

**O‘ZBEKISTON RESPUBLIKASI OLIIY VA
O‘RTA MAXSUS TA‘LIM VAZIRLIGI**

FARG‘ONA DAVLAT UNIVERSITETI

SIRTQI BO‘LIM

IJTIMOIIY-GUMANITAR FANLAR KAFEDRASI

“TASDIQLAYMAN”

O‘quv ishlari bo‘yicha prorektor

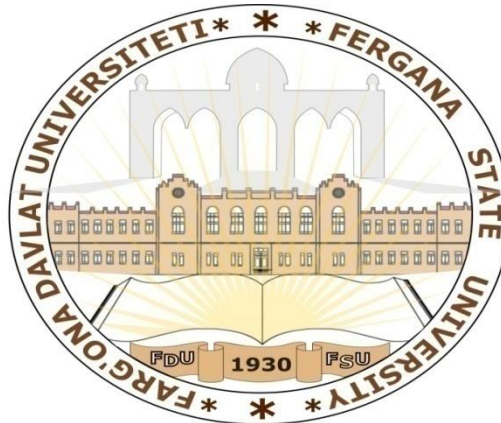
R.Shukurov

2022 yil “__” _____

“KOMPYUTER LINGVISTIKASI”

fanidan

O‘QUV-USLUBIY MAJMUA



FARG‘ONA – 2023

Annotatsiya: ushbu o'quv-uslubiy majmuada – “Kompyuter lingvistikasi” kursining vazifasi, mazmun va mundariyasi, zaruriy ilmiy adabiyotlari, til strukturasi, tarixiy taraqqiyoti, uning jamiyatdagi vazifasi, tilshunoslik fanining asosiy bo'limlari va bu bo'limlarning tekshirish sohalari xususidagi muhim masalalar uslubiy yondashuvlar asosida yoritilgan. O'quv rejasidagi har bir mavzuga oid aniq tavsiyalar berilgan.

Mazkur uslubiy tavsiya Farg'ona davlat universiteti ikkinchi kurs talabalariga mo'ljallangan, u material nuqtayi nazaridan bir necha yildan buyon o'qib kelinayotgan ma'ruzalarga asoslanmoqda.

Tuzuvchilar:

- Zokirov M.T. – Tilshunoslik kafedrasi professori, f.f.n.
Shokirova X.N – Ijtimoiy-gumanitar fanlar kafedrasi dotsenti, f.f.d.

Taqrizchilar:

Hakimov M.X. – Farg'ona davlat universiteti professori, filologiya fanlari doktori.

Abduvaxobov G'. – Tilshunoslik kafedrasi o'qituvchisi.

O'quv-uslubiy majmua Farg'ona davlat universiteti Ilmiy kengashida ko'rib chiqilgan va nashrga tavsiya qilingan (20__ yil _____dagi «___» bayonnoma).

© FarDU, 2023

O'QUV-USLUBIY MAJMUA

TARKIBI

1.	O'QUV MATERIALLARI	4
2.	MUSTAQIL TA'LIM MASHG'ULOTLARI	69
3.	GLOSSARIY	70
4.	ILOVALAR:	
4.1	FAN DASTURI	89
4.2	ISHCHI DASTUR	94
4.3	TARQATMA MATERIALLAR	10
		4
4.4	BAHOLASH MEZONI	11
		9

1. O'QUV MATERIALLARI

Ma'ruza № 1

“Kompyuter lingvistikasi” faniga kirish

Reja

1. Kompyuter lingvistikasi fanining yuzaga kelish asoslari.
2. Kompyuter lingvistikasi va matematik lingvistika.
3. Kompyuter lingvistikasi fanining tekshirish obyekti, maqsad va vazifalari.
4. Kompyuter lingvistikasi va klassik tilshunoslik.

Til sistema sifatida aynan matematikaning o'zidir.

Matematika aniq fan bo'lib, formulalar munosabati orqali reallashadi. Til ham matematika singari aniqlikni talab qiladi va unda ikki yoki undan ortiq elementlarning o'zaro munosabati bir butunlikni tashkil etadi.

Beryozin

Tayanch so'z va iboralar: til algebrasi, matematik lingvistika, kompyuter lingvistikasi, mashina tarjimasi, metatil, sun'iy til, deskriptiv, kvantitativ, muloqotning optimal modeli, analiz, sintez, NLP, injener lingvistikasi, generatsiya, tilshunoslik va adabiyotshunoslik sohalarida kompyuterdan unumli foydalanish, tillarga o'qitish, bilimlarini baholash, matnlarni tahrir qilish, tabiiy tillarning matematik modeli, o'zbek tilining kompyuter uslubi, matnlar korpusi, elektron lug'at, tezaurus.

Ma'lumki, tilshunoslik fani XIX (1816 yilda) asrda mustaqil fan sifatida shakllandi. Shundan boshlab u turli aspektlarda, yo'nalishlarda rivojlanib kelmoqda. Keyingi yillarda barcha fanlardagi kabi tilshunoslikda ham ikki fanning „chorrahasida“ (kesishuvida) yuzaga kelgan fanlar jadal rivojlanmoqda. Jumladan, ana shunday fanlar sirasiga sotsiolingvistika (sotsiologiya va tilshunoslik), psixolingvistika (psixologiya va tilshunoslik), etnolingvistika (etnografiya va lingvistika), neyrolingvistika (nevrologiya va tilshunoslik), matematik lingvistika va kompyuter lingvistikasi fanlarini kiritish mumkin. Bunday holni boshqa fanlar doirasida ham kuzatish mumkin: bioximiya, astrofizika, matematik fizika, matematik logika kabi. Buni fanlar tizimida bir necha fanlarning o'zaro hamkorligi, integratsiyasi deb baholash lozim bo'ladi.

XX asrning 50-yillaridan boshlab tilshunoslikda "**mashina tarjimasi**", "**mashina tilshunosligi**" atamaları qo'llanila boshlandi.¹ Bu asrning buyuk kashfiyoti bo'lgan kompyuter texnologiyalari tilshunoslikka ham kirib kelganining isboti edi.² Mashina tarjimasi yoki avtomatik tarjima deyilganda bir tildagi matnni ikkinchi bir tilga EHM (kompyuter) vositasida, tez vaqt ichida tarjima qilish nazarda tutiladi. Mashina tarjimasining asoschilari kibernetika va matematika sohasi vakillari bo'lib, keyinchalik bu ishda tilshunoslar ham faol qatnasha boshlagan. Shu tariqa mashina tarjimasi g'oyalari butun dunyoda nazariy va amaliy tilshunoslikning rivojlanishida katta ahamiyat kasb etdi. Bu yo'nalish bilan parallel ravishda formal grammatika nazariyasi yuzaga kelib, til va uning alohida aspektlari modelini yaratishga e'tibor qaratildi. Tilning bu jihatlari matematik lingvistika fanida ishlab chiqilib, bu o'z navbatida, kompyuter lingvistikasi fanining yuzaga kelishi uchun poydevor bo'ldi. Demak, shu asosda

¹ Grishman R. Computational linguistics. Cambridge University Press, 1994. – P. 4.

² Po'latov A., Muhamedova S. Kompyuter lingvistikasi. – Toshkent, 2007. – B. 4-5.

tilshunoslikning yangi yo'nalishi bo'lgan kompyuter lingvistikasi va tilshunoslikning bir qator nazariy va amaliy yo'nalishlari vujudga keldi.

Matematik lingvistika fani XX asrning 50-yillarida (1952 yilda) tilshunoslikning alohida yo'nalishi sifatida yuzaga keldi. Bu fanning shakllanishida Kopengagen struktural tilshunoslik maktabi (glossematika)ning asoschisi Lui Yelmslevning g'oyalari o'ziga xos „turtki“ vazifasini o'tagan. U hatto til hosidalarini matematik bayonda tushuntiradigan fanning nomini ham taklif etgan. Olimning fikricha, bu fan „Til algebrasi“ („Lingvistik algebra“) deb atalishi lozim edi.³ Amerikalik tilshunos Noam Chomskiyning formal grammatika, transformatsion grammatika haqiagi qarashlari bevosita matematik lingvistikaning alohida yo'nalish sifatida yuzaga kelishiga sabab bo'lgan.⁴ Mana shunday qarashlar ta'sirida matematik lingvistika fani shakllandi. Matematik lingvistika bu – tabiiy tillarning matematik modellarini (bunday formallashtirilgan til metatil deb ataladi) ishlab chiqish, xususan, sun'iy tillarni yaratish algoritmini tuzish bilan shug'ullanuvchi fandır. Matematik lingvistika oldida turuvchi eng muhim masalalar quyidagilardir:

- tilning aksiomatik nazariyasini ishlab chiqish;
- formal grammatika yaratish;
- tillarning matematik modellarini ishlab chiqish.

Har bir fanning o'z maqsad va vazifalari bo'ladi. Matematik lingvistika fanining asosiy maqsadi tabiiy tillarning matematik modelini ishlab chiqishdir. Ushbu maqsadga erishish uchun fan o'z oldiga quyidagi vazifalarni qo'yadi:

- tabiiy va sun'iy tillarning formal modellari algoritmini ishlab chiqish;
- lisoniy hodisalarni matematik parametrlarda baholash;
- til hodisalarini matematik metodlar yordamida tahlil qilish (*ehtimollar nazariyasi, statistika va kvantitativ metodlarni tatbiq etish*).

Kompyuter lingvistikasi matematik lingvistikaning mantiqiy davomi bo'lib, u amaliy tilshunoslikning eng muhim qismini tashkil etadi. Kompyuter lingvistikasi 1954 yil AQSHda Jorjtaun universitetida mashina tarjimai bo'yicha dunyoda o'tkazilgan birinchi tajriba asnosida yo'nalish sifatida shakllana boshladi, 1960 yilga kelib mustaqil fan sifatida shakllandi. Kompyuter lingvistikasi inglizcha „computational linguistics“ so'zining kalkasidir. XX asrning 80-yillariga qadar bu fan turlicha nomlar bilan atalgan: *hisoblash lingvistikasi, matematik lingvistika, kvantitativ lingvistika, injener lingvistikasi* kabi. Bu fanning asosiy maqsadi lingvistik masalalarni yechishning kompyuter dasturlarini ishlab chiqish, inson va mashina (kompyuter) muloqotini optimallashtirish⁵, tabiiy tilni qayta ishlash (**Natural Language Processing**)dir.⁶ NLP kompyuter lingvistikasida tabiiy tillarning kompyuter analizi va sintezini o'z ichiga oladi. Bunda *analiz* tabiiy tilning kompyuterda morfologik, sintaktik va semantik tahlil yordamida tushunilishiga nisbatan ishlatiladi, *sintez* esa kompyuterda matnning grammatik shakllantirilishi va *generatsiyasi* (hosil qilinishi) demakdir. NLP bo'yicha yaratilgan dasturiy ta'minotlar quyidagilar: [AlchemyAPI](#), [Expert System S.p.A.](#), [General Architecture for Text Engineering \(GATE\)](#), [Modular Audio Recognition Framework](#), [MontyLingua](#), [Natural Language Toolkit \(NLTK\)](#).

³ Нурмонов А. Структур тилшунослик: илдиэлари ва йўналишлари. – Тошкент, 2008. – Б. 145.

⁴ Гладкий А.В., Мельчук И.А. Элементы математической лингвистики. – М.: Наука, 1969. – С.7.

⁵ Новое в зарубежной лингвистике. Вып.24. Компьютерная лингвистика. – М.: Прогресс, 1989. – С.10.

⁶ Jurafsky D., Martin J.H. Speech and Language Processing. – New Jersey, 2000. – P. 2-3.

Kompyuter lingvistikasining asosiy vazifalariga quyidagilar kiradi:

- avtomatik o'qitish tizimini ishlab chiqish;
- bilimlarni tekshirish;
- matnlarni turli jihatdan avtomatik tahrirlash;
- matnlarni avtomatik tarzda morfologik, sintaktik va semantik tahlilini (inglizcha **parsing**) ta'minlovchi tizimlar yaratish (**parser** dasturlar);
- mashina tarjimai uchun mo'ljallangan dasturlarni ishlab chiqish;
- lug'atlarni va kompyuterdagi matnni statistik tahlil qilish;
- lingvistik muammolarni hal qilishga yo'naltirilgan optimal dasturlar yaratish;
- muloqotning kompyuter modelini ishlab chiqish;
- matn strukturasi gipertekst texnologiyasini yaratish;
- elektron lug'atlar va tezauruslar yaratish;
- sun'iy intellekt tizimini ishlab chiqish;
- matnlar korpusini yaratish;
- matnni avtomatik referatlash;
- syujet strukturasi modellash;
- nutqni avtomatik tushunish tizimini ishlab chiqish (inglizcha **ASR – Automatic Speech Recognition**);
- matndan faktlarni avtomatik ajratib olish (inglizcha **fact extraction, text mining**).

Kompyuter lingvistikasida qo'llaniladigan va yaratiladigan lingvistik vositalarni shartli ravishda ikkiga qismga bo'lish mumkin: deklarativ hamda protsedura qismlari. Deklarativ qismga til va nutq birliklari lug'ati, grammatik ma'lumotnomalar, matnlar korpusi kabilarni kiritish mumkin. Protseura qismi esa yuqoridagi lingvistik ta'minot bazasini boshqarish vositalarini (algoritm tuzish, dasturlar yaratish, kompyuter analizi va sintezi kabilari) o'z ichiga oladi.

O'zbek tili materiallari bo'yicha kompyuter lingvistikasiga oid tadqiqot olib borgan olimlar sifatida S.Rizaev va S.Muhamedov, [A.Po'latov](#), [S.Muhamedova](#), [N.Jo'rayevalarni](#) alohida ko'rsatish mumkin. Olimlar asosan statistik tahlil, algoritm tuzish, o'zbek tilining aksiomatik nazariyasi, fe'llarning kompyuter analizi va sintezi kabi yo'nalishlar bo'yicha izlanishlarni amalga oshirganlar. O'zbek tilshunosligidagi kompyuter lingvistikasining boshqa yo'nalishlari esa o'z tadqiqotchilarini kutmoqda. Shunga ko'ra, bugungi kunda o'zbek tilshunosligining kompyuter bilan bog'liq holda hal etilishi lozim bo'lgan quyidagi vazifalarni ta'kidlab ko'rsatish mumkin:⁷

1. O'zbek tilining kompyuter uslubini yaratish.
2. Axborot matnlaridagi qoliplilik, qisqalik standartlarini ishlab chiqish.
3. Kompyuter izohli va tarjima lug'atlarini yaratish.
4. O'zbek tili va adabiyoti darsliklarining elektron versiyalarini ishlab chiqish.
5. Kompyuterda inglizcha-o'zbekcha tarjima dasturlarini ishlab chiqish.
6. Kompyuterdagi matnlarni avtomatik tahrirlash dasturlarini yaratish.
7. O'zbek tili grammatikasining kompyuter modelini yaratish.
8. "Inglizcha-o'zbekcha tarjimon" kompyuter dasturini ishlab chiqish.

⁷ Po'latov A., Muhamedova S. Kompyuter lingvistikasi. – Toshkent, 2007. – B. 5-6.

9. O'zbek tilidagi matnlarni tahrir qilish dasturlarini yaratish.
10. O'zbek tilidagi "Windows" operatsion tizimini yaratish.
11. Inglizcha-o'zbekcha, o'zbekcha-inglizcha kompyuter lug'atlarini ishlab chiqish.

Kompyuter lingvistikasi amaliy tilshunoslikning tarkibiy qismi hisoblanadi, u nazariy tilshunoslik yutuqlaridan foydalanadi, oziqlanadi. Kompyuter lingvistikasi va nazariy tilshunoslik bir-birini to'ldiradi. Kompyuter lingvistikasida til insondan tashqarida mavhum sistema sifatida, L.Yelmslev ta'rifi bilan aytganda "sof munosabatlar tizimi" sifatida tavsiflanadi. Kompyuter lingvistikasi va klassik (mumtoz) tilshunoslik orasidagi farq quyidagi nuqtalarda ko'rinadi:

- Mumtoz tilshunoslikda til inson bilan mutanosiblikda, juftlikda ko'rib chiqiladi. Ya'ni mumtoz tilshunoslik insonga yo'naltirilgan bo'ladi va uning faol ishtirokida qabul qilinadi. Kompyuter lingvistikasi esa tavsiflash jarayonida insonni istisno qiladi va u ko'proq kompyuterga moslashtiriladi.

- Mumtoz tilshunoslik ko'proq tavsifiy (deskriptiv) xarakterga ega hisoblanadi. Kompyuter lingvistikasi esa masalani miqdoriy xarakteristikalar va aniq parametrlar asosida hal qiladi. Demak, mumtoz tilshunoslik ko'proq tavsifiy bayonga asoslansa, kompyuter lingvistikasi miqdoriy (kvantitativ) tavsifga, algoritmlash, modellash, statistik tahlilga asoslanadi.

- Mumtoz tilshunoslik ko'proq nazariy xarakterga ega bo'lib, tilshunoslikning nazariy masalalari bilan shug'ullanadi. Kompyuter lingvistikasi esa amaliy harakterga ega bo'lib, til bilan bog'liq muammolarning amaliy jihatlariga e'tibor qaratadi hamda uni aniq maqsadga yo'naltirilgan dasturlar, metodlar, tizimlar yordamida hal etish bilan shug'ullanadi.

- Mumtoz tilshunoslikning tahlil obyekti tabiiy til hisoblanadi va tahlilda uning mavjud barcha imkoniyatlari e'tiborga olinadi. Kompyuter lingvistikasida esa ko'proq sun'iy tillarga (programmalashtirish tillari, algoritmik tillar)ga tayaniladi, tabiiy tillarning mavjud imkoniyatlari cheklanadi, bunda tabiiy tilga ishlov berilib (**NLP**), kompyuterga moslashtiriladi.

Fanlar hamisha o'zaro uzviy bog'liqlik va hamkorlikda ish ko'radi, ular bir-birisiz yashay olmaydi. Jumladan, kompyuter lingvistikasi fani ham bundan mustasno emas. Kompyuter lingvistikasi mantiq, informatika, statistika, semiotika, kibernetika, ehtimollar nazariyasi kabi turli fan sohalari bilan o'zaro aloqadorlikda ish ko'radi.

Ma'ruza № 2
KOMPYUTER LINGVISTIKASINING ASOSIY YO'NALISHLARI
Reja

1. Kompyuter lingvistikasi va uning asosiy yo'nalishlari.
2. Kompyuter lingvistikasiga oid nazariy va amaliy tadqiqotlar.
3. Lingvistik programmalar va ularning yo'nalishlari.

Tayanch so'z va iboralar: *kompyuter lingvistikasi, mashina tarjimai, avtomatik tahrir, avtomatik o'qitish tizimi, matning statistik tahlili, lingvostatistika, injener lingvistikasi, chastotali lug'at, o'qituvchi lingvistik avtomat, lingvistik ta'min, lingvistik programmalar.*

Kompyuter lingvistikasi fani 1960 yilda AQSHda shakllandi. Mazkur fan XX asrning o'rtalarida paydo bo'lgan matematik lingvistika fani asoslaridan kelib chiqqan. Tabiiy tillar grammatikasi formal modeli kompyuterda tarjima qilish, tillarga o'rgatish, bilimlarni diagnostika qilish, matnlarni tahrirlash kabi amaliy dasturlarni ishlab chiqishga asos bo'lgani holda kompyuter lingvistikasi yo'nalishiga zamin yaratdi.⁸ Kompyuter lingvistikasi fanining asosiy maqsadi lingvistik masalalarni hal qilishning kompyuter dasturlarini yaratishdir.⁹ Uning asosiy vazifalariga tillarga o'qitish, bilimlarni tekshirish, matnlarni turli jihatdan tahrirlash va mashina tarjimai uchun mo'ljallangan dasturlarni ishlab chiqish kabilar kiradi.

Kompyuter lingvistikasi amaliy tilshunoslikning tarkibiy qismi, matematik lingvistikaning mantiqiy davomi bo'lib, u hozirda turli aspektlarda jadal rivojlanmoqda. Kompyuter lingvistikasi fanining asosiy yo'nalishlari sifatida quyidagilarni ajratish mumkin.

- a) Matn tahriri – muayyan matnni kompyuter yordamida xatolarni tog'rilash.
- b) Avtomatik tarjima – berilgan matnni bir tildan boshqa tilga qisqa vaqt ichida tezkor tarjima qilish tizimi.
- c) CALL (Computer Assisted Language Learning) – til o'qitishni avtomatlashtirish - xorijiy tillarning kompyuter yordamida o'qitilish tizimini ta'minlovchi amaliy dasturlar hamda bilimlarni avtomatik baholash mexanizmlari.
- d) Kompyuter leksikografiyasi (Elektron lug'atlar tuzish) – muayyan til leksikonidagi so'zlarni ma'lumotlar bazasi va maxsus programmashtirish tillari yordamida ishlash tizimiga ega bo'lgan dasturlar (kompyuter lug'atlari) yaratish bilan shug'ullanuvchi soha.
- e) Sun'iy intellekt tizimini yaratish.
- f) NLP – tabiiy tilning qayta ishlanishi
- g) Gipertekst texnologiyasi

⁸ Po'latov A., Muhamedova S. Kompyuter lingvistikasi. – Toshkent, 2007. – B.37-38.

⁹ Новое в зарубежной лингвистике. Вып.24. Компьютерная лингвистика. -М.: Прогресс, 1989. - С.10.

h) On-line o'qitish tizimi (virtual sinf va laboratoriyalarda) – kompyuter lingvodidaktikasi

i) Informatsion qidiruv tizimi.

Xorijda kompyuter lingvistikasi bo'yicha olib borilgan tadqiqotlarning salmog'i va qamrovi juda keng. Ayniqsa, AQShda kompyuter lingvistikasi bo'yicha tematik to'plamlar, «Computational Linguistics» («Kompyuter tilshunosligi») nomli jurnal nashr etiladi. AQShda kompyuter tilshunosligi assotsiatsiyasi mavjud bo'lib (**Association of Computational Linguistics**), bu uyushma kompyuter tilshunosligi bilan bog'liq ko'plab tashkiliy va ilmiy ishlarni amalga oshiradi. AQShda ikki yilda bir marta **COLING** konferensiyasi o'tkaziladi. AQShning 10 ta universitetida kompyuter lingvistikasi fakultetlari mavjud. Kompyuter tilshunosligining muammoli masalalari sun'iy intellekt bo'yicha o'tkaziladigan turli xalqaro konferensiyalarda ham ko'rib chiqiladi. Rossiyada har yili kompyuter lingvistikasiga oid “**ДИАЛОГ**” xalqaro konferensiyasi o'tkaziladi.¹⁰ Shuningdek, Niderlandiyada “**Artificial Intelligence**” (“Sun'iy intellekt”) jurnali ham chop etiladi. Mexiko shahrida har yili 10 yildan beri “**Computational Linguistics and Intelligent Text Processing**” nomli xalqaro konferensiya o'tkazib kelinadi.¹¹

Rus kompyuter tilshunosligidagi tadqiqotlarning quyidagi yo'nalishlarini alohida ajratib ko'rsatish mumkin:

- Mashina tarjimasi (MT).
- Avtomatik tahrirlash (AT).
- Til o'rgatish jarayonini kompyuterlashtirish (avtomatik tarzda tilga o'qitish - ATO').
- Statistik tadqiqotlar (ST).

Kompyuter lingvistikasining mashina tarjimasi yo'nalishida amalga oshirilgan tadqiqotlar bugungi kunda barcha kompyuterlarda mavjud bo'lgan ruschadan boshqa tillarga tarjima qiluvchi dasturlarning yaratilishiga asos bo'ldi. Xususan, L.L.Nelyubin va uning shogirdlari tomonidan ishlab chiqilgan so'zma-so'z tarjima qilishning ilmiy strategiyasi (rasmiy-idoraviy uslub asosida) o'sha davr jahon injener lingvistikasi talablari va ilg'or tajribasiga monand bo'lgan edi.¹²

L.L.Nelyubin tadqiqotlarining quyidagi o'ziga xos jihatlarini ta'kidlash lozim: birinchidan, ushbu ishlarda so'z va iboralarning avtomatik lug'ati bilan birgalikda to'liq tarjima qilish imkoniyatini beruvchi morfologik-sintaktik algoritm-freym yaratishga harakat qilingan; ikkinchidan, vujudga keltirilgan tarjima dasturlari asosida o'qituvchi lingvistik avtomat (ОЛА - Обучающий лингвистический автомат) yaratish me'yorlari ishlab chiqilgan. Ya'ni u o'quvchilarga ingliz tilidagi rasmiy hujjatlarni rus tiliga tarjima qilishni o'rgatish imkoniyatini beradi. Professor L.L.Nelyubinning ilmiy faoliyatida mashina

¹⁰ <http://www.dialog-21.ru/>

¹¹ <http://www.CICLing.org>

tarjimasi uchun (ingliz tili bo'yicha) algoritmlar va dasturlar ishlab chiqish asosiy o'rinni egalladi. L.L.Nelyubin mashina tarjimasida ingliz tili leksikasini butunligicha qamrab olmagan, u o'z tadqiqotlarining ob'ekti sifatida AQSh harbiy hujjatlarini belgiladi va bu sohada ko'plab ilmiy asarlar yaratdi. Jumladan, olimning *"Перевод боевых документов армии США"* (1971), *"Частотный англо-русский военный словарь-минимум"* (1974), *"Перевод и прикладная лингвистика"* (1983), *"Компьютерная лингвистика и машинный перевод"* (1991) kabi monografik tadqiqotlari kompyuter lingvistikasi fani bo'yicha fundamental ishlar hisoblanadi.

Shuningdek, Yu.N.Marchuk tadqiqotlarida ham bir tabiiy tildan boshqa tilga bo'lgan tarjimani modellashtirish tamoyillari, tarjima birligi, tarjima jarayonining statikasi va dinamikasi hamda EHM ishtirokidagi tarjimaning texnologik liniyalari tavsiflangan. Olimning nazariy qarashlari va g'oyalari quyidagi asarlarida o'z ifodasini topgan: *"Некоторые принципы автоматизации перевода с немецкого языка на русский"* (1980), *"Проблемы машинного перевода"* (1983), Yu.V.Rojdestvenskiy bilan hammualiflikda yaratgan *"Введение в прикладную филологию"* (1987) monografiyalari, shuningdek, *"Теория и практика машинного перевода"*, *"Модель «текст-текст» и переводные соответствия в теории машинного перевода"* nomli maqolalari. Yu.N.Marchuk ham tarjimani modellashtirish metodlari va uni avtomatlashtirish yo'llarini ishlab chiqqani holda bugungi rus kompyuter lingvistikasiga asos solgan olimlardan hisoblanadi.

L.L.Nelyubin va Yu.N.Marchuklarning mashina tarjimasi bo'yicha olib borgan tadqiqotlaridan ta'sirlangan va ilhomlangan holda ko'plab rus va ukrain olimlari ham kompyuter lingvistikasiga oid qator izlanishlar olib bordilar. Natijada bir nechta monografiyalar yaratildi, o'nlab dissertatsiyalar himoya qilindi, ilmiy maqola va tezislari e'lon qilindi. Xususan, A.M.Kondratovning *"Машинный перевод"* (1963), I.I.Revzin, V.Yu.Rozensveyglarning *"Основы общего и машинного перевода"* (1964), V.Ingvening *"Язык для программирования задач машинного перевода"* (1967), E.I.Korolyovning *"Промышленные системы машинного перевода"* (1991) nomli monografik tadqiqotlari; V.M.Kalininning *"Развитие схемы Пуассона и ее применение для описания статистических свойств речи"*, I.L.Yeshanning *"Опыт статистического описания научно-технического стиля (на материале переводов с румынского языка)"*, A.V.Zubovning *"Переработка текста естественного языка в системе «человек-машина»"*, V.A.Chijakovskiyning *"Фразеология и машинный перевод (опыт составления работы немецко-русского автоматического словаря для публицистических и научных текстов)"*, A.D.Borisevichning *"Англо-русский автоматический словарь оборотов (к проблеме идиоматичности при обращении текста в системе «человек-машина-человек»)"*, V.V.Goncharenkonning *"Лексикографические, лингвостатистические и инженерно-лингвистические вопросы построения*

автоматического словаря (англо-русский автословарь по полупроводникам) ” mavzularidagi nomzodlik dissertatsiyalari; V.I.Gerasimovning “Современное состояние машинного перевода”, G.G.Kotovning “Лингвистика и современное состояние машинного перевода”, T.N.Moloshnayaning “Алгоритм перевода с английского языка на русский”, Yu.V.Vannikovning “Виды адекватности и типологии переводов” sarlavhali maqolalarini alohida ko’rsatish mumkin.

Mazkur tadqiqotlar rus tili bo’yicha jahon kompyuter tarmoqlarida ma’lumotlar bazasi yaratilishiga va buning natijasi sifatida rus tilidagi matnlarni avtomatik tarzda boshqa tillarga tarjima qiluvchi dasturlarning vujudga kelishi uchun asos bo’ldi. O’zbek tilida ham shu kabi ma’lumotlar bazasini yaratish bugungi kundagi eng dolzarb vazifadir.

Rus kompyuter tilshunosligida taraqqiy qilgan yana bir yo’nalish matnlarni avtomatik tahriri yo’nalishdir. Mazkur yo’nalishda ilmiy tadqiqotlar olib borgan olimlar sirasiga R.R.Kotov, V.E.Berzon, V.G.Britvin, I.A.Melchuk, L.I.Belyaeva, V.A.Chijakovskiy, G.G.Belonogov, I.S.Duganova, A.B.Kuznetsov kabilarni kiritish mumkin. Nomi zikr etilgan olimlarning quyidagi tadqiqotlarini alohida ajratib ko’rsatish mumkin: G.G.Belonogov va G.G.Kotovlarning “Автоматизированные информационно-поисковые системы” (1968), G.G.Kotovning “Лингвистические аспекты автоматизированных систем управления” (1977), “Лингвистические вопросы алгоритмической обработки сообщений” (1983), “Прикладная лингвистика и информационная технология” (1987), “Оптимизация речевого воздействия” (1990), G.G.Kotov va B.V.Yakuninning “Язык информационных систем” (1989), I.L.Belyayeva hamda V.A.Chijakovskiyning “Тезаурус в системах автоматической переработки текста” (1983) nomli monografiyalari; V.G.Britvinning “Прикладное моделирование синтагматической семантики научно - технического текста (на примере автоматического индексирования)” mavzusidagi nomzodlik dissertatsiyasi; I.A.Melchukning “Порядок слов при автоматическом синтезе русского слова (предварительные сообщения)”, G.G.Belonogov, I.S.Duganova, A.B.Kuznetsovlarning “Экспериментальная система автоматизированного обнаружения и исправления ошибок в тексте” sarlavhali maqolalari avtomatik tahrir bo’yicha muhim manbalar sanaladi.

Ushbu olimlarning asosiy xizmati shundaki, ular avtomatik tahrir qiluvchi dasturlar uchun lingvistik ta’min yaratib berishgan. Buning natijasida kompyuterlardagi ruscha matnlarni tahrirlovchi dasturlar ishlab chiqilgan. Bu dasturlar bugungi kunda barcha kompyuterlarda mavjud. Ya’ni bunda noto’g’ri yozilgan so’zning tagiga qizil chiziq chiziladi, so’z to’g’ri yozilganidan so’ng chiziq yo’qoladi. Ko’rinadiki, bu dastur asosan imloni tekshirishga mo’ljallangan, unda uslubiy xatolarni aniqlash ko’zda tutilmagan. Kelajakda yaratiladigan dasturlar, umid qilamizki, avtomatik tahrirning bu jihatlarini ham e’tibordan chetda

qoldirmaydi. Ta'kidlash lozimki, bunday dasturlarni ishlab chiqish uchun til birliklarining nafaqat fonetik, semantik, morfologik xususiyatlari, balki aynan sintaktik-valentlik xususiyatlari hisobga olingan lingvistik ta'min yaratish lozim bo'ladi.

Rus tilshunosligida avtomatik tarzda tilga o'qitish yo'nalishida ham izchil tadqiqotlar olib borilgan. Ular oliy va o'rta maktabda til o'qitishni optimallashtirishning psixologik-kibernetik, semiotik, lingvo-statistik, injenerlingvistik va lingvodidaktik asoslarini ishlab chiqishga yo'naltirilgan. Xususan, G.G.Piotrovskiyning *“Компьютеризация преподавания языков”* o'quv qo'llanmasida EHM vositasida o'qitish, ya'ni lingvodidaktik vazifalarni kompyuter yordamida yechish, til o'rgatish jarayonini optimallashtirish, o'qituvchi lingvistik avtomat (OLA) uchun nutqiy materiallarni tayyorlash kabi masalalar tadqiq qilingan.¹³

Tadqiqotlar natijasida ko'plab OLA lar uchun lingvistik ta'min yaratilgan bo'lib, bu dasturlar rus tilini o'qitish jarayonini optimallashtirishga xizmat qiladi. Eng muhimi, rus kompyuter tilshunosligida *“tillarni o'qitishni kompyuterlashtirish”* deyilganda o'qituvchi yoki o'quvchi harakatlariga taqlid qiluvchi ko'ngilochar dasturlarni yaratish tushunilmaydi. Bunda asosiy e'tibor AO'S (avtomatik o'qituvchi sistema)larga tayanuvchi avtomatlashtirilgan o'quv kurslarini vujudga keltirishga qaratiladi. AO'S esa o'quv jarayoni ishtirokchilariga nafaqat taqlid qilishi, balki ularning intellektual faoliyatini modellashtirishi ham zarur bo'ladi.

Rus tilshunosligidagi tillarga o'qitish yo'nalishida lingvistik statistikadan ham keng foydalanilgan. Ma'lumki, ona tilidan tashqari ikkinchi bir tilni o'rganayotganda, avvalo, ushbu tilning lug'at boyligiga murojaat qilinadi. Ammo har bir tilning lug'at boyligida ming-minglab turli so'zlar mavjud bo'lib, ularning hammasini eslab qolish mumkin emasligi tabiiy. Shu sababli o'rganilayotgan tilning dastlab eng asosiy hamda tez-tez qo'llanib turadigan so'zlarinigina o'zlashtirishga kirishiladi va muntazam ravishda bosqichma-bosqich so'z boyligi orttirib boriladi. Buning uchun esa leksikostatistik ma'lumotlar asosiy poydevor vazifasini o'taydi. L.N.Zasorinaning ta'kidlashicha, leksikostatistikaning markaziy muammosi jonli (funksional) tilning statistik qonuniyatlarini va matnning statistik strukturasini aniqlashdir. Matnning statistik strukturasi deyilganda, shartli ravishda, ma'lum matndagi turli so'zlar miqdori bilan shu matnda uning qaytarilish-qaytalanish chastotasi orasidagi munosabat tushuniladi. Shunga ko'ra statistik ma'lumotlarni to'plash, qayta ishlash kabi murakkab jarayondagi barcha ishlarni EHMga yuklash zaruriyati kelib chiqqan holda kompyuter lingvistikasida statistik yo'nalish yuzaga keldi. U rus tilshunosligida avtomatik tarzda tilga o'qitish yo'nalishi bilan hamohang tarzda rivojlanib borgan. Bu yo'nalishda amalga oshirilgan eng asosiy tadqiqotlar quyidagilardir: P.M.Alekseev, L.N.German-Prozorova, G.G.Piotrovskiyning *“Основы статистической оптимизации преподавания иностранных языков”* (1974), N.D.Andreevning *“Статистико-комбинаторные методы в теоретическом и прикладном*

языковедении” (1967), P.M.Alekseevning “Статистическая лексикография” (1975), Ye.A.Kalininaning “Изучение лексико-статистических закономерностей на основе вероятностной модели” (1985).

Matematik va injener lingvistikasi, ya’ni hozirgi kompyuter lingvistikasining yo’nalishlaridagi tadqiqot metodlari turkologiyada ham qo’llanilgan. Turkologiyada bu yo’nalishda amalga oshirilgan tadqiqotlar sirasiga quyidagilarni kiritish mumkin: K.A.Veliyevaning “Формальное описание синтеза азербайджанского слова” (1970), V.Ya.Pinesning “Моделирование структуры азербайджанских глагольных форм в связи с проблемой автоматического словаря” (1970), X.F.Is’hakovaning “Исследования в области формальной морфологии тюркских языков (на материале татарского литературного языка в сопоставлении с турецким и узбекским)” (1972), A.Babanarovning “Разработка принципов построения словарного обеспечения турецко-русского машинного перевода” (1981), M.A.Mahmudovning “Разработка системы формального морфологического анализа тюркской словоформы (на материале азербайджанского языка)” (1982), T.Sadikovning “Моделирование киргизской именной морфологии” (1982) mavzusidagi nomzodlik dissertatsiyalari; M.Ayimbetovning “Проблемы и методы квантитативно - типологического измерения близости тюркских языков (на материалах каракалпакского, казахского и узбекского языков)” (1997) mavzusidagi doktorlik dissertatsiyasi; G.P.Melnikovning “Принципы системной лингвистики в применении к проблемам тюркологии”, A.Babanarovning “Автоматический анализ турецкой словоформы и турецко-русский машинный перевод” nomli maqolalari; M.S.Jikiyaning “Морфологическая структура слов в турецком языке (ранговая структура аффиксальных морфем)” (1975), K.B.Bektaevning “Статистико-информационная типология тюркского текста” (1978) nomli monografiyalari mavjud.

Matematik va kompyuter lingvistikasi Qozog’istonda izchil rivojlangan. Bu yerda dastlabki amalga oshirilgan tadqiqotlar, asosan, statistik yo’nalishga oid bo’lib, keyinchalik qozoq tilshunoslari mashina tarjimasini, avtomatik tahrir, tillarga o’qitish kabi sohalarda ham keng ko’lamli ishlarni amalga oshirishdi. Bunda A. Axbayevning “Статистический анализ лексико-морфологической структуры языка казахской публицистики” (1971), A.X.Jubanovning “Статистическое исследование казахского текста с применением ЭВМ (на материале романа М.Ауэзова «Абай жолы»)” (1973), K.A.Maldibekovning “Частотный словарь казахской детской литературы” (1980), D.A.Baytanayevaning “Информационные характеристики казахского языка” (1985) kabi dissertatsion ishlarni alohida ta’kidlovchi joiz.

Qozoq kompyuter lingvistikasining shakllanishida prof. Q.B.Bektaevning ilmiy faoliyati muhim ahamiyatga ega. Olim matematik va injener lingvistikasiga oid o’nlab asarlar yaratgan.¹⁴ Q.B.Bektaev rahbarligidagi “Lingvostatistika va

¹⁴ Бектаев К.Б., Лукьяненко К.Ф. О законах распределения единиц письменной речи. – СтРААТ, 1971; Бектаев К.Б., Джубанов А. Индексация и кодирование текстовой информации для ввода в ЭВМ. – Алма-Ата, 1973; Пиотровский Г.Г., Бектаев К.Б., Пиотровская А.А. Математическая лингвистика. –М.: Выш.шк, 1977; Бектаев К.Б., Кенесбаев С.К., Пиотровский Г.Г. О инженерной лингвистике // ВЯ. 1973, №2; Бектаев К.Б. Статистика – информационная типология тюркского текста. – Алма-Ата: Наука, 1978; Бектаев К.Б., Садчикова П.В. Опыт-но-промышленная система лексического машинного перевода. –

avtomatlashtirish” guruhining a’zolari M.Avezovning “Abay yo’li” romani (4 kitob) tilining chastotali lug’ati so’zligini kompyuter yordamida 50 soat davomida tuzib chiqqanlar. Mazkur chastotali lug’atda 20000 dan ziyodroq leksema va 60000 so’z shakli (“glossema”) mavjud bo’lib, u 466 000 marta qo’llanilgan. Guruh a’zolari bu lug’atni tuzish uchun atigi sakkiz oy vaqt sarflaganlar. Vaholanki, shu ish oddiy qo’l kuchi bilan bajarilganda, mazkur guruh a’zolari tinimsiz 10 yil, bir tilchi esa kamida 100 yil ishlagan bo’lar edi. Shu tariqa ushbu ulkan ish qozoq tilshunosligida kompyuter lingvistikasining rivojlanishiga asos bo’ldi.

O’zbek tili materiallari bo’yicha kompyuter tilshunosligiga oid tadqiqotlar olib borgan olimlar S.Rizaev va S.Muhamedovlardir. Mazkur olimlar ham asosan kompyuter lingvistikasining statistik tahlil yo’nalishi bo’yicha ishlarni amalga oshirishgan. S.Muhamedov o’zining R.R.Piotrovskiy bilan hammualliflikda yozgan *«Инженерная лингвистика и опыт системно-статистического исследования узбекских текстов»* nomli kitobida lingvistik modellar, modellashtirish va uning umumiy tamoyillari haqida fikr yuritib, o’zbekcha matnlarning kvantitativ modellarini keltiradi.¹⁵ Mazkur asarda sun’iy intellekt yaratish va injener lingvistikasi metodlari bilan o’zbekcha nutqni avtomatik qayta ishlash jarayonida foydalanish uchun zaruriy bo’lgan o’zbekcha matnlarning leksik – morfologik tuzilishining statistik tahlili natijalari keltirilgan. Ya’ni uning *«Статистический анализ лексико–морфологической структуры узбекских газетных текстов»* (nomzodlik dissertasiyasi) va *“O’zbek tilining alfavitli-chastotali lug’ati (gazeta tekstlari asosida)”* asarlari yuqoridagi asarning yozilishida manba bo’lib xizmat qilgan. S.Rizaev *“Kibernetika va tilshunoslik”* asarida tilshunoslikda aniq metodlarning qo’llanilishi va bunda EHMni qo’llash, til va nutq hodisalariga statistik yondashishining sabablari hamda matnlarni avtomatik qayta ishlash va mashina tarjimasini muammolari, shuningdek, harflar chastotasini aniqlashda EHMdan foydalanishga doir ma’lumotlar bergan.

S.Muhamedov va S.Rizaevlarning kompyuter tilshunosligiga oid yaratgan asarlari bugungi kun kompyuter tilshunosligiga asos hisoblanadi. Ammo o’zbek kompyuter tilshunosligini yangicha, zamonaviy texnologiyalar, dasturlar asosida rivojlantirish hamda samarali, ixcham, qulay va arzon kompyuter dasturlari uchun lingvistik ta’minot yaratib berish eng dolzarb vazifa hisoblanadi.

Shuni ta’kidlash lozimki, mazkur vazifalarni amalga oshirishda kompyuter lingvistikasi o’zbek tilshunosligida shu paytgacha yaratilgan birorta ham ilmiy asarni rad etmaydi. Aksincha, u A.G’ulomov, M.Asqarova, A.Kononov, F.Kamol, Z.Ma’rufov, N.Mamatov, Sh.Rahmatullaev, M.Mirtojiev, I.Qo’chqortoev, A.Nurmonov, N.Mahmudov, H.Ne’matov, R.Sayfullaeva, M.Sodiqova, R.Rasulov, E.Begmatov, I.Mirzaev, H.Dadaboev, Yo.Tojiev kabi atoqli o’zbek olimlarining

o'zbek tili leksikasi va grammatikasiga oid yaratgan tadqiqotlariga suyanadi. Ulardagi ilmiy-nazariy fikrlarga asoslanadi, ulardan oziqlanadi.¹⁶

So'nggi yillarda o'zbek kompyuter lingvistikasi asta-sekinlik bilan shakllanmoqda. Ayniqsa, bu sohada A.Po'latov, N.Jo'rayevalarning *“Разработка формальной модели грамматики узбекского языка”*, A.Po'latov, B.Alixo'jayev, N.Jo'rayevalarning *“Разработка программы компьютерного анализа и синтеза глаголов узбекского языка”*, A.Po'latov, S.Muhamedovalarning *“Компьютер tilshunosligida avtomatik tahrir qiluvchi dasturning lingvistik ta'minotini yaratish asoslari”* nomli maqolalari, shuningdek, A.Po'latov va S.Muhamedovalarning *“Компьютер lingvistikasi”* nomli o'quv qo'llanmalari (2007) muhim ahamiyatga molik.

Kompyuter lingvistikasi sohasida turli xil lingvistik muammolarni hal qilishga yo'naltirilgan dastrurlar yaratilgan va hali hanuz bunday amaliy tadqiqotlar amalga oshirilmoqda. Quyida ana shunday lingvistik dasturlar va ularning tavsifi haqida to'xtalamiz. Ular asosan matn tahriri, avtomatik tarjima, analiz va sintez, tabiiy tilning qayta ishlanishi, kompyuter leksikografiyasiga oid dasturlarni o'z ichiga oladi.

10-jadval

1	<u>Link Grammar Parser</u> John Lafferty, Daniel Sleator, Davy Temperley Carnegi Melon University, USA.	Ingliz tilining sintaktik parser dasturi, uning bazasiga 60000 so'z shakllari kiritilgan.
	<u>C#237;bola Oleada</u> loyihalari Computing Research Laboratory (CLR) New-Mexico State University, USA.	Mazkur dastur matnlarning lingvistik analizida ishlatiladi. Bu dastur yordamida 16 ta tildagi matnlarning avtomatik tarjimasi, statistik tahlilini amalga oshirish mumkin. Dastur tarkibida turli xil lug'atlar hamda tezauruslar ham mavjud.
	<u>Russian Morphological Dictionary</u> Sergey Sikorsky	Ruscha morfologik lug'at rus tilidagi matnlarning sintaktik va morfologik analizi uchun ishlatiladi. Unga 120000 so'zning morfologik lug'ati kiritilgan.
	<u>Морфологический парсер Яndex</u>	Ushbu morfologik parserning asosi A.A.Zaliznyak tomonidan yaratilgan <i>“Rus tilining grammatik lug'ati”</i> (110 ming so'zni o'z ichiga oladi) hisoblanadi. Tahlil natijalarini ko'rish uchun «разбор запоса» rejimi tanlanadi.
	<u>Лингвоанализатор</u> Д.В.Хмелев	Matn strukturasi matematik tahlilini ta'minlovchi dastur bo'lib, u on-line rejimida ishlaydi. Bunda mualliflik etaloni matn strukturasi asosida aniqlanadi. Kiritilgan matn analiz qilinadi

		(Русская Фантастика resursidan olingan matnlar analiziga tayaniladi), o'sha matnning muallifi bo'lishi mumkin bo'lgan 3 yozuvchining ismi hamda ularning matnga yaqin keladigan asarlari chiqarib beriladi.
	<u>Система StarLing</u> С.А.Старостин СУБД	Ushbu tizim katta hajmdagi ko'p tilli matnlar bilan, transkripsion belgilar bilan ishlashga mo'ljallangan bo'lib, u DBF-formatida yuklanadigan Ojegovning izohli lug'ati va Zaliznyakning morfologik lug'ati asosida so'zshakllarning analizi va sintezini, Myullerning lug'ati asosida avtomatik tarjimani ta'minlaydi.
	<u>Морфологический анализатор</u> С.А.Старостин.	Rus, ingliz tillaridagi so'zlarning morfologik analizi dasturi so'zlarning asosiy shakllari hamda morfologik ma'lumotini menyuga chiqaradi.
	<u>MonoConc</u> Michael Barlow Dept of Linguistics, Rice university, Texas, USA. 2 версии (MonoConc Pro 2.0 и MonoConc 1.5)	Tanlangan simvolli matnlar korpusi uchun konkordands yaratishda, so'zliklar va indekslar generatsiyasida foydalaniladi.
	<u>WordSmith Tools</u> Mike Scott.	WordSmith 3.0 turli alohida fayllar mvjmuidan tashkil topgan matnlar korpusi uchun konkordans yaratishda foydalaniladi. Bunda berilgan matnli fayllar majmui uchun so'zlar ro'yxatini generatsiya qiladi, katta hajmli matnlarni fragmentlar majmuiga ajratadi, matnlarning paketli tahririni amalga oshiradi.
0	<u>Concordance 2.0.0</u> R.J.C. Watt, 2000.	Win9x/NT/2000 uchun chastotali ro'yxat va konkordanslar yaratish uchun ishlatiladi. Windows tizimida mavjud bo'lgan kodirovkalarda matnlarga ishlov beradi, natijalar HTML-fayllar ko'rinishida saqlanadi.
1	<u>TextAnalyst 2.0</u> Научно-производственный инновационный центр «МикроСистемы».	Matn analizi uchun ishlatiladi, bunda belgilangan matnda kontekstga mos keluvchi tushunchalarning semantik tarmog'i shakllantiriladi, matn fragmentlari bo'yicha mazmunli qidiruv amalga oshiriladi, iyerarxik daraxt ko'rinishida (tema-podtema) matn analiz qilinadi, matn avtomatik referatlanadi. <u>TextAnalyst SDK</u> - ingliz va rus tillari uchun lemmatizatsiya (so'zlarni me'yoriy shaklga keltirish), tushunchalarning chastotali ro'yxatini yaratish, kontekstda so'zlarni qidirish kabi ishlarni bajaradi.

		<u>TextAnalyst Lib</u> – gipertekst texnologiyasi asosida elektron kitoblar yaratishda foydalaniladi.
2	<u>Galaktika-ZOOM</u> корпорация Галактика, Москва.	Informatsiyani analitik qayta ishlash va avtomatik qidiruvda ishlatiladi.
3	<u>Система Пропись 4.0</u> АО Агама	Rus tilidagi matnlarga ishlov berish uchun ishlatiladigan dasturlash paketi bo'lib, bunda orfografiya tekshiriladi, so'zlarning sinonimlar va antonimlarning ro'yxati beriladi, matn grammatik va stilistik jihatdan tekshiriladi, izohli lug'at asosida so'zlarning izohi beriladi, matndagi so'zlarni qidirish va almashtirish hamda matnning statistik tahlilini amalga oshirish imkoniyati ham mavjud.
4	<u>Textual Analysis Computing Tools (ТАСТ)</u> Library Electronic Text Resource Service Indiana University, USA.	Matnni avtomatik joylashtirish, so'zlarning tartibli ro'yxatini yaratish, matndagi so'zlarning uzunligi va chastotasiga ko'ra taqsimlanishi tahlili, konkordans yaratish kabilar uchun mo'ljallangan dastur sanaladi.
5	<u>Худломер</u> Леонид Делицын	Mazkur dastur rus tilidagi matnlarning uslubigv ko'ra avtomatik tasniflashga mo'ljallangan. Bunda 4 xil matnlar korpusi (badiiy asar, publitsistika, ilmiy maqolalar va dialoglar) tahlil etiladi, natijada stilga bog'liq holda matndagi so'zlar uzunligining egri joylashuvi ko'rinadi. Ular tasniflash jarayonida mezon sifatida olinadi. Dastur yordamida kiritilgan matn “РАЗГОВОРНАЯ РЕЧЬ” (OG'ZAKI NUTQ), “ХУДЛО” (BADIY ADABIYOT), “ГАЗЕТНАЯ СТАТЬЯ” (GAZETA MAQOLASI) yoki “НАУЧНАЯ СТАТЬЯ” (ILMIY MAQOLA) deb tasniflanadi.
6	<u>«Штампомер»</u> Леонид Делицын	Bu dastur mualliflik matnlarida tez-tez ishlatiladigan iboralar (shtamplar), o'sha muallif tomonidan ko'p ishlatiladigan so'z va ifodalarni aniqlashga yordam beradi.
7	<u>Свежий взгляд</u> <u>Fresh Eye версия 1.21,</u> 1995 Дмитрий Кирсанов	Bu dastur rus tilidagi matnlarni stilistik jihatdan tekshiradi. Bunda matnda o'zaro yaqin joylashgan fonetik va morfologik jihatdan o'xshash bo'lgan so'zlarni topadi hamda stilistik g'alizlikni, tautologiyani oldini oladi.
18	<u>RussianWord Constructor</u> (RWC) версия 1.0, 1992 Дмитрий Кирсанов.	Bu rus tilidagi she'rga o'xshash matnlarning generatsiyasi uchun yaratilgan tajriba dastur bo'lib, leksikostatistik ma'lumotlarga va lug'atga tayanib ruscha neologizmlarni va okkazonalizmlarni hosil qiladi.
9	<u>Scott Pakins automatic complaint-letter generator</u> Scott Pakin.	Bu dastur muayyan shaxs yoki tashkilotga shikoyat matnini generatsiya qilishga yordam beradi. Dasturga 7000 leksik unurni o'z ichiga olgan lug'at bazasi kiritilgan bo'lib, u ingliz tilida on-line rejimida ishlaydi.

0	<u>Cyrano Server</u> Nando.net.	Bu dastur sevgi maktublarining generatsiyada ishlatilvdigan on-line rejimida ishlaydigan dastur bo'lib, bunda foydalanuvchi o'z ismini, yozish stilini, sevgilisiga murojaat shaklini va yana sevgilisining ayrim atributlarini tanlaydi, dastur ingliz tilida sevgi maktubini yaratadi.
1	<u>Письмовник</u> MediaLingua Ltd., 1995-2002.	Ish hujjatlari va yozishmalarni avtomatik tarzda shakllantiradi hamda ularni rus tilidan ingliz tiliga, ingliz tilidan rus tiliga tarjima qiladi.
2	<u>ALICE</u> . Dr.Richard S.Wallace. ALICE AI Foundation, USA.	Bu dastur gapiruvchi programma bo'lib, u 2000 va 2001 yillarda Lobner mukofotiga (Loebner prize) loyiq topilgan. Dastur AIML (sun'iy intellekt uchun maxsus til texnologiyasiga asoslanadi).
3	<u>ANANOVA</u> Ananova Limited United Kingdom.	Buyuk Britniyadagi virtual diktir asosida ishlaydigan axborot agentligi bo'lib, u orqali kunning yangiliklarini bilib olish mumkin.
4	<u>Sakrament Text-to-Speech Engine v2.0</u> компания "Сахрамент", 2002.	Rus tilidagi nutq sintezi tizimi bo'lib, u Microsoft Speech API (SAPI) 5.1. standartiga to'liq mos keladi, erkak va ayol ovozida matnlarni talaffuz qiladi, foydalanuvchining xohishiga ko'ra intonatsion pauzalar qiladi, nutq temбри va tonini o'zgartiradi.
5	<u>Аки Росс</u> Альберт Товмасын,2002.	Bu dastur rus tilida muloqotni ta'minlaydi. Bunda maxsus plug-in Aki Ross ismli animatsiya ko'rinishidagi qiz-bola suhbatdosh bilan muloqot qilishga imkon beradi. Dasturning bilimlar bazasi juda katta emas, lekin uni AIML tili yordamida to'ldirib borish mumkin yoki dastrur saytidan bilimlar bazasini yangilab turish mumkin.
6	<u>Programmable Artificial Intelligence (PAIv3.01)</u> или <u>Программируемый Искусственный интеллект</u> Claudio Scordino, 2002.	Ushbu dastur foydalanuvchi bilan ingliz tilida muloqot qilishni ta'minlaydi, u C++ tilida Linux, FreeBSD va Windows tizimlarida ishlaydi.

Topshiriqlar

1. Kompyuter lingvistikasi bo'yicha xalqaro konferensiya materiallarini internetdan yuklab oling.
2. Kompyuter linvistikasi fani bilan shug'ullangan olimlar va ularning asarlarini ko'rsating.
3. Kompyuter linvistikasi fanini rivojlantirish uchun nimalar qilish lozim?
4. Kompyuter linvistikasida avtomatik tarjima muamosi bilan kimlar shug'ullangan?
5. O'zbek tilshunosligida kompyuter lingvistikasiga oid qanday tadqiqotlar mavjud?

Adabiyotlar

1. Grishman R. Computational Linguistics. – Cambridge University Press, 1994.
2. Шемакин Ю.И. Начало компьютерной лингвистики. – М.: МГОУ. 1992.
3. Gelbukh A. Computational Linguistics and Intelligent Text Processing. – International Conference. – Mexico, 2003.
4. Новое в зарубежной лингвистике. Вып.24. Компьютерная лингвистика. –М.: Прогресс, 1989.
5. Po'latov A., Muhamedova S. Kompyuter lingvistikasi. – Toshkent, 2007.
6. Нелюбин Л.Л. Перевод и прикладная лингвистика. - М.: Высшая школа, 1983.

Ma'ruza № 3

Lingvistik bilimlar bazasi

Reja

1. Bilimlar bazasi haqida.
2. Korpus tushunchasi va uning mohiyati.
3. Jahon tillarining lingvistik korpusi.
4. O'zbek milliy korpusini yaratish muammolari.

Tayanch so'z va iboralar: *Bilim, bilimlar bazasi, tabiiy model, matematik model, mantiqiy-matematik model, korpus, milliy korpus*

Matn biror millatga tegishli xalqning shunchaki adabiy tiligina emas, balki uning madaniyati va ma'naviyati aks etuvchi murakkab strukturadir. Avtomatlashtirilgan leksikografiyaning o'zi (so'z va so'z birikmasi) matnni tarjima qilib berish uchun yetarli bo'lmaydi. Tillar tipologik jihatdan farqlanadi. Lingvistik ta'minot tuzishda o'rganilgan va amalda sinalgan tadqiqotlardan foydalanish maqsadga muvofiq. Mashina tarjimasi, asosan, axborot (ilmiy, rasmiy, publitsistik) uslubi uchun xosdir. Tarjima modellarini tuzishda ana shu uslubdagi me'yoriy talablar hisobga olinadi. Tarjimaning uch asosiy komponenti mavjud, bular: asos matn, unga qo'yilgan masalalarning to'g'riligi hamda terminologiya masalalaridir. Matnni tarjima qilishda quyidagi masalalar o'rganiladi: tahlil jarayoni, so'zlarning matndagi ma'nolarini tushunish jarayoni hamda matnni tiklash jarayoni. Mashina tarjimasida morfologik va sintaktik tahlil muhim ahamiyatga ega. Morfologik analizda kategoriyalarni tasniflashda nimani emas, balki ularni qanday qilib tasniflash nazarda tutiladi. Xususan, turkcha-ruscha mashina tarjimasi ta'minoti quyidagi qismlardan iborat bo'lgan: so'zshakllari so'zligi va oborotlaridan iborat bo'lgan avtomatik lug'at hamda ot va fe'l so'z turkumiga kiruvchi so'zlarning asosi va qo'shimchalaridan iborat bo'lgan. Turkcha matnlar uchun qo'llanilgan bunday lug'at S. Muhamedovning kvantativ modeli tamoyillari asosida yaratilgan¹⁷.

Ma'lumki, istalgan mustaqil ma'noli leksema ierarxik semantik maydonlarga ega. Bu semantik maydonlar o'rtasidagi munosabat kontekst orqali voqelanadi.

¹⁷ Мухамедов С.А., Пиотровский Р.Г. Инженерная лингвистика и опыт системно- статического исследования узбекских текстов. - Ташкент: Фан, 1986. - С. 141-142.

