьютер саводхонлиги даражасини ошириш, мантиқий тўғри, изчил фикрлашга ўргатиш, уларда назарий билимларни шакллантириш ва ўз йўналишлари бўйича амалда қўллашга оид кўникмаларни ҳосил қилишни назарда тутади.

Шунингдек, ушбу модул филолог мутахассисларда компьютер лингвистикаси бўйича чуқур билимларни шакллантириш, компьютер лингвистикаси тамойилларини ўзлаштириш, математик моделларни тузиш бўйича билимларни тўғри йўналтириш, формаллаштириш бўйича лингвистик таҳлил қилиш кўникмасини ривожлантириш, математик мантиқ асосида фикрлашга ўргатиш, тил ва адабиётга доир компьютер дастурларини яратиш, алгоритм ва компьютер дастурлари тузиш асосларини бериш каби вазифаларни қўяди.

Модул бўйича тингловчиларнинг билими, кўникма ва малакаларига қўйиладиган талаблар

“Компютер лингвистикаси асослари” фани бўйича тингловчи компютер лингвистикаси ва НЛП предмети, вазифалари, соҳалараро компютер

лингвистикасининг йўналишларини ўрганиш, табиий тилларни қайта ишлаш, тилларнинг математик моделини яратиш, автоматик таҳрир ҳақида билимга ега бўлиш керак;

- компьютер лингвистикаси учун табиий тилни қайта ишлаш, тил технологияси, моделлаштириш, формал тил назарияси (Ноам Чомский таълимоти), автоматик таҳлил қилиш мезонлари ва тамойиллари, тиллараро муносабатлар таҳлили, табиий тил қонуниятларининг алгоритмик мезонларини билиш каби кўникмаларига ега бўлиши керак;

- лингвистик билимлар базаси, машина таржимасида тизимлаштириш масаласи, лингвистик технологиялар, синтактик парсинг, компьютер лексикографияси, морфологияси, морфологик теглаш, корпус лингвистикаси, матнни автоматик таҳрир қилиш, математик лингвистикага оид тадқиқотларни ўрганиш, назарий маълумотларни умумлаштириш, компьютер лингвистикаси фанининг истиқболи, долзарб масалаларига доир фикрлар билдириш ва тадқиқот олиб бориш малакаларига ега бўлиши керак.

Модулни ташкил етиш ва ўтказиш бўйича тавсиялар

“Компютер лингвистикаси асослари” модули маъруза ва амалий машғулотлар шаклида олиб борилади.

Модулни ўқитиш жараёнида:

- маъруза дарсларида замонавий компьютер технологиялари ёрдамида презентацион ва электрон-дидактик технологиялардан;

- ўтказиладиган амалий машғулотларда техник воситалардан, блис-сўровлар, тест саволлари, ақлий ҳужум, гуруҳли фикрлаш, кичик гуруҳлар билан ишлаш, коллоквиум ва бошқа интерактив таълим усулларини қўллаш назарда тутилади.

Модулнинг ўқув режасидаги бошқа модуллар билан боғлиқлиги ва узвийлиги

“Компютер лингвистикаси асослари” модулини ўқитиш ўқув режасидаги электрон педагогика асослари ва педагогнинг шахсий, касбий ахборот майдонини лойиҳалаш, филологик тадқиқотларда тизимли таҳлил модуллари билан боғлиқликда ва узвийликда олиб борилади.

Модулнинг андрагогик таълимдаги ўрни

Ҳозирги замон компьютер лингвистикасининг ривожланувчи соҳалари автоматик таржима, автоматик дастур ва тил ўргатиш жараёнини автоматлаштиришдир. Ушбу модулни ўзлаштириш орқали тингловчилар компьютер лингвистикасига доир ахборотни амалда қўллаш, автоматик таржима ва автоматик таҳрирга оид масалаларни баҳолашга доир касбий компетентликка ега бўладилар.

Модул бўйича соатлар тақсимоти

№	Модул мавзулари	Тингловчининг ўқув юкلامаси, соат			
		Хаммаси	Аудитория ўқув юкلامаси жумладан		
			Жами	Назарий	Амалий машғулот
1.	Компютер лингвистикасига кириш	2	2	2	
2.	Машина таржимаси	2	2		2
3.	Лингвистик ресурслар: тезаурус, онтология, корпус	2	2		2
Жами		6	6	2	4

НАЗАРИЙ МАШЎУЛОТЛАР МАЗМУНИ

1-мавзу. Компютер лингвистикасига кириш
 “Компютер лингвистикаси ва асослари” модули доирасида ўрганиладиган асосий масалалар. Компютер лингвистикасининг назарий ва амалий масалалари.

АМАЛИЙ МАШЎУЛОТЛАР МАЗМУНИ

1-амалий машғулот. Машина таржимаси
 Машина таржимаси технологияси ва тизимлари. Машина таржимаси алгоритми. Табiiй тилларнинг машина таржимасида координацияси.

2-амалий машғулот. Лингвистик ресурслар: тезаурус, онтология, корпус

Компютер лексикографиясида маълумотлар базаси. Тезаурус ва онтология тушунчалари. Корпус лингвистикаси.

II. МОДУЛНИ ЎҚИТИШДА ФОЙДАЛАНИЛАДИГАН ИНТЕРФАОЛ ТАЪЛИМ МЕТОДЛАРИ

«Зиг-заг» методи

Методнинг мақсади: Бу метод гуруҳ асосида ишлаш, мавзунини тезкор ва пухта ўзлаштиришга хизмат қилади. Методнинг моҳияти шундан иборатки, бунда мавзунинг турли тармоқлари бўйича турли бўлимлардан ахборот берилади ва вақт тежаланилади. Жамоа бўлиб ишлаш кўникмаси шаклланади. “Зиг-заг” методидан маъруза машғулотида, амалий ва семинар машғулотида мавзу юзасидан билимларни мустақамлаш, таҳлил қилиш ва таққослаш мақсадида фойдаланиш мумкин.

Методни амалга ошириш тартиби:



тренер-ўқитувчи иштирокчиларни 5-6 кишидан иборат кичик гуруҳларга ажратади;



Ҳар бир гуруҳга мавзунинг муайян қисми (1-матн, 2-матн...) берилади ва ўрганиш вазифаси топширилади. Белгиланган вақт мобайнида гуруҳлар матн устида ишлайдилар; вақтни тежаш мақсадида уларнинг орасидан лидерлар тайинланади ва улар матнни гуруҳдошларига сўзлаб берадилар.



Барча гуруҳлар ўзларига берилган матнни пухта ўзлаштиргандан сўнг матнлар гуруҳлараро алмаштирилади ва юқоридаги жараён давом еттирилади.



Шу таҳлитда мавзунинг моҳиятан ёритувчи яхлит матн иштирокчилар томонидан ўзлаштирилади.

Намуна:

1-матн

Бугунги кунда таксономия, тезаурус, онтология каби атамалар кенг қўлланилмоқда. Таксономия – бошқариладиган луғат бўлиб, тармоқли (шажарасимон) тузилган аجدод-авлод (ота-она – фарзанд), бутун – қисм каби бошқа муносабатлар асосида тузилган иерархик мажмуадир.

Тезаурус еса таксономиядан каттароқ сўзлардаги мураккаб муносабатларни ўз ичига олувчи бошқариладиган луғат туридир. Унда ҳам иерархик, ҳам эквивалент муносабатлар акс этади.

Муррейнинг ОЕД (Оxford English Dictionary)си 1928-йилда анъанавий диахрон аспектда тузилган ва унда лексикологик билимларнинг синхроник

Тезаурус (ТС) (тҳесаурос) юнон тилидан олинган термин бўлиб, хазина маъносини беради. Луғатларда, хусусан, ЛИНГВО 5да қуйидаги таснифлар берилган:

1

- 1) идеографик луғат;
- 2) синоним ва антонимлар луғати;
- 3) сўзл. маълумотнома, енсиклопедия

ИИ. информатикада калит терминлар орқали қидирувни амалга оширувчи ҳамда компютерда сақланувчи муайян соҳа бўйича маълумотларнинг тизимлаштирилган мажмуаси.

2. Тилларни автоматик қайта ишлашда омоним, синоним, атоқли отлар, ибораларга оид луғатлар.

Веб-тезауруси ёки WWW- тезауруси Интернет ресурслардан фойдаланиш ҳамда ахборот қидируви учун WWW глобал гиперматнли тизими тезауруслари ҳам луғатларда қайд етилади. Тезаурус ҳақидаги фикрлар билдирилар экан Адам Килгарифф ва Коллин Яллопнинг Тезаурусдаги нима? (Ҳаҳатс ин а тҳераурус?) мақоласига мурожаат етамиз. Унга кўра, ТС бир гуруҳга кирувчи бир-бирига яқин маъноли сўзларнинг ресурси сифатида унинг қуйидаги турлари кўрсатилади:

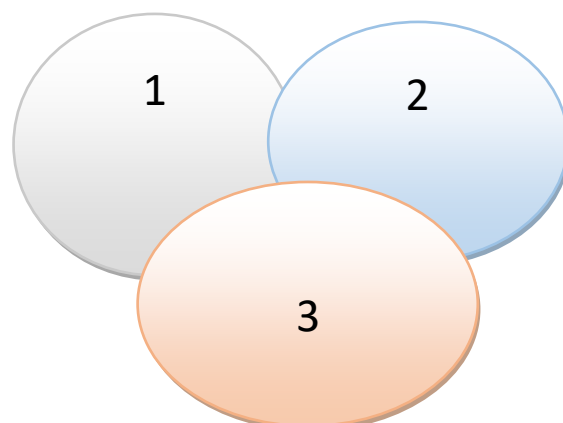
1. Роет ТС:

- Китоб шаклидаги муаллифлар томонидан яртаилган Роет ТС ёзувчилар учу сўз қоллашда ёрдам беради;
- Ворднет ва ЕуроWordНет;
- Ахборот қидирув тизимида ТС фойдаланиш;
- Корпусларини қайта ишлаш орқали ҳосил бўладиган автоматик тезауруслар. Компютер лингвистикасида, шубҳасиз, ТСга оид адабиётларнинг турли хили топилади.

Роет типидagi ТСда нофаол сўзлар, инглизча бўлмаган сўзлар, исмлар, ескирган ва нофразали ифодалар, иборалар йиғилган бўлиб, бу ёзувчига, қолаверса, инглиз тили чет тили бўлган фойдаланувчиларга сўзни тўғри танлашда ва қўллашда катта ёрдам беради. Ундаги лексемалар сўз туркумлари бўйича тузилган бўлиб, боғланган калит сўзлар орқали катта ҳажмдаги семантик корпусда ўз ифодасини топган.

“Вен диаграммаси” методи

Методнинг мақсади: мазкур метод тингловчилар ёки қатнашчиларда мавзуга нисбатан таҳлилий ёндашув, айрим қисмлар негизида мавзунинг умумий моҳиятини ўзлаштириш (синтезлаш) кўникмасини ҳосил қилишга йўналтирилади. Ёзув тахтаси ўзаро тенг тўрт бўлакка бўлинади ва ҳар бир бўлакка қуйидаги диаграмма **чизилади:**

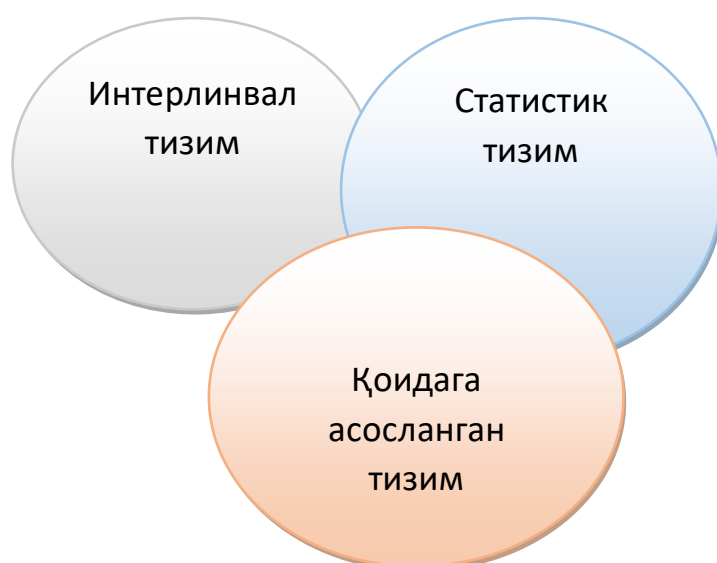


Метод иштирокчилар томонидан ўзлаштирилган ўзаро яқин назарий билимлар, маълумотлар ёки далилларни қиёсий таҳлил қилишга ёрдам беради. Ушбу методдан муайян бўлим ёки боблар бўйича якуний дарсларни ташкил қилишда фойдаланиш самаралидир.