Morfologik tahlilning asosini morfologik va tarjima baza, grammatik qoidalar va istisnoli jadvalga kiritilgan belgilar tashkil qiladi:

- ID->English_Suz
- ID->TranslateUz_Noun
- ID->TranslateUz_Verb
- ID->TranslateUz_Adjective
- ID->TranslateUz_terminology
- ID->TranslateUz_combinator

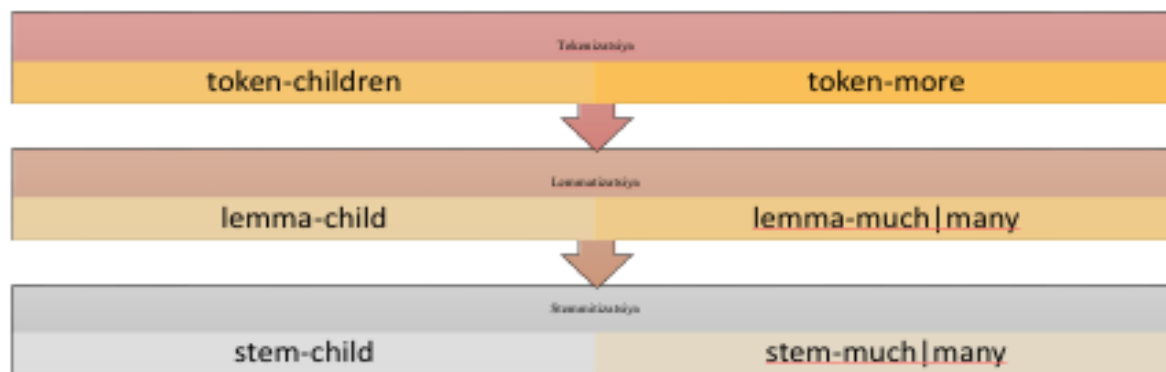
O'zbek tili va ingliz tilining morfologik bazasini tuzishda mezonlar aniqlab olindi. Bunda fe'l so'z turkumining o'zak shakli kiritildi: *o'qimoq* - <o'qi>, *ochmoq*

- <och>. Ushbu holatda ayrim fe'llar boshqa so'z turkumlari bilan omonimlik hosil qilishi mumkin. Shu bois bazadagi ma'lumotlarni dasturiy ta'minotda aniqlay olishi uchun quyidagi belgilar kiritildi:

H_W- agar muayyan so'z faqat bir so'z turkum doirasida omonimlik hosil qilsa; H_L- agar muayyan so'z ikkidan ortiq so'z turkumi doirasida omonimlik hosil qilsa; H_Adj. {H_Adv.}- agar ikki so'z turkumi doirasida bo'lsa, ikkinchi so'z turkum ko'rsatiladi.

Ko'p tadqiqotlarda morfologik tahlilni optimallashtirishga oid bir qator fikrlar o'rta tashlangan¹⁸. Matnlarni avtomatik tarzda morfologik tahlil qilishda tokenizatsiya, lemmatizatsiya, stemming usullaridan foydalaniladi. Tokenizatsiyalash jarayonida tabiiy tildagi nutq birliklari (token, so'zshakllari) alohida tarzda ularning xususiy ma'nolariga ajratiladi. Lemmatizatsiya bosqichida esa so'zshakllarining dastlabki holati aniqlanadi. Stemming hodisasiga binoan unda so'zning o'zagi topiladi. Buni quyidagi sxema orqali tushuntirish mumkin:

¹⁸ Qarang: Belonogov G. G. On the use of a similarity method for automatic processing of textual information [Об использовании метода аналогии при автоматической обработке текстовой информации] Problems of cybernetics [Problemy kibernetiki]. Issue 28. -Moscow: Nauka Publ., 1974; Bolshakov I. A. CrossLexica: A large electronic dictionary of collocations and semantic links between Russian words [КроссЛексика — большой электронный словарь сочетаний и смысловых связей русских слов]. Компьютерная лингвистика и Интеллектуальные Технологии: Труды Международной Конференции "Диалог 2009" [Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference "Dialogue 2009"]. -Moscow, 2009, pp. 45-50.; Gelbukh A. F. Effectively realizable morphologic model of inflective language [Эффективно реализуемая модель морфологии и инфлексивного языка]. Научно-техническая Информация [Scientific and Technical Information], series 2, #1, 1992, PP. 24-31.; Sokirko A. V. Bystroslovar': morphological prediction of new Russian words using very large corpora [Быстрословарь: предсказание морфологии русских слов с использованием больших лингвистических ресурсов]. Компьютерная лингвистика и Интеллектуальные Технологии: Труды Международной Конференции "Dialog 2010" [Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference "Dialogue 2010"]. -Moscow, 2010. -P. 450-45; Zaliznjak A. A. Grammatical dictionary of Russian: ИНА эсион [Грамматический словарь русского языка: Словоизменение]. - Moscow: Русские Язык Публ., 1977.



Matnlarni tahlil etishning bu kabi usullari ba'zi adabiyotlarda kompyuter ma'lumotlar bazasi-korpus yaratishning maxsus protseduralari hamda dasturlari sifatida tilga olinadi¹⁹. Inglizcha matnlarni o'zbek tiliga tarjima qilishda, avvalo, morfoklassifikator kerak bo'ladi. Xususan, ingliz tilida sifat va ravish so'z turkumlariga oid so'zlarni morfologik jihatdan tasniflash va ularning o'zbekcha variantlarini hisobga olish darkor.

Ingliz va o'zbek tillari boshqa-boshqa til oilalariga mansub bo'lganligi uchun ham morfologik kategoriyalarning bir-biriga o'xshamasligi, tabiiy. Masalan, ingliz tilidagi ba'zi olmoshlar o'zbek tilida mavjud emas. Ular boshqa kategoriyalar sifatida qaraladi: *few*, *little* olmoshlari o'zbek tilida *kam*, *oz* so'zlariga mos bo'lib, ravish so'z turkumiga tegishli. Bularning yechimi tegishli lug'atlarga kiritilgan so'zlarning o'zaro kodlanadi. Tarjimaga qo'yiladigan asosiy talablardan biri matnning shakli va mazmunini saqlab qolishdir. Shu kabi muammolarni hal qilishda transformatsion metod samarador usullardan biridir. Transformatsion usulda to'rt jarayon voqealanishi mumkin: *o'rniga qo'yish*, *almashtirish*, *qo'shish*, *tushurib qoldirish*²⁰. Shuningdek, transformatsion usulning adyunsiya, substitutsiya turlari ham mavjud. Unga ko'ra, transformasiya protsessi asosini yadro gaplar deb ataluvchi konstruksiyalar (kernel structures) tashkil qiladi. Bu konstruksiyalar eng kichik yoki sintaksisdagi eng sodda gaplardan iborat²¹. Bunday asosiy bosqichlar analiz va sintez jarayonida o'z ifodasini topadi.

O'rniga qo'yish usuli matndagi so'z va so'z birikmalarini morfologik jihatdan o'rinlashtiradi.

1	2	3	4
<i>I</i>	<i>go</i>	<i>to</i>	<i>school</i>
1	4	3	2
<i>Men</i>	<i>maktab</i>	<i>ga</i>	<i>boraman</i>

Almashtirish ikki hodisa orqali amalga oshiriladi: 1) konkretizatsiya; 2) generalizatsiya. Bu ikki hodisa ham morfologik va sintaktik jihatdan yuzaga

¹⁹ Захаров В.П., Богданова С.Ю. Корпусная лингвистика: учебник для студентов гуманитарных вузов. - Иркутск: ИГЛУ, 2011. - С. 39.

²⁰ Бархударов Л. С. Язык и перевод (Вопросы общей и частной теории перевода). -М.: Междунар. отношения, 1975. -С.190-191.

²¹ Бурунов Ж.Б. Инглиз ва узбек тиллари янги грамматикаси. -Т.: Укитувчи, 1973. - Б. 40.

keladi. Xususan, *and* soʻzi biriktiruv bogʻlovchi vazifasidan (*va, hamda*) tashqari quyidagi maʼnolarda keladi: *boʻlsa, esa, biroq - I shall go and you stay here -Men ketaman, sen esa shu yerda qol.* Ayrim feʼlli birikmalar tarkibida kelib, oʻzbek tilida feʼlning harakat nomiga mos kelishi mumkin: *try and do it* - buni qilishga harakat qiling, *come and see* -koʻrishga keling, *wait and see* -yashasak koʻramiz²² kabi. *Generelizatsiya* hodisasi koʻproq sintaktik birliklarda kuzatiladi. *Soʻzlarni qoʻshish* usulida baʼzi morfologik kategoriyalar matnga qoʻshimcha ravishda qoʻshiladi: *I have a book -Menda kitob bor.* Bunda feʼl va otning oʻzaro munosabati semantik aspektida koʻrinadi. Inson tomonidan tarjima qilinganda, bu tabiiy holat, biroq mashina tarjimasida bu alohida jarayon. *Tushirib qoldirish* hodisasida tarjima jarayonida u yoki bu soʻz tushirib qoldiriladi: *those-ular, these-bular* koʻrsatish olmoshlari lugʻatda koʻplikdagi narasaning predmetlarga nisbatan qoʻllanadi, ammo ular ot bilan kelganda *-lar* tushirib qoldiriladi: *These are mine - Bular meniki* *These children are mine -Bu bolalar meniki.*

Shu jihatdan lingvistik taʼminotda morfologik tasnif jiddiy masalalardan biri. Shu oʻrinda, rus tilidagi *Krossleksika* elektron lugʻatidagi morfologik klassifikatorga 115 guruhga ajratilgan turlanuvchi soʻzlar tasnifi kiritilgan²³. Morfokvalifikator lugʻatda mavjud boʻlmagan soʻzshakllarini ham tuzish imkoniyatiga ega. Lingvistik translyator uch bosqichli tizimdan tashkil topadi: analizator va sintezatordan tarkib topgan *lingvistikprotessor*, grammatika va semantika haqidagi bilimlardan tashkil topgan *lingvistik model*, oʻzaro bogʻlangan deklarativ va protsedurali qismlarning lingvistik translyatsiyasini aks ettiruvchi *assotsiasiv protsedura*²⁴. Koʻplab mashina tarjima tizimlarining moduli *analiz->transfer->sintez* bosqichida voqelanishini kuzatamiz. Matnlarni morfologik tahlil qilishda analizator lingvistik jihatdan taʼminlanadi. Dasturiy taʼminotda grammatik taʼminot ishlab chiqilgan: “terminlarning grammatik guruhga kiruvchi soʻzlar (artikl+ot+predlog) hamda terminlarning funksional elementlaridan (ega+kesim) chiziqli tuzilmasi yoki elementlarning ketma-ketligidan tashkil topgan *zanjirli grammatika*. Unda u yoki bu elementlarning oʻzidan oldingi yoki oʻzidan keyingi ketma-ketligini koʻrsatadi. Shuning uchun bu kesim oldi sintaktik analiz deb yuritiladi. Keyingisi *tashkil etuvchi grammatika* boʻlib, unda grammatik element guruhlari haqida maʼlumot beradi, masalan, ismlar guruhi (ot, artikl, sifat v.b.dan tashkil topuvchi guruh), predlogli guruh (predlog va boshqa soʻz (ism)lardan tashkil topgan guruh) hamda gap darajalarigacha boʻlgan elementlar. Shuningdek, *tobelanishga asoslangan grammatika* boʻlib, unda har bir element oʻzaro tobe munosabatlarga boʻysunadi. Tahlil strategiyasi yuqoridan pastga tomon boʻlib (top-down) gapning markazi feʼl, yaʼni kesim aniqlanadi. *Kontekstli erkin grammatikada* transformatsion qoidalarga amal qilinadi. *Unifiksion grammatika* toʻrt komponent (tarkib): unifikatsiya paketi, leksik tavsif va qoidalar uchun interpretator, yoʻnaltirilgan graflarni qayta ishlash dasturi hamda graf-sxemalar yordamidagi analizatordan iborat boʻladi. Unifikatsion grammatika semantik valentlik bilan sintaktik valentlik, lugʻatga oid tavsif bilan grammatik qoidalarni aniqlashtiradi”²⁵.

²² ABBY Lingvo x5

²³ Болшаков И. А., Болшакова Е. И. Автоматический морфоклассификатор русских именных групп. Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной Международной конференции «Диалог» (2012) Выпуск 11. - С. 81.

²⁴ Марчук Ю.Н. Компьютерная лингвистика - М.: Восток, 2006. -С. 272.

²⁵ Мамедова М.Г., Мамедова З.Ю. Машинный перевод: эволюция и основные аспекты моделирования. - Баку: ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ АКАДЕМИИ, 2005. -С. 69-72.

Ma'ruza № 4

Lingvistik protsessorning grammatik komponenti

Reja

1. Matnning statistik tahlili.
2. Kvantitativ metod va uning tilga tatbiqi.
3. Chastotali lug'atlar.
4. Flektiv tillarda o'zakni aniqlash, polisemantik so'zlarni va sinonimik qatorda dominanta so'zni aniqlash.

Tayanch so'z va iboralar: kvantitativ metod, matematik parametr, statistika, indeks, tanlama, tanlamaning hajmi, absolyut chastota, nisbiy chastota, lisoniy tip, flektiv tillar, agglutinatativ tillar, chastotali lug'atlar, matnning statistik tahlili, Zif qonuni, matn atributsiyasi, anonim va psevdonim matnlar, idiostil, polisemantik so'zlar, sinonimik qator, dominanta, to'plam, lingvostatistika.

Mashina tarjimasini uchun o'zbek tilining formal grammatikasini yaratish muhim

sanaladi. O'zbek tilini tahlil qilish uch bosqichda amalga oshiriladi:

- 1) so'z turkumlari; 2) gap bo'laklari; 3) gap turlari

O'zbek tilida quyidagi so'z turkumlari mavjud:			
Mustaqil so'z		Yordamchi so'z	Alohida so'zlar
Ot	Ravish	Bog'lovchi	Undov so'zlar
Fe'l	Son	Yuklama	Taqlid so'zlar
Sifat	Olmosh	Ko'makchi	Modal so'zlar

Ingliz va o'zbek tillari so'z turkumlarini ajratishda sezilarli farqlar uchraydi. Bunday har xillik mashina tarjimasida lingvistik ma'lumotlarni yaratishda muhim ahamiyatga ega. Zero, tasnifda turli xillik mavjudligi, qolaversa, muayyan grammatik kategoriyalar ayrim hollarda tubdan farqlanishi quyidagi tasnif orqali anglash qiyin emas. Ingliz tilida so'z turkumlarini tasniflashda bosh va ikkinchi darajali bo'laklar quyidagicha guruhlanadi²⁶:

I. Gapning bosh bo'laklari: 1) Noun; 2) Adjective; 3) Verb; 4) stative (turg'un semali so'zlar want, know); 5) Pronoun; 6) Numeral; 7) Adverb; 8) Modal words; 9) Particle (fe'lli frazemaning o'zgarmas qismi); 10) Exclamation

II. Ikkinchi darajali bo'laklar: 1) preposition; 2) conjunction Boshqa

²⁶ Кобрин Н.А., Корнеева Е.А. и др. Грамматика английского языка. Морфология. Синтаксис. Санкт - Петербург, 2008. - С. 7.

adabiyotdagi tasnifda tafovut yanada farqlanadi²⁷:

I. Gapning bosh bo‘laklari: 1) Noun; 2) Adjective; 3) Verb; 4) stative (turg‘un semali so‘zlar want, know); 5) Pronoun; 6) Numeral; 7) Adverb; 8) Modal words; 9) Exclamation

II. Ikkinchi darajali bo‘laklar: 1) preposition; 2) conjunction; 3) particle (fe‘lli frazemaning o‘zgaras qismi); 4) article

Boshqa bir grammatik adabiyotda gap bo‘laklari tasnifi o‘zgacha shaklda beriladi²⁸: Asosiy so‘z guruhlari: 1) Verb; 2) Noun; 3) Adjective; 4) Adverb; 5) Preposition; 6) Determinant (aniqlovchi so‘z), 7) Pronoun; 8) Conjunction. Fe‘l, ot, sifat, ravish - vokabular so‘zlar (vocabulary words), predlog, olmosh, determinant, bog‘lovchi - grammatik so‘zlar (grammatical words)ga bo‘lingan. Detriminantlar quyidagi ichki guruhlarga ajratilgan: artikl, egalik olmoshlari, ko‘rsatish olmoshlari, hisob so‘zlar. Ma‘lumki, lingvistik ta‘minotda so‘z turkumlari bazasini shakllantirish muhim sanaladi. O‘zbek tili so‘z turkumlari ajratilishi quyidagicha: I. Mustaqil so‘z turkumlari: 1) fe‘l; 2) ot; 3) sifat; 4) ravish; 5) son; 6) olmosh; II. Yordamchi so‘z turkumlari: 1) bog‘lovchi; 2) ko‘makchi; 3) yuklama; III. Alohida guruhni tashkil yetuvchilar: 1) modal so‘zlar; 2) undov so‘zlar; 3) taqlid so‘zlar.

Bunday tasniflash har ikki tilning o‘ziga xos jihatlari bilan bog‘liqdir. Mashina tarjimasida so‘z turkumlarini tahlil qilishda quyidagilar inobatga olinishi zarur: 1) har ikki tilda mavjud grammatik kategoriyalarning o‘zaro o‘xshash jihatlari aniqlash; 2) har ikki tilda o‘zaro farqlanuvchi kategoriyalarni muvofiqlashtirish; 3) yuqoridagi har ikki guruhni o‘zaro modellashtirish.

Ingliz tilida, asosan, sakkizta so‘z turkumi (Parts of speech|tagset) mavjudligi qayd etiladi: noun, verb, pronoun, preposition, adverb, conjunction, participle va article (ot, fe‘l, olmosh, predlog, ravish, bog‘lovchi, artikl va mayl). So‘z turkumlari u yoki bu so‘zni qaysi kategoriyaga tegishli ekanligi yoki qaysi paradigmaga ko‘ra tasniflanish, so‘zshakllari va h.k. haqida ma‘lumot beradi. So‘z turkumlarining ikki ochiq va yopiq turga ajratilgan holatlarini uchratamiz²⁹. Ochiq turdagi so‘z turkumlari sirasiga ot, fe‘l, ravish, sifat kabilar kiritiladi, zero boshqa tildan o‘zlashgan yoki o‘zga til ta‘sirida yangi so‘zlar hosil bo‘ladi. Yopiq turdagi

²⁷ Kaushanskaya V.L., Kovner R.L., Kojevnikova O.N., Prokofeva i dr. A grammar of the English language, - M., 2008. - P. 14.

²⁸ John Eastwood Oxford guide to English grammar. Oxford University, 2002. - P. 10.

²⁹ Juravskiy D. Speech and language processing. 2007. - P. 140.

soʻz turkumlariga tildagi yordamchi soʻz turkumlari kiritiladi: predlog, artikl, olmosh, oʻzgarmas soʻzlar (particle), bogʻlovchilar, yordamchi feʼllar (are, must.), sonlar.

Mashina tarjimasida gapning muayyan tarkibini qismlarga ajratishda soʻz birikmalariga ajratish muhim sanaladi. Negaki, sintaktik tahlil soʻzlarning oʻzaro tobelanish asosiga quriladi. Shuningdek, soʻz turkumlaridagi har xillik, qolaversa, soʻz birikmalari tuzilmasidagi ayrim tovufutli jihatlar tarjima jarayoniga taʼsirini koʻrsatadi. Matnlarni sintaktik tahlil qilishda gap qismlari oʻrtasidagi tobelanish bevosita soʻz birikmasi va soʻz birikmasi zanjirini hosil qiladi.

Maʼruza № 5

Morfologik tahlilning avtomatik tizimi

Reja

4. Kompyuter lingvistikasi va uning asosiy yoʻnalishlari.
5. Kompyuter lingvistikasiga oid nazariy va amaly tadqiqotlar.
6. Lingvistik programmalar va ularning yoʻnalishlari.

Tayanch soʻz va iboralar: *kompyuter lingvistikasi, mashina tarjimasida, avtomatik tahrir, avtomatik oʻqitish tizimi, matning statistik tahlili, lingvostatistika, injener lingvistikasi, chastotali lugʻat, oʻqituvchi lingvistik avtomat, lingvistik taʼmin, lingvistik programmalar.*

Morfologik analiz faqat u yoki bu kategoriyalarni tahlil qilish bilan cheklanmaslik kerak, nazarimizda. Ayniqsa, mashina tarjimasida buning oʻzi kifoya emas. Yuqorida tilga olingan baʼzi muammoli holatlar uchun Rodolfe Delmonte leksikon uchun quyidagi lingvistik kategoriyalarni koʻrsatib oʻtadi³⁰: 1) grammatik kategoriyalar; 2) semantik kategoriyalar; 3) diskurs darajasidagi kategoriyalar; 4) sintaktik kategoriyalar; 5) bogʻlangan kategoriyalar; 6) semantik kotseptual kategoriyalar; 7) chegaralangan leksik birliklar; 8) grammatik cheklovlar. Bu tasniflash italyan tili uchun xos boʻlib, uning lingvistik tomonlari hisobga olingan. Matnlarni tahlil qiluvchi avtomatik analizatorni yaratishda dastlabki bosqich sifatida oʻzbek tilining elektron morfologik lugʻatini yaratish masalasi qoʻyildi. Unga koʻra har bir soʻzning qaysi soʻz turkumiga tegishli ekanligi statistik tahlilda koʻrsatildi. Har bir leksemaning muayyan morfologik xususiyatlari jadvallarda aks etdi. Soʻng har bir leksema soʻz turkumlari boʻyicha toʻplamlarga tasniflanadi. Toʻplamlarga ajratishda leksemalarning umumiy semantik, qolaversa, formal jihatlarini inobatga olinadi. Toʻplamlarga ajratilgan har bir soʻzga qoʻshilishi mumkin boʻlgan qoʻshimchalar kombinatsiyasi aniqlanadi. Undan oldin barcha qoʻshimchalar funksional belgilariga koʻra guruhlariga ajratilib, ularning oʻzbek tilidagi barcha mumkin boʻlgan mantiqiy bazasi shakllantiriladi.

Avtomatik tahlil qilishning quyidagi turlari mavjud³¹: 1) soʻzlarning lugʻatdagi asoslariga koʻra tahlili; 2) soʻzlarning lugʻatdagi soʻzshakllariga koʻra analizi; 3)

³⁰ Rodolfe D. Computational Linguistic Text Processing: Lexicon, Grammar, Parsing and Anaphora Resolution. Nova Science Publishers, Inc. -New York, 2008. - P. 4-5.

³¹ Марчук Ю. Компьютерная лингвистика. -М., 2007. - С. 65.

mantiqiy ko'paytirish metodi orqali tahlili; 4) lug'atsiz jadvallar yordamida analiz qilish. Tilshunoslikning tillar tipologiyasi va qiyosiy tahliliga oid tadqiqotlar generativ tilshunoslik aspektida ko'rib chiqiladi. Kompyuter bazasi uchun tilni modellashtirish eng qulay o'rganish usuli hisoblanadi. Kompyuterli morfologiya yuqori darajadagi dasturlar uchun zarur hisoblanib, morfologiya doirasida modellashtirish tadqiqotlarda o'rganilib kelinmoqda. Qariyb o'ttiz yildan oshiq vaqt mobaynida turli dasturlar uchun morfoanalizatorni takomillashtirish bo'yicha ishlar amalga oshirilmoqda³². Zero, statistik metodga asoslangan tizimlar mashina tarjimai uchun yaratilgan bo'lsa-da, tabiiy tillarning imkoniyatlari to'liq qamrab ololmaydi. Ushbu yo'nalishda qoidaga asoslangan tizimning grammatik tahlil qilish usuli "tilga oid bo'lgan polisemiya va omonimiya hodisalarini izchil tadqiq yetishi hamda turli tipga kiruvchi tillarning grammatik kategoriyalari farqli jihatlari bilan birgalikda o'rganilishi"³³ bilan ham ahamiyatlidir.

Morfologik modellashtirishda lingvistik tavsifning nechog'lik to'g'ri va aniq berilishi muhim sanaladi. Morfologik mezonlarning to'g'riligi quyidagi bosqichlarda tekshiriladi: 1) faqat to'g'ri tuzilgan grammatik so'z shakllarining tahlili; 2) grammatik jihatdan noto'g'ri deb topilgan so'z shakllarining qay tarzda aniqlanishi. So'zshakllarning to'g'riligi faqat to'g'ri deb topilgan modellardan aniqlashtiriladi, qolgan holatlar esa uning noto'g'riligini tasdiqlaydi.

Ayrim adabiyotlarda morfologik tavsifning komponentlari sifatida quyidagilar tushuniladi: lug'atning maxsus ob'ektli modeli, lug'at, morfologik qoidalar hamda fonologik va morfofonologik qoidalar. Maxsus lug'atlarda morfem tiplari, barcha turlarining xossalari va tavsifi beriladi. Lug'atdagi har bir morfema faqat grammatik ko'rsatgichga ega bo'lishi kerak. Morfemalarning semantik xususiyatlari ularning strukturasini belgilab beradi. Morfologik qoidalar so'zshakldagi morfemalarning o'zaro bog'lana olish xususiyatlarining chegarasini aniqlay oladi, ya'ni morfemalarning omonimiyaga xosligi shu holatda muayyan turga tegishlilik bilan farqlanadi. Fonologik komponentda morfemalarning allomorfligi natijasi aniqlanib, kontekstda so'zshakllaridagi allomorflar o'zgarishi hamda so'zshakllarining fonotaktik to'g'riligi tekshiriladi.

Tabiiy tillarni qayta ishlashda grammatik tahlil muayyan bosqichlarga ega bo'ladi. Grammatikaning morfologiya va sintaksis bo'limlariga tegishli barcha lingvistik qurilmalar, jumladan, so'z turkumlarining modellarini tuzish, ya'ni kompyuterning o'qish holatiga o'tkazishda tildagi barcha me'yoriy vaziyatlar qamrab olinishi maqsadga muvofiq. E'lon qilingan avvalgi ishlarimizda mashina tarjimai uchun avtomatik tahlil jarayonlari o'rganilgan. Unga ko'ra leksik, mashina tarjimasining dastlabki tokenizatsiya->lemmatizatsiya->stemming bosqichlarida morfologik tahlil amalga oshiriladi³⁴. Tokenizator so'zlarni chapdan o'ngga qarab tekshiradi. Natijada u yoki bu so'zni qaysi so'z turkumiga tegishlilik hamda asos va qo'shimchalar tekshiriladi.

Tarjimon dastur inglizcha-o'zbekcha yo'nalishda olib borilganligi bois, dastlab ingliz tilidagi otga xos barcha kategoriyalar o'zbek tilining kategoriyalari bilan

³² Cerstin M., Michael P. State of the art in computational morphology Workshop on Systems and Frameworks for Computational Morphology, SFCM 2009 Zurich, Switzerland, September 4, 2009 Proceedings, - P. 39.

³³ Zangenfeind R., Sonnenhauser B. Russian Verbal Aspect and Machine Translation // Computational Linguistics and Intellectual Technologies Papers from the Annual International Conference "Dialogue" № 13, 2014. - P. 744745.

³⁴ Abdurakhmonova N.Z. Grammatical analyses in machine translation between English and Uzbek // Proceedings of the I International conference on Computer processing of Turkic Languages (TurkLang.2013), - Astana, 2013. - P. 294-297.

solishtiriladi. Lingvistik ta'minotda ingliz tilidagi ot so'z turkumining quyidagi paradigmalari o'rganiladi:

1. *Atoqli va turdosh ot:*
 - 1.1. Atoqli otlar (PropN):
 - 1.2. Turdosh otlar (CN):
 - 1.2.1. Aniq ot:
 - 1.2.1.1. Jonli;
 - 1.2.1.2. Jonsiz;
 - 1.2.1.3. Mavhum.
 - 1.2.2. *Sanalmaydigan ot:*
 - 1.2.2.1. Mavhum;
 - 1.2.2.2. Modda va material.
 - 1.2.3. Jamlovchi ot:
 - 1.2.3.1. Kishi nomlari (oila, xalq);
 - 1.2.3.2. Jonsiz otlar (shakar).
 2. *Otning tuzilishiga ko'ra turlari:*
 - 2.1. Sodda ot;
 - 2.2. Qo'shma ot;
 - 2.3. Juft ot;
 - 2.4. Qisqartma ot;
 - 2.5. Takroriy ot.
 3. *Son:*
 - 3.1. Birlik;
 - 3.2. Ko'plik.
 4. *Kelishik*

Tilshunoslikda atoqli otlarning (N_p) tasniflashi muayyan turlarga ajralatiladi: *kishi nomlari, geografik nomlar, tashkilot nomlari, mahsulotga qo'yilgan nomlar, lavozim yoki daraja, medallar, taxalluslar* v.h. Bosh harfdagi so'zlar bilan boshlangan atoqli otlarni quyidagicha $N_p = \{A, B, C \dots Z\}$ ifodalasak, u holda N_p N bo'ladi. Ayrim vaziyatlarda ingliz tilidagi atoqli otlar o'zbek tilidagi turdosh otlarning qism to'plami ham bo'lishi mumkin, u holda to'plamlar EngPropNn *UzbCN* ko'rinishida bo'ladi: a) hafta kunlari (Monday); b) oy nomlari (April); c) til va millat nomlar (Uzbek) - bu turdagi otlar o'zbek tilida turdosh ot sanaladi. Grafematik jihatdan otga tegishli so'zlarni nafaqat tahlil qilish, balki lingvistik ta'minotda ularni morfologik jihatdan loyihalash muhim bosqichlardan biridir. Zero, grammatik modellar semantik tahlillarga ham asoslanishi, qolaversa, atoqli otlarning ayrim farqli tomonlarini tasniflashda alohida o'ringa ega. Atoqli otlarni tarjima qilishda, dastlab har ikki tilda barcha atoqli otlarning inglizcha-o'zbekcha lug'ati elektron baza shaklida tuzilgan bo'lishi kerak. Negaki, tarjimadagi ayrim xatoliklar bevosita shu o'rinda hal yetiladi. Masalan, Chomsky => <Xomskiy> yoki <Chomskiy>; Regent Street => <Rejent ko'chasi> yoki <Rejent Ko'cha / ko'cha>. Inglizcha atoqli otlarni o'zbek tiliga tarjima qilishda muayyan vaziyatlar o'rganilishi, ya'ni mavjud "qog'oz" lug'atlarda atoqli otlarning transkripsiyasi mavjudiligi hamda ularning o'zbek tiliga qay tarzda o'zlashtirilishi aniqlangan holda ma'lumotlar bazasida berilishi maqsadga muvofiq. Tadqiqotda o'zbek tilida atoqli ot hisoblanmaydigan yoki aytilishida farqlanuvchi so'zlarning alohida lug'at bazaga kiritildi: English - ingliz tili, Monday -dushanba kabi. Otlarning tub yoki yasama so'z ekanligi u darajada muhim emas. Masalan, *dragonfly* so'zining matn ichidan so'z₂ sifatida ajratib olinishida so'zning qanday komponentdan tashkil topganligi e'tiborga olinmaydi. Zotan, aksariyat flektiv tillarda

morfemadan keyin ko‘makchi morfemalar qo‘shilmaydi (-s, -ed va h.k.). Biroq o‘zbekcha so‘zlarning grammatik tahlilida har bir morfema ahamiyatga ega. Zero, konteksda *tub so‘z Otub so‘z, tub so‘z O yasama so‘z, qo‘shimcha Oqo‘shimcha, tub so‘z Ogrammatik shakllangan so‘z* omonimiyasi hodisasi ko‘p uchraydi: *olmaOol+ma; suv+saOkel+sa, gul+don (ot) Obilim+don (sifat)*.

Ingliz tilida qo‘shma otlar qo‘shilib (bluebell) yoki chiziqcha bilan (mother-in-low) yoziladi. Aksariyat hollarda bu kabi otlar har ikki tilda alohida tarzda qo‘llanishi ham mumkin. Qisqartma so‘zlar ham nazardan chetda qolmasligi maqsadga muvofiq. Qisqartmalar ingliz tilida turli so‘z turkumiga oid bo‘lib ham turdosh otga, ham atoqli otga tegishli bo‘lishi mumkin v.s.-yuqoriga *qarang*, v.v.- *aksincha, ye.x.-masalan; NATO (North Atlantic Treaty Organization)-NATO, ISBN (International Standard Book Number)-ISBN, biroq the UNO, the UN (the United Nations Organization)-BMT (Birlashgan millatlar tashkiloti), PC (Personal computer), ShK (shaxsiy kompyuter)*. Bundan tashqari Internet muloqotida *ONNA = oh no, not again "yo‘q, qaytarish shart emas"* kabi ba‘zi akronimlar ishlatiladi. Qisqartma otlarning o‘zgarishsiz qoluvchi shakllari lug‘atdan tayyor holatda olinishi, o‘zgaruvchi birliklar haqida esa terminografiya va leksikografiyada aniq to‘xtamga kelinishi zarur.

So‘zning morfologik tahlili gapning umumiy strukturaviy modelini, qolaversa, leksemaning sintaktik vazifasini belgilashga yordam beradi. Otga tegishli deb topilgan so‘z tokenizatsiya (morfemalarga ajratish) jarayoniga o‘tkaziladi. Ingliz tilida so‘zga qo‘shilishi mumkin bo‘lgan affikslar turi ko‘p emas. Ular sirasiga ko‘plik qo‘shimchasi (-s, -es) hamda egalik qo‘shimchasi (‘s, s’) kiradi. Ingliz tilidagi otlar son kategoriyasi o‘zbek tilidan biroz farq qiladi. Umuman, otlar birlik va ko‘plik shaklida grammatik ko‘rsatgichga ega bo‘ladi, yasaliş xossalariga ko‘ra ko‘plik kategoriyasi istisnoli holatlarda ham aks etishi mumkin. Shu bois bazada har bir otga tegishli bo‘lgan sanaladigan va sanalmaydigan ko‘plik turlari an‘anaviy qo‘shilishi yoki qo‘shilmasligi yoxud ularning ishtirokisiz hosil qilinuvchi ko‘plik kategoriyasiga mansub otlarni bazada indekslash maqsadida har bir turini shartli ravishda ot turiga ko‘ra belgilab olamiz: *N1, N2, N3, N4, N5, N6, N7, Ng, N9, N10*

1- qoida. *O‘zbek tilida koplik qo‘shimchasi bevosita qo‘shilish imkoniyatiga egaligi bois leksemaning qaysi harf bilan yakunlanishidan qat‘iy nazar umumiy yagona model quyidagicha bo‘ladi:*

N + -lar —> kitob+lar; olma+olmalar

2- qoida. Sanaladigan otlar ko‘plikda ifodalanishiga ko‘ra quyidagi model ko‘rinishida bo‘ladi:

I. Noun+(e)s->umumiy ko‘rinish:

1. Ni-tipidagi otlar. Unga <s> yoki < yes> suffikslari qo‘shilganda fonetik hodisaga uchramaydigan leksemalar kiradi hamda ular quyidagicha modelga ega:

1.1. $N(\text{cons.})+s \longrightarrow \text{so‘z} \longrightarrow N_{\text{count}}+s \longrightarrow \text{flowers (gullar), guys (yigitlar);}$

1.2. $\text{PropN}+\text{KQ} \longrightarrow \text{the Germany}+s \longrightarrow \text{the Germanys (germaniyaliklar);}$

1.3. $(N(\text{vow}+y))+s \longrightarrow \text{day}+s \longrightarrow \underset{\text{zo}}{\text{days (kunlar);}}$

1.4. (CN+y)+s —> stand-by+s—>stand-bys (tayyorgarlik holatlari);

1.5. (N(+o)count+s)—>photo+s—>photos (rasmlar).

2. N₂-tipidagi otlar. Bu turga kiruvchi otlarga <es> suffiksi qo'shilganda quyidagi modellar hosil bo'ladi:

1.6. N(+x|s|sh|as|ch|z)count+es—>box+es—>boxes (qutilar);

1.7. N(+o)_{count}+es—>heroes (qahramonlar), potatoes (pomidorlar).

3. N₃-tipidagi otlar. Bu tipga <s> yoki <es> qo'shilganda leksemalarning oxirgi qismida <f> -> <v> ga o'zgaradi:

3.1. N(-f)_{count}+es—>leaf->leaves (yaproqlar);

3.2. N(-fe)count+s—>knife->knives (pichoqlar).

N₄-tipidagi otlar. Bu turdagi otlar <f> undoshi bilan tugallansa-da, fonetik hodisaga uchramaydi:

3.3. N(-f)count+s=>chief+s=>chiefs (rahbarlar).

3.4. N₅ tipidagi otlar. Bularga <es> qo'shilganda y>i hodisaga uchrovchi ot sinfiga mansub so'zlar kiradi:

3.5. N(-y) count+es=>sky+es=>skies (osmonlar);

3.6. N(-que) _{count}+es=>soliloque+es>soliloques (monologlar).

4. N₆ tipidagi otlar. Bu sinfga *Noun->Noun1* komponentli otlar mansub bo'lib, ko'plik shaklida so'zning o'zagida fonetik o'zgarish bo'luvchi otlar kiradi:

1) Goose->geese (g'ozlar), tooth->teeth (tishlar);

2) Child->children (bolalar), ox->oxen (buqalar).

5. N₇ tipidagi otlar. Bularga *Noun->Noun* ko'rinishli tur tegishli bo'lib, ingliz tilida leksik usulda ko'plik hosil qilinsa-da, hech qanday KQ ifodolovchi suffiks qo'shilmaydi, bu holatda N 2e N1):

One sheep (bitta qo'y) <=>many sheep (ko'p qo'ylar)

A fish (bitta baliq) <=>a few fish (bir qancha baliqlar)

6. N₈ tipidagi otlar. Bu kabi otlarga boshqa tillardan o'zlashgan so'zlar tegishli bo'lib, o'zagida mutloq yoki qisman o'zgarish sodir bo'ladi:

Corpus->corpora; stimulus->stimuli; criterion->criteria; datum->data

II. N₉ tipidagi otlar. Bunday turga qo'shma ot tipiga mansub so'zlar tegishli bo'lib, unda tarkibidagi ko'plik qo'shimchasi turli holatlarda kelishi mumkin:

1. N_{9_both}: Tarkibida *man* yoki *woman* leksemasi bo'lgan qo'shma otlar o'zagida o'zgarish bo'luvchi otlarning ²⁹har ikki qismida KQ beriladi va ular alohida

tasniflanadi: men-servants (*erkak xizmatkorlar*), women-doctors (*ayol shifokorlar*);

2. N₉_first: Qo'shma otlarning birinchi qismiga KQ qo'shiladigan otlar (bunday otlar tarkibida, asosan, predloglar ishtirok etadi):

1) *commanders-in chief* (bosh sardorlar), *mothers-in law* (qaynonalar);

2) birinchi qismga KQ qo'shiladigan otlar: *passers-by* (yo'lovchilar), *lookers-on* (tomoshabinlar);

3. N₁₀_last tipidagi otlarning oxirgi qismiga KQ qo'shiladigan otlar: *forget-me-nots* (*bo'tako'zlar*), *stand-byes* (*tayyorgarliklar*).

III. N₁₁ tipidagi otlar. Bu turdagi otlarning so'nggi qismida <s> qo'shimchasi yaxlitlangan bo'ladi.

Ular birlik ma'nosida kelsa-da, o'zidan keyin keluvchi kesimning ko'plikda kelishini taqozo etadi. Bu qo'shimcha so'zning tarkibida yaxlit holda bo'lgani bois tarjima qilinmaydi: *jeans-djensi*, *scissors-qaychi*, *scales-tarozi* v.h. Ba'zi so'zlar har ikki holatda ham ishlatilishi mumkin, biroq tarjimada ular tafovutga ega: *timeuncoun* (vaqt); *timecount+s* (marta)=>2 times (2 marta). Bu holatda sanaladigan ot sanalmaydigan ot bilan omonimik munosabatga kirishadi. Tarjimada inobatga olinishi kerak bo'lgan yana bir jihati shundaki, otga tegishli so'zlar oldidan keluvchi sanoq sonlardan keyin, albatta, ingliz tilida ko'plik qo'shimchasi qo'shiladi, biroq o'zbek tilida ular tarjima qilinganda birlikda bo'ladi: 2 days ->2 kun, 5 hours -5 soat. O'zbek tilida KQ -lar affiksi bo'lib, qo'shilish bilan bog'liq bo'lgan fonetik hodisa uchramaydi. -lar ko'plik ma'nosidan tashqari uslubiy ma'no hosil qiladi: *dadamlar*, *o'zlari*, *kechalari* v.h. Keyingi tadqiqotlarda o'zbek tilidan ingliz tiliga tarjima qilinganda bunday birikmalar tahlilga tortiladi.

Ingliz tilida artikllar (a, an, the) nafaqat ot so'z turkumidagi leksemalar, balki sifat bilan ham ishlatilishi natijasida otlashish hodisasi yuz beradi: *poor* (*kambag'al*) <-> *the poor* (*kambag'allar*). Otlar tipida bunday so'zlar ham qamrab olinishi maqsadga muvofiq.

Ingliz tilida bosh va qaratqich kelishiklari mavjud. Qaratqich kelishigi qo'shimchasi ikki usulda hosil qilinadi, ya'ni 's va of predlogi orqali.

1) N+es=>mother's vase, director's name-> N+ning=>onaning guldoni, direktorning ismi

2) N+of+{DET}+N=>the page of the book-> (N+ning)+(N+EQ)=> kitobning varag'i

3) NPL+s'=>teachers' ->NS+lar+ning—>o'qituvchilarning

4) N1+'s=>children's->N1+lar+ning—>bolalarning

Shu o'rinda so'zlarning <'s> shaklida qo'tlanilishida chalkashlik seziladi:

She's a doctor=>She is a doctor (*Ushifokor*); *Mother's bag* (*onaning so'mkasi*).

Bu holatda <'s> ishtirokidagi gap bo'laklari modeli quyidagicha bo'ladi:

N|P+'s+NP > Nargiza s at ho^me. (Nargiza uyda) yoki Nargiza s tolerable.

(Nargiza sabrli).

$N|P+ 's+N+VP \Rightarrow$ *My uncle's name is Aziz. (Mening amakimning ismi Aziz) yoki Feruza's impression was good. (Feruza taassuroti ijobiy.)*

5) $[N+ 's]$ gap so'nggida kelib o'zidan oldingi so'zga ishora etsa yoki bunday birikmadan keyin ot so'z turkumi qo'llanilmasa, u holda - niki tarzida tarjima qilinadi: $NP|VP+(NP+ 's)+0 \Rightarrow NP+VP+(NP+niki)+Aff \{-mas, -dir, -mi, -ya, \dots\}$.

Our house is bigger than yours. \Rightarrow Bizning uyimiz siznikidan ko'ra kattaroq(dir).
Is our house bigger than yours? \Rightarrow Bizning uyimiz siznikidan ko'ra kattaroqmi? The door of the room (Xonaning eshigi). Bu holatda egalik qo'shimchasi ingliz tilida ikki xil holatda berilishi mumkin:

1) $N+of+N \rightarrow N_{ning} + Neq$

The picture of the book $\rightarrow$ Kitobning rasmi

2) $P \{PP, both, neither, yeither, none\} + \{of+Det.\} + N|P \rightarrow P$ My book $\rightarrow$

Mening kitobim

Both of you $\rightarrow$ Har ikkingiz

None of the students $\rightarrow$ Talabalarning hech qaysi biri O'zbek tilida egalik olmoshi belgisiz holatda ham ishlatilishi mumkin. Ingliz tilida bir necha marta ishlatilgan egalik olmoshi o'zbek tilida takrorlanmaydi, ya'ni u faqat egalik qo'shimchasi sifatida tarjima qilinadi: *He goes his English course \Rightarrow U (uning) ingliz tili kursiga boradi.* Bunday holat quyidagicha modellashtiriladi: $N- > book|Mushtariy|Earth|UN$.

$Det \rightarrow the|a|an|my|this|all|both|clever$.

$V \rightarrow go|go off|have a chat$

Avtomatik tarzda ingliz tilidagi otlarni qo'shimchasiga ko'ra aniqlash yoki uning ega yoki to'ldiruvchiga tegishli ekanligini topish mushkul. Matnni faqat sintaktik strukturasiga ko'ra aniqlashda Chomskiyning ilgari surgan daraxtsimon gap tahlili usuli samarali hisoblanadi.

$S (NP)+VP (V+NP)$; bu yerda $NP=det+N$ // ushbu holatda NP tarkibida egaga ishora yetuvchi PossP kelsa, u holda sintez bosqichida tushirib qoldirilishi hamda o'zidan keyin keluvchi otga PossPga tegishli bo'lgan egalik qo'shimchasini berishi kerak, shunda tarjima me'yorlashadi: $S (NP)+VP (V+NP) \Rightarrow PP+n_{ing}+N+E_G+PP+n_{ing}+N+E_G+V$ - transformatsion metodning bir ko'rinishi gapda muayyan qismlarning tushirib qoldirilishidir. Yuqoridagi chizmada eng sodda gap qolipi berilgan, biroq gapni, shuningdek, qolgan ikkinchi darajali bo'laklar bilan kengaytirish mumkin, u holda tarmoqlanish yana davom etadi. Natija: *My brother reads his book \Rightarrow Mening akam kitobini o'qiyapti.*

O'zbek tilida egalik va ko'plik qo'shimchalarining qo'shilishiga doir imloviy qoidalar qabul qilingan. Ushbu qoidalarni modellashtirish zarur bo'lgan jihatlarini tahlil qilamiz. Masalan, matn quyidagicha berilgan bo'lsin: *My town is Andijan.*

I. Har bir komponent morfologik sathda aniqlashtiriladi: $PossP+N+is (to be)$

+PN

II. Gapning bosh bo‘laklari oydinlashtiriladi: *S (ega) + P (kesim)*

III. Fe’lgacha bo‘lgan qism ega guruhi hisoblanadi va uning modeli

aniqlashtiriladi:

IV. Ega guruhining tipi aniqlashtiriladi: ot guruhi (NPh) turkumi tahlil qilinadi.

Ot guruhi (NPh) ikkiga bo‘linadi: 1) Ph+Noun->otdan oldin keluvchi birikmalar; 2) *Noun+Ph->* otdan keyin keluvchi birikmalar. Har ikki holatda ham birikmalar otga tobelanadi. NPh gapda barcha gap bo‘laklari bo‘lib kela oladi. NPh =N+ i Prep. + i Article+ i Adj. +Noun-> the man in the black coat (qora kurtkadagi kishi), the book about my life (mening hayotim haqidagi kitob). Bu kabi birikmalar aksariyat hollarda ingliz tilidagi predlog bilan bog‘lanadi. Ph+N ning strukturasi quyidagicha bo‘ladi: Q+Det.+AdjM+NM (noun modifier)+N. Bu kabi birikmalar quyidagi morfologik modelda shakllanadi:

1) {Q (quantifier: one, two.) | M (modifier: noun, adjective, participle, gerund) | D (determiner: Possessive pronoun, article, Pronoun, numeral)} +N—> K (kvantifikator) | M (modifikator: ot yoki sifat) | D (aniqlovchi, PP (kishilish olmoshi) {men, sen+ning(n₀₋₁)} || {u, biz, siz, ular+ning} +N+EQ || {u, biz, siz, ular+ning} +N+EQ // tahlil quyidagi algoritm bo‘yicha olib boriladi:

- egalik qo‘shimchasi shaxs-songa qarab tanlanadi: my ->im|-m, your ->ing|-ng, his, her, its ->-i, our->-miz|-imiz, your ->-ngiz|-ingz, their ->-lari

- o‘zbek tilida egalik qo‘shimchasi qo‘shilganda fonetik hodisaga uchraydigan ot turkumi turlari aniqlanib, muayyan turga biriktiriladi:

1) tovush tushish: N_{tt} {burun, shahar, o‘g‘il, ko‘ngil.} +EQ-> shahar+im—> shahrim;

2) tovush ortirish: N_{to} {parvo, obro‘, mavqe, mavzu, avzo...} +y+EQ-> parvo+im—> parvoyim;

3) tovush almashinish k<g, q<g‘: N_{ta}{yaxshilik, tilak, bilak, kubok, tayoq.}+EQ—>tilak+im—>tilagim;

Shuningdek, ingliz tilidagi ayrim *artikl+ot* komponentli birikmalarda semantik xususiyat yaqqol seziladi: The English - ingliz xalqi; English - ingliz tili. Ot guruhi tarkibida artikl kelganda muayyan qoidalar kiritilishi kerak: 1) o‘zbek tiliga tarjima qilinganda semantik ma’no anglatuvchilar guruhi; 2) tarjimada ma’no anglatmaydigan artiklli otlar guruhi. Noaniq artikl *a, an* quyidagi holatlarda tarjima qilinadi: 1) a|an+Number+N—> a hundred (bir yuz); 2) a|an +Time —> an hour; a year| minute| month (bir soat, bir-yil, bir minut, bir oy); 3) a|an+Number+N+N—> a four floor building (to‘rt qavatli bino).

Sanalmaydigan otlar na birlikka, na ko‘plikka mansub. Ularning oldidan a / an artikli odatda ishlatilmaydi. Biroq artikllar hisob so‘zlar bilan kelganda ularning oldidan artikl kelish holatlari uchraydi: *a pound of butter, a piece of furniture*.

Ularning modeli a|an+NW (numeral word)+P (predlog)+UnCN. Ayrim birikmalar quyidagi tarkibda voqealanadi: Num+(N+s)+UncN=>Num+N+N (bu

holatda o'zbek tilida KQ qo'shilmaydi) two spoons of sugar (ikki qoshiq shakar). So'zma-so'z tarjimaga yo'l qo'ymaslik uchun bu kabi birikmalar ma'lumotlar bazasida berilishi kerak.

Aniq *the* artikl belgilangan qoidalarga ko'ra qo'llaniladi, biroq ayrim hollarda ma'noga ta'sir etgan shakllarni ham inobatga olish zarur. $Art+PN+N|Art+(PN+s) \Rightarrow PN+lar$ (bu holatda artikl atoqli otga tegishli ismni, unga aloqador boshqa ismlarni ham nazarda tutadi). *The Johnson family / the Johnsons (Jonson va uning oila a'zolari)*. Shuningdek, $Art+Adj+V \Rightarrow Adj+lar$ (bu holatda sifatning oldidan aniq artikl ishlatilib, uning aniqlanmishi bo'lmasa otlashib, ko'plik ma'nosida ishlatiladi). *The poor live in this village (Kambag'allar bu qishloqda yashaydilar)*.

Kompyuterli morfologik tahlilda modellashtirish ahamiyatli hisoblanadi. Ayniqsa, sintaktik tahlilda erkin kontekstli daraxtsimon modellashtirish usuli komponentlarni ajratishda qulay. Model u yoki bu ob'ektni o'rganishga xizmat qiluvchi vositadir. Tabiiy tilni butunlay bir qolipga solish to'g'ri emas, biroq mashina tili uchun modellashtirish, tilni formal ko'rinishga keltirish natijaning aniqligiga xizmat qiladi. So'zlarning semantik maydonlari, valentlik xususiyatlari ham inobatga olinsa, maqsadga muvofiq. Tahlilning to'g'ri olib borilishi dasturning algoritmik bosqichini belgilab beradi.

Shartli qisqartmalar: N-noun, Art-article, NPL-noun plural, NSG-noun singular, N1 - irregular plural noun, N(cons.)-Noun ending with consonant, N (vow)- Noun ending with vowel, EQ-egallik qo'shimchalari, KQ-ko'plik qo'shimchasi, PP-possessive pronoun, P-pronoun, CN-compound noun, n_{0-1} - birinchi elementdan ilk elementni olish, | (mantiqiy "yoki" belgisi), Ntt - o'zagida ikkinchi bo'g'in unlisi tushib qoluvchi otlar, N_{to} - so'z o'zagining oxirgi qismiga 'y' undoshi qo'shiluvchi otlar, N_{ta} - so'z o'zagining oxirgi qismidagi undosh muayyan harfga o'zgaruvchi otlar.

Shartli belgilar: i -ixtiyoriy qatnashish, {}-bo'lishi mumkin, $\rightarrow$ ekvivalentlik hosil qiladi, e -tegishli, 0-bo'sh to'plam, // izoh belgisi, |-yoki

Ingliz tilida fe'lli frazemalarning shakl va ma'no jihatdan murakkab kategoriyaga tegishli ekanligini yuqorida tahlil qilgandik. Fe'l o'zbek tilida ham semantik jihatdan alohida e'tibor qaratish lozim bo'lgan kategoriyadir. Jumladan, ko'makchi fe'lli so'z qo'shilmasi (KF) garchi tadqiqotimiz ingliz tilidan o'zbek tiliga tarjima qilishga yo'naltirilgan bo'lsa-da, ularning alohida modellari va semantikasini o'rganmaslikning imkoni yo'q. Quyida KF modellari keltirilgan (MV -asosiy fe'l, HV -ko'machi fe'l): 1) $MV+HV \Rightarrow$ ko'rib qoldi; 2) $MV+HV_{(1)} + HV_{(2)} \Rightarrow$ ko'rsatib bera oldi; 3) $MV+HV_{(1)} + HV_{(2)} + HV_{(3)} \Rightarrow$ aytib berib qo'ya qoldi; 4) $MV+HV_{(1)} + HV_{(1)} \Rightarrow$ o'qib tura tur; 5) $MV_{(1)}+HV_{(1)} \rightarrow$ tura tur; 6) $MV_{(1)}+HV_{(1)} + HV_{(2)} \rightarrow$ turib tura qolgin.

Modelda ko'rsatilganidek, asosiy fe'l ko'makchi fe'l sifatida takrorlanishi mumkin³⁵. Biroq tuzilish jihatdan KFlar qo'shma fe'llarga ham o'xshab ketadi:

³⁵ Абдураъмонова Н. О'закдош феълларнинг кетма-кет қўлланилишига доир / Тилшуносликка илк қадам (то'плам III), - Т., 2007. - Б. 55-59.

- oshib tushmoq->qo'shma fe'l
- oshib ketmoq->Ko'makchi fe'lli so'z qo'shilmasi

Yana bir jihati shundaki, ko'makchi fe'llar qo'shma fe'llar QF va iboralar tarkibida kelib, fe'lga qo'shimcha ma'no yuklashi mumkin:

- Nonushta qilmoq (have a breakfast) —>nonushta qilib berdi(QF).
 - Kapalagi uchib ketmoq (to be afraid) —> kapalagini uchirib yubordi (Ibora);
- Morfologik analizator har bir birlikni kontekstualogik jihatdan to'g'ri aniqlab berishi kerak. O'zbek tilida fe'lga qo'shilishi mumkin bo'lgan qo'shimchalar quyidagilar:

Inkor shakl-N {-ma| -mas| -may}

Shart mayli-CF {-sa}

Zamon -T {-a|-y|-yap|-moqda|-yotir.}

Kesimlik shakllari -PF {-man|-san|-dir.}

Nisbat -VF {-t|-tir|-giz, |-kiz.}

Fe'lining vazifadosh shakllari -NF {-gan|-kan|-qan|-b|-ib...}

Shaxs-son P {-im|-ing|-k|-ngiz|-lar| -man| -san|-k|-ngiz}

Maqsad -PS {-moqchi}

Morfologik lug'atda ajratilgan barcha so'z turkumlariga qo'shilishi mumkin bo'lgan qo'shimchalarning bazasi Excelda tayyorlanib, qo'shimcha olishi mumkin bo'lganlarga 1, olmaydiganlarga 0 berildi:

1		nisbat														
2		l	il	n	in	giz	kiz	qiz	kaz	qaz	gaz	ar	dir	tir	sh	ish
3	och H_A		1				1					1		1		1
4	abadiylash											1		1		1
5	abadiylashtir		1			1										1

Fe'ning faqat o'zak shakli kiritilgani bois, o'zak shaklida u boshqa so'z turkumlari bilan omonim bo'lib qoladi. Bu holatda uning omonimlik hosil qiluvchi varianti bir so'z turkumiga tegishli bo'lsa, H_A (homonym->adjective), agar fe'ning o'zida bo'lsa, H_V (homonym ->Verb), agar bir necha so'z turkum doirasida kuzatilsa, H_L (homonym lists) shaklida berildi. Bunday holatda berilishi morfoanalizni yanada osonlashtirishga xizmat qiladi.

Fe'lga qo'shilishi mumkin bo'lgan qo'shimchalar kombinasiyasi:

- MV+ N+PF=> o'qimayman
- MV+ N+ T+PF=> o'qimadim
- MV+ CF +PF=> o'qisam
- MV+ N +CF +PF=>o'qimasam
- MV+ T+PF=>o'qiyapman
- MV+VF=>o'qittir

- MV+NF=>o'qigan

- MV+PS+PF=>o'qimoqchiman

Qo'shimchalar KFlarning har ikki qismi ham yetakchi, ham ko'makchi fe'lga qo'shilishi mumkin:

- MV+N+NF)+HV=> ko'rmay qoldi

- (MV+ NF)+(HV+ N)=> ko'rib qolmadi

- (MV+N+NF)+(HV+ N)=> ko'rmay qolmadi

88

- (MV+V+NF)+HV =>ko'rsattirib qo'ydi

- (MV+NF)+(HV+ PF)=>ko'rib borsam

- (MV+NF)+(HV+ N+PF)=>ko'rib bormasam

- (MV+NF)+(HV+ VF+N+T)=>eb ko'rgizmadi

Bundan tashqari fe'ning noaniq shakli *edi, ekan, emish* hamda bog'lama vazifasida keluvchi fe'llar ham bo'lib, asosiy fe'lga qo'shimcha ma'no ottenkasini yuklaydi: *bo'lmoq, hisoblanmoq, sanalmoq, deyilmoq*.

a) N|Adj.|Num.|Pron.|Adv.+bo'lmoq=> Agar soat o'n bo'lsa, sizga qo'ng'iroq qilaman (If it is 10 o'clock, I shall call you);

b) yo'q/bor/oz/ko'p/zarur/lozim/kerak+bo'lmoq=> O'ylagan orzularim bir pasta yo'q bo'ldi (My thinking dreams disappeared abruptly).

d) Infinitiv+kerak/lozim/shart/darkor —> Xona tozalanishi kerak (The room needs cleaning. - The room needs to be cleaned).

Ko'rganimizdek, har qaysi tilning grammatik kategoriyasini chuqur tahlil qilmasdan turib matn sintaksisi yoki semantikasi haqida gapirish mumkin emas. Avtomatik tahlilning quyidagi turlari ko'rsatiladi³⁶:

Ko'rganimizdek, grammatik kategoriyalar morfologik tahlilning asosiy qismi hisoblanadi. Ayrim manbalar ko'ra avtomatik tahlil quyidagi bosqichlardan iborat bo'ladi³⁷: 1) leksikondan stem tahlil qilish; 2) leksikondan so'z shakllarini tahlil qilish; 3) mantiqiy usulda tahlil qilish; 4) lug'atsiz jadval yordamida tahlil qilish. Masalan, fe'lga qo'shilishi mumkin bo'lgan qo'shimchalar ketma-ketligiga e'tibor bersak, u holda ulaning tarmoqlari quyidagicha bo'lishi mumkin:

³⁶ Марчук Ю. Компьютерная лингвистика. -Москва, 2006. - С. 65.

³⁷0'ша жойда.

1.	ga	kelishik	C
2.	gan	o'tgan zamon	Tpast
3.	gan	tugallangan zamon	PP
4.	l il	O'zlik nisbat	Vact.
5.	l il	majhul nisbat	^v pass.
6.	a	hozirgi zamon	Tpres.
7.	m	shaxs-son	Pi
8.	i	3-shaxs egalik qo'shimchasi	PossP3
9.	im	1-shaxs egalik qo'shimchasi	PossPi

- Wstem +PP+PossP=>tuzal+gan+l=> anlq nisbat

Ma'lumotlarni tahlil qilishda semantik tahlil ham talab etiladi. Affikslar tur va vazifalarda keladi. Fe'ning bo'lishsizlik shakllari turlicha. Bu shakllar ma'noga o'z ta'sirini ko'rsatadi:

I. MV +ma=>o'qimadi (He did not read)

II.(MV +may) HV => O'qimay qo'ydi (He is not used to read)

III.MV (HV+ma)=>O'qib qo'ymadi (He did not read any more)

IV.(MV +may) (HV +ma)=>O'qimay qo'ymadi - tasdiq ma'no (Surely he read (o'tgan zamon))

V.(MV+PP)+emas=>O'qigan emas (He has never read)

VI.(MV+PP+Poss) yo'q=>O'qigani yo'q (He did not read)

VII. Na MV {CV, MV, VC}=>na habar oldi | na o'qidi | na berib ketdi
(Neither he read)

VIII. (MV +mas) edi=>O'qimas edi (He did not use to read)

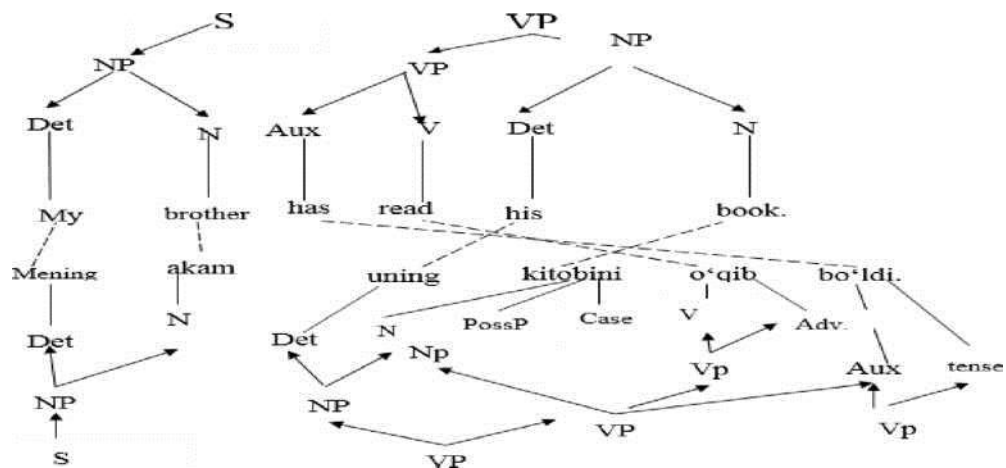
IX.(MV +ma+gan)+ekan+P=>O'qimagan ekanman (I did not read)

X. Na (MV +ma)=>Na o'qimadi (He did not read)

Quyidagi jadvalda grammatik moslashuvni ko'rishimiz mumkin:

Inglizcha -O'zbekcha- Inglizcha tarjima					
	Kecha	ba'	talabalar	darsga	kelmadilar
Yesterday					
some					
of					
the					
students					
did					
not					
come					
to					
the					
lesson					

N.Chomskiyning erkin kontekstga asoslangan sintaktik tahliliga asosan gap qurilmasidagi ishtirokchilar munosabatini kuzatishimiz mumkin:



Tobelanish munosabatiga ko'ra daraxtsimon model gap modelini ifodalaydi. Har ikki tilda gap bo'laklari so'z turkumlari guruhlariga muvofiq tasniflanishi ushbu tasvirda berilgan. Matndagi semantik aloqani o'rganish uchun lingvistik tahlilda grammatik kategoriyalarni o'rganish jarayoni to'xtagani yo'q, chunki morfologik struktura hamma narsani hal qila olmaydi: *read somebody like a book* - *biror kishini judayaxshi tushunmoq*; *book club* - *kitobxonlar to'garagi*; *speak by the book* - *aniq ma'lumotga tayanib gapirmoq*. Yuqorida ko'rganimizdek, *book* so'zi ot turkumi bo'lganda ham muayyan birikmlarda semantik jihatdan ma'no farqlaydi. Ayniqsa, so'z birikmasi va iboralarning omonimiyasi ingliz tilida ko'p kuzatiladi: *It is a piece of cake* - 1) *Bu tort bo'lagi*; 2) *Bu juda oson*. Shunday ekan, lingvistik bazada so'zning barcha imkoniyatlari aks etishi maqsadga muvofiq.

Ma'ruza № 6

Kompyuter lingvistikasida avtomatik tarjima yo'nalishi

Reja

1. Kompyuter lingvistikasida avtomatik tarjima yo'nalishi.
2. Avtomatik tarjima yo'nalishining tarixi. Tarjimon dasturlar.
3. Avtomatik tarjima yo'nalishining zamonaviy holati va istiqboli.

Tayanch so'z va iboralar: *Matn tahriri, avtomatik tarjima, til o'qitishni avtomatlashtirish, elektron lug'atlar, tezaurus lug'atlar, kompyuter leksikografiyasi, kompyuter lingvodidaktikasi, on-line tarjima tizimi, ekvivalent moslik, transformatsion moslik, tarjima protsedurasi, generatsiya, translyatsiya, bilingv bazasi, example-based translation (namunalarga asoslanadigan tarjima), rule-based translation (algoritmga asoslangan tarjima), translation memory (tarjimon xotirasi).*

Informatsiyaning geometrik progressiya asosida keskin ko'payishi ularni boshqa tillarga tezkor tarjima qilinishini ta'minlovchi tizimlarga bo'lgan ehtiyojni keltirib chiqardi. Bunga javoban mashina tarjimasi yoki avtomatik tarjima yo'nalishida amaliy tadqiqotlar yaratila boshlandi. Eng ahamiyatli tomoni shundaki, mashina tarjimasiga oid izlanishlar kompyuter lingvistika fanining yuzaga kelishida tayanch nuqta bo'lib hisoblanadi.

Avtomatik tarjima bo'yicha dastlabki g'oyalar ingliz olimi Charlz Bebbidj tomonidan 1836-1848 yillarda ilgari ~~38~~irilgan. Uning fikricha, mexanik-elektron

mashinalar tillar o'rtasidagi struktur o'xshashliklar asosida kodlashtirilgan avtomatik tarjimani amalga oshirishi mumkin. Uning loyihasi bo'yicha 50 ming ta so'z xotira kuchiga ega bo'lgan elektron mashinalar 100000 ta so'zni avtomatik ravishda tarjima qilishi zarur bo'ladi.³⁸

1949 yilda Ch.Bebbidjing g'oyalaridan ta'sirlangan amerikalik olim Uorren Uiver mashina tarjimasi tizimini yaratishning nazariy asoslarini ishlab chiqdi hamda lingvistika, deshifrovka, programmalashtirish nazariyasi bo'yicha mutaxassislarni hamkorlikka chorladi.³⁹ Shundan so'ng AQSHda MT bo'yicha kollaboratsiyalar, hamkorlik loyihalari yaratildi. Mashina tarjimasi bo'yicha birinchi xalqaro konferensiya 1952 yilda Massachusetts texnologiya institutida o'tkazildi.

1954-yilda AQShdagi Jorjtaun Universitetida dunyoda birinchi marta avtomatik tarjima tajribadan o'tkazildi. U **GAT** tizimi (inglizcha "*Georgetown Automatic Translation*") birikmasining qisqartmasi) deb ataladi. Bu tajribalar IBM-701 kompyuterida o'tkazildi. Bu tajribaning bazasi algoritmlardan iborat bo'lib, 50000 ta so'z (60 ta gap)ni rus tilidan ingliz tiliga tarjima qildi.⁴⁰

Rossiyada mashina tarjimasi yo'nalishining pionerlari D.Yu.Panov, A.A.Lyapunov, I.S.,Muxinlar sanaladi.⁴¹ Shundan so'ng O.S.Kulagina, I.A.Melchuk, Yu.A.Motorin, Yu.N.Marchuk, R.G.Piotrovskiy kabi olimlar ularning izdoshlari sifatida avtomatik tarjima tizimini rivojlantirdilar. Sobiq SSSRda mashina tarjimasi bo'yicha birinchi tajriba 1955 yilda o'tkazildi. Shundan so'ng O.S.Kulagina va I.A.Melchuk rahbarligi ostida Amaliy matematika institutida mashina tarjimasi bo'yicha uchta yirik tajriba tizimi yaratildi: fransuzchdan ruschaga (FR-1), inglizchadan ruschaga va fransuzchadan ruschaga (yangi variantda). 1959 yilda Moskva davlat lingvistik universitetida mashina tarjimasi laboratoriyasi ochildi. Moskvada 1974 yilda INFORM-ELEKTRO institutida fransuzcha-ruscha (ETAP-1), inglizcha-ruscha (ETAP-2) hamda nemischa-ruscha (NERPA) avtomatik tarjima tizimlarini yaratish bo'yicha amaliy tadqiqotlar jonlandi.

1960-yilda sobiq SSSR Fanlar Akademiyasining avtomatik tarjimaga bag'ishlangan maxsus yig'ilishi bo'lib o'tdi. Bu yig'ilishda avtomatik tarjima bo'yicha professor A.Belonogov rahbarligida maxsus laborotoriya tashkil qilindi.

1966-yilda AQShning fan masalalari bo'yicha kongress qo'mitasida avtomatik tarjimaga oid ma'ruza muhokamasi (**Automatic Language Processing Advisory Committee – ALPAC REPORT**) bo'lib o'tdi.⁴² Unda avtomatik tarjima bo'yicha "*Language and Mashines: Computers in Translation and Linguistics*" ("*Til va mashinalar: tilshunoslikda va tarjimada kompyuterlar*") deb nomlangan salbiy ma'nodagi ma'ruza qilindi va yuqori darajada sifatli avtomatik tarjimani ta'minlovchi universal tizim yaratish natijasizligi va imkonsizligi ta'kidlandi. Shuningdek, ma'ruzada ko'rsatilishicha, "tabiiy tarjima" (inson

³⁸ **Po'latov A., Muhamedova S.** Kompyuter lingvistikasi. – Toshkent, 2007. – B. 52.

³⁹ **Баранов А.Н.** Введение в прикладную лингвистику. – М.: Эдиториал УРСС, 2001. –С. 89.

⁴⁰ **Убин И.И.** Современные средства автоматизации перевода: надежды, разочарования и реальность // "Перевод в современном мире". М., ВЦП, 2001 г., стр. 60-69.

⁴¹ **Панов Д.Ю., Ляпунов А.А., Мухин И.С.** Автоматизация перевода с одного языка на другой. Сессия по научным проблемам автоматизации производства. Изд-во АН СССР, 1956.

⁴² **Слокум Дж.** Обзор разработок по машинному переводу: история вопроса, современное состояние и перспективы развития // Новое в зарубежной лингвистике. Вып. XXIV. Компьютерная лингвистика. М., 1989. - С. 358.

tarjimasi) mahina tarjimasidan arzonroq va sifatliroq. Shundan keyin jahon miqyosida avtomatik tarjimaga oid tadqiqotlarni moliyalashtirish keskin kamayib ketdi. Ular faqatgina Yaponiyada davlat tomonidan moliyalashtirildi. Qolgan davlatlarda avtomatik tarjimaga oid loyihalar xususiy yirik kompaniyalar hisoblanagan “Generel Motors”, “Siemens”, “LG”, “Phillips” kompaniyalari bugungi kunda dunyoda avtomatik tarjimaga oid izlanishlarni mablag’ bilan ta’minlab turadi.

AQSH harbiy havo kuchlari buyurtmasi bo’yicha D.Tom rahbarligida **SYSTRAN** avtomatik tarjima sistemasi ishlab chiqildi. Mazkur tizim yordamida bir yilda 2,5 mln. sahifa matn ingliz tilidan nemis, fransuz, ispan, yunon, italyan tillariga hamda fransuz tilidan ingliz, ispan, italyan tillariga tarjima qilinadi. Avtomatik tarjima bo’yicha dunyodagi eng yirik loyiha 1991 yilda AQShda NASA mutaxassisleri tomonidan yaratilgan: **DIANA**. Uning hajmi 10 million ta so’zni tashkil etadi. U dunyoning 6 ta yirik tiliga asoslanadi: ingliz, rus, fransuz, nemis, arab, ispan.

Rus olimlari ham professor Lyapunov rahbarligida “Vista Technology” kompaniyasi “**Retrans Vista**” loyihasini ishlab chiqdi. Uning tarkibiga kiruvchi 3 mln. 400 mingta so’z xotiraga kiritildi. Shunda 1mln. 800 mingtasi ruscha-inglizcha, 1 mln 600 mingtasi inglizcha-ruscha so’zdan iborat.

Bugungi kunda avtomatik tarjimini quyidagi sistemalari mavjud. 40 mln.ta so’z xotira kuchiga ega **YEUROTRA** sistemasini – Yevropa olimlari tomonidan ishlab chiqilgan. U barcha Yevropa tillari bo’yicha avtomatik tarjimini amalga oshira oladi. **SOKRAT** sistemasini 40 mln. ta so’z xotira kuchiga ega. Dunyodagi 40 ta tilda avtomatik tarjimini amalga oshiradi. Ya’ni qaysi tilda Windows operatsion tizimi versiyasi mavjud bo’lsa, shu tilda avtomatik tarjimini amalga oshira oladi. Lingvo sistemasini (lingvo kompyuter lug’ati) – bu lug’atga ham dunyoning 40 tilidagi ma’lumotlari jamlangan. 1990 yillarda Amerikada **Global Link**, Rossiyada **PROMT** avtomatik tarjima tizimlari yaratildi.

Internet tizimining rivojlanishi natijasida avtomatik tarjima global tarmoqda ham tashkil etildi, ya’ni hozirgi kunda **online-tarjima tizimi** ham faol yo’lga qo’yilgan.

Avtomatik tarjimaning kamchiliklari:

1. So’zlar avtomatik tarjima jarayonida to’g’ridan-to’g’ri tarjima qilinadi. Shuning uchun tarjimada grammatik qoidalarga deyarli rioya qilinmagan holda amalga oshiriladi
2. Avtomatik tarjima qilingan matnni uslub jihatdan tahrirlab chiqish zarur bo’ladi.
3. Hozirgi mavjud avtomatik tarjima qiluvchi sistemalar so’zlar birikuvini ham deyarli noto’g’ri tarjima qiladi.
4. Avtomatik tarjima polisemiya, omonimiya, troplar, frazeologizmlar, idiomalarni e’tiborga olmaydi.

Masalan, *The bat looks like a mouse with wings* gapini rus va o’zbek tiliga tarjima qilishda har bir so’zning polisemantik va omonimlik kabi semantik qirralarini e’tiborga olish zarur. Aks holda tarjima muvaffaqiyatli chiqmaydi. Jumladagi *bat* so’zi “ko’rshapalak”, “belkurak”, “to’qmoq”, “kichik ziyofat” kabi ma’nolarni, *look* so’zi esa “qaramoq”, “nigoh”, “ko’rinish” ma’nolarini anglatadi. Bunda mazkur so’zlarni rus yoki o’zbek tiliga tarjima qilinayotganda mos ekvivalentini tanlash uchun kontekst ~~40~~ uning distributsiyasidan (qurshovidan)

kelib chiqish zarur. Ushbu holat, ayniqsa, idiomatik ifodalar qatnashgan matnlar tarjimasida murakkablik tug'diradi. Masalan, *You are an apple of my eyes* jumlasida "Siz ko'zlarimning olmasisiz" tarzida so'zma-so'z tarjima qilinadi, chunki "apple" so'zining denotative ma'nosi "olma"dir. Lekin kontekstdan kelib chiqib tarjimaga yondashilsa, to'g'ri tarjima yuzaga chiqadi. Bunda jumlaning "an apple of my eyes" qismi idioma ekanligini nazarda tutish lozim va u "Siz ko'zimning oq-u qorasisiz" tarzidagi gap hosil bo'ladi.

Tarjimadagi moslik 3 xil bo'ladi:

1) ekvivalent moslik – bunda bir tildagi so'z boshqa bir tildagi so'zga tarjimada to'la muvofiq keladi. Masalan, rus tilidagi **"книга"** so'ziga inglizcha **"book"** so'zi tarjimada to'la mos keladi.

2) variantli moslik – bunda sinonimiya va polisemiya asosida so'zning ma'nosiga mos keluvchi variantlar ham chiqariladi.

3) transformatsion moslik – bu asosan birikmali ifodalar tarjimasida ishlatiladi, bunda mazmunga ta'sir ko'rsatmagan holda shakliy o'zgarishlar asosida tarjima amalga oshiriladi. Masalan, ingliz tilida **"the decision of government"** ifodasi **"the government's decision"** tarzida transformatsiya qilinishi mumkin. Har ikkala holatda ham mazmun (**"hukumat qarori"**) o'zgarmaydi.

Shularni nazarda tutgan holda aytish mumkinki, avtomatik tarjima, ayniqsa, rasmiy til doirasida muvaffaqiyatli amalga oshiriladi. So'zlashuv va badiiy uslubdagi matnlarning avtomatik tarjimasi hali hanuz maksimal darajada hal etilgani yo'q. Tipologik jihatdan o'xshash bo'lmagan, bir sistemaga mansub bo'lmagan tillarning avtomatik tarjimasi masalasi bir qator muammolarni keltirib chiqarmoqda. Demak, lingvistik tipologiyadagi farqlarni izchil o'rganish, lingvistik universalialarni aniqlash, tillardagi frazeologizmlar, idiomalar, paremlar, sinonimlar, omonimlar naborini qat'iy belgilash asosida mexanik tarjimadagi muammolarni bartaraf etish mumkin.

Eng sifatli tarjima insonning aralashuvi bilan yuzaga chiqadi. Chunki tarjimon inson tabiiy tilining nozik qirralarini, semantik, grammatik va pragmatik jihatlarini nazarda tutgan holda tarjimaga yondashadi. Bunda yondashuv quyidagicha bo'lishi lozim: Mexanik tarjima yordamida tanlangan matnning tarjima qilingan dastlabki versiyasiga ega bo'lamiz, bundan so'ng inson tarjima jarayoniga aralashadi, ya'ni u mavjud tarjimani tahrir etadi – qo'shimchalar qo'shilish tartibi, gap bo'laklarining joylashuvi, so'zlarning semalariga asosan to'g'ri tanlanganligi, polisemiya, omonimiya, sinonimiyaning e'tiborga olinishi, jumalarning stilistik jihatdan to'g'riligi kabi parametrlarni tekshiradi.

Rus olimi Yu.I.Shemakin o'zining *"Начало компьютерной лингвистики"* kitobida avtomatik tarjimaning 2 xil modelini va tarjima protsedurasini chizmalar asosida tushuntirib bergan. Birinchi yondashuv "matn-mazmun-matn" modeli va semantikaning formal tavsifiga tayanadi. Ikkinchi yondashuv esa "matn-matn" modeli va tarjima mosligiga asoslangan.

Birinchi yondashuvda tilning semantikasiga maksimal darajada e'tibor bergan holda avtomatik sifatli tarjimaga erishishga harakat qilingan. Tizimning lingvistik asosi *"matndan unng mazmuniga qarab borish"* (*"from surface structure to deep structure"*) hisoblanadi. Bunda tarjima tilning morfologik, sintaktik va semantik tahlili, luga'tlar, grammatik qoidalar, matnlar korpusiga asoslanib amalga oshiriladi.

Ikkinchi yondashuvda matn ortiqchaliklarga ega ko'pbosqichli tizim sifatida baholanadi, bunga ko'ra matnning yuqori sathida ko'proq informatsiya tashuvchi leksik birliklar, quyi sathida esa ozroq informatsiya beruvchi birliklar joylashgan bo'ladi. Bunda so'zma-so'z tarjima amalga oshirilib, ikki tildagi gaplar strukturasi yagona superstrukturaga birlashtiriladi. Bunda avtomatik tarjima modeli ikki holatda bo'ladi: a) *generatsiya holati* tarjimaning ishlashini – lug'atlar jamlash, analiz algoritmlari, aniq bir maqsadga yo'naltirilgan sintezni ta'minlaydi; b) *translyatsiya holati* sistemaning mustaqil ravishda tarjima ustida ishlashini ta'minlaydi.

Ikkinchi yondashuvga ko'ra avtomatik tarjima quyidagi protseduralar asosida ishlaydi:

1-protsedura – qidiruv tizimi bilan bog'liq bo'lib, bunda lug'at bazasidan so'zlar, birikmalar, iboralar izlanadi, so'z-shakllarning morfologik analizi amalga oshiriladi.

2-protsedura – so'zlarning sintaktik vazifasi, grammatik kategoriyalar – zamon, shaxs-son kabilar aniqlanadi.

3-protsedura – bir yoki ko'p ma'noli so'zlar tarjimasi bayon qilinadi.

4-protsedura – jumlani morfologik va sintaktik jihatdan shakllantirish jarayoni amalga oshiriladi.

5-protsedura – transformatsion o'zgarishlar (matn yoki jumla strukturasi uning mazmuniga jiddiy ta'sir etmagan holda o'zgartirish; masalan, Kompyuter yoqildi va musiqa yangradi jumlasini "Kompyuter yoqilgach/ kompyuter yoqilishi bilan/kompyuter yoqilib musiqa yangradi" tarzida o'zgartirish) hosil qilinadi (agar zarurat tug'ilsa).

6-protsedura yakuniy bosqich bo'lib, bunda murakkab grammatik formalarning sintezi amalga oshiriladi.

Mashina tarjimasi strategiyasi sekin-astalik bilan o'zgarib organ. Ilk tajribalarda qo'llanilgan tarjima strategiyasi "*tog'ridan-to'g'ri tarjima strategiyasi*" deb ataladi, bunda faqat tarjima jarayonida so'zma-so'z tarjima mufoviqligi e'tiborga olinadi. Ushbu strategiya tarjima sifatini va samaradorligiga salbiy ta'sir ko'rsatadi. Shuning uchun ushbu strategiyani tanqid qilgan holda "*vositachi til yordamida amalga oshiriladigan tarjima strategiyasi*" ishlab chiqildi. Ayniqsa, bu strategiya grammatik strukturasi va tipologik xususiyatlari bir-biridan farqlanuvchi tillar o'rtasidagi avtomatik tarjima tizimlarida faol tatbiq etiladi.

Avtomatik tarjimada inson hamda kompyuterning ishtiroki quyidagi tarzda bo'ladi: a) *postediting* – matn (mashina-tarjimon) kompyuter yordamida tarjima qilinadi, so'ng inson-muharrir (inson-tarjimon) uni tahrir qiladi; b) *preediting* – bunda inson matnni mashina-tarjimonga moslaydi, so'ng uni kompyuterga havola etadi; c) *interediting* – bunda tarjimada murakkablik tug'diruvchi holatlar duch kelganda inson mashina-tarjimon ishiga aralashadi.⁴³

Matnni bir tabiiy tildan boshqasiga tarjima qilish jarayonidagi murakkabliklar va muammolarga nazariy va amaliy jihatdan yechim topish bo'yicha hali hanuz izlanishlar davom etmoqda. 1984 yilda Yaponiyaning Kioto universiteti professori, mashina tarjimasi bo'yicha yapon davlat dasturi rahbari Makoto Nagao bu

⁴³ **Hutchins J.** Retrospect and prospect in computer-based translation // Proceedings of MT Summit VII "MT in the great translation era". - Tokyo: AAMT, - P.32.

yo'nalishda yangi konsepsiya ishlab chiqdi.⁴⁴ Ushbu konsepsiyaga ko'ra matnlar ilgari inson qo'li bilan tarjima qilingan matnlarga tayanib analogiya tamoyili asosida tarjima qilinishi lozim. Buning uchun kompyuterga kiritilgan katta hajmdagi **biliniv bazasi** (matn va uning tarjimalari) shakllantirilishi zarur. Yangi matnlar tarjimasida biliniv massividan o'xshash matn fragmentlari tanlanadi va tarjimaning muvaffaqiyatli chiqishida foydalaniladi. M.Nagao mashina tarjimasiga bunday yondashuvni **“Example Based Translation”** (“Namunalarga asoslanadigan tarjima”) deb nomladi, ilgari an'anaviy yondashuv **“Rule Based Translation”** (“Qoidalariga asoslanadigan tarjima”) deb ataladi.

Professor M.Nagao o'zi taklif etgan yondashuvning ham mutloq emasligini ta'kidlaydi: *“Mashina tarjimasi tizimi bo'yicha yaratilgan har qanday ishlanmalar, yondashuvlar ertami, kechmi boshi berk ko'chaga (tupikka) kiradi. Bizning yondashuvimiz ham bundan xoli emas, faqat biz bu jarayonni kechiktirmoqchimiz”*.⁴⁵

M.Nagaoning yondashuvi boshqa yangi yondashuvlarning yuzaga kelishiga turtki bo'ldi. Jumladan, amerikalik olim Veb Laynning **“Translation Memory”** (“Tarjimon xotirasi”) yoki **“Sentence Memory”** (“Gap to'plagich”) konsepsiyasi.⁴⁶ Bu yondashuvga ko'ra muayyan matnni bir tildan ikkinchisiga tarjima qilish uchun dastlab uni yuqori malakali tarjimonga ko'rsatiladi. So'ng asl matn va uning tarjimasi kompyuterga kiritiladi, matn fragmentlarga (alohida jumalarga, birikmalarga) bo'linadi, ushbu elementlardan ma'lumotlar bazasi tayyorlanadi. Ma'lumotlar bazasi qidiruv tizimiga yuklanadi, u esa tarjimasi mavjud bo'lgan matn fragmentlarini va alohida jumalarni sifatli tarjima qilishga imkon beradi. Matnning ayrim notanish fragmentlari (bazada tarjimasi mavjud bo'lmagan matn qismlari) inson qo'li bilan tarjima qilinadi. Bunda bazadagi bu fragmentlarga yaqin keladigan tarjimalardan foydalanish mumkin. Matnning inson tomonidan tarjima qilingan qismlari yana ma'lumotlar bazasiga kiritiladi. Mana shu yo'l bilan “tarjimon xotirasi” asta-sekinlik bilan boyitib boriladi va tarjima samaradorligi yuksalib boradi. Lekin ushbu yo'nalishning afzalliklari bilan birga ayrim kamchiliklari ham mavjud. Birinchidan, bu yo'l bilan har qanday turdagi matnni tarjima qilib bo'lmaydi. U asosan muayyan tor yo'nalishdagi bir turdagi matnlarga mo'ljallangan. Ikkinchidan, ma'lum matnning katta fragmenti ikkinchi bir matnning katta qismiga tarjimada mos kelishi yoki yaqin kelishi har doim ham kuzatilavermaydi. Uchinchidan, “tarjimon xotirasi” bazasini yaratish yoki ikki tilli matnlar (biliniv) massivini to'ldirib borish amaliy jihatdan ko'p qiyinchiliklarni keltirib chiqaradi.

Rossiyada professor G.G.Belonogov frazeologik mashina tarjimasi konsepsiyasini asoslab berdi.⁴⁷ Bunga ko'ra matndagi tushunchalar ko'pincha

⁴⁴ Nagao M. A framework of a mechanical translation between Japanese and English by analogy principle, in Artificial and Human Intelligence, ed. A. Elithorn and R. Banerji, p.p. 173-180, North Holland, 1984.

⁴⁵ Qarang: Снокум Дж. Обзор разработок по машинному переводу: история вопроса, современное состояние и перспективы развития // Новое в зарубежной лингвистике, - М.: Прогресс, 1989.

⁴⁶ Webb Lynn E. Advantages and Disadvantages of Translation Memory: a Cost/Benefit Analysis. San Francisco State University. 1992.

⁴⁷ Qarang: Белоногов Г.Г., Зеленков Ю.Г., Кузнецов Б.А., Новоселов А.П., Хорошилов Ал-др А., Хорошилов Ал-сей А. Автоматизация составления и ведения словарей для систем фразеологического машинного перевода текстов с русского языка на английский и с английского на русский // Научно-техническая информация, Серия 2, № 12, ВИНТИ, 1993; Белоногов Г.Г., Кузнецов Б.А. Языковые средства автоматизированных информационных систем. - М.: Наука, 1983; Белоногов Г.Г., Зеленков Ю.Г., Новоселов А.П., Хорошилов Ал-др А., Хорошилов Ал-сей А. Системы фразеологического машинного

alohida so'zlar yordamida emas, balki so'z birikmalari yordamida ifodalanadi. O'z-o'zidan kelib chiqadiki, matnni bir tildan boshqasiga tarjima qilish jarayonida ham mazmunni yuzaga chiqaruvchi asosiy til birliklari ham alohida so'zlar emas, balki tipik vaziyatlarni ifodalovchi frazeologik birikmalar (bu yerda frazeologik birlik keng ma'noda tushuniladi) hisoblanadi. Mazkur yondashuvga muvofiq matnda tez-tez uchraydigan so'zlar, iboralar, so'z birikmalari, nutqiy etiket namunalarning tarjima ekvivalentlari bilan boyitilgan ma'lumotlar bazasi frazeologik mashina tarjimasi tizimining asosini tashkil etadi. Bunda so'zlarning distributsiyasiga eng ko'p e'tibor beriladi. Bu metod polisemantik so'zlarning tarjimada mos keluvchi to'g'ri ekvivalentini tanlashga, omonim so'zlarning kontekstagi holatini aniqlashga, iboralarning ham idiomatik xususiyatini belgilashga imkon beradi.

Mazkur konsepsiyaga asoslangan frazeologik mashina tarjimasining dastlabki versiyasi 1993 yil RETRANS nomi bilan yaratildi. 1998 yilgacha RETRANS MS DOS operatsion tizimi sharoitida ishlashga mo'ljallangan edi. 1998-200 yillarda uning Windows va UNIX operatsion tizimida ham ishlaydigan versiyalari yaratildi. 2001 yilda RETRANS tizimining turli modifikatsiyadagi yangi versiyasi yaratildi: a) MS Word matn muharririda ishlaydi; b) MS Internet Explore'ning Web-brauzerida ishlaydi; c) MS Notepad, MS Wordpadda *Clipboard* (matnni ko'chirib ishlash)ga asoslanib ishlaydi. RETRANS tizimi tabiiy va texnika fanlari, iqtisodiyot, siyosat, qonunchilik hamda harbiy sohaga oid 4 mln. Lug'at maqolasiga ega bo'lgan politematik so'zlik bilan ta'minlangan. Kiritilgan til birliklarining 80 % qismini so'z birikmalari tashkil etadi, ular 2 ta so'z uzunligidan tortib to 17 ta so'z uzunligiga ega bo'lgan birikmalardir. Tizim tarkibida 400 ming lug'at maqolasiga ega bo'lgan qo'shimcha tematik lug'at ham mavjud.

RETRANS tizimining ish tartibini ko'rib chiqaylik. Ruschadan inglizchaga tarjima qilish jarayonida dastlab matn avtomatik tarzda morfologik, so'ng semantik-sintaktik tahlil qilinadi, frazeologik birliklar aniqlanadi. Matnning barcha so'zlari hamda frazeologik birliklari qidiruv obrazlari shakllanadi. Lug'at bazasidagi joylashuv tartibga ko'ra so'z va birikmalar qidiruv obrazlarining xeshkodlari topiladi, ularning dekodlanishi asosida qidiruv natijalari namoyon bo'ladi. Keyingi bosqich matnning har bir fragmentiga mos yagona yoki bir nechta tarjima ekvivalentlari tanlanadi.

Shundan so'ng tarjima ikki rejimda amalga oshiriladi: 1) avtomatik tarjima rejimi; 2) oraliq tarjima natijalarini intellektual tahrirlash imkoniyati mavjud bo'lgan interaktiv rejim. Avtomatik tarjima rejimida ruscha matndagi har bir so'z va birikma uchun faqat bitta (odatda eng asosiy, sinonimik qatordagi dominantlik xususiyatiga ega bo'lgan so'z) inglizcha ekvivalenti tanlanadi va matnning morfologik, semantik-sintaktik analizi asosida sintez jarayoni boshlanadi, so'ng gap strukturasi shakllanadi va tarjima yuzaga chiqadi. Interaktiv rejim ham xuddi avtomatik tarjima singari bo'ladi, bunda faqat foydalanuvchi (tarjimon-muharrir) taklif etilgan tarjima variantlarining semantik jihatdan eng mos keluvchi ekvivalentini o'zi tanlash imkoniga ega bo'ladi, shuningdek, gap strukturasi, so'z tartibini ham lingvistik bilimidan kelib chiqib o'zi tahrir qilishi mumkin.

Shuni alohida ta'kidlash zarurki, yuqoridagi har uchala konsepsiya ham avtomatik tarjima sifati va samaradorligini oshirish, til bilan bog'liq to'siqlarni yo'qotish yo'lidagi muvaffaqiyatli izlanishlar hisoblanadi.

Ma'ruza № 7

Sintaktik tahlilning avtomatik tizimi

Reja

1. Kompyuter leksikografiyasi kompyuter lingvistikasi fanining alohida yo'nalishi sifatida.

2. Kompyuter lug'atlari va ularning ishlash tamoyili.

3. Kompyuter leksikografiyasi sohasining rivojlanishi.

Tayanch so'z va iboralar: *ma'lumotlar ombori, leksikografiya, kompyuter leksikografiyasi, kompyuter lug'atlari, kodlash va dekodlash, lemmatizatsiya, matnlar korpusi, korpus lingvistikasi, virtual matnlar korpusi, parallel matnlar korpusi, on-line lug'atlar, kitobiy lug'atlar, elektron lug'atlar.*

O'tgan asrning 50- yillarning oxiri va 70-yillarning boshlarida Kompyuter lingvistikasi ikki paradigma asosida rivojlandi: simvolik va ehtimollik nazariyasi asosida. Simvolik paradigma ikki yo'nalishda tadqiq etildi: formal til nazariyasiga asoslangan Chomskiy ta'limoti hamda sintaktik analiz (parsing) algoritmi va dinamik dasturlash yordamida pastdan yuqoriga (bottom-up) va yuqoridan pastga (top-down) metodlarga bag'ishlangan lingvistik va kompyuter sohasiga oid ishlarda o'z ifodasini topdi. To'liq sintaktik analiz tizimi Zelig Xarrisning Transformatsiya va Diskurs analiz proekti Pensilvaniya universitetida 1958-59 yillar ichida amalga oshirilgan. Ikkinchi yo'nalish esa sun'iy intellekt sifatida shakllandi. 1956 yilda Mak Karti, Marvin Minskiy, Klaud Shennon va Nazaniel Rochesterlar ikki oylik seminarda ishtirok etishib, sun'iy intellekt terminini qo'llashga kelishib oldilar. Sun'iy intellekt tadqiqotchilarni ehtimollik va statistik algoritmgaga asoslangan ishlarining juda kichik qismini tashkil etsa-da, yangi sohaning asosiy markazi Nivel va Simyonning Mantiqiy nazariyasi va umumiy muammolarni hal qilishga qaratildi.

Mashina tarjimasida matnning sintaktik tahliliga quyidagilar kiradi: 1) modullash (modularity); 2) tranformatsion tahlil (transformational parsing); 4) filtrlash (filtering); 5) tarmog'i orqali bajariladigan sintaktik analiz (augmented transition network parsing-ATN); 6) daraxtsimon qayta qurish (tree transducers); 7) diagramma (charts); 8) semantik tahlil (semantic analysis).

Ingliz va o'zbek tili xususiyatlari bir-biridan farq qiladi. Har ikki til o'ziga xos gap qurilmalaridan iborat. Zotan, ingliz tilida sintaktik tuzilmalar erkin- kontekstli grammatika (context-free grammar) ga asoslanadi (Jurafskiy). Unga ko'ra dastlab gapning bosh bo'laklari aniqlanishi zarur. EKG o'z ichiga quyidagilarni qamrab oladi: terminal belgilar, noterminal belgilar, grammatik qoidalar, boshlang'ich simvollar.

$S \wedge NP VP I + \text{want a morning flight } NP \wedge \text{Pronoun } I \mid \text{Proper-Noun Los Angeles} \mid \text{Det Nominal a} + \text{flight}$

$Nominal \wedge \text{Noun Nominal morning} + \text{flight}$

$\mid \text{Noun flights}$

$VP \wedge \text{Verb do}$

$\mid \text{Verb NP want} + \text{a flight}$

$\mid \text{Verb NP PP leave} + \text{Boston} + \text{in the morning}$

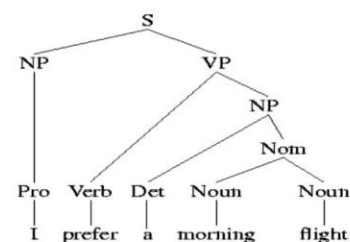
$\mid \text{Verb PP leaving} + \text{on Thursday}$

$PP \wedge \text{Preposition NP from} + \text{Los Angeles}$

Yuqorida ta'kidlanganidek, ingliz tilida sodda gaplar quyidagicha bo'ladi⁴⁸: 1) maqsadiga ko'ra; 2) strukturasiga ko'ra. Gaplar maqsadiga ko'ra o'z navbatida tubandagi to'rt guruhga bo'linadi: 1) darak gap; 2) so'roq gap (umumiy so'roq gap va maxsus so'roq gap) (*yes-no* li so'roq gap, and *wh-li* maxsus so'roq gap); 3) undov gap; 4) buyruq gap.

Darak gap->NP VP (I attend to Yenglish course); buyruq gap-> VP (Show me the meaning to live lonely); so'roq gap ("xa-yo'q"li so'roq gap)-> S ^ Aux NP VP (Do you go to the station?); Wh-li ega maxsus so'roq gap S ^ Wh-NP VP (Who is interviewing now?); Wh- ega bo'lmagan maxsus so'roq gap S ^ Wh-NP Aux NP VP (Where do you want to stay on your holiday?)

Endi bosh bo'laklar qanday so'zlar o'ramida kelishini tahlil qilsak. Bosh so'z (ot)dan oldin keluvchi birliklar determinantlar deyiladi: D-> Article|Pronoun (egalik, ko'rsatish, gumon, belgilash) | Numeral (sonoq son (cardinal), tartib son (ordinal)) | Quantifiers | Adjective. Determinantlar tushunchasi artiklni hisobga olmaganda o'zbek tilida qisman bo'lsa ham aniqlovchilarga to'g'ri keladi. Matn ingliz tilidan o'zbek tiliga tarjima qilinar ekan, awalo ingliz tilining sof lingvistik xususiyatlari inobatga olinishi kerak. Shuningdek, determinantlar (artikl)ni qabul qilmaydigan, sanalmaydigan ot, mavhum ot, atoqli ot, jamlovchi otlar ham mavjud. Bosh so'z (ot) dan oldin keluvchi sifatlar aniqlovchi aniqlanmish xarakteriga ega bo'lib, o'z navbatida sifat ham Adjective phrase (Birikmali sifat - BS) ajraladi. Sifatdan oldin ravish yonma-yon kelishi mumkin: *the least yexpensive fare*, shuningdek, quyidagi tartibda kelish holatlari ham mavjud: NP ^ (Det)-> (Card) (Ord)-> (Quant) ->(AP)-> Nominal - The -> first -> few -> hard-nosed ->ways. Ega (otli birikma)dan so'ng keluvchi gap bo'laklari ham muayyan sintaktik strukturani hosil qiladi. Ingliz tilidagi bunday qurilma o'zbek tilida kengaytirilgan otli birikma tarzida qabul qilinishi mumkin. Otli birikmadan keyin kelib, eganing kengaytiruvchilari sifatida quyidagi grammatik ko'rsatkichlarni ko'rsatish mumkin. Jumladan, predlogli birikmalar (prepositional phrases): *all students in auditory*; fe'lning vazifadosh (non-finite forms) shakllari: *some fruits buying from the shop*; bog'lovchilar (relative clauses): *my student who studies at university*. Egadan keyin keluvchi (post modifier)lar ham mavjud: *a reservation [on flight six oh six] [from Tampa] [to Montreal]* Nominal PP (PP) (PP). Shuningdek, fe'lning noaniq shakli gerundiy (-ing), infinitivli (to+Verb), (VIII--ed): *Which is the aircraft used by this flight?* Otli birikmalar quyidagicha bo'lishi mumkin: Nominal ^ Nominal RelClause; RelClause ^ (who | that) VP. Otli birikmalardan keyin keluvchi gap bo'laklari bog'lovchilar bilan o'zaro birikishi ham mumkin: NP ^ NP and NP; VP ^ VP and VP; S ^ S and S. Fe'llar mustaqil va yordamchi turlarga ajratiladi. Mustaqil fe'llar bir tarkibli fe'l, fe'lli frazemalar, qo'shma fe'llarga bo'linadi: *go - go off-make a seat*-> V- VPH-CF. Mashina tarjimasida matnlarni segmentlarga ajratish muhim sanaladi. Shu bois tahlil jarayonida tokenizatsiya jarayonida ikki bir-biriga o'xshamagan fe'llarni o'zaro muvofiqlashtirish avtomatik rejimda biroz murakkab bosqichlardan biri sanaladi. Ingliz tilida esa fe'l shakllarining leksikon quyidagi maydonlari mavjud bo'ladi:



To'g'ri fe'l	Noto'g'ri fe'l VI	Noto'g'ri fe'l VII	Noto'g'ri fe'l VIII	O'tgan tugallangan (Vn, Vni)	Davoml i	Hozir gi oddiy
play	break	broke	broken	played	breakin g	breaks
live	say	said	said	lived	living	lives

Fe'l qanday morfologik shaklda bo'lsin uning morfem tarkibi orqali qaysi so'z turkumiga tegishliligi aniqlanadi. Ayrim fe'l shakllari fonetik o'zgarishga uchraydi: *tries*=>*try+es*; *stopped*=>*stop+ed*.

Shu bois to'g'ri fe'llarning morfologik kategoriyalari alohida berilganlar (ma'lumotlar) bazasida ko'rsatilishi yoki fe'l maydonlarida muayyan belgilar bilan ajratilishi lozim. Noto'g'ri fe'l shakllari esa o'zi xos jihatlarga ega bo'lib, asosining tuslanishi alohida tiplarni tashkil etadi:

Fe'l shakllari				
ID	I.	II.	III.	IV.
Asos	g ^o	cut	drive	catch
-ing	goin	cutting	drivi	catching
es	goes	cuts	drive s	catches
Past simple	wen	cut	drove	cought
Past participle	gon	cut	drive	cought

Yordamchi fe'llar vazifasiga ko'ra: modal fe'llar, tugallanganlik,

hamda majhul nisbatni ko'rsatuvchi elementlarga bo'linadi: MV(must, *may*, *should*), Paux (*have*, *has*, *had*), Contaux (to be+Ving), PassAux (be+Viii). Zamonni ko'rsatuvchi yordamchi fe'llar ham mavjud: *do*, *did*, *have*, *had*, *will* kabilar. Bunday yordamchi fe'llar gapda turli shakllarda kelishi mumkin: *Should have gone*-> modal perfect; *Should be done*-> modal passive; *Has been raining*->perfect continuous; *Must have been implemented*->Modal perfect passive.

Mashina tarjimasida, avvalo, gap bo'laklarini aniqlash lozim. Shuning uchun gap bo'laklari modellarini aniqlaymiz (Shartli belgilar: {}- ixtiyoriy tanlov, [] - model belgisi, | -bo'lishi mumkin.):

Ega (Subject) o'rnida quyidagi so'z turkumlari kelishi mumkin [Noun| Adjective| Pronoun{Personal, demonstrative, defining, indefinite, negative, possessive, interrogative}| Adjective | Participle | Numeral| Infinitive | Gerund| It]

Kesim (sodda) (predicate) o'rnida quyidagi so'z turkumlari kelishi mumkin: [Verb| phraseological unit| Word combinations Noun| Adjective| Pronoun {Personal, demonstrative, defining, indefinite, negative, possessive, interrogative, reflexive}| Adjective | | Numeral| Infinitive | Participle II | Participle| Adverb]; Kesim (murakkab) o'rnida quyidagilar kelishi mumkin: [{can| may|must|should|would|ought|dare|need+Verb}| phrasal verbs|word compound verbs]

To'ldiruvchi (the object) o'rnida quyidagi so'z turkumlari kelishi mumkin:

[noun (in the common case) | pronoun (personal in the objective case, possessive, defining, reflexive, demonstrative, indefinite) |Adjective | Participle|Infinitive| infinitive phrase| Infinitive construction| Gerund| Gerundial

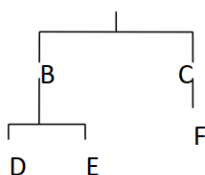
phrase| Gerundial construction]; to‘ldiruvchi quyidagilarga tobelanishi mumkin: [Object-> Verb | Noun | Adjective| Gerund].

Aniqlovchi (the attribute) o‘rnida quyidagi so‘z turkumlari kelishi mumkin: Adjective|Pronoun{ demonstrative, defining, indefinite, possessive, interrogative}| Numeral {cardinal | ordinal} Noun| Prepositional phrase Participles I and II or a participial phrase | Infinitive| Infinitive phrase | Infinitive construction]; aniqlovchi quyidagi tobelanishi mumkin: [Attribute -> Noun | Pronoun].

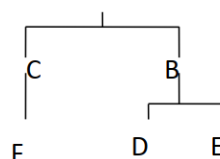
Hol (the adverbial modifier) o‘rnida quyidagi so‘z turkumlari kelishi mumkin: [Adverb | Prepositional phrase | than+ {Noun| Pronoun|Adjective, Infinitive| Participle|Prepositional phrase} | Participle |Participial phrase | Infinitive | Participial phrases | Prepositional phrases].

Inglizcha sodda gaplarini o‘zbek tiliga tarjima qilishda ega va kesim moslashuvi o‘zbek tilida aynan muvofiqlashmaydi. Bu gapni sintaktik tahlil qilganda yaqqol ko‘rinadi. Gap bo‘laklari aniqlangandan so‘ng matn morfologik va sintaktik jihatdan tahlil qilinadi. Buning uchun tahlil usullaridan daraxtsimon qayta qurish turini namuna sifatida keltiramiz. Bu tahlil usuli dastlab TAUM mashina tarjimasi loyihasining Q -tizimida vujudga kelgan. Mazkur jarayon sintaktik tahlil hisoblanib, u quyidan yuqoriga (bottom-up) tahlil metodiga o‘xshash. Uning strukturasi quyidagicha:

1-чизма



2-чизма



Agar 1-chizmaning formulasi $A(B(D,E),C(F))$ bo'lsa, ya'ni gap tashkil etuvchilarining soni qancha va qaysi o'rinda bo'lishidan qat'iy nazar qiymati bir xil bo'ladi: $(A(B(*), C(*)) \rightarrow A(C(*), B(*)))$ (2-chizma). Ushbu chizmani matn orqali tahlil etamiz: *It is hard to live for me in Africa. <-> To live in Africa is hard for me. (Men uchun Afrikada yashash qiyin.)* Ya'ni bir xil mazmun anglatuvchi gaplarning strukturasi o'zgarsa-da, umumiy ma'noga ta'sir etmaydi. Ushbu gap orqali gap bo'laklarining bevosita va bilvosita tobelanish munosabatini ko'rimiz mumkin: *Translation of information between different languages has become an important topic.* Sodda gapshakl modeli quyidagicha bo'ladi: **Noun(1)** + {Prep.} + Noun(2) + prep. + Adjective + Noun(3) + Verb + Adjective + Noun

Tobelanuvchi so'zlarning miqdori qancha bo'lishidan qat'iy nazar hokim so'zning o'rni va mavqei o'zgarishsiz qolaveradi. Tahlildagi asosiy bosqich gapning bosh bo'laklarini aniqlab olishdir. Dastlab gapda ishtirok etayotgan kesim (ot-kesim yoki fe'l-kesim)ni va unga tobelanib keluvchi bo'lak- eganing o'rnini topish kerak. Inglizcha matn o'zbek tiliga, odatda, matnning oxirgi qismidan boshlab tarjima qilinadi. Dasturiy ta'minot uchun so'z valentliklari bazasi zarur. Matn nutqiy jarayonda voqelanar ekan uni sof sintaktik modellarga mos, deb aytish qiyin. Zero, aynan grammatik tamoyillarga istisno holatlarni ko'plab kuzatamiz.

Ma'ruza № 8

Kompyuter leksikografiyasi

Reja

1. Kompyuter leksikografiyasi kompyuter lingvistikasi fanining alohida yo'nalishi sifatida.
2. Kompyuter lug'atlari va ularning ishlash tamoyili.
3. Kompyuter leksikografiyasi sohasining rivojlanishi.

Tayanch so'z va iboralar: *ma'lumotlar ombori, leksikografiya, kompyuter leksikografiyasi, kompyuter lug'atlari, kodlash va dekodlash, lemmatizatsiya, matnlar korpusi, korpus lingvistikasi, virtual matnlar korpusi, parallel matnlar korpusi, on-line lug'atlar, kitobiy lug'atlar, elektron lug'atlar.*

Ma'lumki, leksikografiya tilshunoslik fanining amaliy sohalaridan biri hisoblanadi. U yunoncha "lexikos – so'z", "lug'at", "grapho – yozaman" degan ma'nolarni anglatadi. Leksikografiya bo'limida lug'atlar, ularni tuzish yo'llari va tamoyillari o'rganiladi. Avtomatizatsiya ijtimoiy hayotning barcha jabhalariga kirib kelishi natijasida lug'atlar ham elektron holatda yaratilish imkoniyati yuzaga keldi. Kompyuter yordamida lug'atlar bilan ishlashning optimallasuvi natijasida kompyuter leksikografiyasi yo'nalishi shakllandi. Kompyuter leksikografiyasi amaliy tilshunoslikning muhim tarkibiy qismi bo'lib, unda lingvistik va dasturiy ta'minot asosida ishlaydigan, kodlash va dekodlash prinsipi asosida yaratilgan

kompyuter lug'atlari, ularni tuzish dasturlari, algoritmlari o'rganiladi. Mazkur sohaning rivojlanishi natijasida turli nomlarda elektron lug'atlar yaratilmoqda. Eng mashhur elektron lug'atlar sirasiga CONTEXT, ABBY LINGVO, MULTITRAN, POLYGLOSSUM, MULTILEKS kabilar kiradi.

Leksikografiya sohasiga kompyuter texnologiyalarining tatbiq etilishi lug'at tuzish ishlarini ancha yengillashtiradi. Lug'at tuzish jarayonining naqadar murakkabligini bir tilshunos olim hazilomuz tarzda quyidagicha izohlagan: *“Agar biror kishi og'ir jinoyat sodir qilsa, uni katorgaga surgun qilish shart emas. U lug'at tuzsin, shunda u katorganing barcha azoblarini, qiyinchiliklarini tortgan bo'ladi”*.⁴⁹ Axborot manbalarini yig'ish bo'yicha maxsus dasturlar (Database Software) lug'at asosini tashkil etuvchi barcha ma'lumot va misollarni jamlash va sistematik tarzda ishlov berishga qulaylik tug'diradi. Bundan tashqari boshqa maxsus dasturlar lug'atni tahrir qilish va chop etish borasidagi ancha mashaqqatli mehnatni bir qancha osonlashtiradi. Elektron lug'atlar ham o'z navbatida an'anaviy lug'atlarga qaraganda bir muncha ko'proq afzalliklarga ega. Bugungi kunda masalan, ABBY LINGVO kompaniyasining lug'at va tarjima dasturlari ulardan foydalanuvchi har bir xaridorga o'z lug'atini tuzishga yoki mavjud lug'at so'z boyligini to'ldirib boorish imkonini beradi. Ochiq turdagi internet lug'atlari (**on-line lug'atlar**) ham ko'p hollarda barcha foydalanuvchilarga ushbu lug'atlarni to'ldirish imkonini beradi. Shuningdek, elektron lug'atlarning hajm jihatidan ixchamligi, kompaktligi, boshqa manbalarga (internet tarmog'i orqali, gipermurojaatlar yordamida) ulanish imkoniyati, mavjud matnlar korpuslari yordamida illustrativ misollar tuza olish imkoniyati, multimediali misollar bilan boyitilganligi, so'zlarning sinonimik variantlari, omonimlik xususiyatlari, grammatik ma'lumotlar bilan ta'minlanganligi va ularni tezlik bilan foydalanuvchiga havola etilishi kabi jihatlari elektron lug'atlarning optimalligini ta'minlaydi.

Kompyuter leksikografiyasini elektron matnlar korpusi yoki parallel matnlar korpuslarisiz tazavvur qilish mumkin emas. **Matnlar korpusi (corpus – lotincha “tana” degan ma'noni anglatadi)** bu – elektron holda saqlanadigan ma'lum til birliklari bo'lib, ular tilshunoslar uchun turli xil muammolarni hal etish uchun tatbiq etishda va turli yo'nalishdagi tadqiqotlar uchun zaruriyatga qarab turli shakllarda tuziladi. Bular fonema, grafema, morfemalardan tortib undan kattaroq birliklar leksema, gap va matnlardan (badiiy yoki ilmiy asar, gazeta va jurnal matnlari) tashkil topishi mumkin. Ularning qay tarzda saqlanishiga qarab maxsus dasturlar yordamida har bir kerakli so'z yoki so'z birikmasi uchun darhol uning

⁴⁹ Белоногов Г.Г. Компьютерная лингвистика и перспективные информационные технологии. - М.: Русский мир, 2004. – С. 150.

qo'llanishi bo'yicha misollar topilishi, imlo bo'yicha variantlari, sinonimik qatorlari topilishi mumkin. Matnlar korpusiga oid ilmiy tadqiqotlar salmog'ining ko'payishi natijasida tilshunoslikda **korpus lingvistikasi** yo'nalishi shakllandi.⁵⁰

Kompyuterda yaratilgan birinchi matnlar korpusi **Braun korpusi** (БК, inglizcha *Brown Corpus*, BC) hisoblanadi, u 1961 yilda Braun universitetida yaratilgan, har biri 2000 so'zli 500 ta matn fragmentini o'z ichiga oladi. 1970-yillarda 1 million so'zni o'z ichiga olgan matnlar korpusi asosida rus tilining chastotali lug'ati yaratildi. 1980 yillarda Shvetsiyaning Uppsala universitetida ham rus tilida matnlar korpusi yaratildi. Keyinchalik kompyuter leksikografiyasining rivojlanishi natijasida katta hajmli matnlar korpusiga ehtiyoj tug'ildi. Ya'ni 1 millionta so'z elektron lug'atlar bazasi uchun yetarli emas. Shu asosda yirik hajmli matnlar korpusi yaratila boshlandi. Ko'pgina mamlakatlarda XX asrning 80-yillaridan boshlab bunday korpuslar tuzila boshlandi. Ular turli maqsad va vazifalarga xizmat qiladi. Buyuk Britaniyada **Ingliz tili Banki** (**Bank of English**) hamda **Britaniya Milliy Korpusi** (**British National Corpus**, **BNC**), Rossiyada **Rus tilining mashina fondi** (**Машинный фонд русского языка**) hamda **Rus tilining Milliy Korpusi** (**Национальный корпус русского языка**) loyihalari ishlab chiqildi.⁵¹ Masalan, Rus tilining Milliy Korpusi hajmi hozirgi kunda 149 million so'zdan iborat. Keyingi yillarda internet tizimining rivojlanishi **virtual matnlar korpusining** yuzaga kelishiga olib keldi. Ya'ni internetdagi qidiriv saytlari, elektron kutubxonalar, virtual ensiklopedialar korpus vazifasini bajarmoqda. Korpusning janri va tematik rang-barangligi internetdan foydalanuvchining qiziqishlariga bog'liq. Masalan, ilm-fan doirasida **Wikipedia** katta hajmdagi matnlar korpusi sifatida foydalanilmoqda.⁵²

Korpus lingvistikasida parallel matnlar korpusi ham muhim ahamiyat kasb etadi. **Parallel matnlar korpusi** esa o'z navbatida badiiy asar, qo'llanma, ommaviy axborot vositasi nashrlari, turli xil hujjatlarning ikki yoki undan ko'p tillardagi elektron holdagi ko'rinishlaridir. Masalan, Yevropa Ittifoqi o'zining barcha qonun va hujjatlarini ingliz, fransuz, nemis, ispan va italyan tillarida nashr qiladi hamda ular internet tizimiga barchaga ochiq arxiv sifatida qo'yiladi. Bunday korpuslarning afzalligi shundaki, ular yordamida nafaqat biron bir so'z yoki jumlaning, balki butun boshli matnlarning turli tillardagi variantlarini bilish imkoniyati mavjud. Xuddi mana shu imkoniyat tufayli maxsus **konkordanser dasturlar** ishlab chiqish orqali turli xil ixtisoslik lug'atlari tuzish imkoniyati

⁵⁰ Баранов А.Н. Введение в прикладную лингвистику. – М.: Эдиториал УРСС, 2001. –С. 61.

⁵¹ <http://www.corpus.leeds.ac.uk/list.html>

⁵² <http://www.wikipedia.org/wiki/corpus> linguistics

tug'iladi. Ushbu imkoniyatlar kompyuter leksikografiyasi uchun ulkan ahamiyat kasb etadi.

Kompyuter leksikografiyasiga semantik maydon, semantik tarmoq, semantik to'r hamda freym semantikasining faol tatbiq etilishi natijasida ulkan kompyuter leksikografiya resurslari yaratildi. Shunday yirik leksikografik resurslardan biri **FRAMENET** bo'lib, u internet tizimida on-line rejimida ishlaydi.⁵³ Mazkur tizim Ch.Fillmorning **“Tools for Lexicon Building”** loyihasi asosida Kaliforniya shtati, Berkli shahridagi Xalqaro informatika institutida ishlab chiqilgan. Framenet resursining ma'lumotlar bazasida 10000 ta leksik birlik mavjud, undan 6000dan ortig'i to'liq annotatsiyaga ega. Bundan tashqari ma'lumotlar bazasida 800ta semantik freym ko'rsatilgan, 135000 annotatsiyali gaplar keltirilgan.

Elektron lug'atlar tuzish jarayonida ma'lumotlar ombori, matnlar korpusini yaratish, qidiruv tizimi, kodlash, lingvistik va dasturiy ta'minot unsurlari qatorida **lemmatizatsiya** bosqichi ham mavjud. Lemmatizatsiya bu – so'zning dastlabki, boshlang'ich formasini (lug'atdagi shaklini – lemmasini) tashkillashtirish texnikasi bo'lib, bu jarayon o'sha so'zning boshqa so'z-shakllaridan kelib chiqqan holda amalga oshiriladi. Lemmatizatsiya morfologik tahlil metodi tarkibiga kiradi, u ikki bosqichni o'z ichiga oladi: 1) deklarativ bosqich – bunda muayyan so'zning bo'lishi mumkin bo'lgan barcha shakllari (so'z-shakllar) belgilanadi; 2) protsedura bosqichi – bunda so'z asos va qo'shimchalarga, ya'ni leksemalarga yoki morfemalarga bo'linadi. Lemmatizatsiya so'zlarning grammatik valentligi, qaysi affikslar bilan birika olish imkoniyatini ham belgilab beradi. Masalan, o'zbek tilida so'zlarning lg'atdagi shakli – lemmasi quyidagicha:

- Ot so'z turkumi uchun – bosh kelishik, birlik shakli.

- Fe'l so'z turkumi uchun – harakat nomi shakli

- Sifat so'z turkumi uchun – oddiy daraja shakli

Daftarlarni, daftarlarga, daftarlarning, daftarlardan, daftarlarda → daftar

Yugurdi, yugurgan, yuguryapti, yugurmoqchi → yugurmoq

Ko'kimtir, ko'kish, ko'kroq → ko'k

Kitobiy lug'atlar va elektron lug'atlarning farqi quyidagilarda ko'rinadi:

1. Kitobiy lug'atlarning tuzilishi: a) lug'atning so'zligi shakllantiriladi; b) misollar kartotekalari tuziladi; c) lug'at maqolalari yoziladi; d) lug'atning qolyozma varianti tayyorlanadi; e) qo'lyozma tahrir etiladi; f) muallif tuzatishlar qiladi; g) nashr uchun teriladi; h) lug'at sahifalanadi; i) korrektirovka qilinadi; j) lug'at nashr etiladi.

2. Elektron lug'atlarning tayyorlanishi: a) lug'atning so'zligi shakllantiriladi; b) misollar korpusi (elektron kartotekalari) tuziladi; c) lug'at

⁵³ <http://framenet.icsi.berkeley.edu/>

maqolalari yoziladi; d) lug'at maqolalarini ma'lumotlar bazasiga ko'chiriladi; e) bevosita ma'lumotlar bazasida lug'at matni tahrir etiladi, korrektirovka qilinadi; f) lingvistik ta'minot dasturiy ta'minot bilan uyg'unlashtiriladi; g) elektron lug'at.