Методни амалга ошириш тартиби:

- Ёзув тахтаси ўзаро тенг тўрт бўлакка бўлинади ва ҳар бир бўлакка диаграмма чизилади;
- тингловчиларга мавзуга ёки бобга тегишли бўлган сўзлар, тушунчалар ёки топшириқлар берилади (индивидуал ёки гуруҳли тартибда);
- лидерлар гуруҳ аъзолари томонидан билдирилган фикрларни умумлаштириб, ёзув тахтасида акс етган диаграммани тўлдирдилар.

Намуна: “Машина таржимаси” мавзуси



Гурухлар	Диаграмманинг тартиб рақами	Топшириқлар мазмуни
1-гурух	1-диаграмма	Машина таржимаси тушунчаси
	2-диаграмма	Машина таржимаси технологияси
	3-диаграмма	Машина таржимаси тарихи
2- гурух	1-диаграмма	Машина таржимасига оид тадқиқотлар
	2-диаграмма	Машина таржимаси тизимлари
	3-диаграмма	Машина таржимасининг афзалликлари
3- гурух	1-диаграмма	Машина таржимаси алгоритми
	2-диаграмма	Машина таржимасида лингвистик масалалар
	3-диаграмма	Машина таржимаси методлари

"Инсерт" стратегияси

Стратегиянинг мақсади: кўзда тутилган янги мавзу бўйича иштирокчиларнинг муайян тушунчаларга эгалликларини аниқлаш ва уларда матнга нисбатан тахлили ёндашиш кўникмларини шакллантиришга хизмат қилади мавжуд назарий билимлар ва амалий тажрибаларни тахлил қилиш, таққослаш орқали муаммони ҳал этиш йўлларни топишга, билимларни мустаҳкамлаш, такрорлаш, баҳолашга, мустақил, танқидий фикрлашни, ностандарт тафаккурни шакллантиришга хизмат қилади. Шунингдек, ўз хатолар устида ишлаш, мантикий фикрлашга ёрдам беради.

Стратегияни амалга ошириш босқичлари:

- кичик гурухлар шакллантирилади;
- ҳар бир гурухдан ўзлаштирилиши режалаштиралаётган мавзуга оид иккита фикр сўралади;
- қатнашувчилар навбатма-навбат фикрларини баён этадилар;
- баён этилган фикрлар ёзув тахтасига ёзиб борадилар;
- янги мавзунини ёрутувчи матнлар гурухларга тарқатилади;
- гурухлар матн билан танишиб чиқиб, матн ва ўзлари билдирган фикрларнинг бир-бирига қай даражада мувофиқ келганилигини аниқлайдилар (ўхшашлик ва фарқлар махсус белгилар ёрдамида қайд

этилади)

- гурух азолари шахсий фикрларини билдирадилар ва махсус белгилар сони бирлаштирилади;
- гурух аъзолари орасидан лидерлар танланади;
- лидерлар гурухига натижалар билан таништирадилар;
- гурухларнинг карашлари умумийлаштирилиб, ягона хулоса чиқарилади.

Намуна: "Компютер лингвистикасига кириш"

Махсус белгилар	Белгиларнинг маъноси
Ҳ (ҳа)	Матнда гурухлар томонидан билдирилган фикрлар ўз аксини топган бўлса
Й (йўқ)	Матнда гурухлар томонидан билдирилган фикрлар ўз аксини топмаган бўлса
?	Матн билан танишиш жараёнида саволлар туғилса (тушунмовчиликлар юзага келса)

1. Лидерларнинг ҳисоботидан сўнг гурухларнинг натижаларини ўрганиш чоғида:

Махсус белгиларнинг тартиб рақами	Гурухларнинг тартиби		
	I	II	III
1.			
2.			
3.			

"Кейс-стади" методи

"Кейс-стади" - инглизча соъз боълиб, ("case" - аник вазият, ходиса, "стади" - оърганмок, тахлил килмок) аник вазиятларни оърганиш, тахлил килиш асосида оъкитишни амалга оширишга каратилган метод ҳисобланади. Мазкур метод дастлаб 1921 йил Гарвард университетиде амалий вазиятлардан иктисодий бошқарув фанларини оърганишда фойдаланиш тартибида коълланилган. Кейсде очик ахборотлардан ёки аник вокеа-ходисадан вазият сифатида тахлил учун фойдаланиш мумкин. Кейс харакатлари оъз ичига куйидагиларни камраб олади: Ким (Вхо), Качон

(Вхен), Каерда (Вхере), Нима учун (Вхй), Кандай/ Канака (Хов), Нима-натижа (Вхат).

"Кейс методи" ни амалга ошириш босқичлари

Иш босқичлари	Фаолият шакли ва мазмуни
1-босқич: Компютер морфологияси	<ul style="list-style-type: none"> ➤ якка тартибдаги аудио-визуал иш; ➤ кейс билан танишиш(матнли, аудио ёки медиа шаклда); ➤ ахборотни умумлаштириш; ➤ ахборот таҳлили; ➤ муаммоларни аниқлаш
2-босқич: Морфологик таггинг	<ul style="list-style-type: none"> ➤ индивидуал ва гуруҳда ишлаш; ➤ муаммоларни долзарблик иерархиясини аниқлаш; ➤ асосий муаммоли вазиятни белгилаш
3-босқич: Синтактик парсинг	<ul style="list-style-type: none"> ➤ индивидуал ва гуруҳда ишлаш; ➤ муқобил йечим йўлларини ишлаб чиқиш; ➤ ҳар бир йечимнинг имкониятлари ва тўсиқларни таҳлил қилиш; ➤ муқобил йечимларни танлаш
4-босқич: Формал грамматика	<ul style="list-style-type: none"> ➤ якка ва гуруҳда ишлаш; ➤ муқобил вариантларни амалда қўллаш имкониятларини асослаш; ➤ ижодий-лойиҳа тақдимотини тайёрлаш; ➤ якуний хулоса ва вазият йечимининг амалий аспектларини ёритиш

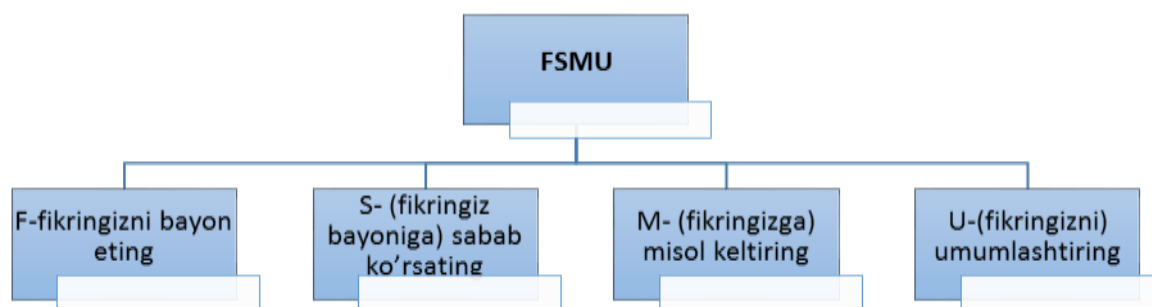
Кейс. Ўзбек тилининг морфологик ва синтактик парсинг жараёнида токен, лемма, стем каби таҳлил жараёнларини аниқлашда тилнинг агглютинативлик хусусиятларининг мавжудлиги.

«ФСМУ» методи

Технологиянинг мақсади: Мазкур технология иштирокчилардаги умумий фикрлардан хусусий хулосалар чиқариш, таққослаш, қиёслаш орқали ахборотни ўзлаштириш, хулосалаш, шунингдек, мустақил ижодий фикрлаш кўникмаларини шакллантиришга хизмат қилади. Мазкур технологиядан маъруза машғулотларида, мустаҳкамлашда, ўтилган мавзунини сўрашда, уйга вазифа беришда ҳамда амалий машғулот натижаларини таҳлил етишда фойдаланиш тавсия етилади.

Технологияни амалга ошириш тартиби:

- қатнашчиларга мавзуга оид бўлган якуний хулоса ёки ғоя таклиф етилади;
- ҳар бир иштирокчига ФСМУ технологиясининг босқичлари ёзилган қоғозларни тарқатилади:



- иштирокчиларнинг муносабатлари индивидуал ёки гуруҳий тартибда тақдимот қилинади.

ФСМУ таҳлили қатнашчиларда касбий-назарий билимларни амалий машқлар ва мавжуд тажрибалар асосида тезроқ ва муваффақиятли ўзлаштирилишига асос бўлади.

Намуна

Фикр: “Машина таржимаси кучли лингвистик таъминот билан таъминланиши зарур”.

Топшириқ: Мазкур фикрга нисбатан муносабатингизни ФСМУ орқали таҳлил қилинг.

III. МАЪРУЗА МАТНЛАРИ

1-мавзу: Компютер лингвистикасига кириш

Режа:

- 1.1 Фаннинг мақсади ва вазифаси.
- 1.2 Компютер лингвистикаси фан сифатида ривожланишида назарий ва амалий тадқиқотлар.
- 1.3 Компютер лингвистикасининг бошқа фанлар билан алоқаси (математика, психология, информатика).
- 1.4 Компютер лингвистикаси йўналишлари

Таянч тушунчалар: еҳтимоллик назарияси, компютер лингвистикаси, машина таржимаси, автоматик таҳрирлаш, моделлаштириш, математик мантиқ, текст ресогнитион, савол-жавоб, ахборот қидирув тизими

1.1. Фаннинг мақсади ва вазифаси

Олимлар компютер лингвистикаси борасида турли фикрларни билдиради. Компютер лингвистикаси информатика учун лингвистик таъминотни яратиш масалалари билан шуғулланади. Информатика фан сифатида замонавий техник воситалар ёрдамида ёзув қонуниятлари, ахборот тўплаш, қайта ишлаш, узатиш ва ундан фойдаланиш масалаларини ҳал қилади¹. Яна бир таърифда компютер лингвистикаси компютер нуқтаи назаридан табиий тилни илмий тадқиқ қилувчи соҳа эканлиги, компютер лингвистлар лингвистик ҳодисаларга оид турларни компютерли моделларини яратиши ҳақида фикрларни бериб ўтади².

¹ Марчук Ю.Н. Компьютерная лингвистика (учеб. пособ.) –Москва: Восток Запад, 2007. –С. 13.

² Sakthi Vel Applications of Computational Linguistics to Language Studies: An Overview // International Journal of Engineering Research in Computer Science and Engineering (IJERCSE) Vol 4, Issue 3, March 2017

Компютер лингвистикаси инженерия соҳаси билан алоқадорликда компютер нуқтайи назаридан ёзма тилни тушуниш ва қайта ишлаш жараёнига тегишли вазифаларни ўрганади. Тил тафаккур ойнаси ва мулоқот воситаси сифатида компютер лингвистикаси тафаккур ва идрокни нутқий вазиятда қандай моделга ега бўлишини компютер орқали таҳлил қилади. Лингвистик компетентга ега бўлган компютерлар дастурий таъминот билан машина ўртасида интерфейс сифатида информацион ресурсларни ўзаро боғлайди.

Компютер лингвистикаси фанининг мақсади табиий тил билан боғлиқ масалаларни компютер технологиялари ва методлари орқали ўрганади, тегишли муаммолар юзасидан ўз тавсияларини беради, методларини ишлаб чиқишдан иборат. Компютер лингвистикаси бир неча фанларнинг интеграллашган занжиридир. Хусусан, матнларни овозлаштиришда физик акустика, машина таржимасида статистик ва математик моделлаштириш, автоматик таҳрирда дастурий таъминот билан боғлиқ бўлган метрикалар, компютер лексикографиясида маълумот базасини лойиҳалаш кабилардан унумли фойдаланилади.

1.1. Компютер лингвистикасининг вужудга келиши ва ривожланиши

Компютер лингвистикаси компютер технологиялари ёрдамида тил билан боғлиқ масалаларни ўрганувчи соҳа еканлиги юқорида қайд этиб ўтдик. Компютер лингвистикаси термини фанда турли номлар билан номланиб келинган. Руслан Митковнинг еътироф этишича, “Сомпутационал лингвистисс” термини 1960-йилларда Миллий Фанлар академиясининг Тилни автоматик қайта ишлаш маслаҳат комитети аъзоси Давид Ҳайс томонидан киритилган³. Машина таржимасидан Компютер лингвистикасининг ажралиб чиқиши 1974-йилга тўғри келади. Сўнгра машина таржимаси ва Компютер лингвистикаси журналининг ўрнига Компютер лингвистикаси номли Америка журнали нашр еттирилади.

³The Oxford handbook of computational linguistics (edited by Mitkov R.) – Oxford, 2003. – P. 25

Компютер лингвистикаси машина таржимасига доир тадқиқотлардан сўнг тилнинг компютерга оид формал моделларини ривожлантириш маҳсули сифатида назарий ва амалий жиҳатдан ривожланиб борди.