Kitobiy lug'atlar tuzilishi sahifalar ketma-ketligiga tayanadigan chiziqlilik tamoyiliga bo'ysunadi. Elektron lug'atlar strukturasi gipertekst texnologiyasiga



asoslangan bo'ladi, bu esa foydalanuvchiga lug'at maqolalarining ixtiyoriy qismiga tezkor murojaat qilish imkonini beradi.

Elektron lug'atlarning ishlash prinsiplarini umumlashtirib quyidagicha izohlash mumkin:⁵⁴ Tilning har bir so'ziga mutanosib keluvchi kod ishlab chiqiladi va qo'llanadi, kodni qayta ishlash jarayonida zaruriy bo'lgan ma'lumotlar, tarjimalar, sinonim, sinonim, antonym va sharhlarga ega bo'lish mumkin.

So'zlarni kodlashtirish quyidagicha amalga oshiriladi:

Ma'lumotlar tekst fayllarga joylashtiriladi, uning har bir elementi 3 qismdan iborat bo'ladi:

1. So'zning tartib raqami.
2. So'z.
3. Kod.

"So'zning tartib raqami" (ya'ni uning adresi). So'zlarning tarjimasini, sinonimlari va antonimlarini ko'rsatish ularning birinchi harflarni kodirovkadagi tartib raqamlari bilan birgalikda keltirish bilan amalga oshiriladi.

"So'z"-mutanosib alfavit harflari bilan yozilgan oddiy so'z.

"Kod"-raqami va harflar ketma-ketligi bo'lib, unda so'z barcha zaruriy morfologik, sintaktik, leksik xususiyatlari hamda ushbu so'zning qaysi so'zga tegishliligi haqidagi ma'lumotlar jamlangan bo'ladi: Grammatik ma'lumot, adreslar, tarjima, sinonim, antonym, mutanosib sharhlar.

Kodlarni yaratish CREATE va CREATE 1 dasturlari orqali amalga oshiriladi. Ular quyidagi tartibda ishlaydi:

⁵⁴ Po'latov A., Muhamedova S. Kompyuter lingvistikasi. – Toshkent, 2007. – B. 58-61.

Soʻz kiritiladi.

1. Grammatik maʼlumotlar yaratiladi-kompyuter soʻz haqida morfologik va sintaktik maʼlumotlarni soʻraydi va ular darhol aniq nisbat asosida shifrlanadi.
2. Tayyor shifr “Grammatik maʼlumotlar” boʻlimiga yozib qoʻyiladi.
3. Tarjimalar sinonimlar, antonimlar adreslari yaratiladi:
 - a) tarjimasi boʻlishi mumkin boʻlgan soʻz soʻraladi;
 - b) kiritilgan soʻz kodning mutanosib boʻlimiga yozib qoʻyiladi;
 - v) kodlashning oxirida tarjimalar (sinonimlar, antonimlar) kodi topiladi va kiritilgan soʻzlar oʻrniga ularning birinchi harflari va tartib raqamlari yoziladi (masalan, “katta” soʻzi oʻrniga-K0083), agar tarjimalar (sinonim, antonimlar) kodi topilmasa, u holda shu soʻzga nisbatan kodlash operatsiyasi amalga oshiriladi va bu bilan bosqich tamomlanadi.
5. Mazkur soʻzga sharh kiritiladi:
 - a) matn kiritiladi;
 - b) maxsus dastur asosida soʻz va uning sharhi orasidagi moslik belgilanadi, soʻngra kiritilgan matn xotiraga yoziladi, ular orasidagi moslik esa soʻz kodida oʻz ifodasini topadi.

Dasturning ishlashi

Dastur ishlay boshlashi bilan ekranda u va uni ishlab chiquvchilari haqida maʼlumotlar paydo boʻladi. Ular bilan tanishilgandan soʻng klavish bosiladi va dastur quyidagi bosqichlarda oʻz ishini davom ettiradi:

I. Ekranga 6 rejimga ega oyna chiqdi.

1.2.3.4.5.6.

1-rejim. **Only translation (faqat tarjima)** – podstrochnikda foydalaniladigan tarjimalarni va mazkur soʻzning qaysi sohaga tegishliligi haqidagi maʼlumotlarni beradi. Mazkur rejimning asosiy funksiyasini Only-tran maxsus protsedurasi amalga oshiradi. Ushbu protsedura avvalo mazkur soʻzning kodining adreslar boʻlimidan tarjimalar adresini izlab topadi. Soʻng uni qayta ishlashni boshlaydi: tarjimaning birinchi harfini olib tekst faylini ochadi (shu harf bilan nomlangan) va berilgan tartib raqamiga koʻra tarjimini topib, uni ekranga chiqaradi.

2. **Grammatical information (grammatik maʼlumotlar)** – soʻzning barcha morfologik va sintaktik xususiyatlari va undan foydalanishdagi ayrim noaniqliklar haqidagi maʼlumotlarni beradi. Gram info maxsus protsedurasi ishlaydi. Grammatik maʼlumotlar kodi boʻlimida jamlangan maʼlumotlarni rasshifrovka qiladi.

3. **List of synonyms (sinonimlar roʻyxati)**

4. **List of antonyms (antonimlar roʻyxati).**

Ushbu rejimlar Syn-List va Ant-list maxsus protseduralar yordamida sinonimlar va antonimlar roʻyxatini beradi. Ular quyidagi tartibda ishlaydi:

sinonim va antonimlar kiritilgan soʻz kodning “sinonimlar adresi” va “antonimlar adresi” boʻlimlaridan topiladi hamda ekranga chiqariladi.

4. **New word formation (yangi soʻzlarni qayta yasash)** – mazkur soʻz bilan bir xil oʻzakka ega boʻlgan barcha soʻzlarni chiqarib beradi. Asosiy ish NWF protsedurasi vositasida amalga oshiriladi:

a) soʻzning oʻzagi ajratiladi;

b) soʻz haqidagi barcha maʼlumotlarga ega boʻlgan holda ushbu oʻzakka old qoʻshimcha va boshqa qoʻshimchalarni qoʻshish orqali oʻzgartirilishi mumkin boʻlgan barcha soʻzlarni chiqarib beradi.

6. **Comments (sharhlar)** – oʻrganilayotgan soʻzni tez esda qolishi uchun “aytib berish” (podskazka) koʻrinishidagi sharhlarni chiqarib beradi.

II. Ish uchun zarur boʻlgan rejim tanlanganidan soʻng soʻz kiritiladi. Ayrim elektron lugʻatlarda leksikon bazasi ozligi sababli, ayrim soʻzlarning xotirada kodi boʻlmasligi ham mumkin. Shuning uchun kirishdan soʻng **Find Word** protsedurasi ish boshlaydi. Uning vazifasi xotiradagi soʻz kodini izlashdan iborat. Agar u topilsa, maxsus protseduralar uzatiladi, aks holda esa quyidagi xabar chiqadi “Sorry, I do not know (Uzr, men buni bilmayman) va keyingi soʻz kiritiladi.

III. Topilgan kod tanlangan rejimning maxsus dasturlari bilan qayta ishlanadi, buning natijasida talab qilingan maʼlumotlar chiqariladi.

IV. Tanlangan rejimlarda ishni davom ettirish haqida soʻraladi. “Ha” javobidan soʻng dastur ishi ikkinchi bosqichda davom ettiriladi. Aks holda keyingi bosqichga oʻtiladi.

V. Dastur ishini tamom qilish haqida soʻraladi.

Kompyuter leksikografiyasi bugungi kunda turli qoʻshimcha imkoniyatlar asosida rivojlanib bormoqda. Xususan, dastlab kompyuter lugʻatlari bir tilli, ikki tilli boʻlgan, hozirda koʻp tilli (uch, toʻrt, olti, oʻn tilli) elektron lugʻatlar yaratilmoqda. Shuningdek, keyingi paytlarda muayyan sohalarga ixtisoslashgan elektron lugʻatlar ham yaratilmoqda. Yaʼni dastlab kompyuter lugʻatlari faqat umumiy leksikon bilangina cheklangan edi, hozirda bundan tashqari fanga oid boʻlgan, qurilish, aviatsiya, avtomobilsozlik, harbiy, diniy, yuridik sohalarga oid soʻzlar bazasi ham eʼtiborga olinmoqda. Hatto soʻzlarning matnda qoʻllanilish imkoniyatlari, distributiv holatlari, birikma holidagi koʻrinishlari ham nazarda tutilgan holda lugʻatlar yaratilmoqda. Bu koʻrsatkichlar kompyuter leksikografiyasi sohasining taraqqiy etayotganidan darak beradi.

Topshiriqlar

1. Tezaurus lug'at deganda qanday lug'at tushuniladi?
2. Kompyuter leksikografiyasining rivojlanish tendensiyasi haqida yozing.
3. O'quv kompyuter lug'atlari haqida gapiring.
4. Kitobiy va elektron lug'atlar orasidagi mushtarak nuqtalar qaysilar?
5. Kitobiy va elektron lug'atlar orasidagi farqli jihatlar qaysilar?

Adabiyotlar

1. Нелюбин Л. Л. Компьютерная лингвистика и машинный перевод. – М.: ВЦП., 1991.
2. Шемакин Ю.И. Начало компьютерной лингвистики. – М.: МГОУ. 1992.
3. Po'latov A., Muhamedova S. Kompyuter lingvistikasi. – Toshkent, 2007.
4. Р. Г. Пиотровский. Текст, машина, человек. – М.: Просвещение, 1989.
5. Баранов А.Н. Введение в прикладную лингвистику. – М.: Эдиториал УРСС, 2001.
6. Белоногов Г.Г. Компьютерная лингвистика и перспективные информационные технологии. - М.: Русский мир, 2004.

Ma'ruza № 9-10

Kompyuter lingvodidaktikasi

Reja

1. Kompyuter lingvodidaktikasi va ta'lim jarayonida kompyuter texnologiyalaridan foydalanishning ahamiyati.
2. Til o'rgatishda avtomatlashtirish.
3. Til o'rgatuvchi o'quv dasturlari va ulardan foydalanish.

Tayanch so'z va iboralar: *til o'qitishni avtomatlashtirish, o'quv dasturlari, "elektron" o'qituvchi, multimediya, virtual kutubxona, CALL, avtomatizatsiya, bixevioristik yondashuv, kognitiv-intellektual yondashuv, avtomatik korrektorlar, speller, audio va vizual effektlar, interaktiv multimedia texnologiyasi, virtual ensiklopedia.*

Inson tafakkuri shu qadar taraqqiy etib bormoqdaki, texnikalashtirish, kompyuterlashtirish jarayoni nafaqat ishlab chiqarishning turli sohalari, balki madaniyat va ta'lim sohalariga ham dadil kirib bormoqda. Kompyuter texnologiyalarining shiddat bilan rivojlanishi ta'lim jarayonini yangi bosqichga ko'tardi, bu o'z o'rnida ta'lim mazmunini, metod va shakllarini qayta ko'rib

chiqish, uni yangi bilim hamda ko'nikmalar bilan yanada boyitish zaruriyatini tug'dirdi. Hozirgi vaqtda oliy ta'lim muassasalarida yangi pedagogik texnologiyalarning ilmiy asoslarini yaratish, ularni tasniflash, metodik ahamiyatini belgilash, ta'lim jarayoniga tatbiq etish kabi dolzarb muammolar ustida ish olib borilmoqda. Ilg'or pedagogik texnologiyalarning kompyuter texnologiyalari bilan uyg'unlashtirib tashkil etilishi va shu asosda multimedia darslarining yaratilishi esa asosiy yo'nalishga aylanib qolmoqda. Multimedia ("ko'p muhitlik" degan ma'noni bildiradi) zamonaviy kompyuterlarning deyarli barchasida mavjud bo'lgan imkoniyatlar majmuini, kompleks tushunchani anglatadi. Multimedia axborotning turli ko'rinishlari – matn, jadval, grafika, ovoz, animatsiya (multiplikatsiya), videotasvir, musiqa yordamida axborotni yig'ish, saqlash va qayta uzatish vazifalarini bajaradi. Multimedia "inson-kompyuter" interaktiv (dialogik) muloqotning yangi, takomillashgan pog'onasi bo'lib, bunda foydalanuvchi juda keng va har tomonlama axborot oladi. Bir qator olimlar, metodistlar tadqiqotlari shuni ko'rsatadiki, ta'lim oluvchi birinchi marta eshitgan mavzusining faqat to'rttdan bir qismini, ko'rgan materialining uchdan bir qisminigina eslab qoladi; ham ko'rib, ham eshitsa, axborotning ellik foizini esda saqlaydi. Interaktiv multimedia texnologiyalaridan foydalanilganda esa, bu ko'rsatkich yetmish besh foizni tashkil etadi. Shu bois ta'lim jarayonida kompyuter texnologiyalaridan unumli foydalanish katta ahamiyatga ega.⁵⁵

XX asr tarixga ilm-fan inqilobi davri, kompyuter, avtomatizatsiya asri sifatida tarixga kirdi. Bu davrda insoniyat o'zining og'ir mehnatini o'z aqli bilan osonlashtirishga erishdi. Ya'ni, ishlab chiqarish jarayonida foydali ish koeffitsienti bir necha barobarga oshdi, mehnat unumdorligi o'sdi. Insoniyat yirik zavod, fabrikalarni, korxonalarni avtomatlashtirishga erishdi. Bunda ilgari manufaktura (qo'l mehnati)ga asoslangan ishlab chiqarish tizimidan avtomatlashgan, kompyuterlashtirilgan tizimga o'tildi. Hozirgi kunda avtomatizatsiya faqatgina ishlab chiqarishda emas, balki jamiyatning deyarli barcha sohalariga kirib kelmoqda. Xususan, ta'lim jarayonida, bilimlarni diagnostika qilishning, xorijiy tillarni o'rgatishning ham samarali avtomatik tizimi ishlab chiqilgan.

Avtomatizatsiya so'zi lotincha "o'zi harakatlanadi" degan ma'noni bildiradi. So'zning ma'nosidan tushunilib turibdiki, tizim o'zi harakatlanishiga asoslangan, turli operatsiyalar bajarishga yo'naltirilgan va shu tamoyilda ishlashiga moslab programmalashtirilgan mexanizmlar yig'indisidir. Hammaga ma'lumki, yirik korporatsiyalarda, aytaylik, avtomobil ishlab chiqaruvchi zavodlarda avtomatizatsiya tizimi keng yo'lga qo'yilgan. Ushbu tizim optimalligi uchun jamiyatning boshqa sohalariga ham asta-sekinlik bilan tatbiq etilmoqda. Hozirda til o'rgatishning avtomatik tizimlari ishlab chiqilgan. Til o'rgatishning

⁵⁵ **Po'latov A., Muhamedova S.** Kompyuter lingvistikasi. – Toshkent, 2007. – B. 64-65.

avtomatlashganligi shu bilan izohlanadiki, u maxsus o'quv dasturlari yordamida amalga oshiriladi.

Til o'rgatishda kompyuterdan foydalanish bo'yicha dastlabki tajribalar XX asrning 60-yillarida Stenford (Russian-Program) va Nyu York universitetlarida (Das Deutche Program) o'tkazildi. XX asrning 70-80-yillarida til o'qitishda kompyuterdan foydalanishga bo'lgan qiziqish kuchaydi va hatto alohida yo'nalishning yuzaga kelishiga olib keldi: **Computer Assisted Language Learning (CALL)**.⁵⁶ CALL bo'yicha 1-xalqaro konferensiya 1985 yil 12-13 noyabrda Budapesht shahrida o'tkazildi. 2-xalqaro konferensiya 1989 yil 15-17 noyabrda bo'lib o'tdi. Mazkur konferensiyalarda kompyuter yordamida xorijiy tillarni o'rgatishning yangi metodlari taklif etildi, o'qituvchi lingvistik avtomatlar xususida qimmatli mulohazalar o'rta tashlandi, o'qitishda audio va video vositalardan foydalanish samarali ekanligi ta'kidlandi.

O'qituvchi lingvistik avtomatlardan foydalanishda 2 xil yondashuv mavjud:

1) bixevioristik yondashuv – mavjud an'anaviy o'qitish tizimining barcha jabhalarini e'tiborga oluvchi o'qitish mexanizmlari, bunda grammatikaning ma'lum bo'limi tugagandan so'ng, uni topshiriqlar asosida baholash, shuningdek, leksik minimumlar bo'yicha bilimni nazorat qilishga yo'naltirilgan dasturiy ta'minotga asoslangan o'qitish tizimi. Bunda bixeviorizmning **“stimul-reaksiya-mustahkamlash”** formulasiga asoslaniladi. O'qitish usullari “stimul” (turtki) vazifasini bajaradi, o'rganuvchi bundan ta'sirlanadi va bilimni mustahkamlaydi.

2) kognitiv-intellektual yondashuv – CALL (CALL Software) uchun universal dasturiy ta'minotlar yaratish, o'qitish tizimini lug'at bazasi, grammatika bo'yicha qo'llanmalar, tizimni **avtomatik korrektorlar, speller** bilan ta'minlash, **audio** va **vizual effektlar** bilan boyitish kabi amaliy ishlar majmuidir.

Til o'rgatishda quyidagi kompyuter dasturlarini qo'llash mumkin:⁵⁷

1. *Taqdimot texnologiyasi*, ya'ni mashg'ulotni ko'rgazma sifatida o'tkazish, darsda kompyuterdan foydalanishning eng oson usulidir. Buning uchun o'qituvchiga birgina kompyuter va multimedia proektori kerak bo'ladi. MS Power Point dasturi yordamida o'tilajak dars uchun kompyuterda o'qituvchining o'zi slaydlar yaratadi va ko'rgazmalar paketini tayyorlaydi. Elektron nusxa – bu o'zbek tili fani bo'yicha mavjud bo'lgan muayyan darslikning kompyuterdagi elektron versiyasi (nusxasi) bo'lib, unda multimedia vositalaridan faqat ovoz va matn qatnashadi.

2. *Elektron multimedia darsligi* – aynan kompyuter vositasida o'quv kursini yoki uning bitta katta bo'limini o'qituvchi yordamida yoki mustaqil o'zlashtirishni ta'minlovchi dasturiy-metodik kompleks.

⁵⁶ <http://www.yazykcoznanie.ru>

⁵⁷ **Po'latov A., Muhamedova S.** Kompyuter lingvistikasi. – Toshkent, 2007. – B. 66-67.

3. *Elektron ensiklopediya* – kuchli qidiruv tizimiga ega bo'lgan ma'lum yo'nalishdagi illyustrasiyalı maqolalar, elektron xaritalar, fotosuratlar, giperhavola(gipermurojaat)lar, audio va videoizohlar, xronologik ro'yxat, foydalanilgan va tavsiya etiladigan adabiyotlar ro'yxati va hokazolardan iborat bo'lgan katta hajmdagi elektron dastur.

4. *Elektron lug'at* – odatdagi lug'atlar asosida tuzilgan, multimedia vositalari bilan boyitilgan dastur. Elektron lug'at turli yo'nalishda tuzilishi mumkin, undan talabaniing so'z boyligini oshirish, tarjima qilishda foydalaniladi.

5. *O'rgatuvchi audiodasturlar, elektron lingafon kurslari* umuman ikkinchi tilni o'rganuvchilar uchun nihoyatda samarali ta'lim texnologiyalaridir. To'g'ri talaffuzni o'rgatuvchi kompyuter dasturlarida animatsiya yordamida tovushning artikulyatsiya o'rni, usuli ko'rsatiladi, etalon nutq eshittiriladi, talabaniing o'z talaffuzi yozib kiritiladi, so'ngra u adabiy til me'yori bilan taqqoslanadi. Bundan tashqari, audiodasturlar qatoriga "Karaoke" tizimini kiritish mumkin. Undan hordiq chiqarishda yoki biron tadbirni o'tkazishda foydalanish juda qulaydir.

6. *Universal test dasturi* o'zbek tili darslariniing biror yirik bo'limi bo'yicha test topshiriqlari va reyting tizimidan iborat bo'lgan kompyuter dasturidir. Uning qulayligi shundaki, birgina dasturiy ta'minot asosida turli xil test sinovlarini o'tkazish, ya'ni faqat matnni o'zgartirib, dasturni har xil mavzularda qo'llash mumkin.

7. *Elektron virtual kutubxona* – multimedia xonasiniing keng imkoniyatlaridan yana biri, elektron nusxa yoki elektron multimedia darsliklari mujassamlangan, tarmoqqa ulangan kutubxona. Internet tarmog'idan o'zbek tilini o'rganuvchi rusiyzabon talabalar unumli foydalanishlari mumkin: mustaqil ta'lim jarayonida ma'lumotlar to'plash, mavzularni o'zlashtirish, berilgan mavzuda referat, insho yozish va boshqalar. Videomateriallar bilan ishlash til o'qitishda alohida ahamiyat kasb etadi. Axborot texnologiyalariniing bu vositasi turli ko'rinishda, usulda qo'llanilishi mumkin. Masofaviy ta'lim – bu masofadan Internet tarmog'i orqali o'qitish usuli bo'lib, uning asosida respublikamizda ma'lum sohalar bo'yicha test-tajribalar o'tkazilmoqda. Mashg'ulotlarda kompyuterdan foydalanishning turli-tuman usullari mavjud. Sanab o'tilgan va boshqa bir qator kompyuter texnologiyalarini til o'qitish jarayonida qo'llash yuqori natijalar berishi shubhasizdir.

Tillarni avtomatik o'qitish tizimida quyidagi turdagi topshiriqlardan foydalaniladi:

a) **Close the gap** – Bo'sh kataklarni to'ldiring topshirig'i. Bunda semantika + grammatika sohalarini o'rganish ko'zda tutiladi;

b) krossvordlar va leksik o'yinlar;

c) elektron pochtdan foydalanib bajariladigan mashq (turli xatlar yoki elektron hikoyalar tuzish asosida - **maze**) **word-hunting**;

d) matnlarning avtomatik tarjimai va tahriri bo'yicha mashqlar;

e) grammatikaning turli bo'limlari bo'yicha diagnostik testlar o'tkazish;

f) CD-ROMda standart o'qitish kurslaridan foydalanish.

Ingliz tilini o'rgatishga mo'ljallangan quyidagi oquv dastrulari mavjud:

TOEFL tizimi – ingliz tilini avtomatik o'qitishga mo'ljallangan eng rivojlangan xalqaro dastur. U dunyoning deyarli barcha mamlakatlarida amal qiladi, u inglizcha Test Of English as a Foreign Language so'zlarining qisqartmasi bo'lib, ingliz tilidan xorijiy til sifatida sinov standarti hisoblanadi. Bu tizim ingliz tilini o'rgatish bilan birgalikda o'zlashtirilgan bilimlarni diagnostika qila olish imkoniyatiga ham ega.

IELTS – TOEFL tizimiga o'xshaydigan avtomatik til o'rgatish va bilimlarni diagnostika qilish tizimi. U grammatika, lug'at, o'qib va eshitib tushunish, to'g'ri jumla qurish kabilar yuzasidan bilimlarni baholaydi.

ENGLISH FOR ALL (Hamma uchun ingliz tili) – O'zbekistonda yaratilgan avtomatik o'qitishga asoslangan o'quv dasturi. Unda o'quvchiga ingliz tilidagi tovushlar va ularning talaffuzi audio effektlar yordamida tushuntiriladi, qisqacha ingliz tili grammatikasi bayon etiladi, mashqlar va testlar yordamida bilimni tekshiradi. Ushbu o'quv dasturining afzal tomoni shundaki, unda o'qitish tizimi video effektlar bilan ham boyitilgan, ingliz tilida multfilmlar kiritilgan va inglizcha qo'shiqlar karaoke bilan ta'minlangan.

WAY TO PERFECTNESS (Mukammallik uchun yo'l) – ingliz tilini avtomatik tarzda o'qitishga mo'ljallangan o'quv dasturi bo'lib, bunda ham multimedia bilan ta'minlangan. Ushbu dasturning yuqoridagilardan farqli tomoni shundaki, unda dastur “elektron” (“virtual”) o'qituvchi bilan ta'minlangan.

REWARD – bir necha yil davomida o'qitishga mo'ljallangan til o'rgatish tizimi.

PROFESSOR HIGGINS - АНГЛИЙСКИЙ БЕЗ АКЦЕНТА - fonetik trenirovka, ingliz tilida aksentsiz talaffuz me'yorlarini o'rgatuvchi tizim.

REPETITOR ENGLISH - Moskvadagi ingliz tilini o'rgatuvchi xalqaro markaz tashabbusi bilan yaratilgan. Mazkur tizim ingliz tilini turli xil zamonaviy metodlar va elektron metodik ishlanmalar paketi yordamida avtomatik tarzda o'rgatishga mo'ljallangan.

EURO PLUS – uch yil davomida ingliz tilini bosqichma-bosqich (Elementary, Intermediate, Upper intermediate, Advanced English) tizimli tarzda o'rgatishga asoslangan avtomatik o'qitish tizimi.

Til o'qitishning avtomatik tizimi yildan yilga rivojlanib yangi effektlar bilan boyib bormoqda. Jahonda kompyuterni o'quv jarayonida qo'llash bo'yicha yetarli

tajriba to'plangan, bu borada olib borilgan muhim ilmiy-metodik tadqiqotlar, ko'pgina ta'limiy-nazorat dasturlari, multimedia darsliklari mavjud. Yaratilgan dastur va boshqa elektron ishlanmalarning sifati, albatta, mualliflarning shaxsiy qarashlariga, ilmiy-metodik saviyasiga va umumiy madaniyatiga bog'liq bo'ladi.

Bu esa nazorat qilinishi qiyin bo'lgan holdir. Masalan, ba'zi ta'limiy-elektron ishlanmalar ko'zlangan maqsadga to'la-to'kis erishishga xizmat qilsa, ba'zilar o'ta qiziqarli, estetik jihatdan yuqori sifatli bo'lishiga qaramay, muayyan mavzu bo'yicha kerakli bilim va ko'nikmalarni egallashga xizmat qilmay, informatika va informasion madaniyat kursini o'rganishgagina yordam beradi.

Bugungi kunda o'qitishni kompyuterlashtirish nazariyasini ishlab chiqish, haqiqatan ham, o'quv faoliyatining samaradorligini oshiradigan, o'rganilayotgan mavzuga nisbatan ijobiy munosabat va qiziqishni shakllantiradigan ta'limiy-elektron ishlanmalarni yaratishning umumiy va xususiy mezonlarini belgilash zarurati paydo bo'ldi. Hozirgi paytda Rossiya, Fransiya, Yaponiya, Kanada va boshqa mamlakatlarda ta'limiy-elektron ishlanmalarni yaratish konsepsiyasini ishlab chiqish bo'yicha ilmiy-metodik tadqiqotlar olib borilmoqda. Konsepsiyada ishlanmalarning mohiyati, ularni yaratish, tekshirish va qo'llash tartibi belgilanadi. Mutaxassislarning fikricha, ta'limiy-elektron ishlanmalar asosan texnik, pedagogik, metodik talablarga javob berishi zarur. Ular shunday talablarga mos kelsagina, ta'lim jarayonida qo'llanilishi mumkin.⁵⁸

Topshiriqlar

1. Qanday til o'rgatuvchi o'quv dasturlarini bilasiz?
2. Til o'rgatuvchi o'quv dasturlarining samaradorligi qanday bo'ladi?
3. Boshqa turdagi o'quv dasturlari haqida yozing.
4. Internet tizimi yordamida (on-line) til o'qitish kurslaridan foydalanganmisiz?
5. CALL tizimining rivojlanishi va istiqbollari qanday deb baholaysiz?

Adabiyotlar

1. Нелюбин Л. Л. Компьютерная лингвистика и машинный перевод. – М.: ВЦП., 1991.
2. Шемакин Ю.И. Начало компьютерной лингвистики. – М.: МГОУ. 1992.
3. Po'latov A., Muhamedova S. Kompyuter lingvistikasi. – Toshkent, 2007.

⁵⁸ Po'latov A., Muhamedova S. Kompyuter lingvistikasi. – Toshkent, 2007. – B. 68.

4. Р. Г. Пиотровский. Текст, машина, человек. – М.: Просвещение. 1989.
5. Белоногов Г.Г. Компьютерная лингвистика и перспективные информационные технологии. - М.: Русский мир, 2004.

Ma'ruza № 11-12

Mashina tarjimasi

Reja

2. Kompyuter lingvistikasida avtomatik tarjima yo'nalishi.
4. Avtomatik tarjima yo'nalishining tarixi. Tarjimon dasturlar.
5. Avtomatik tarjima yo'nalishining zamonaviy holati va istiqboli.

Tayanch so'z va iboralar: *Matn tahriri, avtomatik tarjima, til o'qitishni avtomatlashtirish, elektron lug'atlar, tezaurus lug'atlar, kompyuter leksikografiyasi, kompyuter lingvodidaktikasi, on-line tarjima tizimi, ekvivalent moslik, transformatsion moslik, tarjima protsedurasi, generatsiya, translyatsiya, bilingv bazasi, example-based translation (namunalarga asoslanadigan tarjima), rule-based translation (algoritmgaga asoslangan tarjima), translation memory (tarjimon xotirasi).*

Informatsiyaning geometrik progressiya asosida keskin ko'payishi ularni boshqa tillarga tezkor tarjima qilinishini ta'minlovchi tizimlarga bo'lgan ehtiyojni keltirib chiqardi. Bunga javoban mashina tarjimasi yoki avtomatik tarjima yo'nalishida amaliy tadqiqotlar yaratila boshlandi. Eng ahamiyatli tomoni shundaki, mashina tarjimasiga oid izlanishlar kompyuter lingvistika fanining yuzaga kelishida tayanch nuqta bo'lib hisoblanadi.

Avtomatik tarjima bo'yicha dastlabki g'oyalar ingliz olimi Charlz Bebbidj tomonidan 1836-1848 yillarda ilgari surilgan. Uning fikricha, mexanik-elektron mashinalar tillar o'rtasidagi struktur o'xshashliklar asosida kodlashtirilgan avtomatik tarjimini amalga oshirishi mumkin. Uning loyihasi bo'yicha 50 ming ta so'z xotira kuchiga ega bo'lgan elektron mashinalar 100000 ta so'zni avtomatik ravishda tarjima qilishi zarur bo'ladi.⁵⁹

1949 yilda Ch.Bebbidjning g'oyalaridan ta'sirlangan amerikalik olim Uorren Uiver mashina tarjimasi tizimini yaratishning nazariy asoslarini ishlab chiqdi hamda lingvistika, deshifrovka, programmalashtirish nazariyasi bo'yicha mutaxassislarni hamkorlikka chorladi.⁶⁰ Shundan so'ng AQSHda MT bo'yicha kollaboratsiyalar, hamkorlik loyihalari yaratildi. Mashina tarjimasi bo'yicha

⁵⁹ **Po'latov A., Muhamedova S.** Kompyuter lingvistikasi. – Toshkent, 2007. – B. 52.

⁶⁰ **Баранов А.Н.** Введение в прикладную лингвистику. – М.: Эдиториал УРСС, 2001. –С. 89.

birinchi xalqaro konferensiya 1952 yilda Massachusetts texnologiya institutida o'tkazildi.

1954-yilda AQShdagi Jorjtaun Universitetida dunyoda birinchi marta avtomatik tarjima tajribadan o'tkazildi. U **GAT** tizimi (inglizcha "*Georgetown Automatic Translation*") birikmasining qisqartmasi) deb ataladi. Bu tajribalar IBM-701 kompyuterida o'tkazildi. Bu tajribaning bazasi algoritmlardan iborat bo'lib, 50000 ta so'z (60 ta gap)ni rus tilidan ingliz tiliga tarjima qildi.⁶¹

Rossiyada mashina tarjimasi yo'nalishining pionerlari D.Yu.Panov, A.A.Lyapunov, I.S., Muxinlar sanaladi.⁶² Shundan so'ng O.S.Kulagina, I.A.Melchuk, Yu.A.Motorin, Yu.N.Marchuk, R.G.Piotrovskiy kabi olimlar ularning izdoshlari sifatida avtomatik tarjima tizimini rivojlantirdilar. Sobiq SSSRda mashina tarjimasi bo'yicha birinchi tajriba 1955 yilda o'tkazildi. Shundan so'ng O.S.Kulagina va I.A.Melchuk rahbarligi ostida Amaliy matematika institutida mashina tarjimasi bo'yicha uchta yirik tajriba tizimi yaratildi: fransuzchadan ruschaga (FR-1), inglizchadan ruschaga va fransuzchadan ruschaga (yangi variantda). 1959 yilda Moskva davlat lingvistik universitetida mashina tarjimasi laboratoriyasi ochildi. Moskvada 1974 yilda INFORM-ELEKTRO institutida fransuzcha-ruscha (ETAP-1), inglizcha-ruscha (ETAP-2) hamda nemischa-ruscha (NERPA) avtomatik tarjima tizimlarini yaratish bo'yicha amaliy tadqiqotlar jonlandi.

1960-yilda sobiq SSSR Fanlar Akademiyasining avtomatik tarjimaga bag'ishlangan maxsus yig'ilishi bo'lib o'tdi. Bu yig'ilishda avtomatik tarjima bo'yicha professor A.Belonogov rahbarligida maxsus laborotoriya tashkil qilindi.

1966-yilda AQShning fan masalalari bo'yicha kongress qo'mitasida avtomatik tarjimaga oid ma'ruza muhokamasi (**Automatic Language Processing Advisory Committee – ALPAC REPORT**) bo'lib o'tdi.⁶³ Unda avtomatik tarjima bo'yicha "*Language and Mashines: Computers in Translation and Linguistics*" ("*Til va mashinalar: tilshunoslikda va tarjimada kompyuterlar*") deb nomlangan salbiy ma'nodagi ma'ruza qilindi va yuqori darajada sifatli avtomatik tarjimini ta'minlovchi universal tizim yaratish natijasizligi va imkonsizligi ta'kidlandi. Shuningdek, ma'ruzada ko'rsatilishicha, "tabiiy tarjima" (inson tarjimasi) mahina tarjimasidan arzonroq va sifatliroq. Shundan keyin jahon miqyosida avtomatik tarjimaga oid tadqiqotlarni moliyalashtirish keskin kamayib ketdi. Ular faqatgina

⁶¹ Убин И.И. Современные средства автоматизации перевода: надежды, разочарования и реальность // "Перевод в современном мире". М., ВЦП, 2001 г., стр. 60-69.

⁶² Панов Д.Ю., Ляпунов А.А., Мухин И.С. Автоматизация перевода с одного языка на другой. Сессия по научным проблемам автоматизации производства. Изд-во АН СССР, 1956.

⁶³ Слокум Дж. Обзор разработок по машинному переводу: история вопроса, современное состояние и перспективы развития // Новое в зарубежной лингвистике. Вып. XXIV. Компьютерная лингвистика. М., 1989. - С. 358.

Yaponiyada davlat tomonidan moliyalashtirildi. Qolgan davlatlarda avtomatik tarjimaga oid loyihalar xususi yirik kompaniyalar hisoblanagan “Generel Motors”, “Siemens”, “LG”, “Phillips” kompaniyalari bugungi kunda dunyoda avtomatik tarjimaga oid izlanishlarni mablag’ bilan ta’minlab turadi.

AQSH harbiy havo kuchlari buyurtmasi bo’yicha D.Tom rahbarligida **SYSTRAN** avtomatik tarjima sistemasi ishlab chiqildi. Mazkur tizim yordamida bir yilda 2,5 mln. sahifa matn ingliz tilidan nemis, fransuz, ispan, yunon, italyan tillariga hamda fransuz tilidan ingliz, ispan, italyan tillariga tarjima qilinadi. Avtomatik tarjima bo’yicha dunyodagi eng yirik loyiha 1991 yilda AQShda NASA mutaxasislari tomonidan yaratilgan: **DIANA**. Uning hajmi 10 million ta so’zni tashkil etadi. U dunyoning 6 ta yirik tiliga asoslanadi: ingliz, rus, fransuz, nemis, arab, ispan.

Rus olimlari ham professor Lyapunov rahbarligida “Vista Technology” kompaniyasi “**Retrans Vista**” loyihasini ishlab chiqdi. Uning tarkibiga kiruvchi 3 mln. 400 mingta so’z xotiraga kiritildi. Shunda 1mln. 800 mingtasi ruscha-inglizcha, 1 mln 600 mingtasi inglizcha-ruscha so’zdan iborat.

Bugungi kunda avtomatik tarjimini quyidagi sistemalari mavjud. 40 mln.ta so’z xotira kuchiga ega **YEUROTRA** sitemasi – Yevropa olimlari tomonidan ishlab chiqilgan. U barcha Yevropa tillari bo’yicha avtomatik tarjimini amalga oshira oladi. **SOKRAT** sistemasi 40 mln. ta so’z xotira kuchiga ega. Dunyodagi 40 ta tilda avtomatik tarjimini amalga oshiradi. Ya’ni qaysi tilda Windows operatsion tizimi versiyasi mavjud bo’lsa, shu tilda avtomatik tarjimini amalga oshira oladi. Lingvo sistemasi (lingvo kompyuter lug’ati) – bu lug’atga ham dunyoning 40 tilidagi ma’lumotlari jamlangan. 1990 yillarda Amerikada **Global Link**, Rossiyada **PROMT** avtomatik tarjima tizimlari yaratildi.

Internet tizimining rivojlanishi natijasida avtomatik tarjima global tarmoqda ham tashkil etildi, ya’ni hozirgi kunda **online-tarjima tizimi** ham faol yo’lga qo’yilgan.

Avtomatik tarjimaning kamchiliklari:

1. So’zlar avtomatik tarjima jarayonida to’g’ridan-to’g’ri tarjima qilinadi. Shuning uchun tarjimada grammatik qoidalarga deyarli rioya qilinmagan holda amalga oshiriladi
2. Avtomatik tarjima qilingan matnni uslub jihatdan tahrirlab chiqish zarur bo’ladi.
3. Hozirgi mavjud avtomatik tarjima qiluvchi sistemalar so’zlar birikuvini ham deyarli noto’g’ri tarjima qiladi.
4. Avtomatik tarjima polisemiya, omonimiya, troplar, frazeologizmlar, idiomalarni e’tiborga olmaydi.

Masalan, *The bat looks like a mouse with wings* gapini rus va o'zbek tiliga tarjima qilishda har bir so'zning polisemantik va omonimlik kabi semantik qirralarini e'tiborga olish zarur. Aks holda tarjima muvaffaqiyatli chiqmaydi. Jumladagi *bat* so'zi "ko'rshapalak", "belkurak", "to'qmoq", "kichik ziyofat" kabi ma'nolarni, *look* so'zi esa "qaramoq", "nigoh", "ko'rinish" ma'nolarini anglatadi. Bunda mazkur so'zlarni rus yoki o'zbek tiliga tarjima qilinayotganda mos ekvivalentini tanlash uchun kontekst va uning distributsiyasidan (qurshovidan) kelib chiqish zarur. Ushbu holat, ayniqsa, idiomatik ifodalar qatnashgan matnlar tarjimasida murakkablik tug'diradi. Masalan, *You are an apple of my eyes* jumlasida "Siz ko'zlarimning olmasisiz" tarzida so'zma-so'z tarjima qilinadi, chunki "apple" so'zining denotative ma'nosi "olma"dir. Lekin kontekstdan kelib chiqib tarjimaga yondashilsa, to'g'ri tarjima yuzaga chiqadi. Bunda jumlaning "an apple of my eyes" qismi idioma ekanligini nazarda tutish lozim va u "Siz ko'zimning oq-u qorasisiz" tarzidagi gap hosil bo'ladi.

Tarjimadagi moslik 3 xil bo'ladi:

1) ekvivalent moslik – bunda bir tildagi so'z boshqa bir tildagi so'zga tarjimada to'la muvofiq keladi. Masalan, rus tilidagi "книга" so'ziga inglizcha "book" so'zi tarjimada to'la mos keladi.

2) variantli moslik – bunda sinonimiya va polisemiya asosida so'zning ma'nosiga mos keluvchi variantlar ham chiqariladi. Masalan, rus tilidagi "добиться" so'ziga ingliz tilida mos keluvchi tarjima variantlari quyidagilar:

добиться гл сов

1. land (чего-либо)
2. win
3. gain (чьего-либо расположения)

добиться гл сов (достигнуть, достичь, добиться) - attain

добиться гл сов (добиваться)

1. contest (премии, места в парламенте и т. П.)
2. court

добиться гл сов (выигрывать, добиваться, выиграть) gain

добиться гл сов (достигать, добиваться, достичь) fetch, fetch up

добиться гл сов (обеспечивать, добиваться, обеспечить) instate

добиться гл сов (достигать, добиваться, достичь) achieve

достичь цели – achieve one's purpose

достичь цели – achieve one's aim

добиться гл сов (добиваться, стремиться) seek

добиться гл сов (добиваться) seek after, seek for

иметь большой спрос – be much sought after

иметь успех – be much sought after

быть популярным – be much sought after

добиться гл сов (выпросить, ухитриться получить) wangle

ухитриться получить лишнюю неделю отпуска – wangle an extra week's holiday

добиться гл сов (достигать, достичь, добиваться) obtain

добиться гл сов (добиваться)

3. carry

отстоять свою позицию, добиться своего – carry one's point

4. come at (чего-либо)

как вы это узнали? – how did you come at the information?

Добиться гл сов (обрести, получить) find

добиться гл сов (достигать, достичь, добиваться) get

добиться славы – get glory

добиться гл сов (добиваться) secure

добиться гл сов (получать, получить, добиваться) have

мы получили известие – we had news

ничего не добьешься – there is nothing to be had

3) transformatsion moslik – bu asosan birikmali ifodalar tarjimasida ishlatiladi, bunda mazmunga ta'sir ko'rsatmagan holda shakliy o'zgarishlar asosida tarjima amalga oshiriladi. Masalan, ingliz tilida **“the decision of government”** ifodasi **“the government's decision”** tarzida transformatsiya qilinishi mumkin. Har ikkala holatda ham mazmun (**“hukumat qarori”**) o'zgarmaydi.

Shularni nazarda tutgan holda aytish mumkinki, avtomatik tarjima, ayniqsa, rasmiy til doirasida muvaffaqiyatli amalga oshiriladi. So'zlashuv va badiiy uslubdagi matnlarning avtomatik tarjimasi hali hanuz maksimal darajada hal etilgani yo'q. Tipologik jihatdan o'xshash bo'lmagan, bir sistemaga mansub bo'lmagan tillarning avtomatik tarjimasi masalasi bir qator muammolarni keltirib chiqarmoqda. Demak, lingvistik tipologiyadagi farqlarni izchil o'rganish, lingvistik universalialarni aniqlash, tillardagi frazeologizmlar, idiomalar, paremlar, sinonimlar, omonimlar naborini qat'iy belgilash asosida mexanik tarjimadagi muammolarni bartaraf etish mumkin.

Eng sifatli tarjima insonning aralashuvi bilan yuzaga chiqadi. Chunki tarjimon inson tabiiy tilining nozik qirralarini, semantik, grammatik va pragmatik jihatlarini nazarda tutgan holda tarjimaga yondashadi. Bunda yondashuv quyidagicha bo'lishi lozim: Mexanik tarjima yordamida tanlangan matnning tarjima qilingan dastlabki versiyasiga ega bo'lamiz, bundan so'ng inson tarjima jarayoniga aralashadi, ya'ni u mavjud tarjimani tahrir etadi – qo'shimchalar qo'shilish tartibi, gap bo'laklarining joylashuvi, so'zlarning semalariga asosan to'g'ri tanlanganligi,

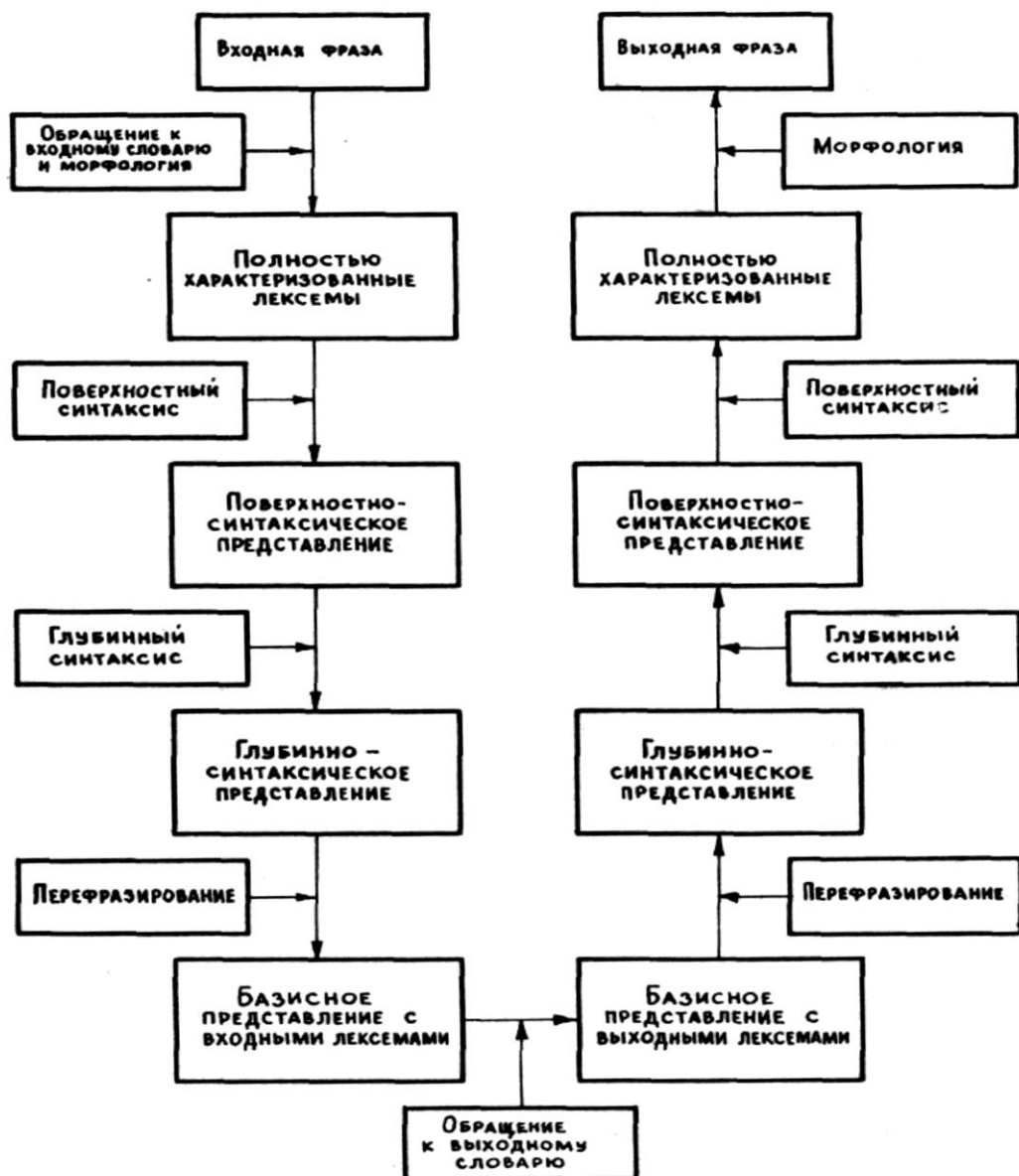
polisemiya, omonimiya, sinonimiyaning e'tiborga olinishi, jumlalarining stilistik jihatdan to'g'riligi kabi parametrlarni tekshiradi.

Rus olimi Yu.I.Shemakin o'zining *“Начало компьютерной лингвистики”* kitobida avtomatik tarjimaning 2 xil modelini va tarjima protsedurasini chizmalar asosida tushuntirib bergan. Birinchi yondashuv “matn-mazmun-matn” modeli va semantikaning formal tavsifiga tayanadi. Ikkinchi yondashuv esa “matn-matn” modeli va tarjima mosligiga asoslangan.

Birinchi yondashuvda tilning semantikasiga maksimal darajada e'tibor bergan holda avtomatik sifatli tarjimaga erishishga harakat qilingan. Tizimning lingvistik asosi *“matndan unng mazmuniga qarab borish”* (*“from surface structure to deep structure”*) hisoblanadi. Bunda tarjima tilning morfologik, sintaktik va semantik tahlili, luga'tlar, grammatik qoidalar, matnlar korpusiga asoslanib amalga oshiriladi. Buni Yu.I.Shemakin quyidagi chizma bilan tushuntirgan.⁶⁴

11-jadval

⁶⁴ **Шемакин Ю.И.** Начало компьютерной лингвистики. – М.: МГОУ. 1992, - С.68.



Ikkinchi yondashuvda matn ortiqchaliklarga ega ko'pbosqichli tizim sifatida baholanadi, bunga ko'ra matnning yuqori sathida ko'proq informatsiya tashuvchi leksik birliklar, quyi sathida esa ozroq informatsiya beruvchi birliklar joylashgan bo'ladi. Bunda so'zma-so'z tarjima amalga oshirilib, ikki tildagi gaplar strukturasi yagona superstrukturaga birlashtiriladi. Bunda avtomatik tarjima modeli ikki holatda bo'ladi: a) *generatsiya holati* tarjimaning ishlashini – lug'atlar jamlash, analiz algoritmlari, aniq bir maqsadga yo'naltirilgan sintezni ta'minlaydi; b) *translyatsiya holati* sistemaning mustaqil ravishda tarjima ustida ishlashini ta'minlaydi.

Ikkinchi yondashuvga ko'ra avtomatik tarjima quyidagi protseduralar asosida ishlaydi:

1-protsedura – qidiruv tizimi bilan bog'liq bo'lib, bunda lug'at bazasidan so'zlar, birikmalar, iboralar izlanadi, so'z-shakllarning morfologik analizi amalga oshiriladi.

2-protsedura – so'zlarning sintaktik vazifasi, grammatik kategoriyalar – zamon, shaxs-son kabilar aniqlanadi.

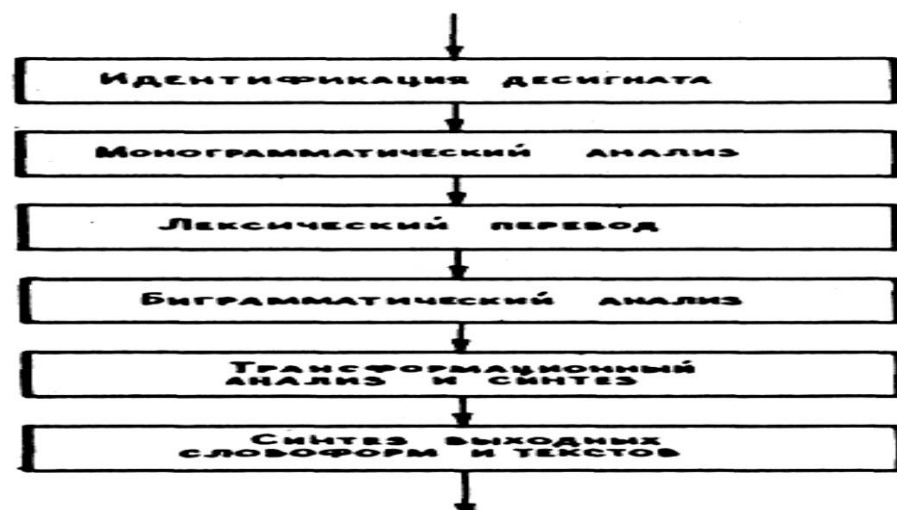
3-protsedura – bir yoki ko'p ma'noli so'zlar tarjimasi bayon qilinadi.

4-protsedura – jumlani morfologik va sintaktik jihatdan shakllantirish jarayoni amalga oshiriladi.

5-protsedura – transformatsion o'zgarishlar (matn yoki jumla strukturasi uning mazmuniga jiddiy ta'sir etmagan holda o'zgartirish; masalan, Kompyuter yoqildi va musiqa yangradi jumlasini “Kompyuter yoqilgach/ kompyuter yoqilishi bilan/kompyuter yoqilib musiqa yangradi” tarzida o'zgartirish) hosil qilinadi (agar zarurat tug'ilsa).

6-protsedura yakuniy bosqich bo'lib, bunda murakkab grammatik formalarning sintezi amalga oshiriladi.

12-jadval



Mashina tarjimasi strategiyasi sekin-astalik bilan o'zgarib organ. Ilk tajribalarda qo'llanilgan tarjima strategiyasi “*tog'ridan-to'g'ri tarjima strategiyasi*” deb ataladi, bunda faqat tarjima jarayonida so'zma-so'z tarjima

mufoviqqligi e'tiborga olinadi. Ushbu strategiya tarjima sifatini va samaradorligiga salbiy ta'sir ko'rsatadi. Shuning uchun ushbu strategiyani tanqid qilgan holda "*vositachi til yordamida amalga oshiriladigan tarjima strategiyasi*" ishlab chiqildi. Ayniqsa, bu strategiya grammatik strukturasi va tipologik xususiyatlari bir-biridan farqlanuvchi tillar o'rtasidagi avtomatik tarjima tizimlarida faol tatbiq etiladi.

Avtomatik tarjimada inson hamda kompyuterning ishtiroki quyidagi tarzda bo'ladi: a) *postediting* – matn (mashina-tarjimon) kompyuter yordamida tarjima qilinadi, so'ng inson-muharrir (inson-tarjimon) uni tahrir qiladi; b) *preediting* – bunda inson matnni mashina-tarjimonga moslaydi, so'ng uni kompyuterga havola etadi; c) *interediting* – bunda tarjimada murakkablik tug'diruvchi holatlar duch kelganda inson mashina-tarjimon ishiga aralashadi.⁶⁵

Matnni bir tabiiy tildan boshqasiga tarjima qilish jarayonidagi murakkabliklar va muammolarga nazariy va amaliy jihatdan yechim topish bo'yicha hali hanuz izlanishlar davom etmoqda. 1984 yilda Yaponiyaning Kioto universiteti professori, mashina tarjimasi bo'yicha yapon davlat dasturi rahbari Makoto Nagao bu yo'nalishda yangi konsepsiya ishlab chiqdi.⁶⁶ Ushbu konsepsiyaga ko'ra matnlar ilgari inson qo'li bilan tarjima qilingan matnlarga tayanib analogiya tamoyili asosida tarjima qilinishi lozim. Buning uchun kompyuterga kiritilgan katta hajmdagi **biliniv bazasi** (matn va uning tarjimalari) shakllantirilishi zarur. Yangi matnlar tarjimasida biliniv massividan o'xshash matn fragmentlari tanlanadi va tarjimaning muvaffaqiyatli chiqishida foydalaniladi. M.Nagao mashina tarjimasiga bunday yondashuvni "**Example Based Translation**" ("Namunalarga asoslanadigan tarjima") deb nomladi, ilgari an'anaviy yondashuv "**Rule Based Translation**" ("Qoidalariga asoslanadigan tarjima") deb ataladi.

Professor M.Nagao o'zi taklif etgan yondashuvning ham mutloq emasligini ta'kidlaydi: "*Mashina tarjimasi tizimi bo'yicha yaratilgan har qanday ishlanmalar, yondashuvlar ertami, kechmi boshi berk ko'chaga (tupikka) kiradi. Bizning yondashuvimiz ham bundan xoli emas, faqat biz bu jarayonni kechiktirmoqchimiz*".⁶⁷

M.Nagaoning yondashuvi boshqa yangi yondashuvlarning yuzaga kelishiga turtki bo'ldi. Jumladan, amerikalik olim Veb Laynning "**Translation Memory**" ("Tarjimon xotirasi") yoki "**Sentence Memory**" ("Gap to'plagich") konsepsiyasi.⁶⁸ Bu yondashuvga ko'ra muayyan matnni bir tildan ikkinchisiga tarjima qilish uchun dastlab uni yuqori malakali tarjimonga ko'rsatiladi. So'ng asl matn va uning tarjimasi kompyuterga kiritiladi, matn fragmentlarga (alohida

⁶⁵ **Hutchins J.** Retrospect and prospect in computer-based translation // Proceedings of MT Summit VII "MT in the great translation era". - Tokyo: AAMT, - P.32.

⁶⁶ **Nagao M.** A framework of a mechanical translation between Japanese and English by analogy principle, in Artificial and Human Intelligence, ed. A. Elithorn and R. Banerji, p.p. 173-180, North Holland, 1984.

⁶⁷ Qarang: **Спокум Дж.** Обзор разработок по машинному переводу: история вопроса, современное состояние и перспективы развития // Новое в зарубежной лингвистике, - М.: Прогресс, 1989.

⁶⁸ **Webb Lynn E.** Advantages and Disadvantages of Translation Memory: a Cost/Benefit Analysis. San Francisco State University. 1992.

jumlalarga, birikmalarga) bo'linadi, ushbu elementlardan ma'lumotlar bazasi tayyorlanadi. Ma'lumotlar bazasi qidiruv tizimiga yuklanadi, u esa tarjimas mavjud bo'lgan matn fragmentlarini va alohida jummalarni sifatli tarjima qilishga imkon beradi. Matnning ayrim notanish fragmentlari (bazada tarjimas mavjud bo'lmagan matn qismlari) inson qo'li bilan tarjima qilinadi. Bunda bazadagi bu fragmentlarga yaqin keladigan tarjimalardan foydalanish mumkin. Matnning inson tomonidan tarjima qilingan qismlari yana ma'lumotlar bazasiga kiritiladi. Mana shu yo'l bilan "tarjimon xotirasi" asta-sekinlik bilan boyitib boriladi va tarjima samaradorligi yuksalib boradi. Lekin ushbu yo'nalishning afzalliklari bilan birga ayrim kamchiliklari ham mavjud. Birinchidan, bu yo'l bilan har qanday turdagi matnni tarjima qilib bo'lmaydi. U asosan muayyan tor yo'nalishdagi bir turdagi matnlarga mo'ljallangan. Ikkinchidan, ma'lum matnning katta fragmenti ikkinchi bir matnning katta qismiga tarjimada mos kelishi yoki yaqin kelishi har doim ham kuzatilavermaydi. Uchinchidan, "tarjimon xotirasi" bazasini yaratish yoki ikki tilli matnlar (bilingv) massivini to'ldirib borish amaliy jihatdan ko'p qiyinchiliklarni keltirib chiqaradi.

Rossiyada professor G.G.Belonogov frazeologik mashina tarjimas konsepsiyasini asoslab berdi.⁶⁹ Bunga ko'ra matndagi tushunchalar ko'pincha alohida so'zlar yordamida emas, balki so'z birikmalari yordamida ifodalanadi. O'z-o'zidan kelib chiqadiki, matnni bir tildan boshqasiga tarjima qilish jarayonida ham mazmunni yuzaga chiqaruvchi asosiy til birliklari ham alohida so'zlar emas, balki tipik vaziyatlarni ifodalovchi frazeologik birikmalar (bu yerda frazeologik birlik keng ma'noda tushuniladi) hisoblanadi. Mazkur yondashuvga muvofiq matnda tez-tez uchraydigan so'zlar, iboralar, so'z birikmalari, nutqiy etiket namunalarining tarjima ekvivalentlari bilan boyitilgan ma'lumotlar bazasi frazeologik mashina tarjimas tizimining asosini tashkil etadi. Bunda so'zlarning distributsiyasiga eng ko'p e'tibor beriladi. Bu metod polisemantik so'zlarning tarjimada mos keluvchi to'g'ri ekvivalentini tanlashga, omonim so'zlarning kontekstagi holatini aniqlashga, iboralarning ham idiomatik xususiyatini belgilashga imkon beradi.

Mazkur konsepsiyaga asoslangan frazeologik mashina tarjimasining dastlabki versiyasi 1993 yil RETRANS nomi bilan yaratildi. 1998 yilgacha RETRANS MS DOS operatsion tizimi sharoitida ishlashga mo'ljallangan edi. 1998-200 yillarda uning Windows va UNIX operatsion tizimida ham ishlaydigan versiyalari yaratildi.