Бугунги кунда «Компютер лингвистикаси»нинг йўналишлари (машина таржимаси, матнларни овозлаштириш, тилга ўқитиш, билимларни баҳолаш, ахборот кидируви, автоматик таҳрирлаш, сентимент анализ, корпус каби) соҳа мутахассислари томонидан ўрганилиб, жиддий ютуқларга еришиб келинмоқда. Чунки компўйтер лингвистикаси билан математиклар, компютер соҳа вакиллари, лингвистлар, таржимонлар биргаликда амалий илмий лойиҳаларда ҳамкорлик қилмоқдалар.

Тил тафаккур билан боғлиқ абстракт ҳодиса, шу билан бирга муайян даражада моделлаштирилган ва тизинмлаштирилган белгилар мажмуи ҳамдир. Унинг техника ёки аниқ фанлар билан боғлиқлигини изоҳлашда унинг назарий асосини ҳам ўрганиш мақсад мувофиқ.

Дастлаб 1952-йили Массачусец техника университетида Машина таржимаси бўйича ўтказилган илк илмий-амалий конференция ҳамда 1954-йил 7-январда Нью-Ёркдаги Жоржтаун университети билан ИБМ компанияси ҳамкорлигида биринчи ИБМ ИИ машина таржимаси дастурининг амалда синовдан ўтказилганлигини Компютер лингвистикасининг яратилишига омил бўлган муҳим воқеалик сифатида қараш мумкин.

Д.Журафский Компютер лингвистикасининг ривожланиш босқичларини қуйидагича даврий босқичларга кўра таснифлайди⁴:

Дастлабки босқич 1940-1950-йилларга тўғри келади. 1940-йилнинг охири 50-йилларнинг бошида икки асосий парадигма: автоматлаштириш ҳамда еҳтимоллик ёки информацион-назарий моделлар асосида тадқиқотлар олиб борилди. Автоматлаштиришнинг ўтган аср 50-йилларида амалда тадбиқ этилиши алгоритмик ҳисоблашнинг Тюринг моделини яратилишига (1936) асос бўлди. Тюринг тести Мак Кулоч –Питц нейрон тизимини вужудга

⁴Jurafskiy D., James H. Martin Speech and language processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, 2007 -P. 9-14

келишига асосий сабаблардан биридир. Ҳисоблаш элементи сифатида нейроннинг соддалашган модели мантикий мулоҳазаларнинг бирлигига айланди. Клининг ишларида (1951, 1956) чекли автомат (фине-стате автоматон) ва ўзгармас (регуляр) ифодаларга оид тадқиқотлар ўз аксини топди. Шаннон (1948) тилларни автоматлаштиришда дискрет Марков жараёнининг амалий еҳтимоллар моделидан фойдаланиб, уни амалга татбиқ этди. Чомский (1956) машинанинг муайян усули сифатида автоматлашган грамматика ёрдамида тилни умумлаштириш орқали тил ҳақидаги фикрларни асослаб берди. Ушбу моделлар ўз навбатида формал тил назариясининг вужудга келишига туртки бўлди. Белгиларнинг кетма-кетлиги сифатида алгебра ва кўплик назарияси формал грамматиканинг шаклланишида муҳим омил вазифасини ўтади. Бу Чомский томонидан киритилган табиий тиллар учун яратилган еркин контекстли грамматика (сонтехт-фрее граммар) назарияси билан боғлиқдир. Бироқ бу мустақил ҳолда Баскус (1959) ҳамда Наурлар (1960) томонидан яратилган АЛГОЛ дастурлаш тилида ўз моҳиятини янада очиқ берди.

Компютер лингвистикаси пайдо бўлишига асос бўлган яна бир омиллардан бири тил ва нутқни қайта ишлаш учун еҳтимоллик алгоритмларнинг ривожлантирилишидир. Шанноннинг ушбу масалаларга бағишланган илмий ишларида мулоқот канали ва тилнинг маълумот ҳажмини ўлчаш учун термодинамиканинг энтропия концепти устуворлик қилди. Олим еҳтимоллик назариясидан фойдаланиб, инглиз тилининг биринчи энтропик ҳажмини аниқлаб бера олган. Овозли спектрограф инструментал фонетика асосида ривожлантирилди. Бу еса дастлабки уринишлар сифатида нутқий ҳолатни автоматик тушуниш босқичига кўтара олди. 1952-йилда Белл лаборатория тадқиқотчилари томонидан сўзлаётган бирор кишининг 10 та рақамдан айримларини таний оладиган статистик тизими яратилди.

Кейинги давр 1957-1970-йиллар орасини ўз ичига олади. 50-йилларнинг охири ва 70-йилларнинг бошларида Компютер лингвистикаси

икки парадигма символик ва еҳтимоллик назариялар асосида ривожланди. Символик парадигма икки йўналишда тадқиқ етилди: формал тил назариясига асосланган Чомский таълимоти, синтактик анализ (парсинг) алгоритми, динамик дастурлаш ёрдамида пастдан юқорига (боттом-уп) ва юқоридан пастга (топ-даун) каби методлар. Тўлиқ синтактик анализ тизими Зелиг Харриснинг трансформация ва дискурс анализига доир лойихаси Пенсилвания университетида 1958-59-йиллар ичида амалга оширилган. Иккинчи йўналиш еса сунъий интеллект сифатида шаклланди. 1956-йил ёз ойида Мак Карти, Марвин Минский, Клауд Шеннон ва Назаниел Рочестерлар икки ойлик семинарда иштирок етишиб, сунъий интеллект терминини қўллашга келишиб оладилар. Сунъий интеллект соҳасидаги тадқиқотчиларнинг илмий ишларида еҳтимоллик назарияси ва статистик алгоритмга асосланган методлар кичик ҳажмни ташкил қилса-да, ушбу соҳада олиб борилган асосий еътибор Нивел ва Симёнлар томонидан илгари сурилган мантиқий назария ва умумий муаммоларга қаратилди. Бунинг натижасида табиий тилни тушунувчи автоматик тизим яратилди. 1960-йилларнинг охирига қадар формал мантиқий тизим изчил ривожлантирилди.

Ушбу еҳтимоллик парадигмалар статистика ва електр муҳандислигида кенг қўлланилди. 1950-йилларнинг охирида оптик характердаги Байесов методи тушуниш муаммосини ҳал етиш учун қўлланила бошланди. Бледсов ва Бровнинг илмий уринишлари натижасида (1959) катта ҳажмдаги луғат ва компютерлаштирилган матнларни ўқий оладиган (тeхт ресогнитион) Байесов тизими яратилди.

1960-йилларда дастлабки трансформацион метод асосида табиий тилни қайта ишлаш жараёнини текширувчи дастурнинг психологик модели тузилди. Шу билан бирга дастлабки интернет тармоғидаги корпус: турли жанрга оид манбалар (газета, бадиий, илмий, роман ва бошқ.)дан олинган 500 та ёзма матн жамланмасидан 1 млн.та сўзнинг Америка инглиз тилиси бўйича BROWN корпуси яратилди.

Компютер лингвистикасининг навбатдаги босқичи 1970-1983-йилларига тўғри келади. Бу даврда тил ва нутқни қайта ишлаш ва бошқа кўплаб тадқиқотлар соҳа ривожига ижобий таъсир етди. Нутқни танувчи дастур алгоритмларини ривожлантиришда еҳтимоллик парадигмаси муҳим рол ўйнади. Айниқса, Яширин Марков Моделидан фойдаланиш ҳамда Желинек, Бал, Мерсер ва уларнинг ИБМдаги ҳамкорлари Томас Ж. Вацон томонидан шовқинли канал модели ва декодлаш тизими ривожлантирилди. Колмерауер ва унинг ҳамкасблари томонидан Қ-тизимли ва трансформацион грамматика учун мантиққа асосланган парадигмалардан фойдаланилди.

Шу даврда табиий тилни қайта ишлаш жараёни Терри Винограднинг ШРДЛУ тизими орқали қайта ривожлантирилди. Бу дастур табиий тилдаги матнни қабул қила олиш имкониятига ега еканлиги билан аҳамиятлидир. Халлидей томонидан систем грамматикага асосланган кенгайтирилган инглиз тили грамматикаси яратилди. Виноградов модели семантик ва дискурс моделнинг таҳлил муаммосини йечишга ёрдам берди. Рогер Шанк ва унинг ҳамкасблари томонидан инсон концептуал билими ва хотирасига асосланган тилни тушуна оладиган бир нечта дастурлар амалда жорий етилди.

ЛУНАР деб номланган савол-жавоб тизимида мантиққа асосланган табиий тилни қайта ишлаш парадигмалари семантик ифода сифатида қўлланилиб, предикатив мантиққа бирлаштирилди. (Woodс, 1967, 1973). Емпиризм ва автоматлашган моделларнинг вужудга келиш даври 1983-1993-йилларга тўғри келади. 1950-йиллар охири ва 1960-йиллар бошида Чомскийнинг формал грамматикасидан фарқли бўлган икки модел ўрганила бошланди. Нутқни танишда еҳтимоллик моделларни қўллаш Томас Вецон томонидан олиб борилди. Еҳтимоллик методлари ва шунга ўхшаш маълумотларни бошқариш тизимлари семантика билан боғлиқ кўп маънолик каби масалаларни йечишда фойдаланилган. Шунингдек, ушбу даврда турли шаклдаги матнларнинг корпуслари яратилган. Бу ресурсларнинг амалда қўлланилиши семантик таҳлил, машина таржимаси, савол-жавоб,

маълумотларни бойитиш каби бир қатор лингвистик масалаларни йечишда ёрдам берди.

Ёндиликда компьютер лингвистикаси доирасида шуғулланувчи соҳа вакиллари икки гуруҳга ажратиш мумкин. Чунки биринчи гуруҳ вакиллари лингвистикага компьютер технологияларини ёрдамида нечоғлик тилни тушуниш ва ундан фойдаланиш бўйича изланишлар олиб борадилар. Иккинчи гуруҳ вакиллари НЛП – табиий тилни қайта ишлаш технологияси бўйича компьютер соҳа вакиллари тилга оид масалалар билан шуғулланади. Буларнинг ҳар иккиси ўзаро еришилган илмий ютуқлардан унумли фойдаланади.

Компютер лингвистикаси термини остида бошқа турли йўналишлар ҳам тушунилмоқда. Масалан, табиий тилни тушуниш (Натурал лангуаге ундерстандинг) деганда, Еман Муслим қуйидаги фикрга тўхталиб ўтади: матнга асосланган иловалар, диалогга асосланган иловалар.

1.1. СОҲАДА ОЛИБ БОРИЛАЁТГАН ТАДҚИҚОТЛАР

Жаҳон компьютер лингвистикасининг ўзига хос ривожланиши натижада турли илмий ёндашувлар ва технологияларнинг илмий мактабларини яратилишига асос бўлди. Компютер лингвистикаси бўйича турли тиллар доирасида мутахассислар томонидан тадқиқотлар олиб борилган. Рус компьютер лингвистикасида матнларни автоматик таҳрирловчи дастур, машина таржимаси, формал тил, табиий тилни моделлаштириш, нутқ синтезатори ва бошқа йўналишларда Р.Р.Котов, В.Е.Берзон, В.Г.Бритвин, И.А.Мелчук, Л.И.Беляева, В.А.Чижаковский, Г.Г.Белоногов, И.С.Дуганова, А.Б.Кузнецов каби олимларнинг бир қатор илмий тадқиқотларини кузатиш мумкин.

Машина таржимаси йўналишида тадқиқот олиб борган Н.Д. Андрейев, И.А. Мелчук, И.И. Ревзин, В.Й. Розенсвейг, Й.Н. Марчук, Р.Г. Пиотровский, Й.А. Моторин, К.Б. Бектайев, А.Н. Беляев, И.К. Белский, А.В. Зубов, Г.Е. Мирам, Л.Л. Нелюбин, В.И. Перебийнос, В.А. Чижаковский, Й.А.Шингарев,

Г.Г.Белоногов, Р.Г.Котов, Бабушкина Н.В., З.Шаляпина, О.Й. Мансурова, А.С. Панина, А.А. Хорошилов (Россия); А.Бут, Р. Ричанс, Ж.Хатчинс, Ж.Аллен, П. Браун (АҚШ); М. Нагао (Япония); А.Вахер (Естония); Франсияда Ж.Астроуни; Р.Синха, А.Жаин (Ҳиндистон); Б.Блейзер, У.Швол, А.Сторрер (Германия) каби тадқиқотчилар томонидан турли йўналишларда амалий натижаларга еришилган. Машина таржимасига оид назарий қарашларда қоидага асосланган алгоритм, фразага асосланган таржима, статистикага асосланувчи дастурий таъминот кибернетик, математик-дастурчи ва тилчилар томонидан амалга оширилган. Хусусан, бу соҳада кўплаб номзодлик ва докторлик диссертациялари ёқланганини алоҳида таъкидлаш лозим. Бу тадқиқотлар рус тилининг назарий ва амалий (барча учун фойдали) жиҳатларини тадқиқ қилишга сезиларли ҳисса қўшган. Айниқса, ўтган асрнинг 60-80- йилларида машина таржимаси бўйича кўплаб илмий тадқиқотлар юзага келди.