⁶⁹ Qarang: Белоногов Г.Г., Зеленков Ю.Г., Кузнецов Б.А., Новоселов А.П., Хорошилов Ал-др А., Хорошилов Ал-сей А. Автоматизация составления и ведения словарей для систем фразеологического машинного перевода текстов с русского языка на английский и с английского на русский // Научно-техническая информация, Серия 2, № 12, ВИНТИ, 1993; Белоногов Г.Г., Кузнецов Б.А. Языковые средства автоматизированных информационных систем. - М.: Наука, 1983; Белоногов Г.Г., Зеленков Ю.Г., Новоселов А.П., Хорошилов Ал-др А., Хорошилов Ал-сей А. Системы фразеологического машинного перевода. Состояние и перспективы развития // Научно-техническая информация, сер. 2, № 12, ВИНТИ, 1998.

2001 yilda RETRANS tizimining turli modifikatsiyadagi yangi versiyasi yaratildi: a) MS Word matn muharririda ishlaydi; b) MS Internet Explorerning Web-brauzerida ishlaydi; c) MS Notepad, MS Wordpadda *Clipboard* (matnni ko'chirib ishlash)ga asoslanib ishlaydi. RETRANS tizimi tabiiy va texnika fanlari, iqtisodiyot, siyosat, qonunchilik hamda harbiy sohaga oid 4 mln. Lug'at maqolasiga ega bo'lgan politematik so'zlik bilan ta'minlangan. Kiritilgan til birliklarining 80 % qismini so'z birikmalari tashkil etadi, ular 2 ta so'z uzunligidan tortib to 17 ta so'z uzunligiga ega bo'lgan birikmalardir. Tizim tarkibida 400 ming lug'at maqolasiga ega bo'lgan qo'shimcha tematik lug'at ham mavjud.

RETRANS tizimining ish tartibini ko'rib chiqaylik. Ruschadan inglizchaga tarjima qilish jarayonida dastlab matn avtomatik tarzda morfologik, so'ng semantik-sintaktik tahlil qilinadi, frazeologik birliklar aniqlanadi. Matnning barcha so'zlari hamda frazeologik birliklari qidiruv obrazlari shakllanadi. Lug'at bazasidagi joylashuv tartibga ko'ra so'z va birikmalar qidiruv obrazlarining xeshkodlari topiladi, ularning dekodlanishi asosida qidiruv natijalari namoyon bo'ladi. Keyingi bosqich matnning har bir fragmentiga mos yagona yoki bir nechta tarjima ekvivalentlari tanlanadi. Tarjima natijalarini quyidagi jadvalda yaqqol ko'rish mumkin.

Matnning rus tilidan ingliz tiliga qilingan tarjima oraliq natijalari fragmenti

Перевод	00001 00002 text translation
текстов	00002
	с 00003 00004 00005 00006 00007 00008 from
	some natural languages into others
одних	00004
естественных	00005
языков	00006
на	00007
другие	00008
-	00009 - / is
сложный	00010 complicated / complex / compound /
	difficult
творческий	00011 00012 creative process
процесс	00012
	, 00013
требующий	00014 requiring / demanding / claiming /
	making the demand
от	00015 from
переводчика	00016 interpreter / translator
не	00017 00018 not only / not just

только 00018
 соответствующей 00019 corresponding / appropriate /
 respective / adequate
 лингвистической 00020 linguistic
 подготовки 00021 training / preparation /
 getting ready / preparing
 , 00022
 но 00023 00024 but as well / but also
 и 00024
 хороших 00025 good / nice / well
 знаний 00026 knowledge / knowing
 тех 00027 those
 предметных 00028 00029 subject field / object areas
 / application domain
 областей 00029
 , 00030
 к 00031 00032 to which
 которым 00032
 переводимые 00033 00034 translated texts
 тексты 00034
 относятся 00035 relate / belong / concern
 . 00036

Shundan so'ng tarjima ikki rejimda amalga oshiriladi: 1) avtomatik tarjima rejimi; 2) oraliq tarjima natijalarini intellektual tahrirlash imkoniyati mavjud bo'lgan interaktiv rejim. Avtomatik tarjima rejimida ruscha matndagi har bir so'z va birikma uchun faqat bitta (odatda eng asosiy, sinonimik qatordagi dominantlik xususiyatiga ega bo'lgan so'z) inglizcha ekvivalenti tanlanadi va matnning morfologik, semantik-sintaktik analizi asosida sintez jarayoni boshlanadi, so'ng gap strukturasi shakllanadi va tarjima yuzaga chiqadi. Interaktiv rejim ham xuddi avtomatik tarjima singari bo'ladi, bunda faqat foydalanuvchi (tarjimon-muharrir) taklif etilgan tarjima variantlarining semantik jihatdan eng mos keluvchi ekvivalentini o'zi tanlash imkoniga ega bo'ladi, shuningdek, gap strukturasi, so'z tartibini ham lingvistik bilimidan kelib chiqib o'zi tahrir qilishi mumkin.

Shuni alohida ta'kidlash zarurki, yuqoridagi har uchala konsepsiya ham avtomatik tarjima sifati va samaradorligini oshirish, til bilan bog'liq to'siqlarni yo'qotish yo'lidagi muvaffaqiyatli izlanishlar hisoblanadi.

Topshiriqlar

1. Avtomatik tarjima bo'yicha qanday yondashuvlar mavjud?
2. Qanday tarjimon dasturlarni bilasiz?
3. Avtomatik tarjimaning ishonchlilik darajasi haqida mustaqil mulohaza yuriting.
4. Avtomatik tarjimaning to'laqonli bo'lishiga to'siq bo'luvchi omillarni sanang.
5. Avtomatik tarjima qanday protseduralar asosida ishlaydi?

Ma'ruza № 13-14

Mashina tarjimasining lingvistik muammolari

REJA:

1. Mashina tarjimasida turg'un birikmalarning o'rni.
2. Ibora va ularning talqinlari
3. Idiomalarning tarjima sohasidagi o'rni

Tilshunoslikda turg'un birikmalar sifatida frazema va idiomalar nazarda tutiladi.

Ingliz tilida frazema terminiga har xil yondashuvlar mavjud. Ayrim adabiyotlarda fe'l yoki ravish bilan keluvchi predloglarning birgalikda muayyan ma'noga tenglashishi⁷⁰, leksikologiyaga bag'ishlangan asarlarda predlog bilan keluvchi ot, sifat, fe'l hamda ravish so'z turkumlarining o'zaro yaxlitlanib yagona ma'noda kelishi tushuniladi⁷¹. "Zamonaviy ingliz tili leksikologiyasi kursi"⁷²da frazeologik birliklar ikki guruhga ajratilgan: frazema va idiomalar. So'z birikmalarida qismlarining biri ikkinchisiga tobe munosabatda bo'ladi: *small hours, three hours, pleasant hours* kabi. Idiomalar tarkib toptiruvchi komponentlarning har biri yaxlitlanib yagona ma'no semasini beradi:

⁷⁰ Лебедева Л.Д. Введение в курс лексикографии. - Москва: Высшая школа, 2008. - С.170-176.

⁷¹ Ginzburg R. S., Khidekel S. S., Knyazeva G. Y., Sankin A. A.. A Course in Modern English Lexicology. - M., 1979. - P. 95-96.

⁷² Yuqoridagi asar. - B. 113.

Mare's nest (alohida tarjima qilinganda *a mare - baytal*, *nest-in- uya* ma'nosida), lekin bir butun idioma sifatida *mare's nest - aldov* yoki *to find a mare's-ne to find a mare's-nest* – "*noo 'rin (bemavrid) gapirib qo 'ymoq*"ni anglatadi. Tilshunoslarning o'ziga xos yondashuvlari natijasida turg'un birikmalar (TB) turlicha nomlar bilan atalgan: *idioma*, *frazologizm*, *frazema*, *turg'un birikma*, *frazologik birliklar* va h.k. Ushbu terminlarning xususiyatlarini hisobga olib, "*turg'un birikma*" termini ob'ektni har tamonlama o'rganishda qo'l keladi, degan fikrni bildiramiz. Ingliz tilida TB lar murakkab tabiatlidir, chunki ikki yoki undan ortiq birliklarning o'zaro semantik, grammatik jihatdan birikuvi natijasida yagona birlik hosil bo'ladi. Ilmiy matnlarda TB larning qo'llanilishi muallifning shaxsiy uslubiga xosdir. Masalan, iqtisodchi M.Dobb va J.M.Keynslarning 100 betlik ma'ruzasida 21ta idioma ishlatilganligi qayd etiladi⁵⁸. Ilmiy matnlar boshqa uslublardan aniqligi, qolipga solinganiligi, termin va undagi so'zlarning bir ma'noga egaligi bilan ajralib turadi. Ilmiy matnlarning o'ziga xos xususiyatlari to'g'risida bir qator ilmiy ishlar bajarilgan. Biroq ilmiy va texnikaviy matnlarni tarjima qilish borasidagi ba'zi mulohazalar ko'pincha bahsli munozaralarga sabab bo'ladi. Masalan, Eva Mastnánning "The nature of scientific / technical texts from viewpoint of translation studies"⁵⁹ tadqiqotida tarjimashunoslikda ilmiy matnlarning tabiati tahlil qilingan, shuningdek, ilmiy matnlarni tarjima qilishda mashina tarjimasidan foydalanish samaradorligi, iqtisodiy tejamkorligi bilan kelajakda muhim o'rin egalashini ta'kidlab o'tadi. Mashina tarjimasida muayyan uslubga oid bo'lgan klisherlar, ya'ni turg'un birikmalarni ajratib olish va ularning mazmuniy tarjimasi bilan maxsus lug'atlarga kiritish lozim. Shuningdek, quyidagi lingvistik modellar keltiriladi:

Phrases	[We+V V+us our +Noun]	[Verb+suffix] +preposition=>Adj	Adverb	Passive voice	Phrasal verbs
---------	--------------------------	------------------------------------	--------	------------------	---------------

⁵⁸ Вершинина М.И. Идиоматическая фразеология в английском языке делового общения: лексикологический, лексикографический и функциональный аспекты: Дисс. ...док. филол. наук. – М., 2000. – С. 24.

⁵⁹ Mastnán E. The nature of scientific/technical texts from viewpoint of translation studies. Bachelor degree diploma. 1984.

much less limited almost all of which very effective much the same	we are coming to realize we have taken it to be the tube has shown us we are beginning to see we deal with we are now speaking	to be attendant on to be conducive to to be destructive of to be incidental to to be responsive to to be tolerant of	sleary complete ly consider ably essential ly fairly greatly significa ntly	the voltage is applied the magnetic field is set up the line is terminated the switch is closed	look for look into look down look up look forward
---	--	--	---	---	---

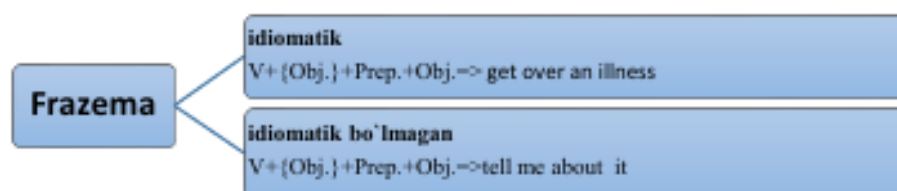
Frazemaga oid soʻz guruhlarining ot (nominal), feʼl (verbal) hamda sifat (adjectival) birikish turlari boʻlib, tuzilish jihatdan ular ikki asosli, uch asosli, toʻrt asosli, besh asosli boʻlishi mumkin: I. Ot fraza: 1) **[sifat+ot]**=> foul play; 2) **[feʼl+ot]**=>have a last word; 3) **[ot+ot]**=>body and soul; II. Ravish fraza: 1) **[predlog+ravish]**=>for ever; 2) **[ravish+ravish]**=>now and again; III. Sifat fraza: 1) **[sifat+sifat]**=>alive and kicking; IV. Feʼl fraza: 1) **[feʼl+feʼl]**=>wait and see; 2) **[feʼl+predlog]**=> Look through.

Turgʻun birikma terminiga oid bir necha qarashlar mavjud. Baʼzi adabiyotlarda verbal (feʼllar) iboralar (verbal expression)⁶⁰, boshqalarda esa iborali feʼllar (phrasal verbs), shuningdek, idiomalar (idioms) tarzda talqin etilganlarni kuzatishimiz mumkin. Iborali feʼllar quyidagi turlarga ajratiladi⁶¹: I. **Verb +preposition (transitive)** =>*Listen to this record Listen!*; II. **Verb+particle (transitive)** =>*Take off your hat. Take yourhat off.*; III. **Verb+particle (intransitive)** => *Sit down!*; IV. **Verb+particle +preposition (transitive)** =>*We've run out of matches. Feʼl+predlog* shaklidagi frazemalarni mazmunij jihatdan uch guruhga ajratish oʻrinli: 1) asos feʼl maʼnosi oʻzgaras frazemalar (hear from); 2) asos feʼl maʼnosi oʻzgaruvchi frazemalar (look for, look after); 3) bir asosli feʼllarning predloglar almashinuvi natijasida maʼno nozikligi bilan farqlanuvchi frazemalar (hear of, hear about). Miqdoriy tarkibiga koʻra frazemalarning quyidagi turlari borligi aniqlandi: 1) feʼl+predlog yoki ravish=>ikki tarkibli frazema; 2) feʼl+ravish+predlog=>uch tarkibli frazema; 3) feʼl+predlog+otli fraza (Verb + particle + noun phrase); 4) feʼl+ otli fraza + predlog (Verb+ noun phrase+particle); 5) feʼl+ravish+predlog+otli fraza (Verb+adverb+ preposition+noun

⁶⁰ Richard A. Spears, Ph.D. Mc. Graw Hill's dictionary of American Idioms and Phrasal verbs. 2005.

⁶¹ Alexander L.G. Longman English Grammar Practice. (for intermediate students). 1998. – P.123.

phrase). Yuqoridagilarni umumlashtirgan holda frazemalar (fe'l frazalar)ning quyidagi turga mansubligini ta'kidlash lozim:



Makoto Nagao konsepsiyasiga qadar professor G.G.Belonogov 1975 yilda frazeologik mashina tarjimasini konsepsiyasini taklif qilgan. Unga ko'ra tarjima qilinayotgan matndagi so'zlar alohida emas, balki frazeologik birikmalar, ifodalanuvchi tushunchalar va turli vaziyatlar o'rtasida munosabatlar tarzida tahlil etilishi kerak. Buning uchun frazeologik mashina tarjimasini tizimida frazemalar, frazeologik birikmalar hamda alohida so'zlar uchun tarjima ekvivalentliklaridan iborat bo'lgan bilimlar ta'minoti bo'ladi. Aynan mana shu konsepsiya bo'yicha yana bir qator olimlarning, xususan, A.A. Xoroshilovning "Teoreticheskie osnovy i metody postroyeniya sistem frazeologicheskogo mashinnogo perevoda", V.A. Chijakovskiyning "Frazeologiya i mashinnyy perevod (opyt sostavleniya raboty nemesko-russkogo avtomaticheskogo slovary dlya publitsisticheskix i nauchnyx tekstov)" mavzusida nomzodlik dissertatsiyasi yoqlandi. Frazeologik birikmalarning to'g'ri shaklda berilishi matn mazmunini tushunishda aniqligini oshiradi. Ingliz tilidagi *take* so'zi orqali inglizcha-o'zbekcha mashina tarjimasini uchun turg'un birikmalarning semantik bazasini tuzish bo'yicha ayrim tahlillarga to'xtalamiz. Lingvistik bazaga quyidagi modellar bo'yicha kiritish mumkin. (So'zlarning konnotativ ma'nolari lug'atda kodlangan bo'ladi.)

- I. [Take+Noun]=> 1) o'lja; 2) tushum, kirim, yig'im; 3) maosh, oylik; 4) yutuq; 5) ijara (er); 6) ijaraga olingan yer; 7) pora.
- II. [Take+Verb]=> 1) olmoq; 2) qabul qilmoq; 3) undirmoq; 4) olib qo'yimoq; 5) ijaraga olmoq; 6) vazifani ado etishga kirishmoq; 7) olmoq; egallamoq (vaqt, joy); 8) qabul qilmoq; 9) yuvintirmoq; 10) to'lamoq
- III. [Take+Preposition (away/for/back/in /off /on /out /over /through /up)]=> take+away=>olib tashlamoq, olib ketmoq
- IV. [Take+Preposition +Noun]=>take on credit– kredit olmoq
- V. [Take+Preposition +Pronoun]=>take with you– o'zi bilan birga olmoq
- VI. [Take+ Adjective+Noun]=>take great care of– ...haqida g'amxo'rlik qilmoq

- VII. [Take+ Adjective(more)+than]=>take more than– ...nafaqat talab qilmoq, balki...
- VIII. [Take+Noun]=>take account of– hisobga olib chiqmoq, ro'yxatga olmoq, hisoblab chiqmoq
- IX. [Take+a/an/the +Noun]=>take a photo– rasmga olmoq
- X. [Take +a/an/the+ Noun+Preposition]=>take the place of– o'zgartirmoq; o'zgarib qolmoq; almashtirmoq; yangilamoq; o'rniga qo'ymoq
- XI. [Take+ Noun+Preposition]=>take care of–qaramoq (economics)
a. Tatable care of– g'amxo'rlik qilmoq
- XII. [Take+ to+Noun]=>take to pieces– ajratib tashlamoq (qismlarga)
- XIII. [Take+Adverb]=>take down– demontaj qilmoq, qismlarga ajratmoq take outside– ...ga olib chiqmoq
- XIV. [Take+ Adverb+Preposition]=> take down in the minute– protokolga kiritmoq
- XV. [Take+ Adverb+ a/an/the +Noun]=>take out a liscence –litsenziya olmoq
- XVI. [Take+ Adverb+ Preposition+Noun]=> take out of production– mahsulot sotib olmoq
- XVII. [Take+as +Noun]=>take as a basis – biror narsaning asosi deb bilmoq
- XVIII. [Take+Verb]=>take leave of-(bilan) – xayrlashmoq

Yuqoridagi namunalardan kelib chiqib shuni aytish mumkinki, *take* fe'lining o'n sakkizta modeli bo'lib, ular turli semalarni ifodalaydi. Biroq keltirilgan modellarda ifodalanayotgan leksema ham denotativ, ham konnotativ ma'no kasb etadi. So'zlarning birikish usullari modellarini berish hamda turg'un birikmalarining lingvistik tabiatini o'rganish mashina tarjima tizimi sifatini yanada oshirishga xizmat qiladi.

Ingliz tilidan o'zbek tiliga tarjima qiluvchi dastur uchun turg'un iboralar, ayniqsa, fe'lli iboralarni chuqurroq tahlil qilish tadqiqotning asosiy maqsadi hisoblanadi. Unga ko'ra tarkibida fe'l mavjud bo'lgan birikmalar modellari aniqlangandan so'ng ularning semantik bazasi tuzildi. Fe'lli frazemalar ingliz tilidagi eng murakkab tarjima birligi bo'lganligi bois u gapning struktural tuzilishi va semasiga bevosita ta'sir ko'rsatadi.

Manbalarda fe'lli frazemalarga quyidagicha ta'rif keltiriladi: "Fe'lli frazemalar ingliz tilida ko'p qo'llaniladigan muhim kategoriyadir. Avvalo, fe'lli frazemalar kundalik iste'molda tez-tez qo'llaniladi, shuningdek, u boshqa til vakillari uchun juda murakkabdir. Chunki uni grammatik va semantik jihatdan to'g'ri qo'llashda qiyinchiliklarni yuzaga keladi. Ikkinchidan, fe'lli frazemalar tilning boyligi hamdir. Ular orqali kishilarning faoliyati va o'zaro munosabati aks etadi."⁶²

⁶² Andreea-Rosalina Olteanu. A holistic approach to phrasal verbs, Editura Sfântul Ierarh Nicolae 2012. – P. 16.

Inglizcha matnlarni o'zbek tiliga tarjima qiluvchi dasturiy ta'minot uchun matn birliklarini kompyuter tarjima qilishda ana shunday TBlarning ekvivalentligini to'g'ri tanlash, lingvistik algoritmini tuzish, qolaversa, quyidagi masalalarni hal etish muhim sanaladi: 1) har bir birlikning lingvistik modeli mukammal holda kiritilgan bo'lishi; 2) muayyan leksik birlikning qiymati so'z birikmasi, termin, turg'un birikma, qo'shma so'z tarzida alohida lug'atlarda aks etishi; 3) dasturiy ta'minot uchun o'zbek tilidagi har bir birlikning ekvivalenti tavsifiy xarakterga ega bo'lishi kerak. Mashina tarjimasida bilimlar bazasi hamda lingvistik ta'minot tarjima faoliyatini yengillashtirish bilan birga matnda anglashilgan umumiy konseptni aniqlashtiradi.

Fe'lli frazemalar (FF) o'timli va o'timsiz (*transitive va intransitive verbs*) turlarga bo'linadi⁶³: 1) **verb+preposition (transitive)** => *Listen to this record. Listen!*; 2) **Verb+particle (transitive)** => *Take off your hat. Take your hat off.*; 3) **Verb+particle (intransitive)** => *Sit down!*; 4) **Verb+particle +preposition (transitive)** => *We've run out of matches.*

Ingliz tilida predloglarning asosiy vazifasi gap bo'laklarini sintaktik munosabatga kirishiga yordam berishdir. Ular fe'lli frazemaning o'zgarma qismi (particle) deb ataladi. Ba'zi o'rinlarda ular o'zi anglatyotgan ma'noni saqlab qolsa, ba'zan semantik yahlitlanib, bir butun leksemani hosil qiladi. Frazemalarni ma'no ko'chirishning o'z va ko'chma ma'no guruhlariga ko'ra ajratish mumkin: a) *drop off* (to'kilmoq- o'z ma'no); b) *drop off* (mudramoq, ko'zi ilinmoq - ko'chma ma'no). Ularni quyidagi namunalarda ham uchratish mumkin: 1. *I think, you have picked up a flu* (kasal bo'lmoq); 2. *I picked up Spain when I was in it* (o'rganib olmoq, bilib olmoq); 3. *I was picked up for speeding my car* (so'roqqa tutilmoq); 4. *I am picked up at 6 p.m.* (olib ketmoq); 5. *Fortunately, economy is starting pick up* (yaxshilanmoq, o'smoq, yuksalmoq). Fe'lli frazemalar shunday nozik xususiyatga egaki, o'zgaruvchi qismning tartibini o'zgartirish yoki boshqa u bilan turdosh bo'lmagan leksemaning kelishi ma'noning umumiy yahlitligini o'zgartirishi mumkin: *blow off* – 1) *uchirib ketmoq, esmoq, esib turmoq*; 2) *g'azablanmoq, jahli chiqmoq*; 3) *vaqtni bekor sarflamoq*; *blow smb. off* – e'tiborsiz qoldirmoq,

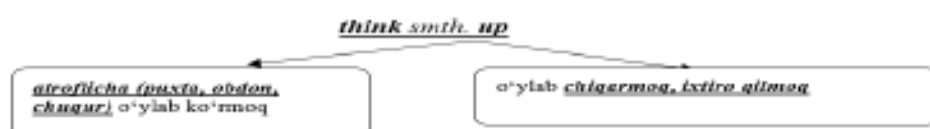
⁶³ Alexander L.G. Longman English Grammar Practice. (for intermediate students) 1998. – P. 116-123.

e'tiborga olmaslik, nazar-pisand qilmaslik; blow smth. off – kutilayotgan joyga bormaslik yoki qatnashmaslik, biror kimsa bilan uchrashuvni qoldirmoq. Demak, bu kabi frazemalarning tarkibiy qismlari ajralish yoki ajralmasligi mumkin: 1) ajraladigan (seperable)– shut down –*shut it down*; 2) ajralmaydigan (inseperable – *go in for*). Shu bois ayrim o'rinlarda fe'lli frazeologik birliklar bilan gumon olmoshlarini (GO) (*smth., smb., somewehere*) birgalikda berish maqsadga muvofiq. Bu esa u yoki bu frazemalarni ajralish yoki ajralmasligini ko'rsatishga xizmat qilsa, boshqa bir tomondan, fe'l nisbatlarining ham o'zgarishiga sabab bo'ladi: **wake up** – *uyg'otmoq; wake smb. up –uyg'otmoq.* Yana bir jihati gumon olmoshlari (GO) fe'lli frazalar tarkibida bo'lsa, ma'noga ta'sir etmasligi yoki umumiy semani o'zgartirib yuborishi mumkin: *I invited him to come to our party but he **turned me down**. –Men uni kechamizga taklif etdim, lekin u meni **rad etdi**; I hope my project will not be **turned down**. –Umid qilamanki, mening loyiham **rad etilmaydi**.* Yana bir misolda esa Golarini o'zgartirish bilan ma'noga ta'sir etganini ko'rish mumkin: **perk smb. up** - *ruhlantirmoq, dalda bermoq, ko'nglini ko'tarmoq; perk smth. up* - 1) *tozalamoq, yangilamoq, musaffo qilmoq; 2) (iqtisodni) jonlantirmoq, ko'tarilishiga sabab bo'lmoq (narxni); perk up -1) (boshini) ko'tarmoq, faollashmoq; 2) (narhnavo haqida) ko'tarilmoq, o'smoq; 3) jonlanmoq, ruhlanmoq, jon kirmoq; 4) (kasallikdan keyin) sog'aymoq, tuzalmoq, sog'ayib ketmoq.* GO lari o'rnida jonli yoki jonsiz ot leksemlarnig qo'llanilishi fe'lli frazemalarda anglashilayotgan ma'no umumiy bo'lsa ham, predloglar bilan farqlanishi mumkin: **agree with smb.** – *biror kishi bilan kelishib olmoq; to agree to* — *biror narsaga kelishmoq, qo'shilmoq; to be angry with smb., to be angry at smth.* –*jahli chiqmoq; to be sorry for smb.; to be sorry about smth.* –*kechirim so'ramoq.* Ingliz tilidagi FF o'ziga xos yana bir jihati o'zbek tiliga tarjima qilishda ko'zga tashlanadi. U yoki bu frazemaning ma'nosi bir leksemaga yoki so'z birikmasiga, hatto, ba'zi o'rinlarda iboralarga teng kelishi mumkin. Masalan, **want smb. or smth. out of smth.**–*kimningdir ketishini yoki chiqishini xohlamoq (so'z birikmasi) yoki want for smth.*–*ehtiyoj sezmoq, muhtoj (hojatmand) bo'lmoq, talab qilmoq (qo'shma fe'l), want for* –*xohlamoq, istamoq (sodda fe'l) tarzida.*

FFlar faqat **fe'l+predlog** tarzida baholash to'g'ri bo'lmaydi, chunki fe'ldan keyin keluvchi predlog umuman fe'lga aloqador bo'lmasligi ham mumkin, ya'ni predlog o'zidan keyingi birlikning tarkibi bo'ladi: *I came from Tashkent – Men Toshkentdan keldim.* Fe'lli frazemalar **fe'l+leksema {predlog}** tarzida ham nomoyon bo'ladi. Bunday birikmaning murakkab tabiati shunda ko'rinadiki, kontekstdagi holatiga qarab ularni erkin birikma yoki barqaror birikma tarzida ham tarjima qilish mumkin.



Ba'zi FFlar o'zbek tiliga tarjima qilinayotganda ularga qo'shimcha leksemalar qo'shilish holatlari ham kuzatiladi:



Biz ingliz tilida mavjud 12 mingga fe'lli frazemalarni yig'ib, ularning o'zbekcha tarjimasining muqobil variantini ma'lumotlar bazasiga kiritdik. Manbalarda FFarning lingvistik modellari quyidagicha ifodalangan⁶⁴: 1) V+ADV=>calm down; 2) V + PREP. => keep on; 3) V+PREP+PREP=> put up with; 4) V+ADV+ PREP.=> change out of smth. (kiyimlarni yechib almashtirmoq); 5) V+N/PRON+ADV =>call smth. down (to smb.); 6) V+ADV+pro/N =>pull down smth.; 7) V+PRON+ADV=> call smth. down (to smb.); 8) V+ N/PRON + PREP=>give something away; 9) V+a / an+Noun=>make a noise; 10) V+a / an+V=>have a smoke

FFlarning yana bir xarakterli tomoni ayrimlari o'tgan zamon shaklida bo'lib, sifat yoki sifatdoshga aylanadi: **bent on doing smth.**; **bereft of smb. or smth.** -1) *ayrilgan, judo bo'lgan*; 2) *bo'sh, quruqdan-quruq, yuksiz*. Shu bois bunday tuzilishdagi birliklar ham lug'atda o'rin olgan bo'lishi lozim. Fe'lli frazemalarning ba'zilarida o'zgaruvchi qismlar (fe'llar) hamda o'zgarmas qismlar (predloglar) o'rtasidagi sinonimiya ham kuzatiladi: **have at smb.**–**go at smb.** (*tashlanmoq, hamla*

⁶⁴ Oxford Phrasal Verbs Dictionary for learners of English. Oxford university press. 2002.

qilmoq, hujum qilmoq); *turn off –turn out* (*o‘chirmoq*). Shuningdek, fe‘lli frazemalardan hosil bo‘lgan otlar ham mavjud bo‘lib, ular yahlit holga kelib qolgan. Bu kabi otlar ikki tuzilishga ega: **1) predlog+fe‘l (offspring) -a) zot zurriyot; b) mahsulot, natija; 2) fe‘l+predlog{ravish} (takeover) -a) o‘lja; b) davlat to‘ntarilishi.**

FFlarning modellari quyidagilar:

- V + oneself +P+smth. => align oneself with smb. or smth.
- V+ oneself +P=> arch (oneself) over
- V + oneself +P+smb.=> attach oneself to smb
- V + P+ smb. or smth.+P smth. =>arrange with smb. about smth.
- V+ smb. or smth. + P+ smb. or smth. => associate smb. or smth. with smb. or smth.
- V+ smth. +P+ smth.=> balance smth. against smth.
- V+ smth. +P=> bail smth. out
- V+ smb. +P+ smth.=> astound smb. with smth.
- V+smb. +P+ smb. or smth.=> bias smb. against smb. or smth.
- V+smb. +P=> beat smb. up
- V +P+smb.=> bet with smb.
- V +P+smb. or smth.=> attend to smb. or smth.
- V +P+ smth.+P+ smth.=>
- V +P+ smb.+P+ smb.or smth.=> book smb. through (to some place)
- V+P+P=> be in for
- V+P+P+it=> be in for it
- V+P+P+smth.=> be off for smth.
- V+P+P+ smb.or smth.=> bound up with smb. or smth.
- V+P+ P+ smth=>bear up (against smth.)

Kompyuterda tahlil qilinayotgan matndagi fe‘lli frazemalarni leksik analiz qilamiz. Leksemani qidirish algoritmiga ko‘ra joriy leksema leksemalar jadvaliga (LJ) joylanadi. LJ da leksemaning tartib raqami, leksema nomi, kodi, turi va tarjimasi beriladi. Joriy leksema (L1) o‘qilgandan so‘ng, bunday leksemaning mavjudligi yoki mavjud emasligini, agar mavjud bo‘lsa, uning qaysi jadvalda joylashgani, kodi qandayligi aniqlanadi. L1 V (verb) jadvalida uchrasa, navbatda leksema (L2) o‘qiladi. L2 leksema proN, Prep, Adv jadvallaridan qidiriladi, agar uchrasa, L1+L2 ulanadi, keyingi L3 leksema o‘qiladi va keyingi qidiruv ishlari davom ettirilladi. Bu holat L1+...+Ln leksemalar ketma-ketligi VpronAdv, VproNPred, VPP, VP, VAdv jadvallarida uchraguncha yoki aksi bo‘lgan holatgacha davom etadi. Mos jadvalda FF uchrasa, L1+...+Ln leksemalar ketma-ketligi bitta leksema, deb qaraladi va L3 ga mos tartibda joylanadi. Agar FF uchramasa, L2 leksema artikl bilan kelish holati (a/an)

aniqlanadi. L2 artikl bo'lsa, keyingi L3 leksema o'qiladi va u Noun yoki V jadvallaridan qidiriladi, L3 ushbu jadvallarda uchrasa, L1+L2+L3 leksemalar ketma-ketligi bitta leksema deb qaraladi va LJ ga mos tartibda joylanadi. Agar FF topilmasa, L1 ning o'zi LJ ga V sifatida joylanadi. Leksemalar jadvalidan namuna:

T/r	Nomi	Kodi	Turi	Tarjimasi
1	I	102	N	Men
2	put off	254	VP	keyinga qoldirmoq
3	my	110	PP	mening
4	meeting	203	N	uchrashuv

1-qadam. Joriy leksemani (L_i) o'qiymiz. L_i bo'sh bo'lmasa, 2-qadam, aks holda ishni to'xtatamiz; **2-qadam.** L_i ni V jadvaldan qidiramiz, uchrasa 3-qadam, aks holda 1-qadam; **3-qadam.** L_{i+1} ni o'qib, uni proN, Prep, Adv jadvallaridan qidiriladi. Jadvallardan birortasida mavjud bo'lsa, 4-qadam, aks holda 6-qadam; **4-qadam.** L_{i+2} leksemani o'qiymiz. L_{i+1} leksema proN jadvalidan topilgan bo'lsa, L_{i+2} leksema Adv, Prep jadvalidan qidiriladi. L_{i+2} ushbu jadvallarda uchrasa 5-qadam, uchramasa L_i ni V sifatida LJ ga joylanadi va 1-qadam; L_{i+1} Adv jadvalida bo'lsa, L_{i+2} leksema Prep, proN jadvalidan qidiriladi, uchrasa 5-qadam, uchramasa $L_i + L_{i+1}$ bitta leksema sifatida LJ ga joylanadi va 1-qadam; L_{i+1} Prep jadvalida bo'lsa, L_{i+2} leksema Prep jadvalidan qidiriladi, uchrasa **5-qadam**, uchramasa $L_i + L_{i+1}$ bitta leksema sifatida LJ ga joylanadi va 1-qadam; **5-qadam.** $L_i + L_{i+1} + L_{i+2}$ ulanib, bitta leksema, deb qaraladi va L_{i+2} ni LJ ga joylanadi va 1-qadamga qaytiladi; **6-qadam.** L_{i+1} ning artikl ekanligi tekshiriladi. Artikl bo'lsa, L_{i+2} Noun yoki V jadvallaridan qidiriladi. Jadvallardan birortasida mavjud bo'lsa 5-qadam, aks holda L_i ni LJ ga mos tartibda joylanadi, 1-qadam.

Shunday qilib, LJ jadval to'ldirib olinadi. Natijada matndagi barcha Verb, Phrasal Verblarni mos qiymatlari bilan birga ajratib olamiz. Va ular ustida kerakli ishlovlarni amalga oshiramiz.

Leksemalar jadvali:

T/r	Nomi	Kodi	Turi
1.	put smth down	102	SPhV
2.	put off	254	PhV
3.	put	110	MV
4.	take a shower	292	CV

Algoritmnin blok sxemasi:



Bu yerda: S_0 – birinchi leksema; S_m – oxirgi leksema; i – joriy leksema; n – $i+1$ chi leksema, j – $i+2$ chi leksema; k – natijaviy leksema; $s1$ – sanagich; $LJ[]$ – leksemalar jadvali.

Tadqiqot natijalari shuni ko'rsatdiki, mashina tarjimasi uchun lug'at yoki grammatik qoidalarining berilishidan tashqari ularning semantik bazasi ham keltirilishi maqsadga muvofiq. Shu bois ingliz tilidagi fe'lli frazemalarning o'zbekcha tarjima lug'ati lingvistik ma'lumotlar bazasiga kiritildi. Endilikda ingliz tilidagi barcha so'z turkumlari bilan bog'liq barqaror birikmalarning o'zbek tiliga tarjimasi bo'yicha ishlar amalga oshirilishi nazarda tutilgan.

O'zbek tilining axborot - kompyuter uslubini yaratish

Reja

1. O'zbek tilining axborot - kompyuter uslubini yaratish.
2. O'zbek tilining axborot - kompyuter uslubini yaratish mezonlari.
3. Axborot - kompyuter uslubida sinonimlar, frazeologizmlar va polisemantik so'zlarning mavqei.

Tayanch so'z va iboralar: *o'zbek tilining axborot - kompyuter uslubi, tilning axborot uslubi nazariyasi, o'zbek tilining axborot - kompyuter uslubini yaratish mezonlari, axborot - kompyuter uslubida sinonimlar, frazeologizmlar va polisemantik so'zlarning mavqei.*

Yunesko ma'lumotlariga ko'ra jahon tillari miqdori 5000 dan ziyod bo'ldi. Til insonlar o'rtasida kommunikatsiya vositasi sanaladi. Bu esa hozirgi globalizatsiya jarayonida til bilan bog'liq o'ziga xos "to'siq" bo'lishini oldini olish uchun tilni optimallashtirish bo'yicha izlanishlar va amaliy tadqiqotlar olib borish dolzarb vazifa ekanligini ko'rsatmoqda. Tilshunoslarning norma va tilni standartlashtirish bilan bog'liq tadqiqotlarida, tilni soddalashtirish va sun'iy tillar yaratish bo'yicha amaliy ishlari bu yo'nalishdagi ishlarning asosini tashkil etadi. Ingliz tili leksikonining 850 ta ko'p ishlatiladigan so'zlarini qamrab oluvchi qisqartirilgan *Basic English* tili bu sohada erishilgan natijalardan biridir.

Bugungi kunda jahon miqyosida kechayotgan integratsiya va globallashtirish jarayonida o'zbek tilini dunyoviy tillar darajasiga olib chiqish hayotiy zaruriyatdir. Mazkur katta, mas'uliyatli, o'ta muhim maqsadni amalga oshirishda kompyuter texnologiyalariga, xususan, kompyuter lingvistikasiga bo'lgan ehtiyoj yana ham oshadi. Chunki aynan kompyuter lingvistikasi o'zbek tilining jahon miqyosiga chiqishida, o'zbek tilining ham dunyoviy tillardan biriga aylanishiga, uni o'rganish va o'rgatish ishlarini optimallashtirishga imkoniyat yaratadi.⁷³

O'zbek tilining kompyuter uslubini yaratish ham tabiiy tilni optimallashtirish muammosining yechimi bo'lib hisoblanadi. Zero jahon andozalari talablariga javob beruvchi, har tomonlama mukammal, qulay va ommabop kompyuter uslubi yaratilganidan so'nggina yuqoridagi muammolarni hal qilish mumkin bo'ladi. Ma'lumki, inson fikrini og'zaki yoki yozma shakllantirishda turli grammatik vositalardan foydalanadi, u o'z nutqida frazeologizm, har xil iboralar, badiiy tasvir vositalarini qo'llaydi, ya'ni har bir insonning nutqi o'ziga xos va takrorlanmasdir. Ammo axborot va texnologiyalar asri bo'lgan XXI asrda kompyuter uslubi uchun

⁷³ Po'latov A., Muhamedova S. Kompyuter lingvistikasi. – Toshkent, 2007. – B. 73-74.

bular ortiqchalik qiladi, ya'ni badiiy uslubning kengligi bugungi kun talablariga, xususan, kompyuter uslubining talablariga javob bermaydi. Chunki ma'lumot uzatishda qisqalik, lo'ndalik, qoliplilik zarurdir. Biz o'zbek tilining barcha imkoniyatlari, ma'no boyligi, asrlar davomida shakllangan lug'at zahirasi, uning go'zalligi Alisher Navoiy hazratlari ta'kidlaganidek, boshqa tillardan ustun jihatlari bilan faxrlanamiz. Ammo yana takroran ta'kidlab aytamizki, kompyuter uslubida badiiy tasvir vositalariga, turli nutqiy jilolar va o'ziga xos neologizmlarga o'rin bo'lishi mumkin emas. Bu uslub badiiy (yoki san'at) uslubi bilan parallel ravishda shakllanib, rivojlanib boraveradi, ammo ularning qo'llanilish sohasi turlicha bo'ladi. Har ikki uslubdan turli soha vakillari foydalanishadi: adabiyot, san'at, madaniyat ahli san'at (badiiy) uslubdan, kompyuter mutaxassislari, reklama, biznes, tabiiy soha vakillari va eng muhimi o'zbek tilida ish yurituvchilar (yurisprudensiya) kompyuter uslubidan foydalanishlari mumkin. Kompyuter uslubi jahon andozalari talablariga monand holda fikrni aniq, qisqa, lo'nda, kompyuter qayta ishlashiga moslashtirilgan bo'ladi. Shunday uslubni yaratmas ekanmiz, Respublikamizda o'zbek tilida to'liq ish yuritishga o'ta olmaymiz. Masalan, bir qiyosiy fakti keltirsak. Yaqinda "Dunyoviy o'zbek tili" kitobi bosmadan chiqdi. Unda o'zbek tilidagi 1 ta fe'lning ("ishlamoq" fe'lining) taxminan 100000 ta shakli (leksik grammatik paradigmalari) rus va ingliz tillariga o'girilgan. Ingliz tilida esa fe'l taxminan 100 ta shakli borligi hammaga ma'lum. Mana sizga qiyos: 100 000 ta va 100 ta. Kompyuter uslubida ushbu 100 000 ta shakldan faqat eng zaruriy eng maqbul (optimal) varianti tanlab olinadi, qolganlari esa imkoniyat tarzida saqlanadi. Xullas, o'zbek tilining kompyuter uslubi yaratishning asosiy mezonlari sifatida quyidagilarni ajratib ko'rsatish mumkin:

- a) aniqlik, qisqalik, lo'ndalik; bir qiymatlilik;
- b) jahon andozalari talablariga moslik;
- v) barcha uchun birdek tushunarliklik;
- g) turli badiiy tasvir vositalaridan xolilik (shu jumladan, sinonimlarning turli variantlari ham kompyuter uslubiga kiritilmaydi, sinonimlik qatorlarning dominantasi tanlab olinadi);
- d) so'zlar asosan bir ma'nolilik kasb etadi (ya'ni uslubga so'zlarning faqat denonativ semalari qamrab olinadi, konnotativ semalarga e'tibor qilinmaydi).
- e) kompyuter talablari qat'iy hisobga olinadi.

Ko'rinadiki, o'zbek tilining kompyuter uslubini yaratish hayotimizda muhim ahamiyatga ega. Zero, kompyuter uslubi aynan kompyuter lingvistikasi bo'yicha mutaxassislar kompyuter dasturlari va filologlar ittifoqida yaratiladi. Kompyuter lingvistikasining ijobiy jihati ham aynan shunda ko'rinadi: u filolog va kompyuter mutaxassislarini birlashtiradi.

Topshiriqlar

1. O'zbek tilida qanday uslublar bor?

2. Uslub deganda nimani tushunasiz?
3. O'zbek tilining axborot-kompyuter uslubi asosan tilning qaysi imkoniyatlarini e'tiborga olmaydi?
4. Til uslubi mezonlari nimalardan iborat?
5. O'zbek tilining axborot-kompyuter uslubi yaratishning asosiy mezonlari nimalardan iborat?

Adabiyotlar

1. Нелюбин Л. Л. Компьютерная лингвистика и машинный перевод. – М.: ВЦП., 1991.
2. Шемакин Ю.И. Начало компьютерной лингвистики. – М.: МГОУ. 1992.
3. Белоногов Г.Г. Компьютерная лингвистика и перспективные информационные технологии. - М.: Русский мир, 2004.
4. Р. Г. Пиотровский. Текст, машина, человек. – М.: Просвещение. 1989.
5. Po'latov A., Muhamedova S. Kompyuter lingvistikasi. – Toshkent, 2007.

Ma'ruza № 15-16

Kompyuter lingvistikasi metodlari

Reja

1. Informatsion jamiyat va kompyuter lingvistikasi.
2. Sun'iy intellekt g'oyasi.
3. Ekspert tizimlar va neyron tarmoqlar.

Tayanch so'z va iboralar: *informatsion jamiyat, sivilizatsiyaning uch to'qlini, istiqbolli yo'nalishlar, sun'iy intellekt, neyron tarmoqlar, ekspert tizimlari, lingvistik protsessorlar, tabiiy tilli interfeys, biokompyuter, semiotik va biologik yondashuv, mashina ijodi, intellektual (ratsional) agent, texnologik singulyarlik, Tyuring testi, bilimlar injenerligi.*

Insoniyat uchinchi ming yillikning boshida axborot asriga qadam qo'ydi. Bu bilan bog'liq holda informatsion jamiyat konsepsiyasi ishlab chiqildi. Bu konsepsiyaning asoschisi amerikalik futurolog olim E.To'ffler hisoblanadi. U jamiyatni turli bosqichlarga bo'lib o'rganishni taklif etib, **sivilizatsiyaning uch to'qlinini** ajratdi: a) birinchi to'qlin – industrial jamiyatgacha bo'lgan davr (qishloq xo'jaligi sivilizatsiya). Bunda asosiy boylik yer va tabiatdan olingan zahiralar hisoblanadi; b) ikkinchi to'qlin – industrial sivilizatsiya. Bunda asosiy boylik ishlab chiqarish vositalari sanaladi; c) uchinchi to'qlin – postindustrial

sivilizatsiya (information jamiyat). Bunda kapital va mehnat o'z o'rnini informatsiya va bilimga bo'shatib beradi.

Minglab yillar davomida rivojlangan inson tafakkuri uchun bugungi kunda dastlabki manba, ya'ni axborot olish hayotiy zaruriyatga aylandi. Shuning uchun jahon bozorida ma'lumot oltindan ham qimmatliroq baholanadi. Mazkur axborotni topish, saqlash, qayta ishlash va boshqalarga yetkazishning qulay usullariga bo'lgan ehtiyoj kun sayin oshib bormoqda. Demak, kimki, mazkur qulay usullarni ishlab chiqsa, ularga egalik qilsa, mana shu odam dunyodagi eng boy odam bo'ladi. Ma'lumki, dunyodagi eng boy-badavlat kishi Bill Geyts hisoblanadi. Nima uchun dunyodagi eng boy kishi kompyuter, ya'ni axborotni qabul qilish, saqlash va qayta ishlash sohasining vakili? Bu savolga Billi Geytsning o'zi shunday javob beradi: *"Kim axborotga ega bo'lsa, u hamma narsaga ega bo'ladi"* yoki uning yana bir mushohadasi dunyoga mashhur: *"Axborot dunyodagi eng katta, eng qimmat moddiy boyliklardan ham ustun turadi, chunki ushbu boyliklar axborot vositasida qo'lga kiritiladi"*.⁷⁴

2003 yil dekabrda Jenevada (Shveytsariya) informatsion jamiyat qurish muammosiga bag'ishlangan Butunjahon Sammiti bo'lib o'tdi. Unda shunday shior o'rta tashlandi: *"Informatsion jamiyat qurish – yangi mingyillikning global chaqirig'idir"*. Sammitda ikki muhim hujjat qabul qilindi: Informatsion jamiyat qurish prinsiplari to'g'risidagi Deklaratsiya⁷⁵ hamda information jamiyat qurish bo'yicha ish Rejasi⁷⁶. Mazkur hujjatlarda belgilanishicha, endilikda jamiyatning taraqqiyoti kompyuter texnologiyalar bilan chambarchas bog'liq bo'lib qoladi. Shu o'rinda alohida ta'kidlash lozimki, kompyuter lingvistikasi informatsion texnologiyalarining istiqbolini belgilovchi eng muhim yo'nalish sanaladi hamda u informatsion jamiyat taraqqiyotida hal qiluvchi ahamiyatga ega bo'ladi.

Olimlar kompyuterlarni intellektual jarayonlar avtomatizatsiyasini yuzaga chiqaruvchi muhim vosita sifatida baholadilar. Qisqa vaqt ichida kompyuterlarning teoremlarni isbotlashda, rasm chizishda, shaxmat o'yinida foydalanilishi sun'iy intellekt bo'yicha izlanishlarning ilk samarasidir. Asta-sekinlik bilan insonning aqliy faoliyatini modellashtiruvchi dasturlarning yaratilishi muayyan sohalardagi murakkab masalalarni yechuvchi ekspert tizimlarining yuzaga kelishiga olib keldi. Lekin shuni ta'kidlash zarurki, hali hanuz inson miyasining analogi bo'la oladigan dastur yoki tizim yaratilgani yo'q.

Kompyuter lingvistikasi fanining eng istiqbolli yo'nalishlaridan biri sun'iy intellekt hisoblanadi. Virtual qomus hisoblangan Wikipedia sahifalarida **sun'iy intellektga (artificial intelligence)** shunday ta'rif berilgan: "Sun'iy intellekt bu – kompyuter yoki robotning shunday xususiyatiki, u asosan insonning aqliy faoliyati bilan bog'liq bo'lgan masalalarni, xususan, fikrlash, nutq mazmunini anglash,

⁷⁴ Po'latov A., Muhamedova S. Kompyuter lingvistikasi. – Toshkent, 2007. – B. 75-76.

⁷⁵ World summit on the information society. Declaration of Principles. Building the Information Society: a global challenge in the new Millennium. Document WSIS-03/GENEVA/DOC/4-E. 12 December 2003.

⁷⁶ World summit on the information society. Plan of Action. Document WSIS-03/ GENEVA/DOC/5-E. 12 December 2003.

ma'lumotlarni umumlashtirish kabilarni yechishga yo'naltirilgan bo'ladi. Bu termin mazkur xususiyatlarga ega bo'lgan sistemalarni ishlab chiqish bilan bog'liq kompyuter texnologiyalarining bir bo'limiga nisbatan ham ishlatilmoqda".⁷⁷ Sun'iy intellekt bo'yicha tadqiqotlar olib brogan olim Jon Makkarti shunday yozadi: "*Sun'iy intellekt – intellektual mashinalar, intellektual kompyuter dasturlar yaratish texnologiyasi va bu haqidagi fan. Intellektual sistema strukturasi 3 asosiy blokni o'z ichiga oladi: bilimlar bazasi, masalani yechuvchi tizim va intellektual interfeys*".⁷⁸

Ekspert tizimlari modelida, intellektual information tizimlarda, ba'zi mashinalarda, hayvonlarda va odamlarda intellektning turlicha darajalari va ko'rinishlari mavjud. Makur turlichalik bilimlar bazasi va qanday operatsiyalarni bajara olishiga bog'liq. Intellektni bunday talqin qilishdan ko'rinib turibdiki, sun'iy intellekt insondagi intellekt tushunchasining aynan imitatsiyasi emas, balki inson intellektual faoliyatining ko'plab qirralarini o'z ichiga olgan, modellashirilgan tizimlar texnologiyasi demakdir. Ayni shu ma'noda "artificial intelligence" terminini birinchi marta **Jon Makkarti 1956 yilda** Darmut universitetidagi konferensiyada qo'lladi va shundan buyon terminning o'ziga bo'lgan tanqidlarga qaramay ishlatilib kelinmoqda. Rossiya sun'iy intellekt assotsiatsiyasi Peterburg bo'limi boshlig'i T.A.Gavrilovning yozishicha, ingliz tilidagi *artificial intelligence* birikmasi fantastik antropomorf bo'yog'iga ega emas, u rus tiliga muvaffaqiyatsiz tarjima qilingan. *Intelligence* so'zi "aql bilan hukm chiqarish, fikr yuritish qobiliyati" ma'nosini anglatadi, u rus tiliga "интеллект" deb noto'g'ri tarjima qilingan.

XX asrning 70-yillariga qadar sun'iy intellekt bo'yicha tadqiqotlar kibernetika hamda informatika doirasida olib borilgan. XX asrning 80-90 yillaridan boshlab sun'iy intellekt juda ko'plab fanlarning o'rganish obyektiga aylandi. Jumladan, neyrolingvistika, psixologiya, informatika, neyrofiziologiya, epistemologiya (falsafadagi bilish haqidagi ta'limot), kognitologiya, kognitiv lingvistika, kompyuter lingvistikasi kabi fanlar ham o'z doirasida sun'iy intelekt muammosi bilan shug'ullanadilar.

Sun'iy intellekt tizimlarni yaratish muammosiga uch xil asosiy yondashuvni ajratib ko'rsatish mumkin:⁷⁹

- Quyi (inglizcha **Top-Down AI**), semiotik yondashuv – ekspert tizimlari, bilimlar bazasi, mantiqiy hukm chiqarish tizimi, yuqori darajadagi psixik jarayonlar (tafakkur, nutq, emotsiya, ijod kabi)ni imitatsiyasi va analogi bo'lgan tizimlarni yaratish;

- Yuqori (inglizcha **Bottom-Up AI**), biologik yondashuv – neyron tarmoqlarini o'rganish, biologik unsurlarga asoslangan holda intellektual

⁷⁷ http://www.wikipedia.org/wiki/artificial_intelligence.html

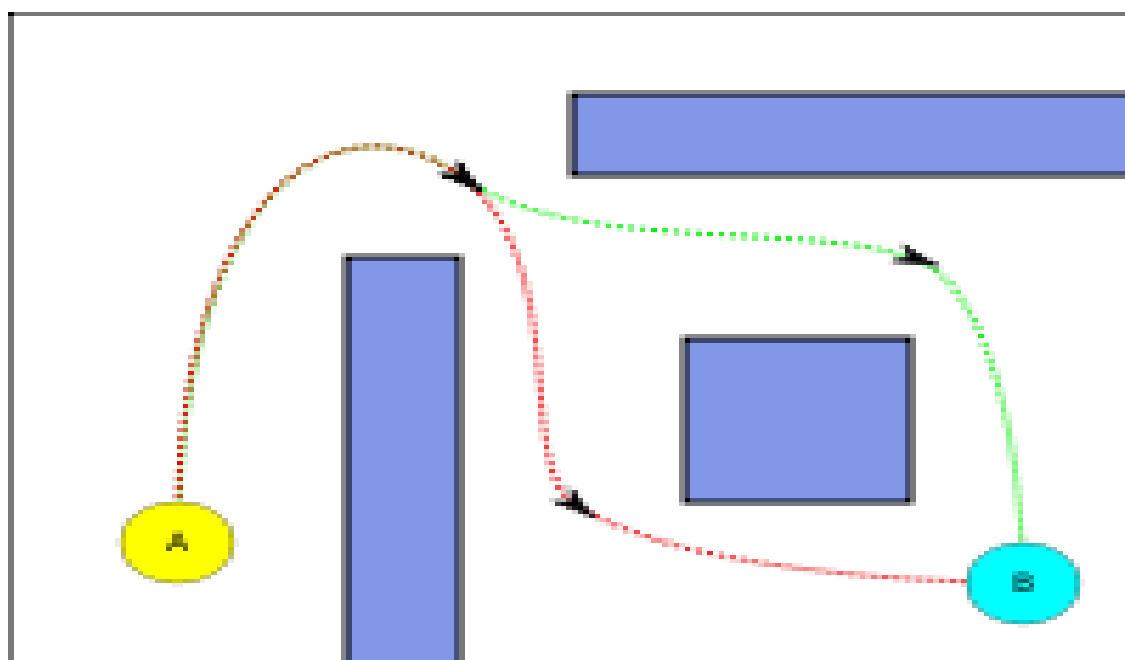
⁷⁸ <http://www-formal.stanford.edu/jmc/whatisai.html>

⁷⁹ **Рассел С., Норвич П.** Искусственный интеллект: современный подход /Artificial Intelligence: a Modern Approach / Пер. с англ. и ред. К. А. Птицына. -2-е изд. - М.: Вильямс, 2006. - С.8.

salohiyatni modellashtirish va natijada neyrokompyuter yoki biokompyuterlarni yaratish.

- Semiotik va biologik yondashuvning sintezi bo'lgan umumlashtiruvchi gibrid tizimli yondashuv – sun'iy intellektning yakuniy maqsadi bo'lishi lozim.

Mazkur yondashuvlardan tashqari yana simvolli, mantiqiy, agentga yo'naltirilgan yondashuvlar ham mavjud. Simvolli yondashuv, asosan, simvolik mantiq qoidalarini tizimga tatbiq etish bilan bog'liq bo'lib, bu ko'proq tizim qoidalarini ishlab chiqishda, hisoblashda (hukmning chin yoki yolg'onligini baholashda) samara beradi. Mantiqiy yondashuv predikatlar tilidan foydalanish asosida sun'iy intellekt tizimini bilimlar bazasining mantiqiy modellari bilan ta'minlash uchun ximat qiladi. 1980 yilda **Prolog** nomli mantiqiy programmashtirish tizimi va tili yaratildi. Prolog tilida yozilgan bilimlar bazasi faktlar naborini hamda mantiqiy xulosa chiqarish qoidalarini belgilaydi. Agentga yo'naltirilgan yondashuv (yoki intellektual, ratsional agentlardan foydalanishga asoslangan yondashuv) 1990 yillarda rivojlandi. Bunga ko'ra intellekt – bu qo'yilgan masalani hisoblash qismi (rejalashtirish qismi) xolos. Bunda intellekt masalani to'laqonli yechish uchun atrofdagi tizimlar munosabatga kirisha oladigan qo'shimcha intellektual (ratsional) agentga ehtiyoj sezadi. Buni quyidagi chizmada yaqqol ko'rish mumkin:⁸⁰



Umuman olganda robototexnika va sun'iy intellekt bir-biri bilan aloqador tushunchalar bo'lib, kelajakda ushbu ikki yo'nalishning integratsiyasi natijasida intellektual robotlar yaratilishi nazarda tutilmoqda, bu esa sun'iy intellektning yana bir yangi yo'nalishi hisoblanishi mumkin. PLEO, AIBO, QRIO kabi o'yinchoq-robotlar intellektual robototexnikaga misol bo'la oladi.

⁸⁰ http://www.wikipedia.org/wiki/artificial_intelligence.html

Sun'iy intellekt tizimi insonning ijodiy tafakkurini modellashtirishni ham o'z ichiga oladi. Shu bilan bog'liq ravishda **mashina ijodi** tushunchasi yuzaga keldi, bunda san'at asarlarini avtomatik tarzda yaratish jarayoni tushuniladi. Bugungi kunda ijodni to'la modellashtirishga erishilgani yo'q, lekin bu sohada sezilarli yutuqlar qo'lga kiritilgan. Jumaladan, kompyuter yordamida musiqali melodiylar yaratilishi, treker-dasturlar, sintezatorlar, sekvensorlar yordamida bir turdagi melodiylar va tovushlar generatsiyasi oqibatida elektron musiqalarning yaratilishi, matnning avtomatik generatsiyasi yordamida she'rlar, ertaklar – umuman matnlar yaratilishi, rasmlar generatsiyasi va kompyuter grafikasi yordamida kino va o'yinlarning yaratilayotganligi (ayniqsa, strategiya va harbiy manyovrlarga asoslangan kompyuter o'yinlarida) kabilar mashina ijodiga misol bo'la oladi.

Sun'iy intellekt tizimi doirasida tatbiq etilgan loyihalar (eksperimentlar) sifatida quyidagilarni ajratib ko'rsatish mumkin:

- **Deep Blue** tizimi – IBM kompaniyasi tomonidan ishlab chiqilgan shaxmat bo'yicha superkompyuter. Deep Blue 1997 yil 11 mayda shaxmat bo'yicha jahon chempioni Garri Kasparov ustidan g'alaba qozondi. “Deep Blue” nomi “Deep Thought” (“Chuqur mulohaza, tafakkur” ma'nosini anglatadi) loyahasiga aloqadordir.

- **MYCIN** tizimi - bir qator kasalliklarga shifokor singari to'g'ri tashxis qo'yishni ta'minlaydi.

- **20Q** — “20 savol” klassik o'yini asosida yaratilgan loyiha bo'lib, u sun'iy intellekt bo'yicha o'tkazilgan dastlabki tajribalardan biridir. Bu o'yin muallifi Robin Byorgenerdir. Bunda foydalanuvchidan biror obyektни topish so'raladi, so'ng foydalanuvchiga 20 ta savol beriladi, berilgan javoblardan kelib chiqqan holda o'sha narsa topiladi. Ayniqsa, bu tizim internet paydo bo'lgandan so'ng juda mashhur bo'ldi (<http://www.20q.net> sayti mavjud).