Компютер лингвистикаси туркологияда изчил ривожлантирилмоқда. Бунга Туркия, Қозоғистон, Татаристон, Қрим, Озарбайжон, Қирғизистон ва бошқа туркий давлатларда олиб борилаётган турли йўналишдаги илмий тадқиқотларни келтириш мумкин. Бу йерда статистик анализ, машина таржимаси, автоматик таҳрир, тилларга ўқитиш каби соҳаларда ҳам кенг кўламли ишлар олиб борилиб, муайян натижаларга еришилган. Жумладан, турк тили бўйича компютер лингвистикасига доир изланишлар К.Офлазер⁵, Е.Адали, Гулсэн Эрийгит⁶, М.Орхун⁷, А.С Тантуғ⁸, Сунейд Тантуғ каби

⁵ Oflazer, K. (1994). Two-level description of Turkish morphology. *Literary and Linguistic Computing*, 9(2).; Oflazer, K. (1999). Dependency parsing with an extended finite state approach. In *Proceedings of ACL'99, the 37th Annual meeting of the Association for Computational Linguistics.*; Oflazer, K. and Kuruoz, I. (1994). Tagging and morphological disambiguation of Turkish text. In *Proceedings of the 4th Applied Natural Language Processing Conference*, p. 144-149. ACL.; Oflazer, K. and Tur, G. (1996). Combining hand-crafted rules and unsupervised learning in constraint-based morphological disambiguation.; In Brill, E. and Church, K., (eds), *Proceedings of the ACL-SIGDAT Conference on Empirical Methods in Natural Language Processing*. Oflazer, K. and Tur, G. (1997). Morphological disambiguation by voting constraints. In *Proceedings of ACL'97/EACL'97, The 35th Annual Meeting of the Association for Computational Linguistics*.

⁶ Gül, sen Eryigit, Joakim Nivre and Kemal Oflazer. 2008. Dependency parsing of Turkish. *Computational Linguistics*, doi: 10.1162/coli.2008.07-017-R1-06-83, 34(3):357-389. (indexed in SCI and SSCI); Umut Sulubacak and Gül, sen Eryigit. 2018. Implementing universal dependency, morphology and multiword expression annotation standards for Turkish language processing. *Turkish Journal of Electrical Engineering & Computer Sciences*, 26(3):1662-1672.

⁷ Orhun, M.,Tantuğ, A.C,Adalı, E. veSönmez,A.C., (2009a).Computational comparison of the Uyghur and

олимпларнинг номлари билан боғлиқ. Ушбу соҳада морфологик анализ, машина таржимаси, синтактик таҳлил, корпус лингвистикаси, савол-жавоб тизими каби йўналишларда улкан натижаларга еришилган.

Татар тилининг компьютерли қайта ишлаш тизими бўйича Дж.Сулейманов, А. Гатиатуллин, Р.Гилмуллин, Б.Ҳакимов каби олимплар томонидан тадқиқотлар олиб борилган.

2005-2010-йилларда Мисрософт учун Windows 95 ОС учун татар тилининг интерфейси ҳамда Татаристон республикаси Фанлар академияси Амалий семиотика институтида операцион тизим учун татарча интерфейс яратилган⁹. Татар тилининг компьютер анализи бўйича кенг доирадаги илмий тадқиқотлар сифатида АББЙ Лингво луғатига татар тилининг киритилганлиги, Windows 95 ь98 ь2000 ьXP, Виста, Вин 7, Вин 8 WordCorp деб номланган автоматик таҳрирловчи дастурининг яратилганлиги, 1990-йилларда татар тили бўйича электрон ўқитиш дастурларининг амалда қўлланганлигини келтириш мумкин. «Татар Телле Заман» татар-онлине мултимедиали электрон дарслиги 2000 та татарча сўз, 2500 та расм ва фотосурат, 11 та кўнгилочар ўйин ва билимларни текширишга йўналтирилган машқлардан иборат татарча талаффуз қилувчи дастурни ўз ичига олади. Ушбу дастурнинг интерфейси рус, татар (кирил ва лотин) ҳамда инглизча тилга мослаштирилган кўп тилли платформага ега¹⁰.

Қирғиз тили бўйича компьютер лингвистикасига доир бир қатор тадқиқотлар олиб борилган. 2000-йилгача қирғиз тилининг тил моделларини ўрганиш ва уларни яратишда фақат амалий жиҳатдан математик аппарат ва

Turkish Grammar. The 2nd IEEE Inter-national Conference on Computer Science and Information Technology, 3, 338-342, Beijing, China. Orhun, M., Tantuğ, A. C., ve Adalı, E., (2009b). Rule Based Tagging of the Uyghur Verbs. Fourth In-ternational Conference on Intelligent Computing and Information Systems. Faculty of Computer & Information Science, Ain Shams University, 811-816, Cairo, Egypt.

⁸ M. Shylov, —Dilmaq: Turkish and Turkmen Morphological Analyzer and Machine Translation Program, Master's thesis, Fatih University, İstanbul Turkey, 2008.; 91 Tantuğ, A. C., Adalı, E., and Oflazer, K. 2006. Computer analysis of the Turkmen language morphology. Advances in natural language processing, proceedings (Lecture notes in artificial intelligence), 4139 . pp. 186-193. 92 Hamzaoglu, G. 1993. Machine translation from Turkish to other Turkic languages and an implementation for the Azeri language. MSc Thesis, Bogazici University, İstanbul 93 Oner, M., 2007, (In Turkish) Tatar Turkcesi; Turk Lehceleri Grameri Ed., Ahmet Ercilasun, Akcag Publications, Ankara, Turkey

9 Д.Ш. Сулейманов К вопросу внедрения татарского языка в киберпространство / Turklang. 2013, -С. 20

10 O'sha joyda. -С. 22

компютер технологияларсиз amalga oshirilgan¹¹. 2007-йилдан бошлаб Қирғизистоннинг бир неча университетларида “Компютер лингвистикаси” кафедраси ташкил етилган бўлиб, унга бириктирилган талабалар Фасебоок, Твиттер, Гоогле ва Википедиа сайтларини қирғиз тилига таржима қилувчи дастурий таъминот яратганлар. Хусусан, Е.Д.Есанов томонидан яратилган (2002) «Тамга-КИТ» дастури 20 та компонентдан иборат бўлиб, қирғиз тилининг ижтимоий-коммуникатив ривожлантиришга мўлжалланган. Шунингдек, 2011-йилда «ҚйрСпелл 2.2 - Проверка орфографии кыргызского языка для MS Оффисе 97-2013 учун қирғиз тили имлосини текширувчи» тил пакети яратилди ¹².

Ўзбекистонда компютер лингвистикасига доир муайян даражада изланишлар олиб борилган. Компютер лингвистикаси шаклланишида ўзбек тили материаллари бўйича лингвостатистик тадқиқотлар олиб борган С.Ризайев ва С.Мухамедовларнинг муҳим ўрни бор. Мазкур олимлар томонидан компютер лингвистикасининг статистик таҳлил йўналиши ривожлантирилган. С.Мухамедов ва Р.Р.Пиотровскийлар томонидан «Инженерная лингвистика и опыт системно – статистического исследования узбекских текстов» номли китобида лингвистик моделлар, моделлаштириш ва унинг умумий тамойиллари ҳамда ўзбекча матнларнинг квантитатив моделлари ўрганилган. Компютер лингвистикаси фани дастлаб математик лингвистика сифатида Ўзбекистонда 2001-йил профессор Абдумажид Пўлатов томонидан олий таълим тизимида фан сифатида киритилиб, лаборатория ташкил етилган. Компютер лингвистикаси лабораториясида турли КЛ муаммоларига қаратилган тадқиқотлар устида дастурчи, математик ва лингвистлардан тузилган илмий гуруҳ фаолият олиб борган. 2000-йилдан буён Компютер лингвистикасининг фан сифатида шаклланишида профессор А.Пўлатовнинг ўрни катта. 2005-2010-йиллар мобайнида Компютер лингвистикасида лингвистик ёндашувлар асосида тил билан боғлиқ масалалар битирув

11 С.Ж. Мусаев, С. Ж.Карабаева, А.И.Иманалиева Проблемы и перспективы развития компьютерной лингвистики в Кыргызстане/ Turklang. 2013, -С. 30

12 О'sha joyda. -С. 33

малакавий ишлари ва магистрлик диссертацияларида ҳимоя қилинган¹³. Хусусан, кўп тилли машина таржимаси тизимини яратишда Х.Аҳмедованинг¹⁴ магистрлик иши, компьютер лингвистикасининг назарий методологиясига илмий материаллар, ўзбек тилини моделлаштиришга оид тадқиқотлар¹⁵, А.Пўлатов¹⁶, С.Муҳамедова¹⁷, А.Раҳимов¹⁸, Н.З.Абдурахмонова¹⁹ларнинг компьютер лингвистикасига доир илмий қарашлари акс етган асарлари еътиборга молик. Бундан ташқари компьютер лингвистикасининг турли йўналишларида М.Ҳақимов, Б. Каримов, С.Жўрайев, Н.Жўрайева, А.Норов²⁰ каби соҳа мутахассислари²¹ томонидан машина таржимаси, тилни математик моделлаштириш каби йўналишларда илмий изланишлар олиб борилган. 2001-2012-йиллар А. Пўлатов асос солган компьютер лингвистикасининг шаклланиш даври бўлган, десак янглишмаган бўламиз.

2014-йилдан бошлаб ЙУРИДА фирмаси ўзбек тилининг тахрирловчи дастури бўйича амалий ишлар олиб бормоқда. Ҳам кирил ва лотин алифбосига асосланган ушбу дастур сўз шакллари ва имлосини текшириш имкониятига ега. Замонавий тахрирловчи дастурлар сингари синтактик ва семантик жиҳатдан текшириш имкониятига еришиш учун соҳа доирасида бир қатор амалий ишлар олиб борилмоқда.

¹³ Abdurahmonova N. Z. Matnlarni kompyuterda ishlash muaommolari: Bitiruv malakaviy ishi. – Toshkent: O‘zMU 2009; Abdurahmonova N.Z. Inglizcha sodda gaplarni o‘zbek tiliga tarjima qilish dasturining lingvistik ta‘minoti. Magistrlik dissertatsiyasi. – Toshkent: O‘zMU, 2011; Abdurahmonova N.Z. Inglizcha matnlarni o‘zbek tiliga tarjima qilish dasturining lingvistik ta‘minoti. Filol. fan.dok-ri.diss-yasi. (PhD) – Toshkent, 2018, 165 b.

¹⁴ Axmedova X. MVC platformasida ko‘p tillik tarjimon muhitini yaratish: Magistrlik dissertatsiyasi. – Toshkent: O‘zMU, 2015.

¹⁵ Пулатов А.К., Джураева Н. Разработка формальной модели грамматики узбекского языка // Узбекский математический журнал. – Т., 2002. – №1. – С.47-54; Пулатов А.К., Алиходжаев Б., Джураева Н. Разработка программы компьютерного анализа и синтеза глаголов узбекского языка // ЎзМУ хабарлари. –Ташкент, 2002. – № 2. – В.17-19.

¹⁶ Пўлатов А. Дунёвий ўзбек тили. – Тошкент, 2008; Пўлатов А. Компьютер лингвистикаси (дарслик). Тошкент: Академнашр, 2011;

¹⁷ Муҳамедова С. Компьютер лингвистикаси (методик қўлланма). – Т., 2007; Пулатов А, Муҳамедова С. Компьютер лингвистикаси (ўқув қўлланма) – Т., 2014; Муҳамедова С. Ўзбек тилидаги ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. – Тошкент, 2006. – 143 Б.

¹⁸ Раҳимов А. Компьютер лингвистикаси асослари. – Т.: Akademnashr, 2011.

¹⁹ Абдурахмонова Н.З. Машина таржимасининг лингвистик асослари. Тошкент: Академнашр, 2012

²⁰ Norov A. Kompyuter lingvistikasi asoslari. – Qarshi, 2017

²¹ Qarang: Ш.А.Назирова, Х.Х.Хомидов, А.И.Аланиязов, К.С.Рахманов, А.З.Махмудов Формализация конструкций предложений узбекского, турецкого и каракалпакского языков / Turklang. 2013. – Astana, 2013. – С. 34-47

2016-йили Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университетининг ташкил етилганлиги бу соҳада ижобий ўзгаришлар ясади. Компютер лингвистикаси фан сифатида филологик йўналишларидаги барча факултетлар учун ўқитилиб, “Математик мантик”, “Дастурлаш асослари”, шунингдек, таржима соҳасида фаннинг мантикий давоми сифатида “Машина таржимаси” фанлари ўқув режага киритилди. Ўзбекча матнларни таржима қилиш дастурининг лингвистик таъминотини яратиш бўйича фалсафа доктори диссертацияси ҳимоя қилинди²², Компютер лингвистикаси лабораторияси қайта ташкил етилди.

Шунингдек, 2017-йил “Марказий Осиё университетларида соҳалараро компютер лингвистикаси магистратура дастурини ривожлантириш” (“Development of an interdisciplinary master program on Computational Linguistics at Central Asian universities 585845-EPP-1-2017-1-ES-EPPKA2-SBHE-JP) номи халқаро Ерасмус+ лойиҳаси қўлга киритилди. Компютер лингвистикасининг халқаро тажриба асосда ривожлантиришда ушбу лойиҳанинг ўрни катта. Ушбу лойиҳада Йевропа давлатларидан Испаниянинг Сантяго де Компостелла ва Коруна университетлари, Португалиянинг Порто университети, Грециянинг Вест Аттиса университети, Полшанинг Адам Мискийевич университети, Қозоғистоннинг Ал-Фаробий номидаги Қозоқ миллий университети, Гумилёв номидаги Йевроосиё университети, Костанай университетлари ва Ўзбекистоннинг Тошкент давлат ўзбек тили ва адабиёти университети, Урганч давлат университети, Ўзбекистон Миллий университети, Самарқанд чет тиллари университети иштирок етмоқда. Ушбу лойиҳада компютер лингвистикаси мутахассисларидан Сарлос Гомез, Пабло Гамало (Испания), Белинда Миа (Португалия), Зигмунт Витулани (Полша), Алтинбек Шариббай, Диана Раҳимова, Алишер Тукайев (Қозоғистон), Мирсаид Арипов, Ғайрат Матлатипов, Муфтоҳ Ҳақимов, Саодат Муҳамедова, Нилуфар

²² Abdurahmonova N.Z. “Inglizcha matnlarni o‘zbek tiliga tarjima qilish dasturining lingvistik ta’minoti” nomli PhD dissertatsiyasi (2018), Mashina tarjimasining dasturiy ta’minoti (monografiya), Toshkent: Muharrir, 2018, 175 b.

Абдурахмоновалар (Ўзбекистон) иштирокида соҳа доирасида ҳамкорлик олиб борилмоқда.

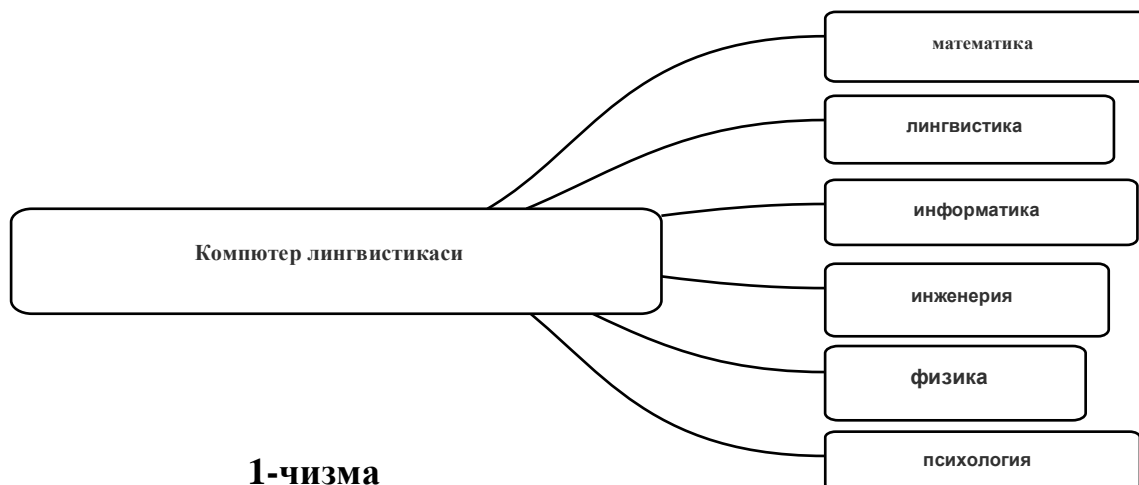
Шунингдек, туркий тиллар бўйича компьютер лингвистикасини ривожлантиришга қаратилган 6-халқаро “Туркланг” конференцияси илк бор 2018-йил (18-20-октябр) Тошкент давлат ўзбек тили ва адабиёти университетида ўтказилди. Ушбу конференция дастлаб 2013-йили Қозоғистоннинг Остона шаҳрида, сўнг анъанавий тарзда Истанбул (Туркия, 2014), Қозон (Татаристон, 2015, 2017), Бишкек (Қирғизистон, 2016) шаҳарларида ташкиллаштирилган. Ушбу конференцияда КЛ соҳасидаги мутахассислар Ешреф Адали, Мурат Орхун (Туркия), Алтинбек Шариббай, Диана Раҳимова (Қозоғистон), Айрат Гатиатуллин, Гилмуллин Ринат, Айдар Хусаинов (Татаристон), Белинда Маиа (Португалия), Олимжон Зокиров (Қирғизистон), Ленера Кубединова (Қрим) каби олимларнинг иштироки компьютер лингвистикасидаги кейинги истиқболли натижаларга еришиш учун асос вазифасини ўтайди. Ендиликда Унитурк семинари доирасида компьютер лингвистикасига доирасида илмий тадқиқотлар олиб борилмоқда. Компютер лингвистикаси соҳасида туркологияда катта ютуқларга еришиб келинмоқда. Компютер лингвистикасининг машина таржимаси, тил технологияси, лингвистик маълумотлар базаси, нутқ технологияси, электрон луғатхунослик (тезаурус, онтология) ва корпусга доир тадқиқотлар билан ушбу конференцияда танишиш мумкин www.turklang.uz, www.turklang.tatar²³.

Хулоса ўрнида шуни айтиш мумкинки, 2018-йилдан бошлаб ўзбек компютер лингвистикаси ҳам назарий ҳам амалий жиҳатдан ривожланиш даврига кирди.

²³ A. Alizade, Z. Guliyeva Building an automatic dictionary in the expert system of teaching scientific vocabulary – P.5-17; A. Bakiyeva, T. Batura Application of the rhetorical structure theory in automatic text processing systems – P.18-29; R. Gilmullin, R. Gataullin Morphological disambiguation in tatar language using purepos –P.30-37; B. Ergesh Identifying the tonality of texts in the kazakh language on the basis of the dictionary of emotional lexis 62-67; N. Prokopiev, D. Suleymanov Automated analysis of natural language question-answer texts in the e-testing system –P.92-98 va boshq. // Turklang. 2017 Kazan, 2017

1.1. Компютер лингвистикасининг бошқа фанлар билан алоқаси (математика, психология, информатика, дастурлаш асослари)

Тил ахборотни сақлаш, узатиш ва алоқа муносабат шаклидир. Компютер лингвистикаси бошқа соҳалар билан ўзаро боғлиқ. Компютер лингвистикаси куйидаги соҳалар билан алоқадор (1-чизма):



1-чизма

Маълумки, соҳанинг ўзи лингвистика ва компютер технологиялари билан боғлиқ. Манбаларда бу икки соҳани кесишуви психология ва мантиқ ёрдамида амалга оширилиши айтилади²⁴.

Ендиликда компютер лингвистикаси киберфазонинг ажралмас қисми саналади. Чунки табиий тил ушбу фазода асосий коммуникацион восита бўлиб, компютер лингвистикаси уни қандай қилиб юзага чиқариш мумкин, деган саволга жавоб излайди. Олимлар томонидан тилга олинаётган “рақамли инсоният” (дигитал ҳуманитй) учун ушбу соҳа катта аҳамиятга эга.

Лингвистик технология дейилганда, матн генерацияси, интерпретацияси, матн таҳлили, нутқни таниш каби вазифаларга мўлжалланган компютер тизимларини ривожлантириш учун лингвистик билимларнинг дастурий иловалари назарда тутилади²⁵.

²⁴ Qarang: Eman M. Muslim An Introduction to Computational Linguistics Advantages&Disadvantages <https://www.researchgate.net/publication/318116302>

²⁵ Moodle.udc.es/course

Маълумотларни қидириш, уларни ажратиш, савол-жавоб тизими каби дастурий таъминотлар фойдаланувчи ва компьютер ўртасидаги алоқани боғлашга хизмат қилади.

Машина ва фойдаланувчилар ўртасида турли тиллар учун яратилган тизимлар сирасига машина таржимаси, компьютер ёрдамида тилга ўқитиш ҳамда лингвистик вазифаларни бажарувчи тизимлар корпус, орфографик текширувчи дастурлар киради.

Компютер лингвистикасининг бошқа соҳалар кесими билан боғлиқ бўлиши компьютер лингвистикасининг нутқ синтезатори (физика, муҳандислик), машина таржимаси (математика, дастурлаш), сентимент анализ ва диалог тизими (психология), ахборот қидируви информатика каби йўналишларида ўз ифодасини топади.

Й.Караулов таъкидлаганидек²⁶, тилга оид билимлар базаси икки асосий концептдан иборат бўлади: 1) тил ҳақидаги исталган маълумот лексикографик шаклда берилиш керак; 2) тил ҳақидаги лексикографик маълумотлар машина форматига келтирилган алгоритмик тизимлаштирилиши лозим.

Компютер лингвистикаси когнитив фанлар ва сунъий интелект ўртасидаги диапасон ҳисобланади. унинг назарий ва амалий компонентлари мавжуд. Назарий компонентлари назарий лингвистика ва когнитив соҳа билан боғлиқ. Ҳануз лингвистик назариялар мураккаб бўлиб, компьютер учун уни осонлаштириш учун қарор чиқара оладилар. Формал лингвистик назариялар учун компьютер моделларни ривожлантириш учун когнитив психология ёрдамида лингвистлар дастурчилар билан ҳамкорлик қиладилар. Амалий комонентлар сунъий интелектга боғлиқ бўлган амалий моделларни ривожлантиради. Шу жиҳатдан “Тил инженерияси” (Лангуаге Енгинееринг) ёки “(Инсон) тил технологияси” ((Хуман) Лангуаге Течнологй). компьютер лингвистикаси терминига синоним бўлади. Амалий компьютер

²⁶ Караулов Ю. Методология лингвистического исследование и машинный фонд русского языка. // машинный фонд русского языка: идея и суждения. –М.: Наука, 1986. –С. 14

лингвистикасининг функцияси бу инсон ва машина ўртасидаги осон мулоқот жараёни осонлаштиришга ёрдам беради²⁷.

1.1. Компютер лингвистикаси йўналишлари

Табиий тилни қайта ишлаш ёки тил ҳақидаги билимлардан иборат дастурни қайта ишлаш ўртасида қуйидагича тафовут мавжуд. *Microsoft office* дастури лингвистик билимлардан фойдаланади, яъни ҳар бир бирликни сўз сифатида ўқисада, унинг имконияти нибатан чекланган. Табиий тилни қайта ишлаш жараёни еса *Microsoft office* имкониятлари билан солиштирилганда анча мураккаб жараён еканлиги амалда исботини топган. Уларга савол-жавоб тизими, машина таржимаси, автоматик аннотациялаш, рефератлаш, нутқий синтезатор каби тизимлар киради.

Табиий тилни қайта ишлаш бўйича маълумотлар²⁸:

Америка ҳаво йўлларидаги *Амтрак* деб номланган автоматик саёҳат-провайдери саёҳат жойи ва манзилга йетиб келиш вақтини автоматик тарзда бошқариб туради.

- Мерседенс-бенз замонавий машиналарда нутқни таниш ва матнни нутққа ўтказувчи технология мавжуд. Бу тизим ҳайдовчиларга муҳит, навигацион тизимлар, кўнгилочар мавзуларда овоз орқали бошқариш имкониятини беради.

- *Блинкx* ва бошқа видео қидирув компаниялари нутқни таниш технологиялари орқали миллион соатли веб-видеоларни ёзувга айлантира олади.

- Google таржимон овозли сўров орқали бошқа тилдаги сўралган маълумотни таржима қилиб (овозли ҳолатдан), сўнгра қидирув тизимига узатади.

²⁷ Eman M. Muslim An Introduction to Computational Linguistics

²⁸ Jurafskiy D., James H. Martin Speech and language processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, 2007 – P. 8.