- **Via Voice** - nutqni tovush orqali tanish tizimi.

- **RoboCup** - har yili soddalashtirilgan futbol bo'yicha robotlar turniri.

Sun'iy intellekt tizimini baholash bo'yicha 2 xil nuqtai nazar mavjud: 1) Sun'iy intelektни inson tafakkuridan ko'ra kuchliroq deb baholash. Ushbu qarash tarafdorlari sun'iy intellekt inson intellektiga qaraganda ko'proq operatsiyalarni qisqa muddatlarda bajarishi mumkin deb hisoblaydilar; 2) sun'iy intellekt inson tafakkuridan, intellektidan kuchsiz. Bu yo'nalishdagi olimlarning fikricha, sun'iy intellektning imkoniyatlari cheklangan bo'ladi, u faqat programmaviy ta'minot asosida muayyan sohalarda ishlaydi, lekin u insonning bilish qobiliyatining, tafakkurining barcha qirralarini qamrab ola olmaydi. Buni Jon Syorl tomonidan o'tkazilgan “Xitoy uyi” nomi bilan mashhur bo'lgan semantik eksperiment ham tasdiqlaydi. Olim bu tajribasini 1980 yilda “*Minds, Brains and Programs*” (“*Tafakkur, miya va programmalar*”) nomli maqolasida e'lon qildi. J.Syorlning fikricha, kompyuter sintaktik qoidalarni algoritm nazariyasi orqali formallashtira

oladi, lekin semantika, tafakkurni to'la modellashtirishga dasturlar ojizlik qiladi.⁸¹ Xuddi shu pozitsiyada turgan olimlardan yana biri Rodjer Penrouz o'zining "*Qirolning yangi aqli*" kitobida tafakkurni formal tizimlar asosida mutloq modellashtirishning imkonsizligini bayon qilgan.⁸²

Ayrim olimlar sun'iy intellekt bo'yicha erishilayotgan yutuqlarning keyingi oqibatlaridan qo'rqmoqdalar. AQSHda Sun'iy intellekt singulyar instituti (SIAI) tadqiqotchisi E.Yudkovskiy insoniyatni sun'iy intellektlar bilan bog'liq global xavf kutayotganidan ogohlantiradi. Olimning fikricha, agar kelajakdagi sun'iy intellekt tizimiga insonlarga nisbatan do'stlik munosabatida bo'lishini ta'minlovchi dastur o'rnatilmasa, butun insoniyat uchun katta xavf paydo bo'ladi.⁸³ Ilmiy fantastik asarlar muallifi Veror Vinj ham sun'iy intellektlar tizimi to'la yaratib bo'lingan paytda jamiyatda keskin o'zgarishlar bo'lishi haqida yozgan. Bu davr ilmiy adabiyotlarda **texnologik singulyarlik** deb atalmoqda. Hatto futurologlarning fikricha, bu davr 2030 yilda boshlanadi.⁸⁴

Fanda hali hanuz sun'iy intellektning aniq va mukammal ta'rifi mavjud emas, chunki fan bu tizimga hali to'liq erishgani yo'q. Mavjud ta'rif va tavsiflarning hammasi bu sohada erishilgan yutuqlarga asoslanib, keltirib chiqarilgan. Falsafada inson intellektining statusi va tabiati haqidaga munozaralar davom etmoqda. Kompyuter "aqllilik" darajasining aniq mezoni ishlab chiqilgani yo'q. To'g'ri, bu sohada ba'zi gipotezalar mavjud. Xususan, "Tyuring testi", "Nyuel-Saymon gipotezasi" kabi.

Sun'iy intellektga doir izlanishlar XX asrning 40-yillarida boshlangan. Sun'iy intellekt muammosiga bag'ishlangan tadqiqotlarning eyforiyasi davrida "*Mashina fikrlay oladimi?*", "*Mashina o'z yaratuvchisidan aqlliroqmi?*" sarlavhalarida maqolalar e'lon qilindi. "Mashina fikrlay oladimi" (1950) sarlavhali maqolaning muallifi ingliz olimi Alan Tyuring mashina imkoniyatini insonning aqllilik darajasi bilan qiyoslash protsedurasini ishlab chiqdi, bu "Tyuring testi" deb nomlanadi. Ushbu empirik test g'oyasi 1950 yilda "Mind" falsafiy jurnalida "*Hisoblash mashinalari va tafakkur*" (*Computing Machinery and Intelligence*) nomli maqolasida e'lon qilindi. Bu testdan ko'zlangan maqsad sun'iy tafakkurning inson tafakkuriga yaqin kelishi mumkinligini aniqlash hisoblanadi. Bunga ko'ra inson bir kompyuter va bir inson bilan o'zaro munosabatga kirishadi. Savollarga javob berish asosida u kim bilan (inson bilanmi yoki kompyuter dasturi bilanmi) gaplashayotganini aniqlashi lozim bo'ladi. Kompyuter dasturining vazifasi esa insonni xatoga yo'l qo'yishiga, noto'g'ri tanlov qabul qilishiga majbur qildirish sanaladi. Barcha ishtirokchilar bir-birini ko'rmaydilar.

TURING TESTidagi savol-javob namunasi:

Q: Please write me a sonnet on the topic of the Forth Bridge.

A: Count me out on this one. I never could write poetry

⁸¹ Searle J.R. Minds, brains, and programs / Behavioral and Brain Sciences, 1980. – P. 417-437.

⁸² Пенроуз Р. Новый ум короля. О компьютерах, мышлении и законах физики. – М.: УПСС, 2005. – P. 14.

⁸³ <http://www.thetech.org/robotics/etics/index.html>

⁸⁴ Kurzweil R. The Singularity Is Near. - N. Y.: Viking, 2005. – P. 3.

Q: Add 34957 to 70764

A: (Pause about 30 seconds and then give answer as) 105621.

Test natijasida A.Tyuring shunday xulosaga keladi: *“Mashina (kompyuter) kelajakda 10 gegabaytli xotira bilan insonni 5 minut ichida 30 %ga chalg’itishi mumkin bo’ladi”*.⁸⁵

Avtomatlashtirilgan ekspert tizimlar mantiqiy masalalarni kompyuter yordamida yechishga imkon beradi. Mazkur tizimlar ma’lum sohadagi bilimlar majmuiga asoslangan implikativ qoidalar – “agar ... sa, u holda ...” qolipidagi mantiqiy qoidalarni o’z ichiga oladi. Masalan, **MYCIN** ekspert tizimi shifokorlarga qon tarkibidagi bir qancha bakterial infeksiyalarga tashxis qo’yishda va ularga mos davolash metodini tanlashda yordam beradi. **MYCIN** tizimining kompyuter dasturi dastlab simptomlar asosida kasalning holati haqida fikr yuritadi, so’ng mikroorganizmning kasallik qo’zg’otish tabiatiga mos keluvchi tashxisni aniqlaydi, bundan so’ng infeksiya sababini aniqlab, bir yoki bir nechta muqobil davolash usullarini tavsiya qiladi.

Tabiiy tilda kompyuter bilan muloqot qilishni ta’minlashga yordam beradigan dasturlar (**lingvistik protsessorlar**) ham yaratilmoqda. XX asrning 70-yillarida AQSHda kompyuter bilan cheklangan muloqotni ta’minlaydigan tabiiy tilli interfeys tizimi (**LIFER** – Language Interface Facility with Ellipsis and Recursion) yaratildi. 1985 yilda “Semantek” korporatsiyasi tomonidan **Q&A** dasturi, “Karnegi grupp” kompaniyasi tomonidan **Language Craft** dasturiy ta’minoti ishlab chiqildi.⁸⁶ Bu turdagi dasturlar hozirda ingliz tilining standartlasgan so’zlashuv uslubida uncha katta bo’lmagan lu’g’at zahirasi bilan ishlamoqda. Tabiiy tilni tushunishga imkon beruvchi dasturlarning qamrovini kengaytirish kompyuter lingvistikasining dolzarb vazifalaridan biridir. Bunday dasturlarni rivojlantirish inson kabi muloqot qila oladigan, erkin kommunikatsiyaga kirisha oladigan sun’iy intellekt g’oyasining istiqbolini belgilaydi.

Sun’iy intellekt yo’nalishidagi tadqiqotlar sirasiga “ekspert tizimlar” yoki “neyron tarmoqlar” bo’yicha olib borilgan amaliy tadqiqotlar ham kiradi. Mazkur ikki yo’nalishga nisbatan “kompyuter intellekti” (“computer intelligence”) termini ishlatilmoqda, bunda inson tafakkurini global modellashtirish jarayoni hamda murakkab masalalarni operativ tarzda yechishga mo’ljallangan avtomatlarning yaratilishi tushuniladi.

Sun’iy intellekt tizimi bo’yicha izlanayotgan mutaxassislar inson tafakkuri, ong-u shuurining barcha qirralarini modellshtirish masalasi imkonsiz qolayotganligini ta’kidlamodalar. Mazkur jihatlar sirasiga insonning ichki dunyosi, intuitsiyasi, ongosti jarayonlari, ilhom kabi sirli fenomenlarni aytish mumkin.

⁸⁵ Jurafsky D., Martin J.H. Speech and language processing. - New Jersey, 2000. – P. 7.

⁸⁶ Анисимов А.В. Компьютерная лингвистика для всех: Мифы. Алгоритмы. Язык. – Киев: Наукова думка, 1991. – С. 10-11.

Inson miyasi murakkab biologik to'ra hisoblanadi, u yuz milliard nerv tolalaridan – neyronlardan tarkib topgan. Neyronlarning bir-biriga uzviy aloqadorligi, zanjirli bog'lanishi natijasida miya faoliyati me'yoriy darajada ishlaydi. Zamonaviy sun'iy neyron tarmoqlar inson miyasidagi neyronlar modeli asosida yaratilmoqda. Hozirgi kunda neyron tarmoqlar bir qator amaliy masalalarni hal etishda keng qo'llanilmoqda. Jumladan, ob-havo prognozini aniqlashda, obrazlarni tanishda (xususan, og'zaki va yozma nutqni aniqlashda), robotlarni boshqarishda, tibbiyotda tashxis qo'yishda ishlatilmoqda. Amerikaning NASA kosmik agentligida neyron tarmoqlarni fazoda robotlarni boshqarish, hududga tasodifan kelib tushgan obyektlarni tutishda foydalanilmoqda. "General Dynamics" kompaniyasi esa dvigatel shovqiniga qarab kemalarni tanishga imkon beruvchi tizim ishlab chiqishdi. AQSH Pensilvaniya universitetida uchuvchi obyektlarni (NUJ) aniqlash uchun neyron tarmoqlar tizimi yaratildi.

Neyron tarmoqlarining ishlash tizimi matematik ifodalar yoki mantiqiy qoidalarga asoslanmaydi, balki namuna (**example based approach**) asosida o'qitishga asoslangan. Bunda muayyan masalani yechish uchun namunalar (misollar), ya'ni o'qituvchi tanlamalar tayyorlanadi va tarmoq shu asosda tekshiradi. Bunda turli xil metodlardan foydalaniladi. Eng keng tarqalgan metod bu dastlabki yechim va sinov natijalarini e'tiborga olgan holda neyron tarmoqlari strukturasini korrektirovka qilishga asoslangan metod (**back propagation method**) bo'lib, unda oldindan ko'plab o'qitish tanlamalari tayyorlab olinadi, u o'z bazasidagi mavjud ma'lumotlarni yangi kiritilayotgan o'qituvchi namuna-tanlamalar bilan solishtiradi. Agar ma'lumotlar bir-biriga mos kelsa, neyron tarmoqlari strukturasida korrektirovka amalga oshirilmaydi. Faqatgina ma'lumotlar bir-biriga muvofiq kelmaganda korrektirovka yuz beradi. Shundan so'ng neyron tarmoqlarida muayyan amaliy vazifani hal etish tizimi to'la shakllanadi.⁸⁷

Sun'iy intellekt tizimi doir amaliy ishlardan biri ekspert tizimlari hisoblanadi. Ekspert tizimi bu – muayyan sohadagi ekspertlik yoki mavjud bilimlarning imitatsiyasi tarzida yaratilgan kompyuter dasturi. Boshqacharoq aytganda ekspert tizimi maxsus sohaga oid barcha bilimlar jamlangan qomus yoki o'ta professional ekspertlarning kompyuter modelidir. Ekspert tizimlari odatiy kompyuter dasturlaridan farq qiladi. Odatiy kompyuter dasturlari qidiruv tizimi, hisoblash va manipulyatsiya tizimigagina ega bo'ladi. Ekspert tizimlari mazkur jihatlardan tashqari xuddi insonlarda bo'lgani kabi dalillarga tayanib muhokama yuritish imkoniyatiga ham egadir. Ekspert tizimlarining asosiy ikkita tarkibiy qismi mavjud: ma'lumotlar ombori va mantiqiy mashina.

Ekspert tizimlari konsepsiyasi Stenford universiteti professori Eduard Feygenbaum nomi bilan bog'liq bo'lib, u 1977 yilda murakkab mantiqiy masalalarni yechishda kompyuter dasturlarining samaradorligi formallashtirish va

⁸⁷ **Anderson J. and Rosenfeld E.** Neurocomputing Foundations of Research, Cambridge: MIT Press, 1998; **Fukushima K.** A Neural Network for Visual Pattern Recognition. IEEE Computer, 1988; **Hopfield J.J.** Neural Networks and Physical Systems with Emergent Collective Computational Abilities. Proceedings of the National Academy of Sciences, 1982.

dasturlash texnikasidan ko'ra muammoli sohaga doir bilimlar bazasiga bog'liq ekanligini asoslab berdi.⁸⁸ Dastlab ekspert tizimlari diagnostika va kasallikni davolash sohasiga tatbiq etildi. Ekspert tizimlari keyinchalik kimyo, biologiya, geologiya sohalarida ham qo'llanila boshlandi.

Ekspert tizimlariga bilimlar bazasini kiritish uchun simvolli ma'lumotlarga ishlov berishga xizmat qiladigan LISP, REFAL, PROLOG kabi sun'iy tillardan foydalanilmoqda. Bu tillar evristik bilimlar va mantiqiy xulosalarni yuzaga chiqarishda eng qulay programmalashtirish tillari hisoblanadi. Bunda yagona qabul qilingan bilimlar bazasi uchun deklarativ formalizm hamda qaror qabul ilish tiimi uchun protsedura tizimi qo'llaniladi.

Ekspert tizimlari ichga ekspertlar hamda bir qator yordamchi programmalar kiradi:

- Savol-javobli dastur foydalanuvchi bilan tabiiy tilda munosabatga kirishishni ta'minlab beradi.
- Bilimlar ombori dasturi foydalanuvchiga bilimlar bazasga o'zgartirishlar kiritish hamda to'ldirib, boyitib borishga imkon beradi.
- Tushuntiruvchi dastur – javoblar tizimini tekshradi.
- Interpretator-dastur muayyan sohadagi terminlarni sharhlash, izohlashga yordam beradi.
- Ijro etuvchi dastur barcha tizimlar ishini ta'minlab beradi.

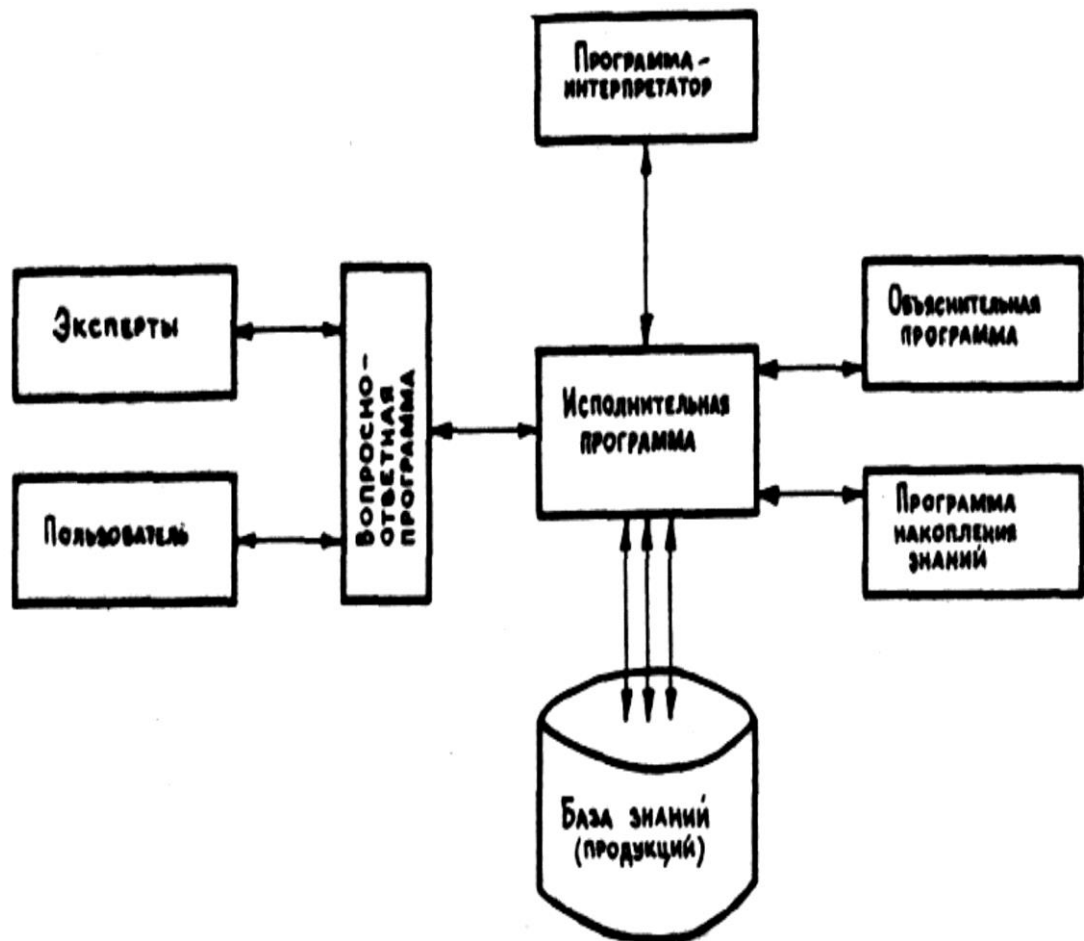
Yuqorida aytilganidek, ekspert tizimlari deklarativ va protsedurali bosqichlar asosida ishlaydi. Deklarativ bosqich bu – ekspert tizimiga kiritilgan bilimlar bazasi bo'lsa, protsedurali bosqich ekspert tizimlarining mavjud bilimlar bazasiga tayanib hukm chiqarish va aniq qaror qabul qilish bosqichini o'z ichiga oladi. Ekspert tizimlarining ishlash prinsipi doir bir misol keltiraylik. Masalan, Qanday sababga ko'ra avtomobildagi benzonasosdan benzin ajralishi qisqarmoqda?

Bu holat bo'icha ekspert tizimlariga kiritilgan deklarativ bilimlar quyidagilar: klapaning tiqilishi, filtrning axlatga to'lganligi, klapan tepa qismining ifloslanganligi, nasos membranasining zararlanganligi. Agar benzonasosdan benzin ajralishi yetarli bo'lmasa, u holda kiritish shlangi axlatga to'lib tiqilib qolgan bo'lishi mumkin kabi.

Iformatsiya shiddat bilan ko'payib borayotgan bugungi davrda ekspert tizimlarining bilimlar bazasini uzluksiz to'ldirib, yangi bilimlar bazasi bilan boyitib borish birlamchi vazifa bo'lib qolmoqda. Chunki vaqt o'tgan sari mavjud bilimlar eskirib, talabga javob bermay qolishi mumkin, shunda ularni yangi va qo'shimcha ma'lumotlar bilan to'ldirish joiz bo'ldi. Bilimlarning mavjud holatlar bilan muvofiqligi hamda bilimlar bazasining boyitilishi ekspert tizimlarining tushuntiruvchi programmalar asosida amalga oshiriladi. Hatto keyingi paytlarda

⁸⁸ Белоногов Г.Г. Компьютерная лингвистика и перспективные информационные технологии. - М.: Русский мир, 2004. – С. 226.

bilimlar bazasi bilan shug'ullanuvchi maxsus yo'nalish ham shakllandi. Bu yo'nalish "**bilimlar injenerligi**" ("Knowledge Engineering") deb ataladi.



Bilimlar injenerligi kognitiv lingvistika, kompyuter semantikasi va kompyuter lingvistikasining yutuqlariga tayanadi. Bunda informatsya va bilimlar strukturasi o'ld bo'lgan **semantik tarmoq**, **freym**, **senariy**, **Word-net** kabi tushunchalar va ularning tatbiqi muhim hisoblanadi. Semantik tarmoq borliqni so'zlar va ularning semantik munosabatlari yordamida modellashtirishga yo'naltirilgan tizim hisoblanadi. **Semantik tarmoq** aniq belgilangan uzellar (obyektlar kesishadigan joy) va dugalar (semantik munosabatlarning biri ikkinchisiga ta'sir ko'rsatadigan nuqta) munosabatdorligi asosida modellashtiriladi. Semantik tarmoq g'oyasi kompyuter texnologiyalaridagi lokal va global tarmoq nazariyasi ta'sirida kelib chiqqan. Uzellar muayyan obyektlarga, muayyan sohaga oid tushunchalarga, dugalar obyektlar o'rtasidagi semantik munosabatlarga aloqador tushunchalardir. Semantik tarmoq ikki qismdan tashkil topadi: 1) mavhum tarmoq; 2) aniq tarmoq. Mavhum tarmoq umumiy tushunchalar orqali ifodalanadigan (zamon, makon, miqdor, sifat kabi) bilimlar fondini o'z ichiga oladi; 2) aniq tarmoq mavhum tarmoq tarkibiga kiruvchi tushunchalarning real obyektlarini va konkret bilimlar fondini o'z ichiga oladi.

Semantik tarmoq tizimi quyidagi munosabatlarni o'rnatish orqali yaratiladi:

1. To'plam va obyekt o'rtasidagi munosabat (muayyan obyekt aniq to'plamga mansub bo'ladi) tasnif munosabati deyiladi. Masalan, "Sharik it hisoblanadi" holati kabi. Ba'zan ushbu munosabat **Member Of, Instance Of** deb ham yuritiladi.

2. Tur va jins o'rtasidagi munosabat **AKO** — «A Kind Of», «Subset Of» («turlaridan biri») deb ataladi. Masalan: "It hayvon hisoblanadi" holati kabi. Sistem leksikologiyada tur uchun **giponim** (it), jins uchun **giperonim** (hayvon), tur-jins munosabati uchun **giponimiya** terminlari qabul qilingan.

3. Butun va qism o'rtasidagi munosabat sistem leksikologiyada **meronimiya** termini bilan ataladi. Masalan, kompyuter – butun; monitor, klaviatura, sichqoncha, protsessor kabilar – qismdir. Butun uchun **xolonim**, qism uchun **meronim** atamaalari ishlatiladi. Demak, motor – avtomobil uchun meronim, avtomobilning o'zi esa motor uchun xolonim sanaladi.

Bilimlar strukturasi oid tayanch tushunchalardan yana biri **freym** hisoblanadi. Freym – (inglizcha frame – “karkas”, “ramka”, “skelet” ma'nolarini anglatadi) sun'iy intellekt tizimida bilimlarni, stereotip (juda ko'p kuzatiladigan, tipik) vaziyatlarni gavdalandirishning bir usuli sanaladi. Ya'ni freymga modellshtirilayotgan hodisa, vaziyat va jarayonlarni umumlashtirilgan holda joylashtirish mumkin. Freym — mavhum obraz modeli bo'lib, muayyan obyekt, voqea-hodisa, vaziyat va jarayonlarning formal tavsifi demakdir (masalan, restoranga borish, shifokor qabulida bo'lish, ma'ruzlarga qatnashish, tug'ilgan kun, futbol o'yini kabi tipik vaziyatlar). Dastlab “freym” termini XX asrning 70-yillarida Marvin Minskiy tomonidan olamning kompyuter modelini o'z ichiga oluvchi bilimlar strukturasi nisbatan ishlatilgan.⁸⁹ Shundan so'ng freym semantikasi konsepsiyasi (**frame semantics**) amerikalik tilshunos Charlz Fillmor tomonidan rivojlantirildi. Mazkur konsepsiyaga ko'ra so'zlar, so'z birikmalari, gaplar, matn ma'nosi **sahnalar** bilan uyg'unlashtirilgan, shartlangan bo'ladi. Bir freym boshqa bir freymning davomi yoki tarkibiy qismi bo'lishi mumkin. Ch.Fillmorning qarashlariga L.Tenyening aktantlar (fe'lining bo'sh o'rinlarini to'ldiruvchilar) haqidagi g'oyasi katta ta'sir ko'rsatgan. Fillmorning **Case (keishiklar)** g'oyasiga asosan aktantlar (fe'lining hamrohlari) fe'l yordamida ifodalanadigan holat yoki harakat munosabatlaridagi bajariladigan “rollar” sifatida talqin qilinadi. Ch.Fillmor ajratgan kelishiklar quyidagilar:

1. **Agent (A)** – harakat bajaruvchisi (inglizcha “agent” so'zidan olingan).
2. **Tema (T)** – predmet (inglizcha “theme” so'zidan olingan)
3. **Manba (S)** – harakat obyektining dastlabki holati (inglizcha “source” so'zidan olingan).
4. **Maqsad (G)** – harakat ijrosidan keyingi yakunniy holati (inglizcha “goal” so'zidan olingan).

⁸⁹ **Minsky M.** A Framework for Representing Knowledge / The Psychology of Computer Vision. McGraw-Hill, New York (U.S.A.), 1975.

5. **Vosita (I)** – harakat uchun qurol-vosita (inglizcha “instrument” so’zidan olingan).

6. **Usul (W)** – harakat usuli (inglizcha “way” so’zidan olingan).

Freymlar sun’iy intellekt tarkibiga kiruvchi ekspert tizimlarida faol ishlatiladi. Freymning quyidagi turlari mavjud: freym-namunalar, freym-strukturalar, freym-rollar, freym-senariylar, freym-vaziyatlar. Freymlar tizimi birlashib **semantik tarmoq**ni hosil qiladi. Freym kompyuter xotirasida bilimlar xazinasini va faktlar omboriga tayanib olamni maxsus shakllantiruvchi model vazifasini bajaradi. Freym **slot** deb atalgan bir yoki bir necha elementlarning majmui hisoblanadi. Slotlarda obyektlarga xos belgilar ma’nosi qayd etilgan bo’ladi.

17-jadval

“Stol” freym fragmentining jadval ko’rinishidagi reprezentatsiyasi

Slot nomi	Slot mazmuni
Oyoqchalar miqdori	To’rtta, undan ko’proq bo’lishi ham mumkin
Materiali	Yog’och, plastmassa, shisha
Yuzasi, sirti	Oval, kvadrat, aylana
Tumbaning (tagkursining) mavjudligi	Ixtiyoriy tarzda
Vazifalari	Ovqatlanish uchun, dars tayyorlash uchun, ish uchun va hokazo
va boshqalar	

Freym-senariy ba’zi harakat, voqea-hodisalarning harakterli unsurlarini o’z ichiga olgan tipik strukturalar hamda vaziyatlarning rivojlanishi, ketma-ketligi (algoritm bosqichlari) demakdir. Senariy bu oddiy voqea-hodisalar zanjiri emas, balki ushbu harakatlar ketma-ketligida kauzal bog’liqlik tamoyili ishlaydi, ya’ni har bir harakat natijasi shartli sanaladi, uning ta’siri va oqibatida navbatdagi ish-harakat sodir bo’lishi mumkin. Har bir senariyda rol bajaruvchilar mavjud. Senariylar real hayotda uchraydigan vaziyatlarni aks ettiruvchi holatlar asosida yuzaga keladi. Masalan, restoranga tashrif buyurish senariysini olib qaraylik.

Senariy: restoran

Rollar: tashrifchi, ofitsiantka, xo’jayn, kassir.

Maqsad: Qorinni to’yg’izish uchun ovqat olish.

Sahna I. Kirish

Restoranga kirildi

Bo’sh o’rindiqlar bor joyga nazar tashlanadi

O’tirish uchun joy tanlanadi

Tanlangan joyga boriladi

O’tiriladi

Sahna II. Buyurtma

Taomnoma (menyu) olinadi

Menyu o'qiladi

Nimani buyurtma qilish haqida kelishiladi

Ofitsiantkadan tanlangan taomni keltirish so'raladi

Sahna III. Taom

Ovqat olinadi

Ovqat yeyiladi

Sahna IV. Chiqish.

Hisob so'raladi

Chek olinadi

Kassirning oldiga boriladi

Pul to'lanadi

Restorandan chiqiladi

Ko'pchilik mutaxassislarning fikricha, ekspert tizimlari va neyron tarmoqlari sun'iy intellekt bo'yicha izlanishlarda o'zaro musobaqalashuvchi ikki muqobil yo'nalishlardir. Har ikkala yo'nalishning o'ziga xos afzalliklari va kamchilik tomonlari mavjud. Shuni e'tiborga olgan holda aytish mumkinki, neyron tarmoqlari va ekspert tizimlarini birgalikda ishlashini ta'minlovchi gibrid tizimlar ishlab chiqilishi lozim. Shunda qo'yilayotgan masalaning xarakteridan kelib chiqqan holda mos tizim tanlash va muammoga optimal yechim topish imkoniyati yuzaga keladi. Agar masala "qoyida bo'yicha" (implikatsiya – if ..., then... - agar...-sa, u holda...) hal etilsa, ekspert tizimlariga murojaat qilsih zarur. Agar masalada ko'pgina empirik ma'lumotlar mavjud bo'lsa, neyron tarmoqlaridan foydalanish maqsadga muvofiq. Yoki obrazlarni tanishda (masalan, moliyaviy holatni aniqlashda) asosan neyron tarmoqlaridan, obraz aniqlangandan keyin bajariladigan mantiqiy amallarda esa ekspert tizimlaridan foydalanilsa, yuqori samaradorlikka erishish mumkin.

Topshiriqlar

1. Kompyuter lingvistikasi fanining zamonaviy yo'nalishlari haqida Internet ma'lumotlarini toping.
2. Lingvistik kompyuter dasturlarning ishlash samaradorligini oshirish uchun nimalar qilish lozim?
3. Ekspert tizimlari qaysi sohalarda ishlatilmoqda?
4. Sun'iy intellekt tizimi bo'yicha erishilgan yutqlarga nimalarni kiritish mumkin?
5. Sun'iy intellekt tizimini yaratish yuzasidan qanday yondashuvlar mavjud?

Adabiyotlar

1. Нелюбин Л. Л. Компьютерная лингвистика и машинный перевод. – М.: ВЦП., 1991.
2. Шемакин Ю.И. Начало компьютерной лингвистики. – М.: МГОУ. 1992.
3. Белоногов Г.Г. Компьютерная лингвистика и перспективные информационные технологии. - М.: Русский мир, 2004.
4. Рассел С., Норвиг П. Искусственный интеллект: современный подход /Artificial Intelligence: a Modern Approach / Пер. с англ. и ред. К.А. Птицына. 2-е изд. - М.: Вильямс, 2006.
5. Searle J.R. Minds, brains, and programs / Behavioral and Brain Sciences, 1980.
6. Jurafsky D., Martin J.H. Speech and language processing. - New Jersey, 2000.
7. Minsky M. A Framework for Representing Knowledge / The Psychology of Computer Vision. McGraw-Hill, New York (U.S.A.), 1975.
8. Анисимов А.В. Компьютерная лингвистика для всех: Мифы. Алгоритмы. Язык. – Киев: Наукова думка, 1991.
9. Po'latov A., Muhamedova S. Kompyuter lingvistikasi. – Toshkent, 2007.

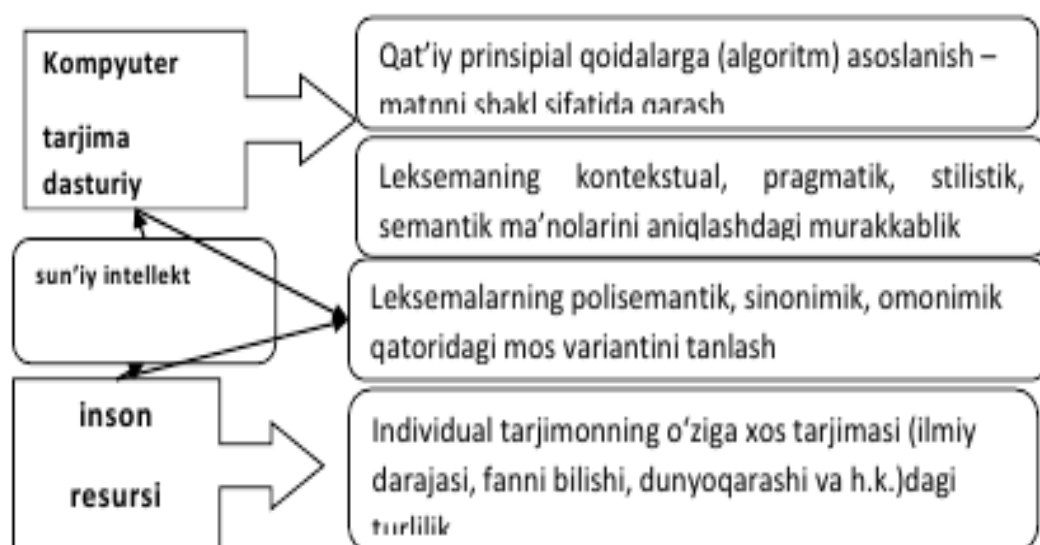
Ma'ruza № 17-18

Kompyuterda matnlarni qayta ishlash tizimlari

REJA:

1. Matnning bugungi kundagi ko'rinishlari
2. Elektron matnlar bazasi
3. Matnlarni qayta ishlashga mo'ljallangan dasturlar
4. XXI asrda yaratilgan matn tiplari

Mashina tarjimasining tadqiqot yo‘li yarim asrdan oshgan bo‘lsa-da, ba’zi bir muammolarni yechishda hanuz ojizlik qilmoqda. **Til->vaziyat->nutq** tabiiy tilni qayta ishlashda ba’zi muammolarni keltirib chiqargan.



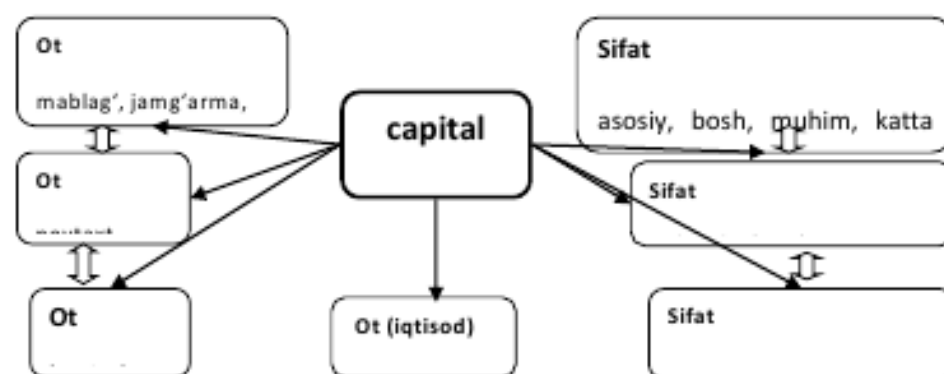
Faylasuf Bar Xillel quyidagilarni mashina tarjimasining dastlabki rivojlanish bosqichlaridagi asosiy muammolar sifatida ko‘rsatgan⁹²: 1) operasion sintaksis (Operational Syntax); 2) tabiiy tillarning ichki tarjimasi (Intertranslatability of natural languages); 3) idiomalar (Idioms); 4) universal sintaktik kategoriyalar (Universal syntactic categories). Quyidagilarni mashina tarjimasidagi umumiy muammolar sifatida ko‘rsatish mumkin:

⁹² <http://www.mt-archive.info/Bar-Hillel-1953.pdf>/[From: *Philosophy of Science*. Vol.20 (1953). – P.217 – 225]



Ma'lumki, ingliz va o'zbek tillari tipologik jihatdan boshqa oilalarga mansubligi bois, qator muammolarni yuzaga keltiradi: 1) inglizcha matnlardagi leksemalarning polisemik va omonimik xususiyatga egaligi; 2) affikslarning turli funksionalligi; 3) ba'zi bir frazemalarning lingvokulturologik kasb etishi; 4) matn tarkibidagi leksemalarning sinonimik qatordan tarkib topishi va ularning uslubiy tomondan bog'liqligi; 5) ilmiy matnlardagi terminlar masalasi va h.k.

Tanlab olingan va cheklangan leksemalarning lingvistik modulini berishda mana shu vazifalardan kelib chiqish maqsadga muvofiq. *Capital* so'zi ham boshqa so'z turkumini, ham omonimiyani hosil qilgan:



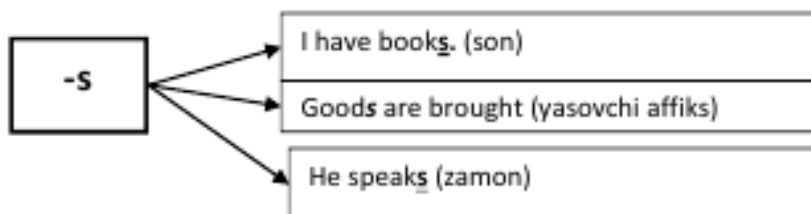
Birdetal qayd etishicha(1987), Vebster muallifligidagi «Ingliz tilining izohli lug'ati» (7-nashri)da 60000 so'zning 21488 tasi, ya'ni umumiy miqdorning 40% ni ikki va undan ortiq ma'noga ega polisemantik so'zlar tashkil etadi. Masalan, birgina *run* – *yugurmoq* so'zi 29 ma'noga ega bo'lib, 125 ta ichki semantik maydonlarga ega.⁹³ Leksikografiyada lug'at tuzuvchilarning polisemiyaga bo'lgan qarashlari, yondashuvlari turlicha bo'lgani bois omonimiya hodisasi bilan

⁹³ Yael Ravin, Claudia Leacock Polysemy. Theoretical and computational approaches. Oxford University press. 2000. – P.13.

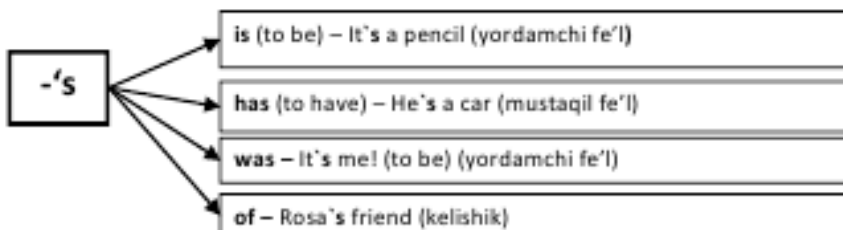
polisemiya ni farqlashda ayrim chalkashliklar kuzatiladi. Ingliz tilida polisemiya hodisasi nisbatan yaxshi o'rganilgan. Biroq avtomatik tarjima uchun tanlangan va chegaralab olingan so'zlarning to'liq lingvistik ta'minotida ularning barchasi o'rin olishi hamda o'zbek tilidagi tarjimasida xususiyatlari ochib berilishi kerak.

XX asrning 60-yillarida Bar Xillel namuna sifatida *pen* so'zining «*ruchka*» va «*o'yinchoq qutisi*» ma'nolari ega bo'lishiga tayanib, *mashina tarjimasini 80 % xato* degan fikrni bildirgan. Endilikda olib borilgan izlanishlar natijasida matnni avtomatik tarjima qilishning 80% muammosi yechilgan bo'lib, qolgan qismi mutaxassislar tomonidan o'rganilmoqda. Mazkur tadqiqotlar sirasida Lesk (1986), Gutri (1991), Vilks (1993), Blek (1988), Zernik (1991), Xerst (1991), Gale (1992), Miller, Lekok (1998), Vurxis (1998), Adam Kilgarri (1997) va boshqa olimlarning ishlarini eslatish o'rinli⁹⁴.

Dastlab affiksali omonimiyani tahlil qilamiz:



Shuningdek, -s suffiksining 's shakli ham bo'lib, ular o'zaro farqlanadi:



Navbatdagi suffiks:

⁹⁴ O'sha joyda.

Biror fe'l bilan birikib kelgan predlog so'zning mazmuniga ta'sir qiladi, masalan, *for* predlogi bir qancha ma'no ifodalaydi: 1) *-ga, uchun (atalganlik)*=> *This letter is for you (Bu xat sizga yoki siz uchun)*; 2) *uchun, tarafdor*=> *I am for your proposal (Men sizning taklifingiz tarafdoriman)*; 3) *uchun (sabab)*=> *He was rewarded for his bravery (U jasurligi uchun mukofatlandi)*; 4) *davomida (ish-harakatning davom etish muddatini ifodalaydi)*=> *I have been to Andijon for two weeks (Men Andijonda uch hafta bo'ldim)*; 5) *for predlogi bilan keluvchi turg'un birikmalar*=> *to ask for (so'ramoq), a contract for (uchun shartnoma)*; 6) *iboralar*=> *for ever (abadiy, mangu)*.

Yu.V.Rojdestvenskiy mashina tarjimasiga sun'iy intellektning eng asosiy muammo⁹⁶ sifatida o'rinli baho beradi. Chunki tabiiy tilning imkoniyatlari lisoniy va nolisoniy omillar ta'sirida voqelanadi. Xullas, mashina tarjimasida mukammal ishlangan lingvistik ta'minot dasturning sifatini belgilaydi. Shuningdek, matnni tahlil etishda uning grammatik xususiyatlarini to'liqroq o'rganilishi lozim. Inglizcha-o'zbekcha mashina tarjimasini yaratishda ikki tilli lug'atlardagi so'zlarning koefitsentlarini belgilashda ularning kontekstologik xususiyatlari, albatta, inobatga olinishi zarur.

⁹⁶ Рождественский Ю.В., Волков А.А., Марчук Ю.Н. Введение в прикладную филологию. – М.: МГУ, 1988. – С. 116.

2. Mustaqil ta'lim tashkil etishning shakli va mazmuni.

“Kompyuter lingvistikasi” bo'yicha talabaning mustaqil ta'limi shu fanni o'rganish jarayonining tarkibiy qismi bo'lib, uslubiy va axborot resurslari bilan to'la ta'minlangan.

Talabalar auditoriya mashg'ulotlarida professor-o'qituvchilarning ma'ruzasini tinglaydilar, misollar topadilar. Auditoriyadan tashqarida talaba darslarga tayyorlanadi, adabiyotlarni konspekt qiladi, uy vazifa sifatida berilgan topshiriqlarni bajaradi. Bundan tashqari ayrim mavzularni kengroq o'rganish maqsadida qo'shimcha adabiyotlarni o'qib referatlar tayyorlaydi xamda mavzu bo'yicha testlar tuzadi. Mustaqil ta'lim natijalari reyting tizimi asosida baxolanadi.

Uyga vazifalarni bajarish, qo'shimcha darslik va adabiyotlardan yangi bilimlarni mustaqil o'rganish, kerakli ma'lumotlarni izlash va ularni topish yullarini aniqlash, internet tarmokdaridan foydalanib ma'lumotlar tuplash va ilmiy izlanishlar olib borish, ilmiy tugarak doirasida yoki mustaqil ravishda ilmiy manbalardan foydalanib ilmiy maqola va ma'ruzalar tayyorlash kabilar talabalarning darsda olgan bilimlarini chuqurlashtiradi, ularning mustaqil fikrlash va ijodiy qobiliyatini rivojlantiradi. Shuning uchun xam mustaqil ta'limsiz o'quv faoliyati samarali bo'lishi mumkin emas.

Uy vazifalarini tekshirish va baholash amaliy mashg'ulot olib boruvchi o'qituvchi tomonidan, konspektlarni va mavzuni o'zlashtirish darajasini

tekshirish va baxolash esa ma'ruza darslarini olib boruvchi o'qituvchi tomonidan xar darsda amalga oshiriladi.

“Kompyuter lingvistikasi” fanidan mustaqil ish majmuasi fanning barcha mavzularini kamrab olgan va quyidagi 6 ta katta mavzu kurinishida shakllantirilgan.

Talabalar mustaqil ta'limining mazmuni va hajmi

	Mustakil ta'lim mavzulari	Berilgan topshiriklar	Bajarish muddati	Xajmi (soatda)
	Kompyuter lingvistikasining asosiy yo'nalishlari	Adabiyotlardan konspekt kilish. Masalalar echish. Mustakil topshiriklarni bajarish.	4.5.6 - xaftalar	10
	Kompyuter lingvistikasi fanining rivojlanish istiqbollari	Adabiyotlardan konspekt kilish. Individual topshiriklarni bajarish	7.8.9 - xaftalar	10
	Tilshunoslikning zamonaviy yo'nalishlari va kompyuter lingvistikasi.	Adabiyotlardan konspekt kilish. Individual topshiriklarni bajarish	16,17,18- xaftalari	8
Jami				28

3. GLOSSARIY

Aksiomatik nazariya - ilmiy nazariyaning shunday qurilishiiki, unda bu nazariya asosida shunday boshlang'ich postulat va aksiomalar qo'yiladiki, qolgan barcha qoidalar (teoremlar) undan kelib chiqadi. U miloddan avvalgi VII-VI asrlarda shakllangan, asoschisi matematik olim Pifagordir. Aksiomatik nazariyaning quyidagi asosiy tamoyillari mavjud:

- ziddiyatsizlik – aksiomalar tizimining shunday xossasiki, unga ko'ra, sistemaning hech qanday ikkita holati bir-biriga zid kelmasligi kerak. Bunda aksiomalar tizimining o'zida rost va yolg'on orasidagi farqni aks ettirishi kerakligi nazarda tutiladi.

- to'lalilik – aksiomalar sistemasiga qandaydir jumlaning qo'shganimizda, sistemada ziddiyat sodir bo'ladi.

- bog'liqsizlik – berilgan aksiomalar sistemasining u yoki bu sistemalari shu sistemaning boshqa sistemalari yordamida isbot qilinishi mumkin emas.

Algoritm - (lotincha *algorithmus* – “ibtido” ma'nosini anglatadi; bu so'zning etimologiyasi buyuk alloma Al-Xorazmiy ismi bilan bog'liq. U Yevropada buzilgan shaklda “Algorithmi” tarzida ishlatilgan) ma'lum bir tipga oid hamma masalalarni yechishda ishlatiladigan jarayonlar tizimining muayyan tartibda bajarilishi haqidagi aniq qoida yoki qoidalar tizimi.

Algoritmik tillar programmashtirishga asoslangan sun'iy tillar bo'lib, undagi matnlar xuddi tabiiy tildagidek aniq interpretatorga – insonga yoki avtomatik qurilmaga yo'nalgan bo'lishi lozim, toki u axborotni qabul qilsin va tushunsin. Agar bu jarayon (qabul qilish va tushunish) amalga oshmasa, u holda algoritmik tillar va kompyuter orasida vositachi sifatida **translyator-dasturlar** (yoki **kompilyatorlar**)dan foydalaniladi. Algoritmik tillar asosan **imperativ** xarakterga ega bo'ladi, ya'ni asosan buyruqlar tizimiga tayanadi (begin - boshla, end – tugalla kabi). Dastlabki shunday til 1958 yilda Parijda yaratilgan va “**Algol**” deb atalgan (Uning to'liq nomi inglizcha – “Algorithmic language”). Hozir jahonda 500 dan ortiq algoritmik tillar mavjud.

Analiz (tabiiy tilning kompyuterda morfologik, sintaktik va semantik tahlil yordamida tushunilishi) va **sintez** (kompyuterda matnning grammatik shakllantirilishi va generatsiyasi (hosil qilinishi) – kompyuter lingvistikasining tayanch tushunchalari. Morfologik analiz morfemalarga ajratish hamda grammatik ma'lumotlar tahlilini o'z ichiga oladi. Morfologik sintez esa matnga avtomatik ishlov berish jarayonida yoki uning so'nggi bosqichida (natijalarni inson hukmiga havola qilishda) so'zlarning turli shakllarini chiqarish, analiz natijalarini umumlashtirishdan iborat. Morfologik analiz va sintez jarayoni kompyuterda kodlash-dekodlash tamoyili asosida ishlaydi. Morfologik analizning mantiqiy davomi sintaktik analiz hisoblanadi. Sintaktik analiz morfologik analiz ustiga quriladi, bunda gap strukturasi turli formal modellar asosida tahlil etiladi.

Avtomatik tarjima (mashina tarjimasi) – bir tildagi matnni ikkinchi bir tilga EHM (kompyuter) vositasida, tez vaqt ichida tarjima qilish tizimi. Mashina tarjimasi bo'yicha birinchi xalqaro konferensiya 1952 yilda Massachusets texnologiya institutida o'tkazildi. 1954-yilda AQShdagi Jorjtaun Universitetida dunyoda birinchi marta avtomatik tarjima tajribadan o'tkazildi. U **GAT** tizimi (inglizcha “*Georgetown Automatic Translation*” birikmasining qisqartmasi) deb ataladi. Bu tajribalar IBM-701 kompyuterida o'tkazildi. Bu tajribaning bazasi

algoritmalaridan iborat bo'lib, 50000 ta so'z (60 ta gap)ni rus tilidan ingliz tiliga tarjima qildi.

SYSTRAN avtomatik tarjima tizimi yordamida bir yilda 2,5 mln. sahifa matn ingliz tilidan nemis, fransuz, ispan, yunon, italyan tillariga hamda fransuz tilidan ingliz, ispan, italyan tillariga tarjima qilinadi. **DIANA** avtomatik tarjima tizimi 10 million ta so'z bazasiga asoslanib dunyoning 6 ta yirik tilida avtomatik tarjimani ta'minlaydi. 40 mln.ta so'z xotira kuchiga ega **YEUROTRA** sitemasi barcha Yevropa tillari bo'yicha avtomatik tarjimani amalga oshira oladi. **SOKRAT** sistemasi ham 40 mln. ta so'z xotira kuchiga ega bo'lib, dunyodagi 40 ta tilda avtomatik tarjimani amalga oshiradi.

Internet tizimining rivojlanishi natijasida avtomatik tarjima global tarmoqda ham tashkil etildi, ya'ni hozirgi kunda **online-tarjima tizimi** ham faol yo'lga qo'yilgan.

Avtomatik tahrir - kompyuter lingvistikasining asosiy yo'nalishlaridan biri bo'lib, u kompyuterdagi matn muharrirlari (**Microsoft Word, Excel, Wordpad, Lexicon**) rivojlanishi bilan bog'liqdir. Avtomatik tahrir matndagi mexanik xatolarni avtomatik tarzda to'g'rilanishi va xato ekanligi haqidagi signallarning userga taklif etilishi kabilar bilan bog'liq. Buni aniq misollar bilan tushuntiriladigan bo'lsa, Microsoft Word dasturi bilan ishlar ekanmiz, unda so'zlarning yozilishi bilan bog'liq xatolar bo'lsa, ularning ostiga qizil yoki ko'k rangda ostiga chizilishi, katta harf bilan yozilishi kerak bo'lgan so'zlarning avtomatik bosh harfga o'tkazilishi, hujjatga avtomatik numeratsiya qo'yilishi kabilar avtomatik tahrir erishgan yutuqlar hisoblanadi.

Avtomatik tahrir yo'nalishi XX asrning o'rtalariga kelib rivojlandi. U matn muharrir dasturlari bilan birgalikda yangicha imkoniyatlar bilan, qo'shimcha effektlar bilan rivojlanmoqda. Shu o'rinda alohida ta'kidlash zarurki, avtomatik tahrirlash tizimi faqat matn muharrirlari dasturlari bazasiga qo'shilgan tillarda amal qiladi, masalan, ingliz, rus, nemis, fransuz tillarida.

Avtomatik tahrirlash tizimi juda ko'p amaliy vazifalarni o'z ichiga oladi. Bu vazifalarning eng asosiylarini umumlashtirib sanab o'tmoqchimiz:

- orfografiyani tekshirish;
- grammatika va stilistikani tekshirish;
- kontekstli qidiruv;
- formatlash;
- avtomatik referatlash – matndagi kalit so'zlar asosida uning umumiy mundarijasini yaratish;
- matn ichiga jadval, rasm, turli figuralar, klip-artlar, grafik sxemalar, diagrammalar qo'shish, giperhavolalar bilan ta'minlash;
- matndagi abzatslar, belgilar, so'zlarning aniq miqdorini aniqlash (statistikasini ko'rsatish);
- tezaurus xizmati – matndagi muayyan so'zning sinonimik variantlarini aniqlash;

- transliteratsiyalash – bir alifbodan boshqa alifboga o'tkazish (masalan, kirill alifbosidan lotin alifbosidagi harfga almashtirish).

Ms Word dasturida orfografiyani tekshirish imkoniyati cheklanganligi bois hozirgi kunda maxsus Spell Checker dasturlar yaratilmoqda. Bunday dasturlar sirasiga quyidagilarni kiritish mumkin: SpellLink Multilingual Pro (ingliz, ispan, italyan, nemis, fransuz, portugal, ukrain tillaridagi so'zlarning to'g'ri yozilishini tekshiradi), Orfo SpellLink English, SpellLink French, SpellLink Italian, SpellLink Spanish, Ruta 1.0 (rus va ukrain tillaridagi so'zlarning imlosini tekshirishga yordam beradi), Asutype, Spellex, Spell Catcher kabi.

Bevosita ishtirokchilar (immediate constituents) sintaktik modeli hisoblanadi. Bevosita ishtirokchilar metodiga muvofiq tayanch nuqta konstruksiya hisoblanadi. Konstruksiya tarkibidan ishtirokchilar va bevosita ishtirokchilar ajratiladi.⁹⁰ Konstruksiya atamasi ostida ma'noli qismlarning ketma-ket munosabatidan tashkil topgan butunlik tushuniladi. Ishtirokchilar deb kattaroq konstruksiya tarkibiga kirgan so'z yoki konstruksiyaga aytiladi. Muayyan konstruksiyaning bevosita shakllanishida ishtirok etgan bir yoki bir nechta ishtirokchilarga bevosita ishtirokchilar deyiladi. Masalan, *A'lochi talabalar o'z vazifalarini aniq va puxta bajaradilar*. Bu jumla konstruksiya hisoblanadi, uning tarkibidagi barcha ma'noli so'zlar ("va" bog'lovchisidan tashqari) ishtirokchilar, bevosita aloqaga kirishgan so'zlar bevosita ishtirokchilar deyiladi. Bunda *a'lochi* so'zi *talabalar* so'zi bilan, *aniq*, *puxta* so'zlari *bajaradilar* so'zi bilan, *o'z* so'zi *vazifalarini* so'zi bilan o'zaro zich bog'langandir. Demak, bevosita ishtirokchilar ssintaktik modelida dastlab so'z birikmalari ajratiladi, so'ngra ular gap ichida birlashtiriladi.

Bilimlar injenerligi (Knowledge Engineering) - bilimlar bazasi bilan shug'ullanuvchi maxsus yo'nalish bo'lib, u sun'iy intellekt tizimining tarkibiy qismi sanaladi. Bilimlar injenerligi kognitiv lingvistika, kompyuter semantikasi va kompyuter lingvistikasining yutuqlariga tayanadi. Bunda informatsiya va bilimlar strukturasiga oid bo'lgan **semantik tarmoq**, **freym**, **senariy**, **Word-net** kabi tushunchalar va ularning tatbiqi muhim hisoblanadi.

g) – til o'qitishni avtomatlashtirish - xorijiy tillarning kompyuter yordamida o'qitilish tizimini ta'minlovchi amaliy dasturlar hamda bilimlarni avtomatik baholash mexanizmlari. Til o'rgatishda kompyuterdan foydalanish bo'yicha dastlabki tajribalar XX asrning 60-yillarida Stenford (Russian-Program) va Nyu York universitetlarida (Das Deutche Program) o'tkazildi. XX asrning 70-80-yillarida til o'qitishda kompyuterdan foydalanishga bo'lgan qiziqish kuchaydi va hatto alohida yo'nalish - **Computer Assisted Language Learning (CALL)**ning yuzaga kelishiga olib keldi. CALL bo'yicha 1-xalqaro konferensiya 1985 yil 12-13 noyabrda Budapesht shahrida o'tkazildi.

O'qituvchi lingvistik avtomatlardan foydalanishda 2 xil yondashuv mavjud:

⁹⁰ **Nurmonov A.** Lingvistik tadqiqot metodologiyasi va metodlari. – Toshkent: Akademnashr, 2010. – B. 88.

1) bixevioristik yondashuv – mavjud an’anaviy o’qitish tizimining barcha jabhalarini e’tiborga oluvchi o’qitish mexanizmlari, bunda grammatikaning ma’lum bo’limi tugagandan so’ng, uni topshiriqlar asosida baholash, shuningdek, leksik minimumlar bo’yicha bilimni nazorat qilishga yo’naltirilgan dasturiy ta’minotga asoslangan o’qitish tizimi. Bunda bixeviorizmning **“stimul-reaksiya-mustahkamlash”** formulasiga asoslaniladi. O’qitish usullari “stimul” (turtki) vazifasini bajaradi, o’rganuvchi bundan ta’sirlanadi va bilimni mustahkamlaydi.

2) kognitiv-intellektual yondashuv – CALL (CALL Software) uchun universal dasturiy ta’minotlar yaratish, o’qitish tizimini lug’at bazasi, grammatika bo’yicha qo’llanmalar, tizimni **avtomatik korrektorlar, speller** bilan ta’minlash, **audio** va **vizual effektlar** bilan boyitish kabi amaliy ishlar majmuidir.

Ekspert tizimi – muayyan sohadagi ekspertlik yoki mavjud bilimlarning imitatsiyasi tarzida yaratilgan kompyuter dasturi. Boshqacharoq aytganda ekspert tizimi maxsus sohaga oid barcha bilimlar jamlangan qomus yoki o’ta professional ekspertlarning kompyuter modelidir. Ekspert tizimlarining asosiy ikkita tarkibiy qismi mavjud: ma’lumotlar ombori va mantiqiy mashina. Ekspert tizimlari konsepsiyasi Stenford universiteti professori Eduard Feygenbaum nomi bilan bog’liq bo’lib, u 1977 yilda murakkab mantiqiy masalalarni yechishda kompyuter dasturlarining samaradorligi formallashtirish va dasturlash texnikasidan ko’ra muammoli sohaga doir bilimlar bazasiga bog’liq ekanligini asoslab berdi.

Elektron ensiklopediya – kuchli qidiruv tizimiga ega bo’lgan ma’lum yo’nalishdagi illyustrasiyalı maqolalar, elektron xaritalar, fotosuratlar, giperhavola(gipermurojaat)lar, audio va videoizohlar, xronologik ro’yxat, foydalanilgan va tavsiya etiladigan adabiyotlar ro’yxati va hokazolardan iborat bo’lgan katta hajmdagi elektron dastur.

Elektron lug’atlar - leksikografiya sohasiga kompyuter texnologiyalarining tatbiq etilishi natijasida yaratilgan lug’atlar, ular odatdagi lug’atlar asosida tuzilgan, multimedia vositalari bilan boyitilgan dastur. Eng mashhur elektron lug’atlar sirasiga CONTEXT, ABBY LINGVO, MULTITRAN, POLYGLOSSUM, MULTILEKS kabilar kiradi.

Elektron lug’atlarning tayyorlanishi quyidagicha: a) lug’atning so’zligi shakllantiriladi; b) misollar korpusi (elektron kartotekalari) tuziladi; c) lug’at maqolalari yoziladi; d) lug’at maqolalarini ma’lumotlar bazasiga ko’chiriladi; e) bevosita ma’lumotlar bazasida lug’at matni tahrir etiladi, korrektirovka qilinadi; f) lingvistik ta’minot dasturiy ta’minot bilan uyg’unlashtiriladi; g) elektron lug’at.