-Пеарсонга ўхшаш маърифий нашриётларда ЕТСга ўхшаш текширув тизими мавжуд бўлиб, автоматлашган тизим орқали минглаб талабаларнинг иншоларини таҳлил қила олади ва уларни компетенциявий даражалай олади.

-анимацион характердаги интерактив тьюторлар афазия ва Паркинсон касаллигига учраган болаларни терапевт сифатида ўқита олади

-матн таҳлили билан шуғулланувчи компаниялар Нилсон Бузметрикс, Умриа каби тизимлар веб-саҳифалардаги ижтимоий фикрлар, муҳокамалар ва муносабатларни таҳлил қилади.

Ендиликда компьютер лингвистикасининг функциялари кенгайиб бормоқда. Шунга кўра унинг турли йўналишлари вужудга келган:

- машина таржимаси
- матн генерацияси
- савол-жавоб тизими
- нутқни овозлаштириш
- матн корпуси
- маълумот қидируви
- компьютер лексикографияси
- сентимент анализ
- матн тайёрлаш:

а) табиий тил матнлардаги сўзларни автоматик ажратиш;

б) нутқий вазиятни текшириш;

с) грамматикани текшириш;

д) услубни текшириш;

- маълумотлар базаси ва бошқа тизимлар учун табиий тил интерфейсларини яратиш

- матнларни янгилаб, бойитиб бориш
- табиий тилни тушуниш
- матнларини мавзули қисқартириш ва ҳ.к.

Маскуарие университети профессори Марк Жонсон компьютер технологиялар ёрдамида тил билан шуғулланиш соҳани иккига ажратади:

И. Компютер лингвистикаси (СЛ):

- 1) тилни тушуниш (лангуаге сомпрехенсион)
- 2) тилга оид машғулотлар тайёраш
- 3) тилдан фойдаланиш

ИИ. Табиий тилни қайта ишлаш (НЛП):

- 1) маълумот кидириш
- 2) мавзунни аниқлаш ва хужжатларни таснифлаш
- 3) хужжатларни қисқартириш
- 4) сентимент анализ (идрокий анализ – кидирув сўровининг

боғлиқлигига кўра матнларни цнифлаш)

- 5) машина таржимаси
- 6) нутқни таниш
- 7) матннинг автоматик таҳлили (текст мининг, глубинный анализ

текста, раскопка текста глубинный анализ текста, раскопка текста):

- a) маълумотларни ажратиб олиш (информатион экстрастион);
- б) матнларни токенлаш, стемларга ажратиш ва матнларни теглаш (кодлаш);
- с) хужжатлар жамланмасини гуруҳлаш (кластерлаш);
- д) гуруҳланган тўпламларни (кластерларни) таснифлаш.

Яна бир Сактҳи Вел компютер лингвистикасининг йўналишларини таснифлашда қуйидаги чизма орқали изоҳ беради²⁹(2-чизма). Лекин ушбу чизмада негадир морфологик анализ еътибордан четда қолган. Морфологик анализ синтактик ва семантик анализ учун асос вазифасини ўтайди. Диққатимизни тортган нарса тадқиқотчи ТТИ-табиий тил интерфейси (натурал лангуаге интефасе), ТТТ-табиий тилни тушуниш (натурал лангуаге ундерстандинг), ТТГ-табиий тил генерацияси (натурал лангиаге генератион), ТТЎ-табиий тилни ўрганиш (натурал лангуаге леарнинг) ўзаро боғланишини муайян даражада тушунтиришга ҳаракат қилган.

²⁹ Sakthi Vel S. Applications of Computational Linguistics to Language Studies: An Overview // International Journal of Engineering Research in Computer Science and Engineering (IJERCSE) Vol 4, Issue 3, March 2017

IV. АМАЛИЙ МАШҒУЛОТ УЧУН МАТЕРИАЛЛАРИ, ТОПШИРИҚЛАР ВА УЛАРНИ БАЖАРИШ БЎЙИЧА ТАВСИЯЛАР

1-амалий машғулот. Машина таржимаси

Машина таржимаси технологияси ва тизимлари. Машина таржимаси алгоритми. Машина таржимасида табиий тилларнинг парадигматик муносабатлари

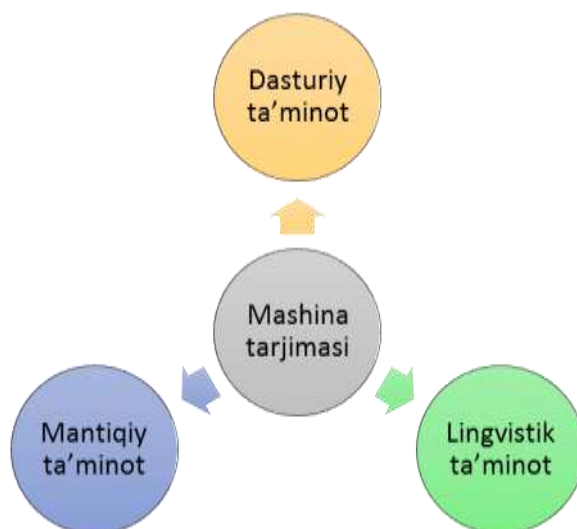
1-вазифа. Машина таржимаси технологияси ва тизимлари

Машина таржимаси тизимларини сананг ва уларнинг ўзаро фарқларини айтинг.

2-вазифа. Машина таржимаси алгоритми

Алгоритм нима? Машина таржимасининг алгоритми қандай босқичлардан иборат?

3-вазифа. Машина таржимасини амалга оширишда қуйидаги таъминотларнинг моҳиятини тушунтиринг



Компьютер лугатларининг маълумотлар базасини яратишда феълга хос қайси хусусиятлар ўз аксини топади.

4-вазифа. Табиий тилларнинг парадигматик муносабатларини қуйидаги схема бўйича тўлдириг:

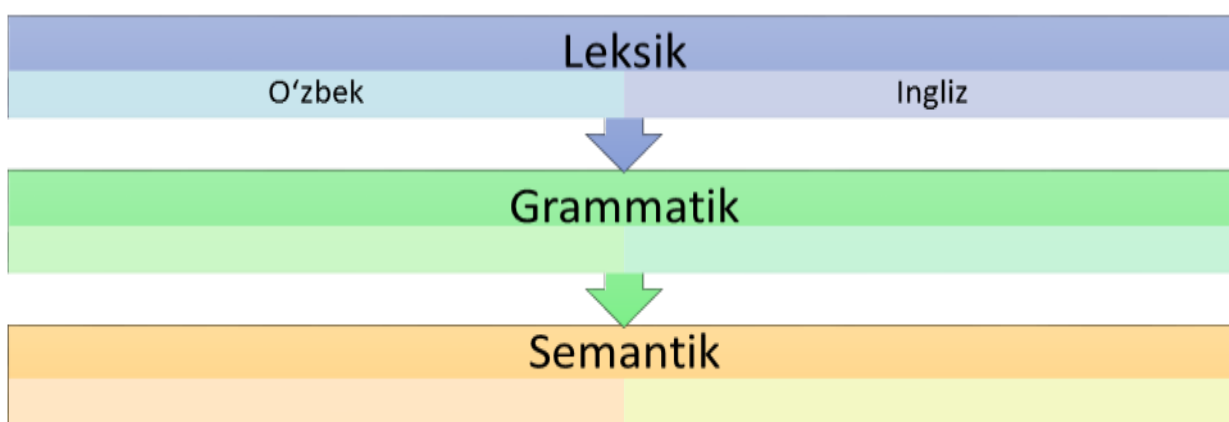


5-вазифа. Online режимда ишловчи машина таржимаси тизимларига мисол келтиринг, уларнинг ишлаш технологиясини солиштиринг.

Масалан: Гоогле транслатор, Яндекс.ру, Дилманс...

6-вазифа. Машина таржимаси тизимларининг муайян матн устида таржима сифатини солиштиринг.

7- вазифа. Машина таржимасида лингвистик муаммолар ҳақида гапиринг. Ушбу жадвални тўлдиринг:

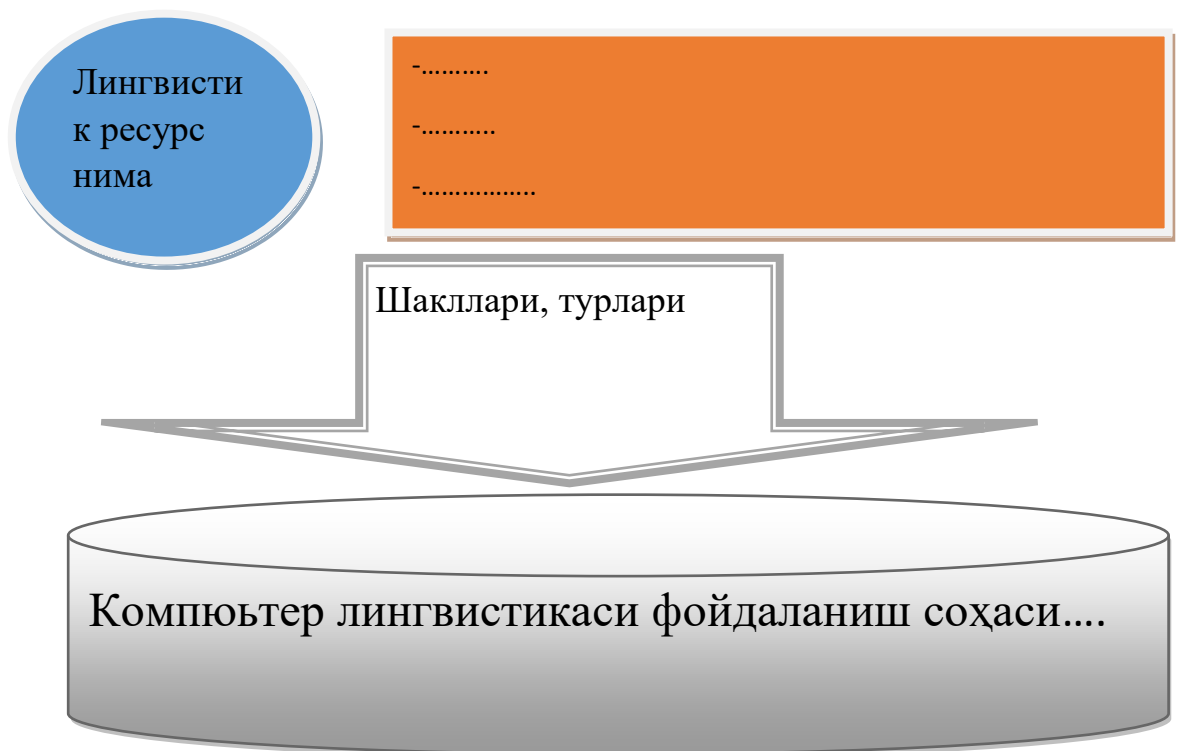


8-вазифа. Машина таржимасида терминологик масалалар. Ушбу мавзу юзасидан қуйидаги саволларга жавоб беринг:

1. Терминография тушунчасига изоҳ беринг.
2. Машина таржимасида терминология қай даражада муҳим?
3. Қандай турдаги машина таржимаси тизимлари терминологик маълумотлар базани бойитишда ташқи очиқ платформадан ҳам фойдаланади?

2- амалий машғулот

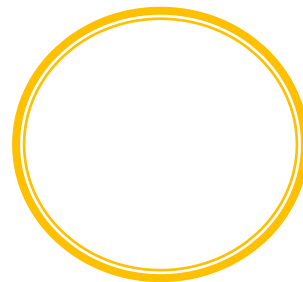
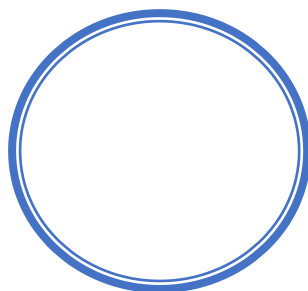
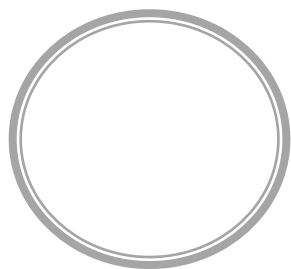
1-вазифа. Лингвистик ресурслар: тезаурус, онтология, корпус
Лингвистик ресурслар: тезаурус, онтология, корпус Компютер лексикографиясида маълумотлар базаси. Тезаурус ва онтология тушунчалари. Корпус лингвистикаси.
Компютер лексикографияси бўйича ушбу жадвални тўлдилинг.



2-вазифа. Тезаурус ва онтология

Тезаурус ва онтологиянинг фарқли ва умумий жиҳатлари:

Кейс



3-вазифа. Корпус лингвистикаси

Туркий тиллар бўйича яратилган қуйида кўрсатилган елестрон корпуслардан фойдаланган ҳолда ушбуи жадвални тўлдилинг:

<http://altay2.gasu.ru/>
<http://web-corpora.net/bashcorpus/search/>
<http://kazcorpus.kz/>
<http://korpus.juls.savba.sk/QIRIM/>
<http://www.corpus.antat.ru>
<http://www.corpus.tatar/>
<http://www.tuvacorporus.ru>
<http://www.tnc.org.tr/>
<http://khakas.altaica.ru/>
<http://corpora.iea.ras.ru/corpora/>
<http://adictsakha.nsu.ru/corpora/corp>

Корпус номлари							
структураси							
жанри							
ҳажми							
имкониятлари							
ютуқ томони							
камчилиги							
Умумий жиҳатлари				Фарқли жиҳатлари			
Корпус номлари							

V.БИТИРУВ ИШЛАРИ УЧУН МАВЗУЛАР.

- 1.«Эмоционал-экспрессив ва услубий хосланган лексемалар» модулини ўқитишда муаммоли таълим технологиясидан фойдаланиш
- 2.«Сифат» модулини ўқитишда концептуал жадвал график органайзеридан фойдаланиш
- 3.«Морфема ва унинг турлари» модулини ўқитишда манбалар устида ишлаш
- 4.«Компьютер луғатларининг маълумотлар базаси сифатидаги аҳамияти» модулини ўқитишнинг дидактик таъминотини такомиллаштириш
- 5.«Келишик категориясининг прагматик таҳлили» модулини ўқитишда концептуал жадвалдан фойдаланиш
- 6.Она тили дарсликларида алоқа-муносабат шакллари бўйича машқ ва топшириқлар (янги тавсиялар)
- 7.«Ўзбек адабий тил меёрларининг бугунги ҳолати» модулини ўқитишда таълимнинг амалий методларидан фойдаланиш
- 8.«Ўқитувчи нутқ сифатлари» модулини ўқитишда ҳамкорлик технологиясидан фойдаланиш
- 9.Адабиёт дарсларида аудио, видео материаларидан, дарслик ва қўлланмаларнинг электрон версия-ларидан фойдаланиш йўллари
- 10.«Ўзбек тилидаги ундов сўзларнинг маъно хусусиятлари ва нутқий қўлланиши» модулини ўқитишда лойиҳа таълим технологиясидан фойдаланиш

VI. КЕЙСЛАР БАНКИ

1-кейс. Ўзбек тилининг WordNetини тузишда қандай босқичлар амалга оширилади:

Sinonimlar: _____

Antonimlar: _____

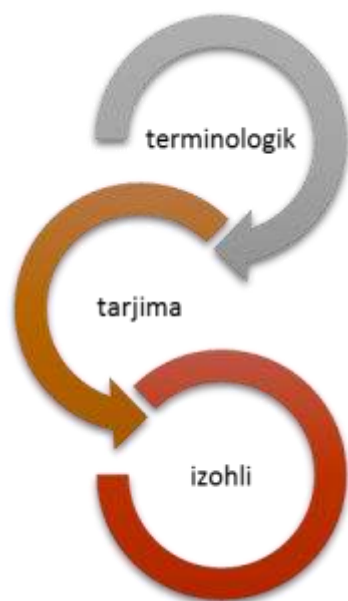
Giponimlar: _____

2-кейс. Ўзбек тили грамматикасининг онтологияси Ехселда яратинг.

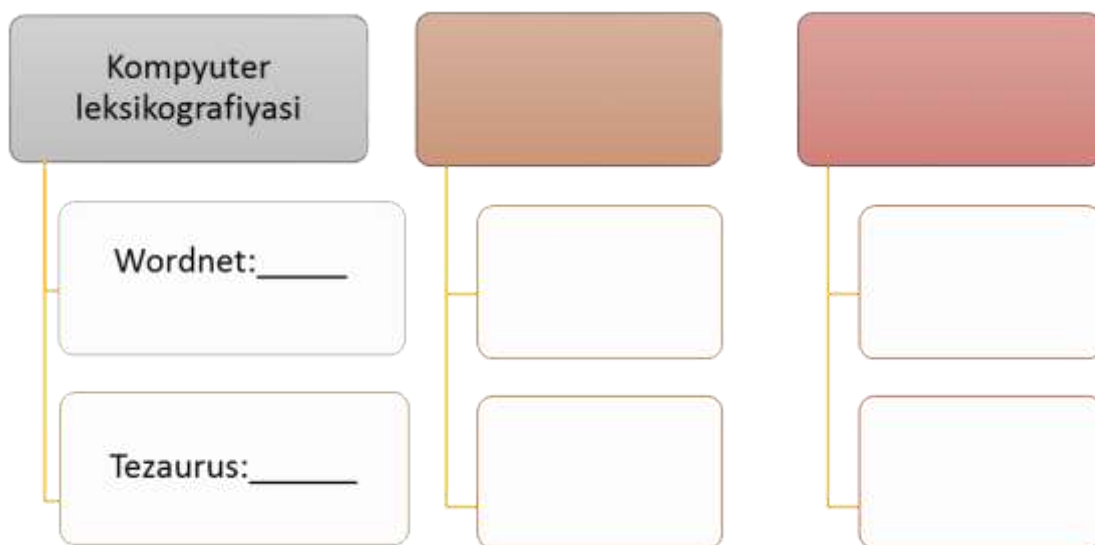
				Uzbek					
Tag	Name_Russian	Name_English	Part of speech	properties	types	definition	questions	example	
1 N	Имя существительное	Noun	От сўз туркуми			шахс, предмет, ходиса, жой номини билдирадиган сўзларга айтилади.	Ким? Нима? Қаер?		
	По составу	By composition	11	таркиби					
SIMP	Простые существительное	Simple Noun		111	одний				
CMPL	Сложные существительное	complex Noun		112	мураккаб				
FUSW	слитные(сложные) существительное	Fused words			1121	Қўшма сўз таркиби даг икки ёки унлан		Қўрғонтепа, контакта	
PAIR	парные существительное	Pair Noun			1122	Жуфт сўз икки сўзнинг жуфтлашиб янги маъно хосил қилиниги		бола-чака, ота-она	

3-кейс. Компютер луғатшунослиги.

Китоб луғатларни электрон луғатга ўтказинг. Ассесс дастуридан фойдаланиб, лингвистик майдонлар қандай яратилиши ҳақида маълумотга ега бўлишга ҳаракат қилинг.



4-кейс. Маъруза, амалийда ўзлаштирилган маълумотлар бўйича терминларни қуйидаги диаграммага жойлаштинг.

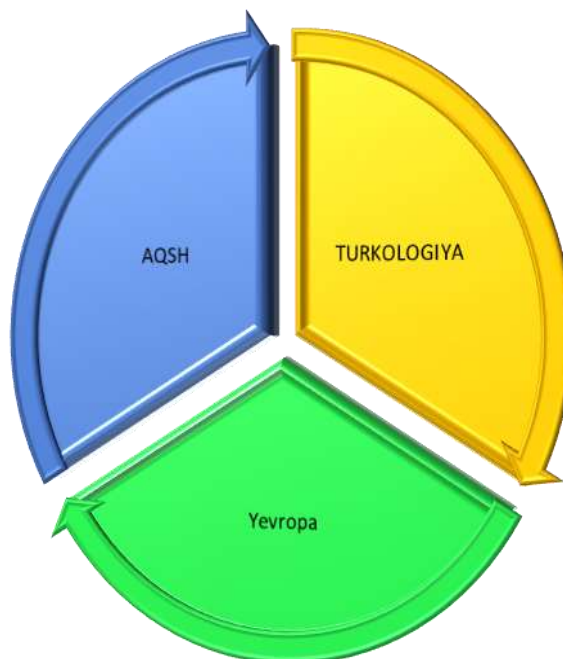


5-кейс. Муайя мавзу ёритилган икки тилда яратилган тўртта матнни (жами саккизта) қуйидаги диаграммаларга жойлаштинг. Уларни лингвистик жоҳатдан таҳлил қилинг. Гап моделларини аниқланг.

Матн (Ўзбекча)	Матн (Инглизча)	Лингвистик таҳлил

6-кейс. Интернетдан фойдаланиб муайян маттни услубига кўра таржима қилинг. Қайси сайт сифатли таржима қила олиши ва ундаги афзалликлар ҳақида фикрингизни баён етинг.

7-кейс. Машина таржимаси бўйича амалга оширилган тадқиқотларни ушбу жадвал бўйича тўлдириб чиқинг.



Тингловчилар мустақил ўқишлари учун модул хусусиятларини ҳисобга олган ҳолда қуйидаги шакллардан фойдаланиб тайёрлаш тавсия етилади:

- Ноам Чомский таълимоти ва синтактик структуралар
- Ўзбек тилининг формаллаштириш меъёрлари
- Ўзбек тилининг морфологик луғатини такомиллаштириш
- Морфологик таҳлилда қўшимчалар комбинацияси
- Ўзбек тили синтактик қурилмаларини моделлаштириш
- Луғатлар типологияси
- Моделлаштиришда морфологик ва синтактик қурилмаларнинг берилиши
- Таржимон луғатларининг семантик майдонлари таҳлили
- Корпус лингвистикаси

VII. ГЛОССАРИЙ

<p>COMPUTATIONAL MORPHOLOGY</p> <p>Компютер морфологияси</p>	<p>The area of computational linguistics covering automatic morphological analysis and generation, typically through finite-state methods.</p>	<p>Мавжуд методлар орқали автоматлашган морфологик анализ ва генерациясини қамровчи компютер лингвистикаси соҳаси</p>
<p>Лемматизация</p>	<p>The process of grouping the inflected forms of a word together under a base form, or of recovering the base form from an inflected form, e.g. grouping the inflected forms `run;`runs;`running,`ran' under the base form `run.</p>	<p>Сўзнинг тусланган шаклини унинг ўзак шакли билан биргаликда гуруҳлаш жараёни. Ёки тусланган шаклидан ўзак шаклини ажратиб олиш. Масалан: тусланган шакллар гуруҳи `рун;`рунс;`руннинг,`ранъ бўлса, ўзак шакли рун дир.</p>
<p>Токенизация</p>	<p>The process of segmenting text into linguistic units such as words, punctuation, numbers, alphanumerics etc.</p>	<p>Сўзлар, пунктуация, сонлар, алифбо рақами каби лингвистик бирликлар ичидаги матнни сегментлаш жараёни.</p>
<p>Токенизер</p>	<p>A software program that performs text tokenization and determines boundaries for individual tokens (words, numbers, punctuation) in text.</p>	<p>Матнда индивидуал белгилар (сўзлар, сонлар, пунктуация) учун чегараларни аниқлайдиган ва матн токенизатион (белгилар) тақдим қилувчи компютер дастури</p>

Генератив грамматика	generative grammar is sometimes used in linguistics to mean a grammar of a formal language; the origin of the term is this use of an automaton to define a language by generating all possible strings.	Баъзан лингвистика илмида формал тил грамматикасини англатувчи генеративе грамматика. Бу атама дастлаб ҳамма ехтимолий боғланишлар генерацияси орқали тилни автоматик таърифлашда фойдаланилган.
Корпус corpus	collection of linguistic data, either written texts or a transcription of recorded speech. Typically, corpora have to be quite large to be of any linguistic use (upwards of 100,000 tokens).	Ёзиб олинган овоз ёки ёзма матнлардан иборат лингвистик маълумотлар йиғиндиси. Одатда корпуслар тилнинг исталган маълумотини жамлагани учун жуда катта (100000 белгидан ошиқ) бўлади.
Ontology ОНТОЛОГИЯ	An inventory of the objects or processes in a domain, together with a specification of some or all of the relations that hold among them, generally arranged as a hierarchy.	Соҳавий ийеархия сифатида умумий белгиланган баъзи ёки ҳамма боғланишларни ушлаб турувчи хусусийликка ега жараёнлар ва тўлдирувчи қурилмалар мажмуи.
Тахоному таксаномия	<u>A network of transitive and disjunctive subtype or subsumption relations with the purpose of classifying.</u>	Объектнинг иархик таснифланиши ёки таснифлаш асосида асослаш.
hyponym	A word that has a more restricted meaning than another word with which it is in a hyponymy relation. For example, 'sparrow'	Умумлаштирувчи сўздан фарқли равишда хусусий маъно касб етувчи бир неча сўзлар. Масалан: чумчуқ ва канарейка қуш сўзининг хусусий

ГИПОТИМ	and `canary' are hyponyms of bird.	таркибий қисмидир.
Hypernym Умумлаштирувчи сўз	A word that has a more general meaning than another word with which it is in a hyponymy relation. Also called superordinate. For example, `bird' is a hypernym of 'sparrow' and `canary':	Аниқ бир маънони ифодаламайдиган бир неча сўзлар учун умумий ном бўлган ва умумлаштириш характериға еға сўз. Масалан: қуш сўзи чумчуқ ва канарейка учун умумий номдир
Concordance Конкорданс	A list showing all the occurrences and contexts of a given word or phrase, as found in a corpus, typically in the form of a KWIC index	Контекста берилган ва корпусда топилган барча сўз ва фразалар кўрсатилган рўйхат ўзига хос KWIC индекси.
corpus-based dictionary Корпусға асосланган луғат	A dictionary based on analysis of a large collection of texts in machine-readable form, usually by a team of lexicographers using concordances and other interactive tools.	Лексикографлар жамоасининг муофиқлиги ва бошқа интерактив қурилмалардан фойдаланиб машина ўқий оладиган шаклдаги матнларнинг катта тўпламининг таҳлилиға асосланган луғат.
COMPUTATIONAL LEXICOGRAPHY (I) Компьютер лексикографияси	Use of computers to assist in the compilation of dictionaries. (ii) use of existing dictionaries for a purpose for which they were not designed, namely applications in natural language processing by computer.	Луғатлар компликациясида компьютер ёрдамидан фойдаланиш. Илова сифатида тақдим етилмаган луғатлардан компьютер воситасида табиий тилни қайта ишлаш жараёнида фойдаланиш.