Elektron lug’atlar an’anaviy kitobiy lug’atlarga qaraganda bir muncha afzalliklarga ega. Elektron lug’atlarning hajm jihatidan ixchamligi, kompaktligi, boshqa manbalarga (internet tarmog’i orqali, gipermurojaatlar yordamida) ulanish imkoniyati, mavjud matnlar korpuslari yordamida illustrativ misollar tuza olish imkoniyati, multimediali misollar bilan boyitilganligi, so’zlarning sinonimik variantlari, omonimlik xususiyatlari, grammatik ma’lumotlar bilan

ta'minlanganligi va ularni tezlik bilan foydalanuvchiga havola etilishi kabi jihatlari elektron lug'atlarning optimalligini ta'minlaydi.

Elektron multimedia darsligi – aynan kompyuter vositasida o'quv kursini yoki uning bitta katta bo'limini o'qituvchi yordamida yoki mustaqil o'zlashtirishni ta'minlovchi dasturiy-metodik kompleks bo'lib, bunda tizim matnli fayllar, turli xil animatsiyalar, audio va videoeffektlar bilan boyitilgan bo'ladi.

Example Based Translation (“Namunalarga asoslanadigan tarjima”) -1984 yilda Yaponiyaning Kioto universiteti professori, mashina tarjimasi bo'yicha yapon davlat dasturi rahbari Makoto Nagao tomonidan yaratilgan yangi konsepsiya. Ushbu konsepsiyaga ko'ra matnlar ilgari inson qo'li bilan tarjima qilingan matnlarga tayanib analogiya tamoyili asosida tarjima qilinishi lozim. Buning uchun kompyuterga kiritilgan katta hajmdagi **bilingv bazasi** (matn va uning tarjimalari) shakllantirilishi zarur. Yangi matnlar tarjimasida bilingv massividan o'xshash matn fragmentlari tanlanadi va tarjimaning muvaffaqiyatli chiqishida foydalaniladi. M.Nagao mashina tarjimasiga bunday yondashuvni **“Example Based Translation”** (“Namunalarga asoslanadigan tarjima”) deb nomladi, ilgari an'anaviy yondashuv **“Rule Based Translation”** (“Qoidalariga asoslanadigan tarjima”) deb ataladi.

Formallashtirish – bilimlar majmuasini, mazmunini muayyan shakllar yordamida sun'iy til belgilari orqali ifodalashdir. Grammatikani formallashtirish tendensiyasi XX asrning o'rtalarida kuchaydi. Bu bevosita tilning strukturasini yanada aniq parametrlarda ifodalashga intilish, uni kompyuterga moslashtirishga urinish bilan bog'liqdir. Formal grammatika nazariyasi matematik mantiq, matematik linvistikaning rivojlanishi natijasida yuzaga kelgan. Grammatikani formallashtirish (lotincha forma – “shakl”, “tashqi ko'rinish”) konsepsiyasining asoschisi amerikalik tilshunos Noam Chomskydir. Formal grammatika abstrakt simvollar orqali bayon etiladigan grammatika hisoblanib, unda til qoidalari, so'z yasalishi, birikma va gap qurilishi turli xil lingvistik modellar, struktur sxemalar yordamida tavsiflanadi. Formal grammatika 3 ta birlik asosida ish ko'radi:

1. Word-classes – so'z turkumlari (tub va yasama so'zlar).
2. Phrases – gap bo'laklari.
3. Sentences – gap konstruksiyalari sxemalari, modellari.

Formal grammatika turlari - N.Chomsky qoidalarining ifodalanish turiga qarab formal grammatikani 4 tipga ajratgan va bu **Chomsky ierarxiyasi** deb nomlanadi:

1. **0 grammatika (unrestricted phrase-structure grammar)** – juda murakkab generativ grammatika, bunda $G=(N, \Sigma, P, S)$ belgilari yordamida ish ko'riladi. N , Σ - alifbo (N – noterminal simvollar, u grammatikadagi simvollarini o'z ichiga oladi; Σ - metatilning terminal simvollar, u tildagi so'zlarni qamrab oladi); S – noterminal to'plamning boshlang'ich simvoli, P – qayta ishlab chiqilgan qoidalar. Bunda terminal va noterminal simvollarini birlashtiruvchi V belgisi qabul qilingan, u inglizcha Vocabulary (lug'at) so'zidan olingan ($V=T \cup N$). Bu

grammatika o'ta abstraktlashgan matematik modellarga, informatika nazariyasiga tatbiq etilgan.

2. **1 grammatika (context-sensitive grammar)** – kontekstga bog'liq grammatika bo'lib, bunda simvollar zanjiri kontekst bilan aniqlanishi mumkin. Bu grammatika tabiiy tillar unsurlarining generatsiyasida faol qo'llaniladi.

3. **2 grammatika (context-free grammar)** – kontekstga bog'lanmagan grammatika bo'lib, u programmalashtirish tillari unsurlarining generatsiyasida (ifodalar, buyruqlar tizimida) ishlatiladi.

4. **3 grammatika (regular grammar)** – regulyar grammatika deb ataladi, u juda sodda va cheklangan grammatika bo'lib, tilning sodda unsurlari uchun ishlatiladi (miqdor, konstantalar, o'zgaruvchilar uchun).

Frazeologik mashina tarjimasi – avtomatik tarjimaning sifatini oshirish uchun Rossiyada professor G.G.Belonogov tomonidan ishlab chiqilgan konsepsiya. Bunga ko'ra matndagi tushunchalar ko'pincha alohida so'zlar yordamida emas, balki so'z birikmalari yordamida ifodalanadi. O'z-o'zidan kelib chiqadiki, matnni bir tildan boshqasiga tarjima qilish jarayonida ham mazmunni yuzaga chiqaruvchi asosiy til birliklari ham alohida so'zlar emas, balki tipik vaziyatlarni ifodalovchi frazeologik birikmalar (bu yerda frazeologik birlik keng ma'noda tushuniladi) hisoblanadi. Mazkur yondashuvga muvofiq matnda tez-tez uchraydigan so'zlar, iboralar, so'z birikmalari, nutqiy etiket namunalari tarjima ekvivalentlari bilan boyitilgan ma'lumotlar bazasi frazeologik mashina tarjimasi tizimining asosini tashkil etadi. Bunda so'zlarning distributsiyasiga eng ko'p e'tibor beriladi. Bu metod polisemantik so'zlarning tarjimada mos keluvchi to'g'ri ekvivalentini tanlashga, omonim so'zlarning kontekstagi holatini aniqlashga, iboralarning ham idiomatik xususiyatini belgilashga imkon beradi.

Freym - bilimlar strukturasiga oid tayanch tushunchalardan biri hisoblanadi. Freym – (inglizcha frame – “karkas”, “ramka”, “skelet” ma'nolarini anglatadi) sun'iy intellekt tizimida bilimlarni, stereotip (juda ko'p kuzatiladigan, tipik) vaziyatlarni gavdalandirishning bir usuli sanaladi. Ya'ni freymga modellshtirilayotgan hodisa, vaziyat va jarayonlarni umumlashtirilgan holda joylashtirish mumkin. Freym — mavhum obraz modeli bo'lib, muayyan obyekt, voqea-hodisa, vaziyat va jarayonlarning formal tavsifi demakdir (masalan, restoranga borish, shifokor qabulida bo'lish, ma'ruzlarga qatnashish, tug'ilgan kun, futbol o'yini kabi tipik vaziyatlar). Dastlab “freym” termini XX asrning 70-yillarida Marvin Minskiy tomonidan olamning kompyuter modelini o'z ichiga oluvchi bilimlar strukturasiga nisbatan ishlatilgan. Shundan so'ng freym semantikasi konsepsiyasi (**frame semantics**) amerikalik tilshunos Charlz Fillmor tomonidan rivojlantirildi. Mazkur konsepsiyaga ko'ra so'zlar, so'z birikmalari, gaplar, matn ma'nosi **sahnalar** bilan uyg'unlashtirilgan, shartlangan bo'ladi. Bir freym boshqa bir freymning davomi yoki tarkibiy qismi bo'lishi mumkin.

Freymlar sun'iy intellekt tarkibiga kiruvchi ekspert tizimlarida faol ishlatiladi. Freymning quyidagi turlari mavjud: freym-namunalar, freym-strukturalar, freym-

rollar, freym-senariylar, freym-vaziyatlar. Freymlar tizimi birlashib **semantik tarmoq**ni hosil qiladi. Freym kompyuter xotirasida bilimlar xazinasini va faktlar omboriga tayanib olamni maxsus shakllantiruvchi model vazifasini bajaradi. Freym **slot** deb atalgan bir yoki bir necha elementlarning majmui hisoblanadi. Slotlarda obyektlarga xos belgilar ma'nosi qayd etilgan bo'ladi.

Freym-senariy - ba'zi harakat, voqea-hodisalarning harakterli unsurlarini o'z ichiga olgan tipik strukturalar hamda vaziyatlarning rivojlanishi, ketma-ketligi (algoritm bosqichlari) demakdir. Senariy bu oddiy voqea-hodisalar zanjiri emas, balki ushbu harakatlar ketma-ketligida kauzal bog'liqlik tamoyili ishlaydi, ya'ni har bir harakat natijasi shartli sanaladi, uning ta'siri va oqibatida navbatdagi ish-harakat sodir bo'lishi mumkin. Har bir senariyda rol bajaruvchilar mavjud. Senariylar real hayotda uchraydigan vaziyatlarni aks ettiruvchi holatlar asosida yuzaga keladi. Masalan, restoranga tashrif buyurish senariysi.

Generativ va transformatsion grammatika - formallashtirishning negizi hisoblangan grammatika bo'lib, 1950 yillarda dastlab Z.Harrisning til strukturasidagi transformatsiya xususidagi qarashlari ta'sirida yuzaga kela boshladi, bu yo'nalish Z.Harrisning shogirdi N.Chomskiy tomonidan rivojlantirildi. Unga ko'ra dastlabki yadro strukturalar birikma strukturalar qoidalar tizimi orqali tug'iladi. Ya'ni har qanday til grammatikasi **yadro strukturalar (kernel structures)** yig'indisidir.

Generativ grammatikada yadro strukturalardan (**operandlar**) turli shakliy o'zgarishlar asosida asosiy ma'noni saqlagan holda ikkilamchi sintaktik qurilmalarning hosil qilinishi (**transformatsiyalar = transformalar**), bu jarayondagi vositalar (**operatorlar**) va **transformatsiya hosil qilish modellari** tadqiq etiladi. Masalan, Salim keldi – Salimning kelishi – Salim kelgach – Salim kelganda – Salim kelishi bilan...; Anvar xatni yozdi – Xat Anvar tomonidan yozildi. Ushbu misollarda transformatsiya uchun asos bo'lgan gap (*Salim keldi* yoki *Anvar xatni yozdi* kabi gaplar) **operand**, transformatsiya natijasi bo'lgan hosila gap (*Salim kelgach*, *Salim kelib* yoki *xat Anvar tomonidan yozildi* kabi sintaktik strukturalar) **transforma** yoki **transformand**, operanddan transformandni hosil qiluvchi vosita (ravishdsh, sifatdash, harakat nomi yoki majhul nisbat shakllari) **transformatsiya operatori** deb atalai.

Gipertekst - kompyuter lingvistikasida inqilobiy o'zgarish yasagan texnologiya. Gipertekst g'oyasi AQSH prezidenti F.Ruzveltning fan va ta'lim bo'yicha maslahatchisi Vannevar Bush nomi bilan bog'liq bo'lib, u foydalanuvchiga matn va uning fragmentlarini turli assotsiativ munosabatlar asosida bog'lashga imkon beruvchi "**MEMEKS**" tizimini nazariy jihatdan asoslab berdi. U 1945 yilda "The Atlantic Monthly" jurnalida "Memeks" deb atalgan kelajakdagi xayoliy qurilma haqida "**As We May Think**" ("**Biz o'ylaganimiz kabi**") nomli maqolasini e'lon qildi. V.Bush o'sha qurilmani kutubxonadagi har qanday hujjatlar, yozishmalar, kitoblarni displeyda ko'rsatib bera oladigan, cheklanmagan mikrofilmlarga ulangan elektromexanik stol sifatida tasvirlaydi.

Kompyuter texnikasining yo'qligi bois loyihaning amalda qo'llanishi bir oz keyinga surildi. "Gipertekst" termini 1965 yil T.Nelson tomonidan ilmiy iste'molga olib kirilgan. 1992 yilda birinchi marta gipertekst World Wide Web (xalqaro to'r)ga tatbiq etildi.

Gipertekst bu shunday matnki, u kompyuterda boshqa matnlarga giperhavolalar yordamida bog'lashga imkon beradi. Bunda u o'quvchiga faqat matnlar ustida emas, balki turli xil jadvallar, sxemalar, rasmlar, videoroliklarda ham erkin navigatsiya qilish imkonini beradi. Asosan, gipertekst World Wide Web (Xalqaro o'rgimchak to'ri) strukturasi ko'rsatuvchi tushuncha bo'lib, u orqali internet tizimida informatsiya tarqalishida qulaylik, osonlik, optimallik ta'minlanadi. Internet olami virtual olam sanalgani uchun biz undagi birinchi olamdan ikkinchi olamga hech qanday to'siqlarsiz o'ta olamiz. Mana shu virtuallikni bizga ta'minlab beruvchi tizim gipertekst tizimidir. Demak, gipertekst texnologiyasi bizga information massivni tezkor ko'rib chiqish (**brauzing**), muayyan informatsiya yoki matn fragmentlariga tezkor havola qilish, gipertekst bo'yicha erkin **navigatsiya** qilish, gipertekstni yangi informatsiyalar bilan boyitib borish kabi optimal imkoniyatlar yaratib beradi.

Grammatika - yunoncha so'z bo'lib, "o'qish va yozish san'ati" ma'nosini bildiradi. Bu tushuncha dastlab xat-savod, imloni o'rgatuvchi fan ma'nosida ishlatilgan. Keyinroq grammatika tilning fonologik, morfologik, sintaktik va semantik belgilarini tavsiflovchi til qoidalari majmuasi ma'nosiga ko'chgan. Grammatika haqidagi ilk qarashlar falsafa, mantiq, **germenevtika** (diniy matnlarni sharhlash, talqin etish) qobig'ida bo'lgan. Bunda tilning paydo bo'lishi, til va tafakkur munosabati, lingvistik belgi motivatsiyasi, diniy matnlar mazmunini to'g'ri tushunish bilan bog'liq masalalar o'rganilgan, til strukturasi esa yetarli darajada tadqiq etilmagan. Faqatgina mil.avv.II-I asrlarda shakllangan Aleksandriya grammatika maktabining eng yirik vakil Dionisiy Traks (bizning eramizgacha bo'lgan 170-90 yillarda yashagan) o'zigacha bo'lgan tadqiqotchilarning ishlari va tajribalarini o'rganib, ulardan foydalanib, rimliklar uchun "**Grammatika san'ati**" ("**Grammatike techne**") nomli sistemalashtirilgan dastlabki yunon grammatikasini yaratdi. Shundan buyon grammatikani falsafadan mustaqil soha sifatida ajralib chiqdi.

Antik davrda grammatika **preskriptiv** (to'g'ri gapirish, to'g'ri jumla tuzishni ko'rsatuvchi yo'riqnoma) xarakterga ega bo'lgan. XVIII asr oxirlarida grammatikani tushunish keskin o'zgardi. Bu davrda bobo til konsepsiyasini ilgari surgan qiyosiy-tarixiy grammatika shakllandi. Unda yevropa tillarining qadimgi sanskrit tili bilan tarixiy aloqasini aniqlash yordamida hind-yevropa tillari shajarasini rekonstruksiya qilish bo'yicha qiyosiy aspektda tadqiqotlar olib borildi.

XX asr boshlarida Ferdinand de Sossyur grammatikani tavsiflashda sinxroniya va diaxroniyani farqladi. Bunda grammatikani tarixiy (diaxron) o'rganishni emas, balki sinxron (muayyan vaqtdagi holati) tadqiq qilish lozim deb hisobladi. XX asrning 50-yillarida Noam Chomsky grammatikani tavsiflashning

yangi metodini tavsiya etdi. Bu tilshunoslikda “formal grammatika nazariyasi” deb ataladi. Olimning tadqiqot natijalari kompyuter lingvistikasining shakllanishiga asos bo’ldi. N.Chomskyning formal grammatika haqidagi qarashlari “*Syntactic Structures*” (1957) hamda “*Aspects of the Theory of Syntax*” nomli kitoblarida yoritilgan.

Grammatikaga bo’lgan yondashuvlar turlicha bo’lganligi bois tilshunoslik tarixida turli grammatikalar yaratilgan. Jumladan, nazariy grammatika va amaliy grammatika; an’anaviy (klassik) grammatika va noklassik grammatika; sinxron va diaxron grammatika; xususiy va umumiy (universal) grammatika; kategorial grammatika, tobelik grammatikasi, leksik-funksional grammatika, konstruktiv grammatika, transformatsion grammatika, kontekstdan xoli grammatika (context-free grammar) kabi.

Implikativ universalisiya - Amerikalik olim Jozef Grinberg tomonidan implikatsiyalar (mantiqiy shartli hukmlar) asosida aniqlangan universalisyalari. Olim 45 ta implikativ universalisyaning aniqlagan. Masalan:

1. Agar til uzluksiz affikslarga ega bo’lsa, unda yo prefiksatsiya yoki suffiksatsiya yaxshi rivojlangan bo’ladi.
2. Agar til istisnoli tarzda suffiksali bo’lsa, u holda u til posleloglarga, kelishiklarga boy bo’ladi.
3. Agar til istisnoli tarzda prefiksali bo’lsa, u holda u til predloglarga boy bo’ladi.
4. Agar til VSO modelida bo’lsa, u tilda predloglar dominantlik qiladi.

Informatsion qidiruv tizimlari - formal belgilarga ko’ra informatsiyani qidirishga mo’ljallangan avtomatik tizim. IQLar quyidagi bosqichlar asosida ishlaydi: informatsiyani jamlash, informatsiyani klassifikatsiya qilish, hujjatning qidiruv obrazi (HQO)ni yaratish, hujjatning qidiruv obrazi va hujjatlarni saqlash, qidirish va natijani chiqarish.

“Informatsion qidiruv” (inglizcha “**information retrieval**”) termini 1948 yilda Kelvin Mur tomonidan fanga kiritilgan. Informatsion qidiruv tizimi bu ma’lumotlarni qisqa vaqt birligi ichida qidirib topish, kerakli hujjatlarni ma’lumotlar bazasidan topishni optimallashtirish tizimidir. Bu yo’nalish ilmiy-texnikaviy axborotlarning ko’payib ketishiga javoban 1960 yilning oxirlarida yuzaga keldi. Axborot ko’paygan sari uni saralash, keraklisini qidirib topish muammosi yuzaga keldi. Mazkur muammoni bartaraf etish uchun olimlar information qidiruv tizimlari g’oyasini taklif etdilar. Bu yo’nalish internet tarmog’i yuzaga kelgandan so’ng yanada rivojlandi. Masalan, bugun hammaga ma’lumki, internetda **Google, Yandex, Rambler, Yahoo** kabi qidiruv tizimlari mavjud.

Informatsion qidiruv tillari - IQLda qidiruvni ta’minlash uchun yaratilgan maxsus tillar. Informatsion qidiruv tillari bu formal til bo’lib, u IQLda saqlangan hujjatlarning mazmunini tavsiflash va so’rov uchun mo’ljallangan. Informatsion qidiruv tilida hujjatlarga ishlov berish protsedurasi **indekslash** deb ataladi. Indekslish natijasida har bir hujjatga IQLdagi ta’rif shakli yoziladi. Mazkur ko’p bosqichli jarayonda deskriptorli information qidiruv tillari ishlatiladi. Hujjat

mavzusi deskriptorlarning birlashuvi bilan yoziladi. **Deskriptorlar** sifatida so'zlar, terminlar, muayyan sohadagi boshlang'ich tushuncha va kategoriyalardan foydalaniladi. Deskriptorlarni kalit so'zlar deb nomlash ham mumkin. Deskriptorli til bilan ishlaydigan tizimlardan biri M.Taube tomonidan yaratilgan **UNITERM** tizimidir. Bu tizimda deskriptorlar sifatida hujjatning kalit so'zlari – unitermlar ishlatiladi. Demak, muayyan sohaning terminologiyasi, ya'ni deskriptorlar qidiruv tizimi va tezauruslarning asosini tashkil etadi.

Informatsion qidiruv tezauruslari (lotincha “thesaurus” – “xazina” degan ma'noni anglatadi) - matnning asosiy mazmun-mundarijasini aks ettiradigan kalit so'zlar, terminlar, tayanch tushunchalar bazasi asosida kompyuterda tuzilgan lug'at. Tezauruslar ensiklopedik va izohli lug'atlardan farqli ravishda til birlilarining matnda qo'llanish chastotasi va matnning predmet mundarijasini aks ettirish darajasiga qarab tuzilda. Bunda terminning mazmuniy tavsifi orqali bazadagi ma'lumotlar qidirilishiga mo'ljallangan bo'ladi. Terminlar tezaurusga qat'iy semantik prinsiplar asosida kiritiladi, bunda ularning ***gipo-geperonimik (tur-jins), holo-meronik (butun-bo'lak), iyerarxik (pog'onali)*** munosabatlari hamda assotsiativ mazmuniy bog'lanishlari e'tiborga olinishi lozim. Shu o'rinda ta'kidlash zarurki, keyingi paytlarda qidiruv tizimlari bazasidagi tezauruslar giperhavolalar bilan ham ta'minlanmoqda, bu esa userga bir muncha qulayliklar yaratadi. Ya'ni qidirilayotgan terminlarga aloqador bo'lgan boshqa yaqin tushunchalar bilan bog'liq ma'lumotlarni ham oson topish imkonini beradi.

Kodlash va **dekodlash** – informatsiyani mashina (kompyuter) tushunadigan tilga o'tkazish va bunga teskari bo'lgan kompyuter tushunadigan tildan inson tushunadigan tilga aylantirish jarayoni.

Kompyuter leksikografiyasi (elektron lug'atlar tuzish) – muayyan til leksikonidagi so'zlarni ma'lumotlar bazasi va maxsus programmalashtirish tillari yordamida ishlash tizimiga ega bo'lgan dasturlar (kompyuter lug'atlari) yaratish bilan shug'ullanuvchi soha. Kompyuter leksikografiyasi bugungi kunda turli qo'shimcha imkoniyatlar asosida rivojlanib bormoqda. Xususan, dastlab kompyuter lug'atlari bir tilli, ikki tilli bo'lgan, hozirda ko'p tilli (uch, to'rt, olti, o'n tilli) elektron lug'atlar yaratilmoqda. Shuningdek, keyingi paytlarda muayyan sohalarga ixtisoslashgan elektron lug'atlar ham yaratilmoqda. Ya'ni dastlab kompyuter lug'atlari faqat umumiy leksikon bilangina cheklangan edi, hozirda bundan tashqari fanga oid bo'lgan, qurilish, aviatsiya, avtomobilsozlik, harbiy, diniy, yuridik sohalarga oid so'zlar bazasi ham e'tiborga olinmoqda. Hatto so'zlarning matnda qo'llanilish imkoniyatlari, distributiv holatlari, birikma holidagi ko'rinishlari ham nazarda tutilgan holda lug'atlar yaratilmoqda.

Kompyuter lingvistikasi - matematik lingvistikaning mantiqiy davomi bo'lib, u amaliy tilshunoslikning eng muhim qismini tashkil etadi. Kompyuter lingvistikasi 1954 yil AQSHda Jorjtaun universitetida mashina tarjimasini bo'yicha dunyoda o'tkazilgan birinchi tajriba asnosida yo'nalish sifatida shakllana boshladi, 1960 yilga kelib mustaqil fan sifatida shakllandi. Kompyuter lingvistikasi inglizcha

“computational linguistics” so’zining kalkasidir. XX asrning 80-yillariga qadar bu fan turlicha nomlar bilan atalgan: *hisoblash lingvistikasi*, *matematik lingvistika*, *kvantitativ lingvistika*, *injener lingvistikasi* kabi. Bu fanning asosiy maqsadi lingvistik masalalarni yechishning kompyuter dasturlarini ishlab chiqish, inson va mashina (kompyuter) muloqotini optimallashtirish, tabiiy tilni qayta ishlash (**Natural Language Processing**)dir.

Kompyuterli modellashtirish - bugungi kunda barcha fanlarda, xususan, kompyuter lingvistikasida ham samarali metod hisoblanadi. Kompyuter modellashtirish quyidagi asosiy bosqichlardan iborat:

- Masalaning qo’yilishi, modellashtirish obyektining aniqlashtirilishi
- Konseptual (tushunhaviy, fikriy) modelning ishlab chiqilishi, tizim asosiy unsurlarining ajratib olinishi
- Formalizatsiya, ya’ni matematik model bosqichi; algoritm yaratilishi hamda dastur ishlanishi
- Kompyuter eksperimentlarini o’tkazish
- Natijalarning tahlili va talqini.

Sun’iy intellekt tizimi doirsidagi tabiiy tilli interfeys, ekspert tizimlari, neyron tarmoqlar, lingvoanalizatorlar, gapiruvchi avtomatlar – barchasi kompyuter modellashtirish natijasi hisoblanadi.

Kontekstga bog’liq bo’lmagan grammatika (Context-free grammar) – matndan xoli bo’lgan grammatika, chunki undagi qoidalar universal hisoblanib, ma’lum matnga tayanib tahlil etilshga asoslanmaydi, balki u avvaldan empirik tarzda belgilab olinadi. Hozirda ushbu grammatika sun’iy programmalash tillariga (Pascal, C++, Delphi, HTML) faol tatbiq etilgan, shuningdek, tabiiy tillarni formallashtirish jarayonida ham grammatikani bayon qilishning eng optimal usuli hisoblanadi.

Lemmatizatsiya – so’zning dastlabki, boshlang’ich formasini (lug’atdagi shaklini – lemmasini) tashkillashtirish texnikasi bo’lib, bu jarayon o’sha so’zning boshqa so’z-shakllaridan kelib chiqqan holda amalga oshiriladi. Lemmatizatsiya morfologik tahlil metodi tarkibiga kiradi, u ikki bosqichni o’z ichiga oladi: 1) deklarativ bosqich – bunda muayyan so’zning bo’lishi mumkin bo’lgan barcha shakllari (so’z-shakllar) belgilanadi; 2) protsedura bosqichi – bunda so’z asos va qo’shimchalarga, ya’ni leksemalarga yoki morfemalarga bo’linadi. Lemmatizatsiya so’zlarning grammatik valentligi, qaysi affikslar bilan birika olish imkoniyatini ham belgilab beradi. Masalan, o’zbek tilida so’zlarning lug’atdagi shakli – lemmasi quyidagicha:

- Ot so’z turkumi uchun – bosh kelishik, birlik shakli.
- Fe’l so’z turkumi uchun – harakat nomi shakli
- Sifat so’z turkumi uchun – oddiy daraja shakli

Daftarlarni, daftarlarga, daftarlarning, daftarlardan, daftarlarda → daftar

Yugurdi, yugurgan, yuguryapti, yugurmoqchi → yugurmoq

Ko’kimtir, ko’kish, ko’kroq → ko’k

Lingvistik protsessorlar - tabiiy tilda kompyuter bilan muloqot qilishni ta'minlashga yordam beradigan dasturlar. XX asrning 70-yillarida AQSHda kompyuter bilan cheklangan muloqotni ta'minlaydigan tabiiy tilli interfeys tizimi (**LIFER** – Language Interface Facility with Ellipsis and Recursion) yaratildi. 1985 yilda “Semantek” korporatsiyasi tomonidan **Q&A** dasturi, “Karnegi grupp” kompaniyasi tomonidan **Language Craft** dasturiy ta'minoti ishlab chiqildi. Bu turdagi dasturlar hozirda ingliz tilining standartlasgan so'zlashuv uslubida uncha katta bo'lmagan lu'g'at zahirasi bilan ishlamoqda.

Lingvistik universal - barcha tillarga yoki dunyo tillarining mutlaqo ko'pchilik qismiga xos umumiy qonuniyatlar. 1941 yilda Aginskiy tomonidan “**The Importance of language universals**” sarlavhali maqola e'lon qilindi. 1961 yilda Nyu York shahrida universalialarga bag'ishlangan xalqaro konferensiya o'tkazildi. Unda J.Grinberg, J.Jenkins, Ch. Osgudlar “**Lingvistik universalialar haqida memorandum**” mavzusida chiqish qildilar. Konferensiyada Ch.Osgud shunday yozadi: “*Biz ushbu konferensiyada qonsiz inqilobning guvohi bo'lamiz*”.

Fonologiya, morfologiya, sintaksis, semantika sohalarida aniqlangan lingvistik universalialar 2000 yilga kelib 1600 tani tashkil etdi. Germaniyaning Konstanz universitetida lingvistik universalialarning bazasi (**Universals Archive**) yaratildi va u har yili yangi ma'lumotlar bilan boyitib boriladi. Mazkur baza internetda <http://www.ling.uni-konstanz.de/pages/proj/sprachbau.htm> saytiga joylangan.

Matematik lingvistika - XX asrning 50-yillarida (1952 yilda) yuzaga kelgan amaliy tilshunoslikning alohida yo'nalishi bo'lib, unda tabiiy tillarning matematik modellarini (bunday formallashgan til metatil deb ataladi) ishlab chiqish, xususan, sun'iy tillarni yaratish algoritmini tuzish bilan shug'ullanuvchi fandır. Bu fanning shakllanishida Kopengagen struktural tilshunoslik maktabi (glossematika)ning asoschisi Lui Yelmslevning g'oyalari o'ziga xos „turtki“ vazifasini o'tagan. U hatto til hosidalarini matematik bayonda tushuntiradigan fanning nomini ham taklif etgan. Olimning fikricha, bu fan „Til algebrasi“ („**Lingvistik algebra**“) deb atalishi lozim edi. Amerikalik tilshunos Noam Chomskiyning formal grammatika, transformatsion grammatika haqiagi qarashlari bevosita matematik lingvistikaning alohida yo'nalish sifatida yuzaga kelishiga sabab bo'lgan.

Matnlar korpusi (corpus – lotincha “tana” degan ma'noni anglatadi) – elektron holda saqlanadigan ma'lum til birliklari bo'lib, ular tilshunoslar uchun turli xil muammolarni hal etish uchun tatbiq etishda va turli yo'nalishdagi tadqiqotlar uchun zaruriyatga qarab turli shakllarda tuziladi. Til birliklarining qay tarzda saqlanishiga qarab maxsus dasturlar yordamida har bir kerakli so'z yoki so'z birikmasi uchun darhol uning qo'llanishi bo'yicha misollar topilishi, imlo bo'yicha variantlari, sinonimik qatorlari topilishi mumkin. Matnlar korpusiga oid ilmiy tadqiqotlar salmog'ining ko'payishi natijasida tilshunoslikda **korpus lingvistikasi** yo'nalishi shakllandi.

Kompyuterda yaratilgan birinchi matnlar korpusi **Braun korpusi** (БК, inglizcha *Brown Corpus*, BC) hisoblanadi, u 1961 yilda Braun universitetida yaratilgan, har biri 2000 so'zli 500 ta matn fragmentini o'z ichiga oladi. 1970-yillarda 1 million so'zni o'z ichiga olgan matnlar

korpusi asosida rus tilining chastotali lug'ati yaratildi. 1980 yillarda Shvetsiyaning Uppsala universitetida ham rus tilida matnlar korpusi yaratildi. Buyuk Britaniyada **Ingliz tili Banki** (**Bank of English**) hamda **Britaniya Milliy Korpusi** (**British National Corpus, BNC**), Rossiyada **Rus tilining mashina fondi** (**Машинный фонд русского языка**) hamda **Rus tilining Milliy Korpusi** (**Национальный корпус русского языка**) loyihalari ishlab chiqildi. Masalan, Rus tilining Milliy Korpusi hajmi hozirgi kunda 149 million so'zdan iborat. Keyingi yillarda internet tizimining rivojlanishi **virtual matnlar korpusining** yuzaga kelishiga olib keldi. Ya'ni internetdagi qidiriv saytlari, elektron kutubxonalar, virtual ensiklopedialar korpus vazifasini bajarmoqda. Korpusning janri va tematik rang-barangligi internetdan foydalanuvchining qiziqishlariga bog'liq. Masalan, ilm-fan doirasida **Wikipedia** katta hajmdagi matnlar korpusi sifatida foydalanilmoqda.

Matnning statistik tahlili – tilni ehtimollik belgilariga tahlil qilish, til hodisalarini statistik xarakteristikalarini asosida baholash demakdir. Til birligining gapda yoki matnda qo'llanish chastotasi (sur'ati) ehtimollik deyiladi. Til hodisalari haqidagi haqiqatlar lingvostatistik tadqiqotlar yordamida aniqlanadi. Masalan, o'zbek tili uchun qaysi bo'g'in strukturasi tipik ekanligini, gap modellaridan qaysi biri eng asosiy konstruktsiya ekanligini, bosh kelishikdagi otlarning asosan qaysi gap bo'lagi vazifasida kelishini, tanlangan matnda so'z turkumlarining qay darajada ishlatilishini aniqlash uchun lingvostatistik tahlillarga ehtiyoj seziladi.

Matnning statistik tahlili matn atributsiyasi muammosiga ham oydinlik kiritishda muhim tadqiqot usuli hisoblanadi. Ayniqsa, badiiy matnlarni ng statistik tahlili natijasida o'sha matn muallifi ham aniqlanmoqda. Bunda matnda tez-tez ishlatiluvchi til birliklari (otlar, sifatlar, kalit so'zlar, fe'llar, grammatik shakllar, jumla qurilishi, bir so'z bilan aytganda, yozuvchining idiosstilini – o'ziga xos uslubini ko'rsatuvchi vositalar) lingvostatistik tahlil yordamida aniqlanadi. Turli matnlardan olingan dalillarning qiyosiy tahlili bizga o'sha matnning mazmun-mundarijasini, matn tuzilgan davrni, dalillarni argumentlash xarakterini va hatto mualliflikni aniqlashga imkon beradi. Badiiy asarlarning syujet qurilishiga, matn strukturasi qantitativ metodlarning tatbiq etilishi natijasida F.Dostoyevskiy, L.Tolstoy, M.Sholoxov kabi yozuvchilarning matnlari atributsiyasi amalga oshirilgan. Keyingi paytlarda anonim (muallifi noma'lum) va psevdonim (soxta muallifi) matnlarning aniqlanishi yuzasidan ham tadqiqotlar amalga oshirilmoqda.

Mashina ijodi - san'at asarlarini avtomatik tarzda yaratish jarayoni demakdir. Bugungi kunda ijodni to'la modellashtirishga erishilgani yo'q, lekin bu sohada sezilarli yutuqlar qo'lga kiritilgan. Jumaladan, kompyuter yordamida musiqali melodiyalar yaratilishi, treker-dasturlar, sintezatorlar, sekvensorlar yordamida bir turdagi melodiyalar va tovushlar generatsiyasi oqibatida elektron musiqalarning yaratilishi, matnning avtomatik generatsiyasi yordamida she'rlar, ertaklar – umuman matnlar yaratilishi, rasmlar generatsiyasi va kompyuter grafikasi yordamida kino va o'yinlarning yaratilayotganligi (ayniqsa, strategiya va harbiy manyovrlarga asoslangan kompyuter o'yinlarida) kabilar mashina ijodiga misol bo'la oladi.

Metagrammatika – yunoncha metha “keyin”, grammatike – “o'qish va yozish san'ati”, metaqoidalarga asoslangan grammatika; tavsifiy grammatika,

formal grammatika, transformatsion grammatika bosqichlaridan so'ng yuzaga kelgan, qayta yozilgan grammatika; programmalash tillarini tavsiflash uchun tabiiy til grammatikasi modeli asosida ishlab chiqilgan metaqoidalar to'plami. Lingvistik universalialar g'oyasi ta'sirida metagrammatika haqida qarashlar, nazariy va amaliy tadqiqotlar yuzaga keldi. Metagrammatika haqidagi dastlabki g'oyalar Dekart, Leybnits kabi matematiklar tomonidan ilgari surilgan ratsional (universal) grammatika yaratish, Klod Lanslo va Antoniy Arnoldlarning (1660 yil Parij) "Por-Royal grammatikasi" bilan bog'liqdir. Algol 68 uchun ishlangan metaqoidalar Winjgaarden-grammar yoki W-grammar deb nomlangan. Keyinroq 1981 yil olimlar Gazdar, Pullman, Saglar ham metaqoidalar ishlab chiqdi, u **Generalized Phrase Structure Grammar (GPSG)** deb ataladi.

Model - (lotincha "modelus" so'zidan olingan bo'lib, "nusxa", "andaza", "o'lchov", "me'yor" ma'nolarini anglatadi) tabiiy fanlar yoki umuman fanda shunday moddiy qurilma, grafik, sxema, umuman, bilish vositasi sifatida tushuniladiki, u muayyan original-obyekt haqidagi ma'lumotlar majmui tarzida yuzaga kelgan hosila-obyekt demakdir. Boshqacharoq tushuntirganda, model tabiiy obyektlarning imitatsiyasidir (o'xshashi, taqlidiy ko'rinishidir), u o'zbek tilidagi qolip, andoza so'zlariga mos keladi. Masalan, globus yerning modeli, o'yinchoq mashina haqiqiy mashinaning modeli sifatida baholanishi mumkin.

Modellarni shartli ravisda quyidagicha tasniflash mumkin: 1) tabiiy modellar – o'rganilayotgan obyekt bilan bir turda bo'ladi va undan faqat o'lchamlari jarayonlarining tezligi va ba'zi hollarda yasalgan materiali bilan farq qiladi; 2) matematik modellar – prototipdan (asl nusxadan) jismoniy tuzilishi bilan farq qiladi, lekin prototip bilan bir xil matematik tasvirga ega bo'ladi; 3) mantiqiy-matematik modellar – belgilardan iborat bo'lib, abstrakt model hisoblanadi va tafakkur jarayonini o'rganishda qo'llaniladi; 4) kompyuterli modellar – matematik va mantiqiy modellashtirish metodlariga asoslanib kompyuterda algoritm va dasturlardan foydalanib yaratiladigan modellar.

Modellashtirish - barcha fanlar uchun xos bo'lgan umumilmiy metod hisoblanadi va u quyidagi tamoyillarga amal qiladi:

- deduktivlik – mantiqiy xulosa chiqarishga asoslangan bo'ladi, xususiylikdan umumiylikka tamoyilida bo'ladi;

- tafakkur eksperimentidan foydalanish;

- evristik funksiyaga ega bo'lishi. Ya'ni u yangi g'oyalar bera olishi va uni amaliyotda sinab ko'rish imkoniyati bo'lishi zarur;

- eksplanatorlik xususiyatiga, ya'ni tushuntirish kuchiga ega bo'lishi kerak. Shundagina model eski nazariya tushuntirib bera olmagan muammoni hal qiladi, obyektning ilgari kuzatilmagan, ammo kelajakda amalga oshishi mumkin bo'lgan tomonini kashf qiladi.

- modelni ideallashtirilgan obyekt sifatida talqin qilish.

Modellashtirish jarayoni uchta unisurni o'z ichiga oladi:

- Subyekt (tadqiqotchi)

- Tadqiqot obyekti
- O'rganuvchi subyekt va o'rganiluvchi obyekt munosabatini aks ettiruvchi model.

Modellashtirish tilshunoslikda strukturalizm yo'nalishi ta'sirida faol tatbiq qilina boshlandi. Gap strukturasini modellashtirish g'oyasi XX asrning 50-yillarida amerikalik tilshunos Charlz Friz tomonidan olg'a surildi (1952 yilda "The Structure of English" nomli tadqiqotini yaratdi). Ch.Friz o'z konsepsiyasini distributiv model deb nomlagan. Bunga ko'ra gap muayyan so'z turkumlariga oid bo'lgan so'zlar zanjri hisoblanadi va tahlilda morfologiya bazasiga tayaniladi. Masalan, "*The young man painted the door yesterday*" jumlasini distributiv model asosida quyidagicha tahlil etiladi: **D 3 I^a 2-d D I^b 4**. Bu yerda **D** – otning aniqlovchisi (inglizcha determiner), **3** – sifat, **I** – birlikdagi ot, **2-d** – o'tgan zamon shaklidagi fe'l, **4** – ravishni anglatadi. Demak, bu modelda turli so'z turkumlariga mansub so'z shakllarining nutq zanjiridagi distributsiyasi (tarqalishi, qurshovi) gap strukturasini modellashtirishning asosiy mezonini sanaladi.

Neyron tarmoqlar - inson miyasidagi neyronlar modeli asosida yaratilgan tizim bo'lib, u sun'iy intellektning muhim tarkibiy qismi hisoblanadi. Hozirgi kunda neyron tarmoqlar bir qator amaliy masalalarni hal etishda keng qo'llanilmoqda. Jumladan, ob-havo prognozini aniqlashda, obrazlarni tanishda (xususan, og'zaki va yozma nutqni aniqlashda), robotlarni boshqarishda, tibbiyotda tashxis qo'yishda ishlatilmoqda. Neyron tarmoqlarining ishlash tizimi matematik ifodalar yoki mantiqiy qoidalarga asoslanmaydi, balki namuna (**example based approach**) asosida o'qitishga asoslangan. Bunda muayyan masalani yechish uchun namunalar (misollar), ya'ni o'qituvchi tanlamalar tayyorlanadi va tarmoq shu asosda tekshiradi.

On-line lug'atlar - ochiq turdagi internet tizimida ishlaydigan lug'atlar. Bu lug'atlarning aktivlanishi uchun albatta internetga ulangan bo'lishimiz lozim. Bunday lug'atlarning afzalligi shundaki, foydalanuvchilarga ushbu lug'atlar bazasini to'ldirib bo'rish, boyitish imkonini beradi.

Parallel matnlar korpusi - badiiy asar, qo'llanma, ommaviy axborot vositasi nashrlari, turli xil hujjatlarning ikki yoki undan ko'p tillardagi elektron holdagi ko'rinishlaridir. Masalan, Yevropa Ittifoqi o'zining barcha qonun va hujjatlarini ingliz, fransuz, nemis, ispan va italyan tillarida nashr qiladi hamda ular internet tizimiga barchaga ochiq arxiv sifatida qo'yiladi. Bunday korpuslarning afzalligi shundaki, ular yordamida nafaqat biron bir so'z yoki jumlaning, balki butun boshli matnlarning turli tillardagi variantlarini bilish imkoniyati mavjud. Xuddi mana shu imkoniyat tufayli maxsus **konkordanser dasturlar** ishlab chiqish orqali turli xil ixtisoslik lug'atlari tuzish imkoniyati tug'iladi. Ushbu imkoniyatlar kompyuter leksikografiyasi uchun ulkan ahamiyat kasb etadi.

Semiotika yoki **semiologiya** (yunoncha semiotike, ya'ni "belgilash", "tasvirlash", "ifodalash" degan ma'nolarni anglatadi) belgilar sistemasi haqidagi

fan bo'lib, unda axborotni saqlovchi, uzatuvchi turli xil belgi va belgilar tizimlarining yaratilishi, vazifalarini, belgilar sistemasiga kiruvchi hodisalarni, ularning tabiati, mohiyati, turlari, ko'rinishlari, amal qilishi kabilarni tekshiradi. Semiotika tilshunoslikda, xususan, kompyuter lingvistikasida ham muhim metodologik ahamiyat kasb etadi. Chunki til ham belgi sanaladi, semiotikada tabiiy tillar ham, sun'iy tillar ham, hatto yo'l belgilari, audio va videosignallar ham tadqiq etiladi. Semiotika fani XIX asrning oxirlarida vujudga keldi, uning asoschisi amerikalik olim Charlz Sanders Pirs (1839 – 1914 y.) hisoblanadi. Semiotikada belgilar tizimi to'rt aspekt asosida o'rganiladi: 1) *sintaktika* - belgilar strukturasini tekshiradi; 2) *semantika* - belgilarning mazmuniy tomonini tekshiradi; 3) *sigmatika* – belgi va uni aks ettiruvchi obyekt o'rtasidagi munosabatni o'rganadi; 4) *pragmatika* - belgilarning insonga bo'lgan munosabati, foydaliligi, amaliy jihatlarini tadqiq etadi.

Semantik tarmoq - borliqni so'zlar va ularning semantik munosabatlari yordamida modellashtirishga yo'naltirilgan tizim hisoblanadi. **Semantik tarmoq** aniq belgilangan uzellar (obyektlar kesishadigan joy) va dugalar (semantik munosabatlarning biri ikkinchisiga ta'sir ko'rsatadigan nuqta) munosabatdorligi asosida modellashtiriladi. Semantik tarmoq g'oyasi kompyuter texnologiyalaridagi lokal va global tarmoq nazariyasi ta'sirida kelib chiqqan. Uzellar muayyan obyektlarga, muayyan sohaga oid tushunchalarga, dugalar obyektlar o'rtasidagi semantik munosabatlarga aloqador tushunchalardir.

Sun'iy intellekt - intellektual mashinalar, intellektual kompyuter dasturlar yaratish texnologiyasi va bu haqidagi fan. Intellektual sistema strukturasida 3 asosiy blokni o'z ichiga oladi: bilimlar bazasi, masalani yechuvchi tizim va intellektual interfeys. Sun'iy intellekt insondagi intellekt tushunchasining aynan imitatsiyasi emas, balki inson intellektual faoliyatining ko'plab qirralarini o'z ichiga olgan, modellashtirilgan tizimlar texnologiyasi demakdir. Ayni shu ma'noda **“artificial intelligence”** terminini birinchi marta **Jon Makkarti** [1956 yilda](#) Darmut universitetidagi konferensiyada qo'lladi.

Tobelik daraxti (shajara) - L.Tenyer g'oyalari ta'sirida shakllangan tobelik grammatikasi modeli bo'lib, u kompyuter lingvistikasining avtomatik analiz dasturlarida ishlatiladi. Bunda gap asosan daraxt ko'rinishida tasvirlanadi, jumla tarkibidagi so'zlar bir-biri bilan tobelik munosabatida bo'ladi (*bevosita dominatsiya munosabati*). Kesim daraxtning ildizida joylashgan bo'ladi va dominant bo'lak hisoblanadi, boshqa so'zlar (gap bo'laklari) unga bevosita yoki bilvosita (boshqa so'zlar yordamida) tobelangan bo'ladi.

Translation Memory (“Tarjimon xotirasi”) yoki **Sentence Memory** (“Gap to'plagich”) – avtomatik tarjima samaradorligini oshirish maqsadida amerikalik olim Veb Layn tomonidan ishlab chiqilgan konsepsiya. Bu yondashuvga ko'ra muayyan matnni bir tildan ikkinchisiga tarjima qilish uchun dastlab uni yuqori malakali tarjimonga ko'rsatiladi. So'ng asl matn va uning tarjimasi kompyuterga kiritiladi, matn fragmentlarga (alohida jummalarga, birikmalarga) bo'linadi, ushbu elementlardan ma'lumotlar bazasi tayyorlanadi. Ma'lumotlar bazasi qidiruv

tizimiga yuklanadi, u esa tarjimasi mavjud bo'lgan matn fragmentlarini va alohida jumalarni sifatli tarjima qilishga imkon beradi. Matnning ayrim notanish fragmentlari (bazada tarjimasi mavjud bo'lmagan matn qismlari) inson qo'li bilan tarjima qilinadi. Bunda bazadagi bu fragmentlarga yaqin keladigan tarjimalardan foydalanish mumkin. Matnning inson tomonidan tarjima qilingan qismlari yana ma'lumotlar bazasiga kiritiladi. Mana shu yo'l bilan "tarjimon xotirasi" asta-sekinlik bilan boyitib boriladi va tarjima samaradorligi yuksalib boradi. Lekin ushbu yo'nalishning afzalliklari bilan birga ayrim kamchiliklari ham mavjud. Birinchidan, bu yo'l bilan har qanday turdagi matnni tarjima qilib bo'lmaydi. U asosan muayyan tor yo'nalishdagi bir turdagi matnlarga mo'ljallangan. Ikkinchidan, ma'lum matnning katta fragmenti ikkinchi bir matnning katta qismiga tarjimada mos kelishi yoki yaqin kelishi har doim ham kuzatilavermaydi. Uchinchidan, "tarjimon xotirasi" bazasini yaratish yoki ikki tilli matnlar (bilingv) massivini to'ldirib borish amaliy jihatdan ko'p qiyinchiliklarni keltirib chiqaradi.

"Tyuring testi" - "Mashina fikrlay oladimi" (1950) sarlavhali maqolaning muallifi ingliz olimi Alan Tyuring tomonidan ishlab chiqilgan empirik test. Olim mashina imkoniyatini insonning aqllilik darajasi bilan qiyoslash protsedurasini ishlab chiqdi. Ushbu g'oya 1950 yilda "Mind" falsafiy jurnalida *"Hisoblash mashinalari va tafakkur"* (*Computing Machinery and Intelligence*) nomli maqolasida e'lon qilindi. Bu testdan ko'zlangan maqsad sun'iy tafakkurning inson tafakkuriga yaqin kelishi mumkinligini aniqlash hisoblanadi. Bunga ko'ra inson bir kompyuter va bir inson bilan o'zaro munosabatga kirishadi. Savollarga javob berish asosida u kim bilan (inson bilanmi yoki kompyuter dasturi bilanmi) gaplashayotganini aniqlashi lozim bo'ladi. Kompyuter dasturining vazifasi esa insonni xatoga yo'l qo'yishiga, noto'g'ri tanlov qabul qilishiga majbur qildirish sanaladi.

Zif qonuni - matnning statistik tahlilida keng foydalaniladi, ayniqsa, katta hajmdagi matnlarda har bir so'zning qanchalik tez takrorlanishini (chastotasini) va ularning takrorlanish tezligini hisoblashda **Zif qonuni** amal qiladi. Bunda so'zning chastotasi (f), uning ro'yxatdagi holati qatori (r) deb belgilanadi. Zif qonuniga ko'ra matnda eng ko'p ishlatilgan so'z $r=1$, undan pastroq chastotada kuzatilgan soz $r=2$, undan keyingisi $r=3$ kabi tarzida belgilanadi. Bundan kelib chiqadiki, matndagi ixtiyoriy so'z uchun (f) chastotaga (r) ning ko'paytmasi natijasi o'zgarmas miqdor - konstanta (C) gat eng bo'ladi ($r \times f = C$), bu yerda C matn hajmiga bog'liq.

Chastotali lug'atlar - matnda so'zlarning ishlatilish chastotasi (qay darajada ishlatilishi, tez-tez ishlatilishi)dan kelib chiqqan holda tayyorlanadigan so'zliklar. Chastotali lug'atlar statistik leksikografiyaning rivoji natijasida yuzaga kelgan. Bunday lug'atlar ayniqsa, til o'qitish jarayonida leksik minimumlarni aniqlashda muhim ahamiyatga ega. So'zlarning chastotasini aniqlash linvostatistik tahlillar yordamida amalga oshiriladi. Chastotali lug'atlar kompyuterning lingvistik informatsion bazasi hisoblanadi.

**O‘ZBEKISTON RESPUBLIKASI OLİY VA O‘RTA MAXSUS TA‘LIM
VAZIRLIGI**

FARG‘ONA DAVLAT UNIVERSITETI

Ro'yxatga olindi:

№ _____

2022 y. ____ ____

“TASDIQLAYMAN”

O'quv ishlari bo'yicha prorektor

“ ____ ” _____ 2022

KOMPYUTER LINGVISTIKASIGA KIRISH FANINING ISHCHI O‘QUV DASTURI

Bilim sohasi: 100000 - Gumanitar soha
Ta’lim sohasi: 120000 – Gumanitar fanlar

Fanning ishchi o'quv dasturi o'quv, ishchi o'quv reja va o'quv dasturiga muvofiq ishlab chiqildi.

Tuzuvchilar:

Zokirov M – FarDU o'zbek tilshunosligi dostenti, f.f.n.
Shokirova X. –Ijtimoiy-gumanitar fanlar kafedrasida dotsenti, f.f.d.

Taqrizchi:

Hakimov M. – FarDU o'zbek tilshunosligi kafedrasida professori, f.f.d.;

Fanning ishchi o'quv dasturi o'zbek tilshunosligi kafedrasining 2022 yil “ ”
dagi “” - son yig'ilishida muhokamadan o'tgan va fakultet kengashida muhokama
qilish uchun tavsiya etilgan.

Kafedra mudiri: _____M.Isroilov.

Fanning ishchi o'quv dasturi filologiya fakulteti kengashida muhokama
etilgan va foydalanishga tavsiya kilingan (2019 yildagi -sonli bayonnoma).

Sirtqi bolim boshlig'i : _____A.Abdumalikov

Kelishildi: O'quv-uslubiy
boshqarma boshligi _____

Kirish

Oliy ta'limning Davlat ta'lim standartiga ko'ra "Filologiya" ta'lim sohasida o'qitiladigan "Matematik lingvistika" fani dasturi tilshunoslikda zarur bo'ladigan: matematik model, matematik mantiq, aksiomatik nazariya, masina tarjimasini, kompyuter lug'ati, avtomatik tahrir, formal grammatika kabi tushunchalarini o'z ichiga olgan bulimlaridan tashkil topgan.

Fanning maqsad va vazifalari

Matematik lingvistika fani til haqidagi fanlar majmuasiga taalluqli bo'lib, talabalar uni III semestrda o'rganishadi.

"Matematik lingvistika" fanining bosh muhim vazifasi, talabalarga fanning o'ziga xos xususiyatlari, tilning nazariy asosi, fan sifatida shakllanishi, ushbu fan yo'nalishida olib borilayotgan ilmiy tadqiqotlar haqida mukammal ma'lumot berish va shu asosda tilshunoslik nazariyasiga daxldor nozik masalalar xususida fikrlay olish qobiliyatini shakllantirishdan iborat.

Fan bo'yicha talabalarning bilimiga, o'quviga va ko'nikmasiga qo'yilgan talablar

- kompyuter lingvistika predmeti, vazifalari;
- kompyuter lingvistikaning ijtimoiy va tabiiy fanlar orasidagi o'rnini va mohiyati;
- kompyuter lingvistikaga doir tadqiqotlar;
- aksiomatik nazariyaning tilshunoslikka tatbiqi;
- matematik mantiq elementlari;
- til o'rgatish jarayonini avtomatlashtirish; tabiiy tillarni qayta ishlash, tillarning matematik modelini yaratish;
- avtomatik tahrir **haqida tasavvurga ega bo'lish kerak;**
- matematik mantiq asoslarini;
- matematik mantiq asoslarining tilshunoslikka tatbiqini;
- tarjima algoritmini;
- avtomatik tahrir dasturining ishlash tamoyilini;
- o'zbek tilining matematik modellarini;
- o'zbek tilining kompyuter uslubini;
- o'zbek tili grammatikasi ta'limi, bilimlarni baholash dasturlariga oid me'yoriy va amaliy ma'lumotlarni yig'ishni;
- o'zbek tili grammatikasining formallashtirilgan qismi uchun model tuzishni **bilishi va ulardan foydalana olishi kerak;**
- matematik mantiq asoslari, tabiiy tillarning matematik modeli, o'zbek tilining matematik modeliga doir masalalarni yoritish;
- axborot uslubi, o'zbek tilining kompyuter uslubi xususiyatlarini farqlash;
- o'zbek tili grammatikasining ma'lumotlar bazasini yaratish; matnning avtomatik tarjima qilish; matnning avtomatik tahrir qilish;

- kompyuter lingvistikaga oid tadqiqotlarni o‘rganish, nazariy ma’lumotlarni umumlashtirish;
- kompyuter lingvistika fanining istiqboli, dolzarb masalalariga doir fikrlar bildirish va tadqiqot olib borish *ko‘nikmalariga ega bo‘lishi kerak.*

O‘quv rejasidagi boshqa fanlar bilan bog‘liqligi

Mazkur fan o‘quv rejasidagi “Til nazariyasi”, “Hozirgi zamon o‘zbek tili”, “Oliy matematika” va “Informatika va axborot texnologiyalari” kabi fanlar bilan uzviy bog‘liqdir. Fanni o‘rganishda “Til nazariyasi”, “Hozirgi zamon o‘zbek tili”, “Tilshunoslik nazariyasi” va “Informatika va axborot texnologiyalari” kurslarida olingan nazariy va amaliy bilimlar zarur bo‘lsa, o‘z navbatida bu fandan olingan bilimlar “Tarjima nazariyasi”, “Kognitiv tilshunoslik”, “Korpus lingvistikasi”, “Psixolingvistika” va “Sotsiolingvistika” fanlarining umumiy nazariyasini chuqur o‘zlashtirishda ko‘prik vazifasini o‘taydi.

Fanni o‘qitishda zamonaviy axborot va pedagogik texnologiyalar

O‘quv jarayoni bilan bog‘liq ta’lim sifatini belgilovchi xolatlar quyidagilar: yuqori ilmiy-pedagogik darajada dars berish, muammoli ma’ruzalar o‘qish, darslarni savol-javob tarzida qiziqarli tashkil qilish, ilgor pedagogik texnologiyalardan va multimedia vositalaridan foydalanish, tinglovchilarni undaydigan, o‘ylantiradigan muammolarni ular oldiga qo‘yish, talabchanlik, tinglovchilar bilan individual ishlash, erkin muloqot yuritishga, ilmiy izlanishga jalb qilish.

“Kompyuter lingvistikasi” kursini loyihalashtirishda quyidagi asosiy konseptual yondoshuvlardan foydalaniladi:

Shaxsga yunaltirilgan ta’lim. Bu ta’lim o‘z mohiyatiga ko‘ra ta’lim jarayonining barcha ishtirokchilarini to‘laqonli rivojlanishlarini kuzda tutadi. Bu esa ta’limni loyihalashtirilayotganda, albatta, ma’lum bir ta’lim oluvchining shaxsini emas, avvalo, kelgusidagi mutaxassislik faoliyati bilan bog‘liq o‘qish maqsadlaridan kelib chiqqan holda yondoshilishni nazarda tutadi.

Tizimli yondoshuv. Ta’lim texnologiyasi tizimning barcha belgilarini o‘zida mujassam etmog‘i lozim: jarayonning mantiqiyliqi, uning barcha bo‘g‘inlarini o‘zaro bog‘langanligi, yaxlitligi.

Faoliyatga yo‘naltirilgan yondoshuv. Shaxsning jarayonli sifatlarini shakllantirishga, ta’lim oluvchining faoliyatini aktivlashtirish va intensivlashtirish, o‘quv jarayonida uning barcha qobiliyati va imkoniyatlari, tashabbuskorligini ochirishga yo‘naltirilgan ta’limni ifodalaydi.

Dialogik yondoshuv. Bu yondoshuv o‘quv munosabatlarini yaratish zaruriyatini bildiradi. Uning natijasida shaxsning o‘z-o‘zini faollashtirishi va o‘z-o‘zini kursata olishi kabi ijodiy faoliyati kuchayadi.

Hamkorlikdagi ta’limni tashkil etish. Demokratik, tenglik, ta’lim beruvchi va ta’lim oluvchi faoliyat mazmunini shakllantirishda va erishilgan natijalarni baholashda birgalikda ishlashni joriy etishga e’tiborni qaratish zarurligini bildiradi.

Muammoli ta’lim. Ta’lim mazmunini muammoli tarzda taqdim qilish orqali ta’lim oluvchi faoliyatini aktivlashtirish usullaridan biri. Bunda ilmiy bilimni

Ob'ektiv qarama-qarshiligi va uni hal etish usullarini, dialektik mushohadani shakllantirish va rivojlantirishni, amaliy faoliyatga ularni ijodiy tarzda qo'llashni mustaqil ijodiy faoliyati ta'minlanadi.