<p>COMPUTER-ASSISTED LANGUAGE LEARNING</p> <p>Kompyuter yordamida til o‘rganish</p>	<p>Any use of computers to provide language instruction or to support language learning.</p>	<p>Тилни ўрганишга ёрдам берувчи тил билимлари ва уни ўргатувчи йўриқнома билан таъминланган компьютер дастури.</p>
<p>CONTEXT (I)</p> <p>Контекст</p>	<p>In corpus linguistics and text linguistics, the parts of a text that surround a particular word or phrase, often providing clues to its meaning; (ii) in discourse analysis and pragmatics, the non-linguistic world situation in which a particular utterance is uttered, sometimes affecting its form and meaning; (iii) in phonetics, the speech sounds that surround a particular phoneme and affect its realization.</p>	<p>Лингвистик корпус ва матн корпусида маънони англатишда калит вазифасини бажарувчи сўз ва фразаларга ега матн парчаси; прагматис ва дискурс анализида талаффуз қилинганда базида сўз шакли ва маъносига таъсир қиладиган лингвистик бўлмаган ҳолатлар; фонетикада маъно англатувчи нутқ товушлари тузилмаси</p>
<p>corpus annotation</p>	<p>corpus annotation See annotation.</p>	<p>Корпуснинг лингвистик хусусиятлари, боғланиш усуллари шарҳлаш жараёни.</p>
<p>Dictionary</p>	<p>A collection of words and phrases with information about them. Traditional dictionaries contain explanations of spelling, pronunciation, inflection, word class (part of speech), word origins, word meaning, and word use. However, they do not provide much information about the</p>	<p>Сўз ёки фраза ва улар ҳақидаги маълумотлар йиғиндиси. Ананавий луғатлар сўзнинг шакли, талаффузи, тусланиши, сўз туркуми, етималогияси, маъноси ва қўлланилиш доираси киради. Бироқ, уларда маъно ва фойдаланиш билан боғлиқ маълумот кўп берилмайди.</p>

Луғат	relationships between meaning and use. A dictionary for computational purposes (often called a lexicon) rarely says anything about word origin, and may say nothing about meaning or pronunciation either.	Ҳисоблаш мақсадидаги луғатларда (лексикон) деб ҳам аталади) сўзнинг келиб чиқиши ва маъно ва талаффузи ҳақида деярли сўз юритилмаслиги мумкин.
Discourse Дискурс	An extended sequence of sentences produced by one or more people with the aim of conveying or exchanging information	Бир ёки бир неча кишининг маълумот алмашиш ёки узатиш учун айтган гапларини кетма-кетлигини кенгайтириш.
discourse model Дискурс модели	For a natural language interaction system, an encapsulation of the events and objects mentioned in the discourse so far and information as to how they relate to one another.	Табиий тилнинг ички тизимида дискурсдаги воқеа ва объектларнинг инкасуляциясига урғу берилиши ва уларнинг бир-бирига қандай боғланганлиги ҳақидаги маълумот.
discourse plans Дискурс режалари	In dialogue, task-independent behavioural patterns regulating exchanges between agents in general. Compare domain plans.	Диалогда умумий тарзда иштирокчилар орасида мустақил вазифа характеридаги наъмуналарни мунтазам алмашиниши
discourse segment Дискурс сегменти	Any separable part of a discourse. A discourse segment will typically be concerned with a sub-part of the topic covered by the discourse as a whole.	Дискурснинг бирор бир ажраладиган қисми. Дискурс сегментининг ўзига хослиги бутун дискурсни қамрам олган мавзунинг маълум бир қисми билан шуғулланади.

VIII. АДАБИЁТЛАР РЎЙХАТИ

I. Ўзбекистон Республикаси Президентининг асарлари.

1. Мирзиёев Ш.М. Танқидий таҳлил, қатъий тартиб-интизом ва шахсий жавобгарлик – ҳар бир раҳбар фаолиятининг кундалик қонидаси бўлиши керак. Ўзбекистон Республикаси Вазирлар Маҳкамасининг 2016 йил якунлари ва 2017 йил истиқболларига бағишланган мажлисидаги Ўзбекистон Республикаси Президентининг нутқи // Халқ сўзи. 2017. 16 январь, №11.

2. Мирзиёев Ш.М. Эркин ва фаровон, демократик Ўзбекистон давлатини биргаликда барпо этамиз. Ўзбекистон Республикаси Президенти лавозимига киришиш тантанали маросимига бағишланган Олий Мажлис палаталарининг кўшма мажлисидаги нутқ. Т., 2016. 56 б.

3. Мирзиёев Ш.М. Танқидий таҳлил, қатъий тартиб-интизом ва шахсий жавобгарлик – ҳар бир раҳбар фаолиятининг кундалик қонидаси бўлиши керак. Мамлакатимизни 2016 йилда ижтимоий-иқтисодий ривожлантиришнинг асосий якунлари ва 2017 йилга мўлжалланган иқтисодий дастурнинг энг муҳим устувор йўналишларига бағишланган Вазирлар Маҳкамасининг кенгайтирилган мажлисидаги маъруза, 2017 йил 14 январь. Т.: Ўзбекистон, 2017. 104 б.

4. Мирзиёев Ш.М. Қонун устуворлиги ва инсон манфаатларини таъминлаш-юрт тараққиёти ва халқ фаровонлигининг гарови. Ўзбекистон Республикаси Конституцияси қабул қилинганининг 24 йиллигига бағишланган тантанали маросимдаги маъруза. 2016 йил 7 декабрь. Т.: Ўзбекистон, 2017. 48 б.

5. Мирзиёев Ш.М. Буюк келажагимизни мард ва олижаноб халқимиз билан бирга курамиз. Мазкур китобдан Ўзбекистон Республикаси Президенти Шавкат Мирзиёевнинг 2016 йил 1 ноябрдан 24 ноябрга қадар Қорақалпоғистон Республикаси, вилоятлар ва Тошкент шаҳри сайловчилари вакиллари билан ўтказилган сайловолди учрашувларида сўзлаган нутқлари ўрин олган. Т.: Ўзбекистон, 2017. 488 б.

II. Норматив- ҳуқуқий ҳужжатлар

3. Ўзбекистон Республикасининг Конституцияси (Ўн иккинчи чақирик Ўзбекистон Республикаси Олий Кенгашининг ўн биринчи сессиясида 1992 йил 8 декабрда қабул қилинган Ўзбекистон Республикасининг 1993 йил 28 декабрдаги, 2003 йил 24 апрелдаги, 2007 йил 11 апрелдаги, 2008 йил 25 декабрдаги, 2011 йил 18 апрелдаги, 2011 йилдаги 12 декабрдаги, 2014 йил 16 апрелда қабул қилинган қонунларига мувофиқ киритилган ўзгартиш ва кўшимчалар билан). Т., 2014.

4. Ўзбекистон Республикасининг “Таълим тўғрисида”ги Қонуни / Ўзбекистон Республикаси Олий Мажлисининг Ахборот-номаси. 1997. 9-сон, 225-модда.

5. Кадрлар тайёрлаш Миллий дастури / Ўзбекистон Респуб-ликаси Олий Мажлисининг Ахборотномаси. 1997. 11-12-сон, 295-модда.

6. Ўзбекистон Республикаси Президентининг 2010 йил 28 июлдаги “Таълим муассасаларининг битирувчиларини тадбир-корлик фаолиятига жалб этиш борасидаги кўшимча чора-тадбирлар тўғрисида”ги Фармони.

7. Ўзбекистон Республикаси Президентининг 2010 йил 2 ноябрдаги “Олий малакали илмий ва илмий-педагогик кадрлар тайёрлаш тизимини янада такомиллаштириш чора-тадбирлари тўғрисида”ги ПҚ-1426-сонли Қарори.

8. Ўзбекистон Республикаси Президентининг 2011 йил май-даги “Олий таълим муассасаларининг моддий-техник базасини мустаҳкамлаш ва юқори малакали мутахассислар тайёрлаш сифатини тубдан яхшилаш чора-тадбирлари тўғрисида”ги ПҚ-1533-сонли Қарори.

9. Ўзбекистон Республикаси Президентининг 2012 йил 24 июлдаги “Олий малакали илмий ва илмий-педагог кадрлар тайёрлаш ва аттестациядан ўтказиш тизимини янада такомил-лаштириш тўғрисида”ги ПФ-4456-сонли Фармони.

10. Ўзбекистон Республикаси Вазирлар Маҳкамасининг 2012 йил 28 декабрдаги “Олий ўқув юртидан кейинги таълим ҳамда олий малакали илмий ва илмий педагогик кадрларни аттестациядан ўтказиш тизимини такомиллаштириш чора-тадбирлари тўғрисида”ги 365-сонли Қарори.

11. Ўзбекистон республикаси Президентининг 2016 йил 13 майдаги “Алишер Навоий номидаги Тошкент Давлат ўзбек тили ва адабиёти университетини ташкил этиш тўғрисида”ги ПФ-4797 сонли Фармони.

12. Ўзбекистон Республикаси Президентининг 2017 йил 20 апрелдаги “Олий таълим тизимини янада яхшилаш ва такомиллаштириш тўғрисида”ги ПҚ-2909 сонли Қарори.

13. Ўзбекистон Республикаси Президентининг 2019 йил 27 августдаги “Олий таълим муассалари раҳбар ва педагог кадрларининг узлуксиз малакасини ошириш тизимини жорий этиш тўғрисида”ги ПФ-5789 сонли Фармони

14. Ўзбекистон Республикаси Вазирлар Маҳкамасининг 2019 йил 23 сентябрдаги “Олий таълим муассалари раҳбар ва педагог кадрларининг малакасини ошириш тизимини янада такомиллаштириш бўйича кўшимча чора-тадбирлар тўғрисида”ги №-797 сонли Қарори

III. Махсус адабиётлар

1. Журафский Д., Ж.Х. Мартин. Speech and Language Processing: An introduction to natural language processing, computational linguistics, and speech recognition. 2006 (Электрон вариант)
2. Митков Р. The Oxford handbook of Computational linguistics. Oxford university press, 2003 (Электрон вариант)
3. Пулатов А. Компьютер лингвистикаси. – Тошкент: Академнашр, 2011.
5. Раҳимов А. Компьютер лингвистикаси асослари. – Тошкент: Академнашр, 2011.
6. Абдурахмонова Н.З. Компьютер лингвистикаси (Электрон дарслик), 2019

7. Абдурахмонова Н.З. Машина таржимасининг лингвистик таъминоти. – Тошкент: Мухаррир, 2018.
7. www.ruscorpora.ru
8. [www.philol.mmsu.ru / rus / chairis / genis / collabor / marc 1/htm](http://www.philol.mmsu.ru/rus/chairis/genis/collabor/marc1/htm).
9. [www.ipcs.math.msu.ru. / rus / mlogic.htm](http://www.ipcs.math.msu.ru/rus/mlogic.htm).
10. www.uzcl.com

IV.Электрон таълим ресурслари

1. www.ziyonet.uz
2. www.edu.uz
3. Infocom.uz электрон журнали: www.infocom.uz
4. www.tdpu.uz
5. www.pedagog.uz
6. www.nuuz.uz
7. www.bimm.uz
8. www.literature.uz
9. www.kutubxona.uz
- 10.[http:// www.dissert.h10.ru/pedagog.html](http://www.dissert.ru/pedagog.html)
- 11.<http://www.istedod.uz/>
- 12.www.scopus.com