Axborotni taqdim qilishning zamonaviy vositalari va usullarini qo'llash - yangi kompyuter va axborot texnologiyalarini o'quv jarayoniga qo'llash.

O'qitishning usullari va texnikasi. Ma'ruza (kirish, mavzuga oid, vizuallash), muammoli ta'lim, keys-stadi, pinbord, paradoks va loyihalash usullari, amaliy ishlar.

O'qitishni tashkil etish shakllari: dialog, polilog, muloqot hamkorlik va o'zaro o'rganishga asoslangan frontal, kollektiv va guruh.

O'qitish vositalari: o'qitishning an'anaviy shakllari (darslik, ma'ruza matni) bilan bir qatorda - kompyuter va axborot texnologiyalari.

Kommunikatsiya usullari: tinglovchilar bilan operativ teskari alokaga asoslangan bevosita o'zaro munosabatlar.

Teskari aloka usullari va vositalari: kuzatish, blits-so'rov, oraliq va joriy va yakunlovchi nazorat natijalarini tahlili asosida o'qitish diagnostikasi.

Boshkarish usullari va vositalari: o'quv mashg'uloti bosqichlarini belgilab beruvchi texnologik xarita kurinishidagi o'quv mashg'ulotlarini rejalashtirish, qo'yilgan maqsadga erishishda o'qituvchi va tinglovchining birgalikdagi harakati, nafakat auditoriya mashg'ulotlari, balki auditoriyadan tashkari mustaqil ishlarning nazorati.

Monitoring va baholash: o'quv mashg'ulotida ham butun kurs davomida ham o'qitishning natijalarini rejali tarzda kuzatib borish. Kurs oxirida test topshiriqlari yoki yozma ish variantlari yordamida tinglovchilarning bilimlari baholanadi.

“Kompyuter lingvistikasi” fanini o'qitish jarayonida kompyuter texnologiyasidan, “Power point” dasturlaridan foydalaniladi. Ayrim mavzular bo'yicha talabalar bilimni baholash test asosida va kompyuter yordamida bajariladi. “Internet” tarmog'idan foydalaniladi, tarqatma materiallar tayyorlanadi, test tizimi hamda tayanch so'z va iboralar asosida oraliq va yakuniy nazoratlar o'tkaziladi.

“Kompyuter lingvistikasi” fanidan mashg'ulotlarning mavzular va soatlar bo'yicha taqsimlanishi:

O'zbek tili ta'lim yo'nalishlari uchun.

t /r	Mavzular nomi	Ja mi soat	Ma' ruza	Amali y mashg'ulot	Musta qil ta'lim
1.	Kirish. Kompyuter lingvistikasi faninig mazmun-mohiyati	2	2		
2.	Matematik mantiq asoslari	4	2	2	
3.	Tilshunoslikda modellashtirish metodidan foydalanish	14	2	2	10

4	Kompyuter lingvistikasida matematik statistika va kvantativ metod	4	2	2	
5	Kompyuter lingvistikasi fanining tarixi va zamonaviy holati.	4	2	2	
6	Kompyuter lingvistikasining asosiy yo'nalishlari	14	2	2	10
7	Kompyuter lingvistikasida avtomatik tarjima yo'nalishi	4	2	2	
8	Kompyuter leksikografiyasi	2		2	
9	Til o'rganish jarayonlarini avtomatlashtirish	4	2	2	
1	Avtomatik tahrir, information qidiruv tizimlari, gipertekst texnologiyasi	4	2	2	
1	Kompyuter lingvistikasi fanining rivojlanish istiqbollari	10	2		8
Жами:		66	20	18	28

Asosiy qism: Fanning uslubiy jihatdan uzviy ketma-ketligi

Asosiy qismda (ma'ruza) fanni mavzulari mantiqiy ketma-ketlikda keltiriladi. Har bir mavzuning mohiyati asosiy tushunchalar va tezislari orqali ochib beriladi. Bunda mavzu bo'yicha talabalarga DTS asosida etkazilishi zarur bo'lgan bilim va ko'nikmalar to'la qamrab olinishi kerak.

Asosiy qism sifatiga qo'yiladigan talab mavzularning dolzarbligi, ularning ish beruvchilar talablari va ishlab chiqarish ehtiyojlariga mosligi, mamlakatimizda bo'layotgan ijtimoiy-siyosiy va demokratik o'zgarishlar, iqtisodiyotni erkinlashtirish, iqtisodiy-xukukiy va boshqa sohalaridagi isloxlarning ustuvor masalalarini qamrab olishi hamda fan va texnologiyalarning so'ngi yutuqlari e'tiborga olinishi tavsiya etiladi.

Ma'ruza mashg'ulotlari

Kirish. Kompyuter lingvistikasi fanining mazmun-mohiyati. kompyuter lingvistikasining predmeti, maqsad va vasifalari, tabiiy tillarning matematik modellarini qurish, lingvistik muammolarni hal qiluvchi kompyuter dasturlarini ishlab chiqish masalalari

Qo'llaniladigan ta'lim texnologiyalari: *dialogik yondoshuv, muammoli ta'lim. Bingo, blist, ajurali arra, nilufar guli, menyu, algoritm, munozara, o'z-o'zini nazorat.*

Adabiyotlar: A1; A2; A3; Q1; Q5; Q6; Q9; Q10 DO.

Matematik mantiq asoslari. Matematik mantiqning tilshunoslikdagi ahamiyati. Mulohazalarni hisoblashning formal nazariyasi. Mantiq funksiyalari va ularning berilishi. Mulohaza. Absolut chin (yolon) mulohaza. Qiymatlar satri. Inkori, konyunksiya, dizyunksiya, ekvivalensiya mantiqiy amallari. Sheffer amali.

Qo'llaniladigan ta'lim texnologiyalari: *dialogik yondoshuv, muammoli ta'lim. Blist-surov, zig-zag usuli, munozara, BBB, Insert, uz-uzini nazorat.*

Adabiyotlar: A1; A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Tilshunoslikda modellashtirish metodidan foydalanish. Model haqida tushuncha. Tilshunoslikning formal nazariyasi. Shartli grammatika tushunchasi.

Qo'llaniladigan ta'lim texnologiyalari: *dialogik yondoshuv, muammoli ta'lim. Pogona, qadamba-qadam metodi, Venn diagrammasi, T-sxemasi, o'z-o'zini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Kompyuter lingvistikasida matematik statistika va kvantativ metod. Metod haqida tushuncha. Kompyuter lingvistikasida matematik statistika. Kompyuter lingvistikasida kvantativ metod.

Qo'llaniladigan ta'lim texnologiyalari: *dialogik yondoshuv, muammoli ta'lim. B/B/B jadvali, munozara, Venn diagrammasi, T-sxema, uz-uzini nazorat*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Kompyuter lingvistikasi fanining shakllanishi va zamonaviy holati. Kompyuter lingvistikasi fanining shakllanishi va zamonaviy holati. kompyuter lingvistikasi va an'anaviy tilshunoslik. Matematik va kompyuter lingvistikasi fanini yuzaga keltirgan omillar haqida. kompyuter lingvistikasi tushunchasi, uning sinonimlari injener lingvistikasi, statistik, lingvistika, hisoblash lingvistikasi, matematik lingvistika. Amerika, Evropa, Rossiya tilshunosligida, shuningdek, turkologiya va o'zbek tilshunosligidagi kompyuter lingvistikasiga oid tadqiqotlar.

Qo'llaniladigan ta'lim texnologiyalari: *dialogik yondoshuv, muammoli ta'lim. B/B/B jadvali, munozara, Venn diagrammasi, T-sxema, uz-uzini nazorat*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Matematik lingvistika fanining asosiy yonalishlari. Tilshunoslikda matematik statistika metodlarini qo'llash, mashina tarjimasini, avtomatik tahrir, tillarga o'qitish, bilimlarni tekshirish.

Qo'llaniladigan ta'lim texnologiyalari: *dialogik yondoshuv, muammoli ta'lim. Ajurali arra, bumerang, 3x3usuli, munozara, uz-uzini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Avtomatik tarjima yonalishi. Fanda mashina tarjima nazariyasining asoslanishi. Kompyuter tarjimasini algoritmini tuzish. «Kiruvchi va chiquvchi» til so'lari haqida grammatik ma'lumot berishi masalalari. Berilgan so'lardan hosil qilinadigan so'zlar ro'yxatini taqdim etish algoritmi. So'larni eslab qolish uchun grammatik sharhlar berish.

Qo'llaniladigan ta'lim texnologiyalari: *dialogik yondoshuv, muammoli ta'lim. Blite, 4 x 4 usuli, munozara, uz-uzini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Til o'rganish jarayonlarini avtomatlashtirish. Til o'rgatuvchi algoritmlar uchun nutqiy materiallarni tayyorlash. Til o'rgatuvchi o'quv dasturlari yaratish. O'rganilayotgan til leksik minimumni aniqlash.

Qo'llaniladigan ta'lim texnologiyalari: *dialogik yondoshuv, muammoli ta'lim. Bingo, blit, ajurali arra, nilufar guli, menyu, algoritm, munozara, o'z-o'zini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Matematik lingvistikaning avtomatik tahrir yo'nalishi. Matbuot va nashriyot ishlarida kompyuter texnologiyalarining ahamiyati. Matnlarni avtomatik

tahlil qilish haqida umumiy tushuncha berish. Mexanik xatolarni ajratish tamoyillari. Imloviy xatolarni aniqlash tamoyillari. Dialektal xatolarni belgilash tamoyillari. Kompyuterda xatolarni topish algoritmlarini tuzish. Aniqlangan xatolarni korreksiyalash algoritmlarini tuzish masalalari. Rasmiy va ilmiy nutq matnlarini tahrirlash algoritmlari.

Orfografik hamda mantiqiy kompyuter xatolari, struktur lingvistika, tadqiqot modellari, tilni formallashtirish, algoritm, universal qonuniyatlar, til modellari, xatolarni topish va tuzatish algoritmlari.

Qo‘llaniladigan ta’lim texnologiyalari: *dialogik yondoshuv, muammoli ta’lim. Integrativ, munozara, uz-uzini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Kompyuter lingvistikasi fanining rivojlanish istiqbollari va muammolari. O‘zbek tili grammatikasini rivojlantirish istiqbollari. Lingvistik masalalarning tabiiy tilda tog‘ri qo‘yilishi. Masalaning matematik tilda to‘g‘ri qo‘yilishi. Matematik va kompyuter lingvistikasi fanining istiqbollari.

Qo‘llaniladigan ta’lim texnologiyalari: *dialogik yondoshuv, muammoli ta’lim. Blite, 4 x 4 usuli, munozara, uz-uzini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

“Kompyuter lingvistikasi” fani bo‘yicha ma’ruza mashg‘ulotlarining kalendar tematik rejasi

/r	Ma’ruza mavzulari	So at
	Kirish. Kompyuter lingvistikasi faninig mazmun-mohiyati	2
	Matematik mantiq asoslari	2
	Tilshunoslikda modellashtirish metodidan foydalanish	2
	Kompyuter lingvistikasida matematik statistika va kvantativ metod	2
	Kompyuter lingvistikasi fanining tarixi va zamonaviy holati.	2
	Kompyuter lingvistikasining asosiy yonalishlari	2
	Kompyuter lingvistikasida avtomatik tarjima yonalishi	2
	Til o‘rganish jarayonlarini avtomatlashtirish	2
	Avtomatik tahrir, information qidiruv tizimlari, gipertekst texnologiyasi	2
	Kompyuter lingvistikasi fanining rivojlanish istiqbollari	2
	Jami	20

Amaliy mashg‘ulotlarning tavsiya etiladigan mavzulari

Mulohazalar formal nazariyasi, mantiq funktsiyalari. Mulohaza. Absolut chin (yolon) mulohaza. Mantiq funktsiyalari va ularning berilishi. Tabiiy tillarning matematik modellarini yaratish, tilning, xususan, o‘zbek tilining funkstional doirasini kompyuter vositasida kengaytirish omillarini belgilash.

Qo‘llaniladigan ta’lim texnologiyalari: *dialogik yondoshuv, muammoli ta’lim. Integrativ, munozara, uz-uzini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Lingvistikaning formal nazariyasi. Model haqida tushuncha. Tilshunoslikning formal nazariyasi. Shartli grammatika tushunchasi.

Soʻzlarni formallashtirishning oʻziga xos xususiyatlari. Mustaqil soʻz turkumlarini formallashtirishning oʻziga xos xususiyatlari, YOrdamchi soʻz turkumlarini formallashtirish. Formal grammatikada tub soʻzlarning berilishi.

Gap boʻlaklarini formallashtirishning oʻziga xos xususiyatlari. Soda gaplarni formallashtirishning oʻziga xos xususiyatlari, qoʻshma gaplarni formallashtirish. Formal grammatikada ega va kesimning aks ettirilishi. Ikkinchi darajali boʻlaklar: aniqlovchi, toʻldiruvchi, holni formallashtirish. Soʻz birikmasining bitishuv, moslashuv va boshqaruv aloqali turlarini formallashtirish. Sodda gap konstruktsiyalari: ega sostavli va kesim sostavli gaplar, uyushiq boʻlakli gap, bir bosh boʻlakli gaplarni modellashtirish, qoʻshma turlari - bolangan qoʻshma gap, bolovchisiz qoʻshma gap, ergashgan qoʻshma gap turlarini modellashtirish.

Qoʻllaniladigan taʼlim texnologiyalari: *dialogik yondoshuv, muammoli taʼlim. Integrativ, munozara, uz-uzini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Oʻzbek tili grammatikasining aksiomatik nazariyasi. Oʻzbek tili grammatikasining aksiomatik nazariyasi. Oʻzbek tili grammatikasining aksiomalari. Oʻzbek tili grammatikasining keltirib chiqish qoidalari. Oʻzbek tili grammatikasining aksiomatik modelini rivojlantirish yoʻllari. Oʻzbek tili grammatikasini rivojlantirish istiqbollari.

Qoʻllaniladigan taʼlim texnologiyalari: *dialogik yondoshuv, muammoli taʼlim. Ajurali arra, bumerang, 3x3usuli, munozara, uz-uzini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Lingvistik masalalarni matematik usullar bilan echish muammolari. Lingvistik masala mohiyatining tabiiy tilda toʻri aks ettirilishi: sinonimiya va koʻp maʼnolilik holatlarining bartaraf etilishi. Masalani matematik tilda toʻri qoʻyilishi.

Qoʻllaniladigan taʼlim texnologiyalari: *dialogik yondoshuv, muammoli taʼlim. Pogona, qadamba-qadam metodi, Venn diagrammasi, T-sxemasi, oʻz-oʻzini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Kompyuterda lingvistik masalalarni yechish yoʻllari. Lingvistik masalalarini oʻrganishda kompyuter texnologiyalarini tadbiq etish. Masalani yechish jarayonining algoritmini tuzish. Masalani yechish dasturini tuzish jarayoni. Komp'yuterda masalani yechish bosqichi.

Qoʻllaniladigan taʼlim texnologiyalari: *dialogik yondoshuv, muammoli taʼlim. Bingo, blist, ajurali arra, nilufar guli, menyu, algoritm, munozara, oʻz-oʻzini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Tabiiy tillarning kompyuter lugʻatlari.

Tabiiy tillarning - oʻzbek, rus, ingliz tillarining kompyuter luatlari. Kompyuter luatida tub soʻzlarning berilishi. Yasama soʻzlarning berilish usullari. Atoqli otlar (antroponimlar, zoonimlar) va geografik nom (toponim)larning kompyuter luatlarini tuzish masalalari. Barqaror birikmalar (frazemalar, idiomalar, murakkab otlar)ning kompyuter luatida berish muammolari.

Qo'llaniladigan ta'lim texnologiyalari: *dialogik yondoshuv, muammoli ta'lim. Bingo, blist, ajurali arra, nilufar guli, menyu, algoritm, munozara, o'z-o'zini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Matnlarni tahrirlash algoritmlari.

Matbuot va nashriyot ishlarida kompyuter texnologiyalarining ahamiyati. Matnlarni avtomatik tahlil qilish haqida umumiy tushuncha berish. Mexanik xatolarni ajratish tamoyillari. Imloviy xatolarni aniqlash tamoyillari. Dialektal xatolarni belgilash tamoyillari. Kompyuterda xatolarni topish algoritmlarini tuzish. Aniqlangan xatolarni korrektsiyalash algoritmlarini tuzish masalalari. Rasmiy va ilmiy nutq matnlarini tahrirlash algoritmlari.

Orfografik hamda mantiqiy kompyuter xatolari, struktur lingvistika, tadqiqot modellari, tilni formallashtirish, algoritm, universal qonuniyatlar, til modellari, xatolarni topish va tuzatish algoritmlari.

Qo'llaniladigan ta'lim texnologiyalari: *dialogik yondoshuv, muammoli ta'lim. Integrativ, munozara, uz-uzini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Kompyuter lingvistikasi fanining rivojlanish istiqbollari.

O'zbek tili grammatikasini rivojlantirish istiqbollari. Lingvistik masalalarning tabiiy tilda tog'ri qo'yilishi. Masalaning matematik tilda to'g'ri qo'yilishi. Matematik va kompyuter lingvistikasi fanining istiqbollari.

Qo'llaniladigan ta'lim texnologiyalari: *dialogik yondoshuv, muammoli ta'lim. Blite, 4 x 4 usuli, munozara, uz-uzini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Lingvistik kompyuter dasturlarini takomillashtirish yo'llari.

Lingvistik masalalarini o'rganishda kompyuter texnologiyalarini tadbiq etish. Lingvistik masala mohiyatining tabiiy tilda to'ri aks ettirilishi: sinonimiya va ko'p ma'nolilik holatlarining bartaraf etilishi. Masalani matematik tilda to'ri qo'yilishi. Masalani yechish jarayonining algoritmini tuzish. Masalani yechish dasturini tuzish jarayoni. Komp'yuterda masalani yechish bosqichi.

Qo'llaniladigan ta'lim texnologiyalari: *dialogik yondoshuv, muammoli ta'lim. Bingo, blist, ajurali arra, nilufar guli, menyu, algoritm, munozara, o'z-o'zini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

Tilshunoslikning zamonaviy yo'nalishlari.

Tilshunoslik fanlari ilmiy paradigmasining o'zgarishi. Qiyosiy-tarixiy tilshunoslik, sistem-struktur tilshunoslik, antropostentrik tilshunoslik. Yangi paradigmadagi matematik lingvistikaning o'rni.

Qo'llaniladigan ta'lim texnologiyalari: *dialogik yondoshuv, muammoli ta'lim. Pogona, qadamba-qadam metodi, Venn diagrammasi, T-sxemasi, o'z-o'zini nazorat.*

Adabiyotlar: A1;A2; A3; Q1; Q2; Q4; Q7; Q10 DO.

«Kompyuter lingvistikasi» fanidan amaliy mashg'ulotlar bo'yicha tuzilgan kalendar-tematik

REJA

	Mavzuning nomi	So at
	Mulohazalar formal nazariyasi, mantiq funktsiyalari	2
	Lingvistikaning formal nazariyasi	2
	O'zbek tili grammatikasining aksiomatik nazariyasi	2
	Lingvistik masalalarni matematik usullar bilan yechish yo'llari	2
	Tabiiy tillarning kompyuter lug'atlari	2
	Matnlarni tahrirlash algoritmlari	2
	Kompyuter lingvistikasi fanining rivojlanish istiqbollari	2
	Kompyuter lingvistikasi dasturlarini takomillashtirish yo'llari	2
	Tilshunoslikning zamonaviy yo'nolishlari	2
	Jami	18

Сухбат саволлари

1. Тилшуносликда қандай оқимларни биласиз?
2. Психолингвистика, социолингвистика, этнолингвистика, математик лингвистика, компьютер лингвистикаси каби йўналишлар ҳақида фикрларингизни билдириңг?
3. Мазкур йўналишларнинг пайдо бўлиш сабаблари ҳақида нима айта оласиз?

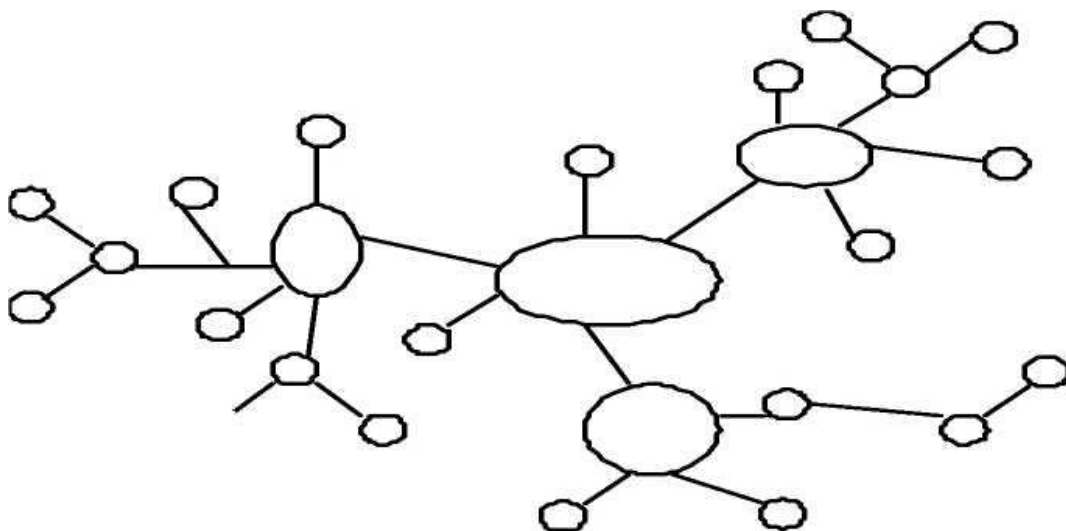
Нима учун XX асрнинг 50-йилларида фанлар ўртасида интеграция кучайди?

Х-ХІ асарларда Беруний, ибн Сино, Хоразмий, Фарғоний ва бошқа олимларнинг бир йўла бир неча фанлар билан шуғуллангани маълум. Улар энциклопедист олимлар ҳисобланади. XX асрга келиб, олимларнинг фаннинг тор соҳаларига ихтисослашуви кучайди, масалан, тилшунослик соҳасида фонетист олимлар, лексикология мутахассислари, синтаксистлар каби. Шу билан бирга, фаннинг бошқа жабҳаларида бунга зид парадоксал ҳолат ҳам кузатила бошланди. Айрим муаммоларни ҳал қилишда 2 ҳатто 3-4 фаннинг натижаларини бирга қўллаш зарурати туғилди. Икки фан чегарасида янги фан соҳалари юзага кела бошлади: биофизика, биохимия, бионика, математик физика, математик биология, математик тилшунослик каби.

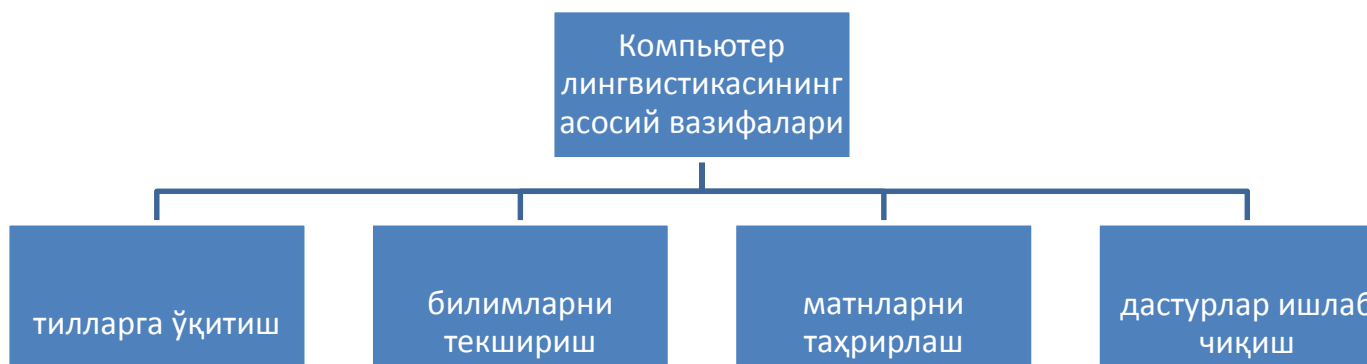
XX асрнинг ўрталаридан фанларда математик методларнинг тадбиқи эраси бошланди. Математика. физиология. техника фанлари ютуқларининг кўшулуви натижасида кибернетика фани юзага келди. Кибернетика туфайли математик методларни тадбиқ қилиш имкониятлари кенгайди. Математик методлар ҳатто энг ижтимоий-гуманитар фанлар бўлган соҳаларга ҳам кириб борди.

3-илова.

XX асрнинг 50-йилларида тилшуносликда янги қандай оқимлар пайдо бўлган эди?



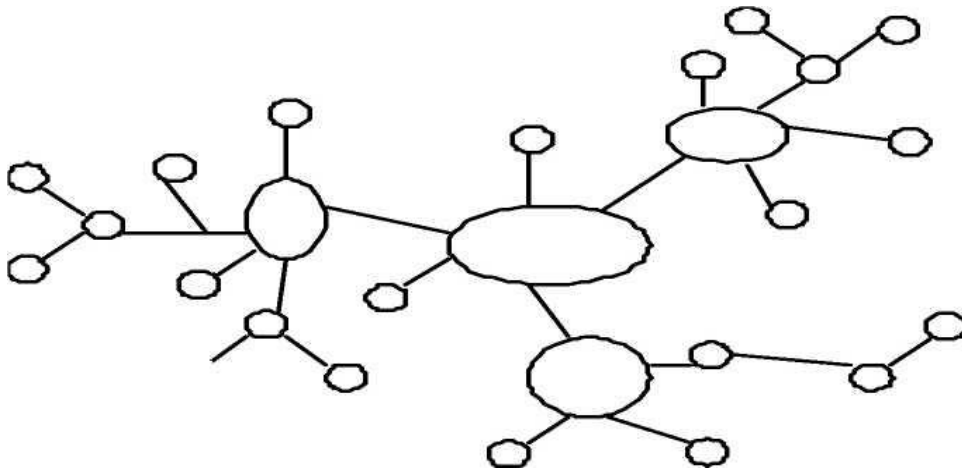
4-илова



- Математик лингвистика ва компьютер лингвистикаси деганда нимани тушунаси?
- Нима деб ўйлайсиз, нима учун 50-йилларда тилшуносликда математик методларни қўллай бошлашган?
- Математик моделлаштириш деганда нимани тушунаси?
- Мантиқ функцияларининг математик мантиқдан фарқи борми?

“Мантиқ” сўзига фикр билдиринг..... (Фикрлар хужуми)

2-илова



2-илова

Кўп йиллар давомида тилшуносликда кузатиш методлари ягона метод ҳисобланиб келинди. Аммо бу методлар ёрдамида лингвистик ходисаларнинг ички моҳияти очилмайди. Бу метод ёрдамида сўз шакллари, гап тузилиши ва бошқа структур хусусиятларни ўрганиш мумкин бўлади. Тил ва нутқнинг тузилиши мураккаб бўлганлиги сабаблари бу метод билан (кузатиш) уларни ўрганиб бўлмайди. Чунки:

Тил – умумийлик хусусиятига эга.

Тил – абстракт.

Тил – имконият.

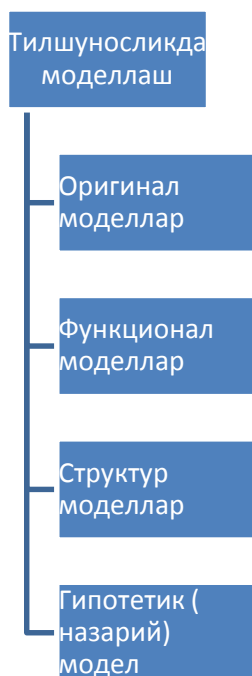
Нутқ – хусусийлик хусусиятига эга.

Нутқ – конкрет.

Нутқ – воқелик.

Моделлаштириш методидида тадқиқотчи объектнинг ўзини эмас балки унинг моделини

Даниэль дефо асарлари



Маъруза машғулоти докладлари мавзулари.

1. Тилшунослик тарихида компьютер лингвистикасининг тутган ўрни.
2. Анъанавий тилшунослик ва компьютер лингвистикаси.
3. Дунё тилшунослигида компьютер лингвистикасига оид тадқиқотлар.
4. Туркология ва ўзбек тилшунослигида соҳага оид тадқиқотлар.

1-илова

Автоматик таржима бўйича дастлабки ғоялар инглиз олими **Чарльз Беббидж** томонидан 1836-1848 йилларда илгари сурилган. Унинг фикрича, механик-электрон машиналар кодлаштирилган *автоматик таржимани* амалга ошириши мумкин бўлади. Унинг лойиҳаси бўйича 50минг та сўз хотира кучига эга бўлган электрон машиналар

100000 та сўзни автоматик равишда таржима қилиши зарур бўлади. **Биббиджинг** ғоялари орадан 100 йил ўтгачгина амалга ошди.

1954-йилда АҚШдаги Жорж Даун Университетида дунёда биринчи марта автоматик таржима байия тажриба бўлди. Бу тажрибалар IBM-701 компьютерида ўтказилди. Бу тажрибанинг базаси алгоритмлардан иборат бўлиб 50000 та сўзни рус тилидан инглиз тилига таржима қилди.

1960-йилда собиқ СССР фанлар Академиясининг *автоматик таржимага* бағишланган махсус йиғилиши бўлиб ўтди. Бу йиғилишда *автоматик таржима* бўйича профессор А.Белоногов раҳбарлигида махсус лаборатория ташкил қилинди.

Қозғоғистонда профессор А. Бектаев раҳбарлигида махсус ишчи гуруҳ тузилиб, «Абай йўли» романини лингвостатистикаси ҳисоблаб чиқилди.

1966-йилда АҚШнинг фан масалалари бўйича конгресс қўмитасида *автоматик таржимага* оид маъруза муҳокамаси бўлиб ўтди. У автоматик таржима бўйича салбий маънодаги маъруза эди.

Шундан кейин жаҳон миқёсида *автоматик таржимага* оид тадқиқотларни молиялаштириш кескин камайиб кетди. Улар фақатгина Японияда давлат томонидан молиялаштирилади. Қолган давлатларда *автоматик таржимага* оид лойиҳалар хусусий йирик компаниялар ҳисобланаган «**General Motors**», «**Siemens**», «**LG**», «**Phillips**» компаниялари бугунги кунда дунёда автоматик таржимага оид изланишларни маблағ билан таъминлаб туради.

2-

Илова

Автоматик таржима бўйича дунёдаги энг йирик лойиҳа 1991 йилда АҚШда **NASA** мутахассислари томонидан яратилган: **DIANA**. Унинг ҳажми 10 миллион та сўзни ташкил этади. У дунёнинг 6 та йирик тилига асосланади: *инглиз, рус, француз, немис, араб, испан*.

Рус олимлари ҳам профессор **Лякунов** раҳбарлигида «**Vista Tehnoloji**» **компанияси «Retrans Vista»** лойиҳасини ишлаб чиқди. Унинг таркибига кирувчи 3 млн. 400 мингта сўз хотирага киритилди. Шунда 1млн. 800мингтаси русча-инглизча, 1млн 600мингтаси инглизча-русча сўздан иборат.

Бугунги кунда *автоматик таржимани* қуйидаги системалари мавжуд. 40 млн.та сўз хотира кучига эга **EVROTRA** ситемаси – Европа олимлари томонидан ишлаб чиқилган. У барча Европа тиллари бўйича *автоматик таржимани* амалга ошира олади. **Сократ системаси** 40 млн. та сўз хотира кучига эга. Дунёдаги 40 та тилда *автоматик таржимани* амалга оширади. Яъни қайси тилда **Windows операцион** версияси мавжуд бўлса, шу тилда *автоматик таржимани* амалга ошира олади. **Лингво** системаси (лингво компьютер луғати) – бу луғатга ҳам дунёнинг 40 тилидаги маълумотлари жамланган.

3-
илова

1. Қандай тил ўргатувчи ўқув дастурларини биласиз?
2. Тил ўргатувчи ўқув дастурларининг самарадорлиги қандай бўлади?
3. Бошқа турдаги ўқув дастурлари ҳақида ёзинг.

4-илова

АТ йўналишининг замонавий ҳолати ва истиқболлари мавзусида эссе ёзиш.

Визуал материаллар

1-илова

Автоматик таржима бўйича дастлабки ғоялар инглиз олими **Чарльз Беббидж** томонидан 1836-1848 йилларда илгари сурилган. Унинг фикрича, механик-электрон машиналар кодлаштирилган *автоматик таржимани* амалга ошириши мумкин бўлади. Унинг лойиҳаси бўйича 50минг та сўз хотира кучига эга бўлган электрон машиналар 100000 та сўзни автоматик равишда таржима қилиши зарур бўлади. **Бebbиджнинг** ғоялари орадан 100 йил ўтгачгина амалга ошди.

1954-йилда АҚШдаги Жорж Даун Университетида дунёда биринчи марта автоматик таржима байиё тажриба бўлди. Бу тажрибалар IBM-701 компьютерида ўтказилди. Бу тажрибанинг базаси алгоритмлардан иборат бўлиб 50000 та сўзни рус тилидан инглиз тилига таржима қилди.

1960-йилда собиқ СССР фанлар Академиясининг *автоматик таржимага* бағишланган махсус йиғилиши бўлиб ўтди. Бу йиғилишда *автоматик таржима* бўйича профессор А.Белогонов раҳбарлигида махсус лаборатория ташкил қилинди.

Қозоғистонда профессор А. Бектаев раҳбарлигида махсус ишчи гуруҳ тузилиб, «Абай йўли» романини лингвостатистикаси ҳисоблаб чиқилди.

1966-йилда АҚШнинг фан масалалари бўйича конгресс қўмитасида *автоматик таржимага* оид маъруза муҳокамаси бўлиб ўтди. У автоматик таржима бўйича салбий маънодаги маъруза эди.

Шундан кейин жаҳон миқёсида *автоматик таржимага* оид тадқиқотларни молиялаштириш кескин камайиб кетди. Улар фақатгина Японияда давлат томонидан молиялаштирилади. Қолган давлатларда *автоматик таржимага* оид лойиҳалар хусусий йирик компаниялар ҳисобланаган «**Generel Motors**», «**Siemens**», «**LG**», «**Phillips**» компаниялари бугунги кунда дунёда автоматик таржимага оид изланишларни маблағ билан таъминлаб туради.

Автоматик таржима бўйича дунёдаги энг йирик лойиҳа 1991 йилда АҚШда **NASA** мутахассислари томонидан яратилган: **DIANA**. Унинг ҳажми 10 миллион та сўзни ташкил этади. У дунёнинг 6 та йирик тилига асосланади: *инглиз, рус, француз, немис, араб, испан.*

Рус олимлари ҳам профессор **Лякунов** раҳбарлигида «**Vista Tehnoloji**» компанияси «**Retrans Vista**» лойиҳасини ишлаб чиқди. Унинг таркибига кирувчи 3 млн. 400 мингта сўз хотирага киритилди. Шунда 1млн. 800мингтаси русча-инглизча, 1млн 600мингтаси инглизча-русча сўздан иборат.

Бугунги кунда *автоматик таржимани* қуйидаги системалари мавжуд. 40 млн.та сўз хотира кучига эга **EVROTRA** ситемаси – Европа олимлари томонидан ишлаб чиқилган. У барча Европа тиллари бўйича *автоматик таржимани* амалга ошира олади. **Сократ системаси** 40 млн. та сўз хотира кучига эга. Дунёдаги 40 та тилда *автоматик таржимани* амалга оширади. Яъни қайси тилда **Windows операцион** версияси мавжуд бўлса, шу тилда *автоматик таржимани* амалга ошира олади. **Лингво** системаси (лингво компьютер луғати) – бу луғатга ҳам дунёнинг 40 тилидаги маълумотлари жамланган.

3-
илова

1. Қандай тил ўргатувчи ўқув дастурларини биласиз?
2. Тил ўргатувчи ўқув дастурларининг самарадорлиги қандай бўлади?
3. Бошқа турдаги ўқув дастурлари ҳақида ёзинг.

4-илова

АТ йўналишининг замонавий ҳолати ва истиқболлари мавзусида эссе ёзиш.

илова

Б/Б/Б техникасини қўллаш бўйича кўрсатма

1. Маъруза режасига мос ҳолда 2-устунни тўлдилинг.
2. Ўйланг, жуфтликда ҳал этинг ва жавоб беринг, ушбу саволлар бўйича нимани биласиз,
- 3-устунни тўлдилинг.
3. Ўйланг, жуфтликда ҳал этинг ва жавоб беринг, ушбу саволлар бўйича нимани билиш керак, 4-устунни тўлдилинг.
4. Маърузани эшитинг.
5. 5-устунни тўлдилинг.

Б/Б/Б жадвали

	Тушунча	Биламан “+”, Билмайман	Билдим “+”,
1.	Тил ўргатувчи алгоритимлар		
2.	Тил ўргатувчи ўқув дастурлари		
3.	Билимларни баҳолаш		
4.	Маълумотлар базасини		

2-илова

Ҳозирги вақтда олий таълим муассасаларида янги педагогик технологияларнинг илмий асосларини яратиш, уларни таснифлаш, методик аҳамиятини белгилаш, таълим жараёнига татбиқ этиш каби долзарб муаммолар устида иш олиб борилмоқда. Илғор педагогик технологияларнинг компьютер технологиялари билан уйғунлаштириб ташкил этилиши ва шу асосда мультимедиа дарсларининг яратилиши эса асосий йўналишга айланиб қолмоқда.

Мультимедиа («кўп муҳитлик» деган маънони билдиради) замонавий компьютерларнинг деярли барчасида мавжуд бўлган имкониятлар мажмуини, комплекс тушунчани англатади. Мультимедиа ахборотнинг турли кўринишлари – матн, жадвал, графика, овоз, анимация (мультипликация), видеотасвир, мусиқа ёрдамида ахборотни йиғиш, сақлаш ва қайта узатиш вазифаларини бажаради. Мультимедиа «инсон-компьютер» интерактив (диалогик) мулоқотнинг янги, такомиллашган поғонаси бўлиб, бунда фойдаланувчи жуда кенг ва ҳар томонлама ахборот олади.

Электрон мультимедиа дарслиги – айнан компьютер воситасида ўқув курсини ёки унинг битта катта бўлимини ўқитувчи ёрдамида ёки мустақил ўзлаштиришни таъминловчи дастурий-методик комплекс.

Электрон энциклопедия – кучли қидирув тизимида эга бўлган маълум йўналишдаги иллюстрацияли мақолалар, электрон хариталар, фотосуратлар, гиперхавола (гипермуружаат)лар, аудио- ва видеоизохлар, хронологик рўйхат, фойдаланилган ва тавсия этиладиган адабиётлар рўйхати ва ҳоказолардан иборат бўлган катта ҳажмдаги электрон дастур.

Электрон лугат – одатдаги луғатлар асосида тузилган, мультимедиа воситалари билан бойитилган дастур. Электрон лугат турли йўналишда тузилиши мумкин, ундан талабанинг сўз бойлигини ошириш, таржима қилишда фойдаланилади.

Ўргатувчи аудиодастурлар, электрон лингафон курслари умуман иккинчи тилни ўрганувчилар учун ниҳоятда самарали таълим технологияларидир.

Универсал тест дастури ўзбек тили дарсларининг бирор йирик бўлими бўйича тест топшириқлари ва рейтинг тизимидан иборат бўлган компьютер дастуридир. Унинг қулайлиги шундаки, биргина дастурий таъминот асосида турли хил тест синовларини ўтказиш, яъни фақат матнни ўзгартириб, дастурни ҳар хил мавзуларда қўллаш мумкин.

Электрон виртуал кутубхона – мультимедиа хонасининг кенг имкониятларидан яна бири, электрон нусха ёки электрон мультимедиа дарсликлари мужассамланган, тармоққа уланган кутубхона.

Интернет тармоғидан ўзбек тилини ўрганувчи русийзабон талабалар унумли фойдаланишлари мумкин: мустақил таълим жараёнида маълумотлар тўплаш, мавзуларни ўзлаштириш, берилган мавзуда реферат, иншо ёзиш ва бошқалар. Шунингдек, электрон почта - ахборотни бир компьютердан иккинчисига юборишни таъминловчи тизим орқали ўқитувчи ёки тенгдошлари билан мулоқотда бўлиши мумкин.

Видеоматериаллар билан ишлаш тил ўқитишда алоҳида аҳамият касб этади.

Масофавий таълим – бу масофадан Интернет тармоғи орқали ўқитиш усули бўлиб, унинг асосида республикада маълум соҳалар бўйича тест-тажрибалар ўтказилмоқда.

Эссе - бу муаллифнинг шахсий нуқтаи назарини ёзма равишда эркин ифода этиш шакли; қандайдир предмет бўйича умумий ёки дастлабки дунёқарашни ўз ичига олади

6. Маъруза режасига мос ҳолда 2-устунни тўлдилинг.

7. Ўйланг, жуфтликда ҳал этинг ва жавоб беринг, ушбу саволлар бўйича нимани биласиз, 3-устунни тўлдилинг.

8. Ўйланг, жуфтликда ҳал этинг ва жавоб беринг, ушбу саволлар бўйича нимани билиш керак, 4-устунни тўлдилинг.

9. Маърузани эшитинг.

10. 5-устунни тўлдилинг.

Б/Б/Б жадвали

	Тушунча	Биламан “+”, Билмайман	Билдим “+”,
	Автоматик таҳрир		
	Информацион қидирув тизимлари		
	Гипертекст технологияси		

Топшириқлар

1. Автоматик таҳрирнинг прагматик жиҳатларини тушунтиринг.
2. Автоматик таҳрирнинг ишончлилик даражаси нималарга боғлиқ?
3. Автоматик таҳрирда маълум салбий жиҳатларини тушунтиринг.
4. Гипертекст технологияси ва анъанавий китоб босиш технологиясини солиштиринг.
5. Қандай информацион қидирув тизимларини биласиз?
6. Дескриптор нима?

3-илова

ТЕКСТ	ГИПЕРТЕКСТ
<i>Тугалланганлик</i>	Тугалланмаганлик
<i>Чизиқлилик</i>	Ночизиқлилик
<i>Аниқ муаллифнинг мавжудлиги</i>	Аниқ муаллифнинг мавжуд эмаслиги
<i>Бир томонлама йўналганлик</i>	Кўп томонлама йўналганлик
<i>Бир жинслилик (бир хилда тузилганлик)</i>	Бир жинсда тузилмаганлик (турфа хиллилик)
<i>Ёпиқ структурага эгалик</i>	Очиқ структурага эгалик

ЭССЕ

Эссе - таклиф этилган мавзуга 50 дан 100 гача сўз ҳажмидаги иншо.

Эссе - бу муаллифнинг шахсий нуқтаи назарини ёзма равишда эркин ифода этиш шакли; қандайдир предмет бўйича умумий ёки дастлабки дунёқарашни ўз ичига олади

Беш дақиқалик эссе

Беш дақиқалик эссе - ўрганилаётган мавзу бўйича олинган билимларни умумлаштириш, мушоҳада қилиш мақсадида ўқув машғулотида охирида 5 дақиқа оралиғида олиб борилади

Маъруза машғулоти юзасидан топшириқлар.

1. Компьютер лингвистикаси фанининг замонавий йўналишлари ҳақида Интернет маълумотларини топинг.
2. Лингвистик компьютер дастурларининг ишлаш самарадорлигини ошириш учун нималар қилиш лозим?
3. Эксперт тизимлар қайси соҳаларда ишлатилмоқда?
4. Сунъий интеллект тизими бўйича эришилган ютуқларга нималарни киритиш мумкин?
5. Сунъий интеллект тизимини яратиш юзасидан қандай ёндашувлар мавжуд?

2-илова

1. Мулоҳазалар формал назарияси, мантик функциялари нима?
2. Тилнинг ифода ва мазмун томонлари нимадан иборат?

3. Ушбу тушунчаларнинг мазмунини ёритинг.

семиотика, семиология, белгилар системаси,
тилнинг ифода томони, тилнинг мазмун томони

	Сигнификат ва денотат		
--	-----------------------	--	--

1. Битта гап билан саволга жавоб шакллантиринг.

Лексикография, лексикология, этимология, ономазиология, диалектология ўртасида қандай боғлиқлик бор?

2. Луғатшунослик фани асосчиларининг тарихий шажарасини тузинг.

3. Ушбу тушунчаларнинг мазмунини ёритинг.

- луғатдаги сўзларнинг маъно қатламларига ажралиши;
- луғат таркибининг ижтимоий-шева қатлами
- луғат таркибининг актив ва пассив лексикаси қатлами
- тил луғат бойлигининг кенгайиши

Ўқув топшириқлар

1-илова

Гуруҳ билан ишлаш қоидалари

Гуруҳ аъзоларининг ҳар бири

- ўз шерикларининг фикрларини ҳурмат қилишлари лозим;
- берилган топшириқлар бўйича фаол, ҳамкорликда ва масъулият билан ишлашлари лозим;
- ўзларига ёрдам керак бўлганда сўрашлари лозим;
- ёрдам сўраганларда кўмак беришлари лозим;
- гуруҳни баҳолаш жараёнида иштирок этишлари лозим;
- “Кемага тушганнинг жони бир” шиорини яхши билишлари лозим;

2-илова

1. “Гап турлари”ни асос қилган ҳолда жадвал тузинг.

2. Гап бўлақларидаги бош ва иккинчи даражали бўлақларни қамраб олган схема тузинг.

3. Ушбу тушунчаларнинг мазмунини ёритинг.

кичик синтаксис, гап предикативлик хусусияти, гап турлари, гап бўлақлари, бош бўлақлар

Гуруҳ билан ишлаш қоидалари

Гуруҳ аъзоларининг ҳар бири

- ўз шерикларининг фикрларини ҳурмат қилишлари лозим;
- берилган топшириқлар бўйича фаол, ҳамкорликда ва масъулият билан ишлашлари лозим;
- ўзларига ёрдам керак бўлганда сўрашлари лозим;
- ёрдам сўраганларда кўмак беришлари лозим;
- гуруҳни баҳолаш жараёнида иштирок этишлари лозим;
- “Кемага тушганнинг жони бир” шиорини яхши билишлари лозим;

1. Қуйидаги саволга жавобни шакллантиринг.

Дунёда нечта тиллар бор? Табиий тил, сунъий тил, математик тил деганда қандай тилни тушунаси?

2. Берилган саволга жавоб беринг:

Морфологик типологиянинг генетик типологияга алоқаси борми?

3. Ушбу тушунчаларнинг мазмунини ёритинг.

Лингвистик масала, табиий тил, сунъий тил, математик тил, алгоритм, дастур,

Гуруҳ билан ишлаш қоидалари

Гуруҳ аъзоларининг ҳар бири

- ўз шерикларининг фикрларини ҳурмат қилишлари лозим;
- берилган топшириқлар бўйича фаол, ҳамкорликда ва масъулият билан ишлашлари лозим;
- ўзларига ёрдам керак бўлганда сўрашлари лозим;
- ёрдам сўраганларда кўмак беришлари лозим;
- гуруҳни баҳолаш жараёнида иштирок этишлари лозим;
- “Кемага тушганнинг жони бир” шиорини яхши билишлари лозим;

1. Битта гап билан саволга жавобни шакллантиринг.

Лексикографияга оид терминлардан мисоллар келтиринг.

2. Лексикологияга оид терминларни схема билан изоҳлаб беринг.

3. Ушбу тушунчаларнинг мазмунини ёритинг.

компьютер луғатида туб сўзларнинг берилиши, ясама сўзларнинг берилиш усуллари, атокли отлар, антропонимлар, зоонимлар, топонимлар, баркарор бирикмалар, фраземалар,

4.4 “KOMPYUTER LINGVISTIKASI” FANIDAN TALABALAR BILIMINI BAXOLASH MEZONI.

“Kompyuter lingvistikasi” fani bo'yicha reyting jadvallari, nazorat turi, shakli, soni xamda xar bir nazoratga ajratilgan maksimal ball, shuningdek joriy va oraliq nazoratlarining saralash ballari xakidagi ma'lumotlar fan bo'yicha birinchi mashg'ulotda talabalarga e'lon kilinadi.

Fan bo'yicha talabalarining bilim saviyasi va o'zlashtirish darajasining Davlat ta'lim standartlariga muvofiqligini ta'minlash uchun quyidagi nazorat turlari o'tkaziladi:

- **joriy nazorat (JN)** - talabaning fan mavzulari bo'yicha bilim va amaliy kunikma darajasini aniqlash va baxolash usuli. Joriy nazorat fanning xususiyatidan kelib chikkan holda amaliy mashg'ulotlarda og'zaki surov, test o'tkazish, suxbat, nazorat ishi, kollektivum, uy vazifalarini tekshirish va shu kabi bopuqa shakllarda o'tkazilishi mumkin;

- **oraliq nazorat (ON)** - semestr davomida o'quv dasturining tegishli (fanlarning bir necha mavzularini uz ichiga olgan) bo'limi tugallangandan keyin talabaning nazariy bilim va amaliy kunikma darajasini aniqlash va baxolash usuli. Oraliq nazorat bir semestrda ikki marta o'tkaziladi va shakli (yozma, og'zaki, test va xokazo) o'quv faniga ajratilgan umumiy soatlar xajmidan kelib chikkan x,olda belgilanadi;

- **yakuniy nazorat (YaN)** - semestr yakunida muayyan fan bo'yicha nazariy bilim va amaliy kunikmalarni talabalar tomonidan o'zlashtirish darajasini baxolash usuli. Yakuniy nazorat asosan tayanch tushuncha va iboralarga asoslangan “Yozma ish” shaklida o'tkaziladi.

ON o'tkazish jarayoni kafedra mudiri tomonidan tuzilgan komissiya ishtirokida muntazam ravishda o'rganib boriladi va uni o'tkazish tartiblari buzilgan xollarda, **ON** natijalari bekor kilinishi mumkin. Bunday xollarda **ON** qayta o'tkaziladi.

Oliy ta'lim muassasasi raxbarining buyrugi bilan ichki nazorat va monitoring bo'limi raxbarligida tuzilgan komissiya ishtirokida **YaN** ni o'tkazish jarayoni muntazam ravishda o'rganib boriladi va uni o'tkazish tartiblari buzilgan xollarda, **YaN** natijalari bekor kilinishi mumkin. Bunday xollarda **YaN** qayta o'tkaziladi.

Talabaning bilim saviyasi, kunikma va malakalarini nazorat kilishning reyting tizimi asosida talabaning fan bo'yicha o'zlashtirish darajasi ballar orqali ifodalanadi.

«Kompyuter lingvistikasi» fani bo'yicha talabalarining semestr davomidagi o'zlashtirish kursatkichi 100 ballik tizimda baholanadi.

Ushbu 100 ball baxolash turlari bo'yicha quyidagicha taqsimlanadi: Ya.N.-30 ball, kolgan 70 ball esa J.N.-35 ball va O.N.-35 ball kilib taqsimlanadi.

Ball	Baho	Talabalarining bilim darajasi
86-100	A'lo	Xulosa va qaror kabul kilish. Ijodiy fikrlay olish. Mustaqil mushoxada yurita olish. Olgan bilimlarini amalda kullay olish. Moxiyatini tushuntirish. Bilish, aytib berish. Tasavvurga ega bulish.
71-85	Yaxshi	Mustaqil mushoxada kilish. Olgan bilimlarini amalda kullay olish. Moxiyatini tushuntirish. Bilish, aytib berish. Tasavvurga ega bulish.
55-70	Konikarli	Moxiyatini tushuntirish. Bilish, aytib berish Tasavvurga ega bulish.
0-54	Konikarsi	Aniq tasavvurga ega bulmaslik. Bilmaslik.

• Fan bo'yicha saralash bali 55 ballni tashkil etadi. Talabani saralash balidan past bo'lgan o'zlashtirishi reyting daftarchasida kayd etilmaydi.

• Talabalarining o'quv fani bo'yicha mustaqil ishi joriy, oraliq va yakuniy nazoratlar jarayonida tegishli topshiriklarni bajarishi va unga ajratilgan ball ar dan kelib chikkan x,olda baxolanadi.

Talabani fan bo'yicha reytingi quyidagicha aniqdanadi: $R = \frac{V \bullet O'}{100}$

bu erda: V- semestrda fanga ajratilgan umumiy o'quv yuklamasi (soatlarda);
O'-fan bo'yicha o'zlashtirish darajasi (ballarda).

• Fan bo'yicha joriy va oraliq nazoratlarga ajratilgan umumiy ballning 55 foizi saralash ball hisoblanib, ushbu foizdan kam ball tuplagan talaba yakuniy nazoratga kiritilmaydi.

• Joriy **JN** va oraliq **ON** turlari bo'yicha 55bal va undan yuqori balni tuplagan talaba fanni o'zlashtirgan deb hisoblanadi va ushbu fan bo'yicha yakuniy nazoratga kirmasligiga yul kuyiladi.

• Talabani semestr davomida fan bo'yicha tuplagan umumiy bali xar bir nazorat turidan belgilangan qoidalarga muvofiq tuplagan ballari yigindisiga teng.

• **ON** va **YaN** turlari kalendar tematik rejaga muvofiq dekanat tomonidan tuzilgan reyting nazorat jadvallari asosida o'tkaziladi. **YaN** semestrning oxirgi 2 xaftasi mobaynida o'tkaziladi.

• **JN** va **ON** nazoratlarda saralash balidan kam ball tuplagan va uzrli sabablarga ko'ra nazoratlarda katnasha olmagan talabaga qayta topshirish uchun, navbatdagi shu nazorat turigacha, sunggi joriy va oraliq nazoratlar uchun esa yakuniy nazoratgacha bo'lgan muddat beriladi.

• Talabani semestrda **JN** va **ON** turlari bo'yicha tuplagan ballari ushbu

nazorat turlari umumiy balining 55 foizidan kam bulsa yoki semestr yakuniy joriy, oraliq va yakuniy nazorat turlari bo'yicha tuplagan ballari yigindisi 55 balidan kam bulsa, u akademik karzdor deb hisoblanadi.

- Talaba nazorat natijalaridan norozi bulsa, fan bo'yicha nazorat turi natijalari e'lon kilingan vaktdan boshlab bir kun mobaynida fakultet dekaniga ariza bilan murojaat etishi mumkin. Bunday holda fakultet dekanining takdimnomasiga ko'ra rektor buyrugi bilan 3 (uch) a'zodan kam bulmagan tarkibda apellyatsiya komissiyasi tashkil etiladi.

- Apellyatsiya komissiyasi talabalarining arizalarini kurib chikib, shu kunning uzida xulosasini bildiradi.

Baxolashning urnatilgan talablar asosida belgilangan muddatlarda o'tkazilishi xamda rasmiylashtirilishi fakultet dekani, kafedra muduri, o'quv-uslubiy boshkarma xamda ichki nazorat va monitoring bo'limi tomonidan nazorat qilinadi.

Yakuniy nazorat "Yozma ish" shaklida belgilangan bulsa, u holda yakuniy nazorat 30 ballik "Yozma ish" variantlari asosida o'tkaziladi.

Agar yakuniy nazorat markazlashgan test asosida tashkil etilgan bo'lib fan bo'yicha yakuniy nazorat "Yozma ish" shaklida belgilangan bulsa, u holda yakuniy nazorat quyidagi jadval asosida amalga oshiriladi

Yakuniy nazoratda "Yozma ish"larni baxolash mezonlari

Yakuniy nazorat "Yozma ish" shaklida amalga oshirilganda, sinov ko'p variantli usulda o'tkaziladi. Har bir variant 2 ta nazariy savoldan iborat. Nazariy savollar fan bo'yicha tayanch so'z va iboralar asosida tuzilgan bo'lib, fanning barcha mavzularini uz ichiga kamrab olgan.

Xar bir nazariy savolga yozilgan javoblar bo'yicha o'zlashtirish kursatkichi 0-15 ball oraligida baxolanadi. Talaba maksimal 30 ball tuplashi mumkin.

Yozma sinov bo'yicha umumiy o'zlashtirish kursatkichini aniqlash uchun variantda berilgan savollarning xar biri uchun yozilgan javoblarga kuyilgan o'zlashtirish ballari kushiladi va yigindi talabaning yakuniy nazorat bo'yicha o'zlashtirish bali hisoblanadi.

REYTING BAHOLASH JADVALI

Reyting baholash

Semestr	-	III
Maksimal bal	-	100
Oraliq baholash	-	20
Joriy baholash	-	30
Mustaqil ta'lim	-	20
Yakuniy baholash	-	30

	N		semestr xaftalari
--	---	--	-------------------

	azorat turi	ami ball	Nazariy o'qish																			aku iy axo lash
											0	1	2	3	4	5	6	7	8	9	0	
	J oriy baholash	0									0				0					0		
	O raliq baholash	0						0					0									
	M ustaqil ta'lim	0										0							0			
	Y akuniy baholash	0																				0

JN uchun qo'yiladigan ballar mezonlari

O'zlashtirish bahosi		Talabani bilimi va faoliyatiga qo'yiladigan talablar	
-10	9 A'lo	Mavzu bo'yicha berilgan toshiriqlarga ijodiy yondoshgan, mustaqil fikrlaydigan, mavzuni yoritishda ishonchli yangi ma'lumotlardan, internet saytlaridan foydalangan, fikrni aniq, tushunarli ifodalay oladigan, toshiriqlarni bajarish bo'yicha 86% va undan yuqori natijaga erishgan talabalarga qo'yiladi.	
-8	7 Yaxshi	Mavzu bo'yicha berilgan toshiriqlarga ijodiy yondoshgan, mavzuni yoritishda yangi ma'lumotlardan foydalangan, fikrni aniq, tushunarli ifodalay oladigan, toshiriqlarni bajarish bo'yicha 71-85 % natijaga erishgan talabalarga qo'yiladi.	
	6 Qoniqarli	Mavzu bo'yicha berilgan toshiriqlarni bajargan, fikrni aniq, tushunarli ifodalay olmaydigan, toshiriqlarni bajarish bo'yicha 55-70 % natijaga erishgan talabalarga qo'yiladi.	
-5	0 Qoniqarsiz	Mavzu bo'yicha berilgan toshiriqlarni bajarmagan, fikrni aniq, tushunarli ifodalay olmaydigan, toshiriqlarni bajarish bo'yicha 54 % dan kam natijaga erishgan talabalarga qo'yiladi.	

Talabalar bilimini reyting tizimi asosida baholashda ON uchun qo'yiladigan ballar mezonlari

all	B	O'zlashtirish bahosi	Talabani bilimi va faoliyatiga qo'yiladigan talablar
1 8-20		A'lo	O'quv fani bo'yicha aniq bilim va tasavvurga ega bo'lgan, mustaqil ijodiy fikrlay oladigan, ON toshiriqlaridan kamida 86% ball to'lagan talabaga qo'yiladi.

1 5-17	Yaxshi	O'quv fani bo'yicha bilim va tasavvurga ega bo'lgan, mustaqil fikrlay oladigan, ON toshiriqlaridan 71-85% ball to'lagan talabaga qo'yiladi.
1 1-14	Qoniqarli	O'quv fani bo'yicha tasavvurga ega bo'lgan, ON toshiriqlaridan kamida 55-70% ball to'lagan talabaga qo'yiladi.
0 -10	Qoniqarsi z	O'quv fani bo'yicha tasavvurga ega bo'lmagan, ON toshiriqlaridan 54% dan kam ball to'lagan talabaga qo'yiladi.

**Talabalar bilimini reyting tizimi asosida baholashda
YaN uchun qo'yiladigan ballar mezon**

B all	O'zlash tirish bahosi	Talabaning bilimi va faoliyatiga qo'yiladigan talablar
2 6-30	A'lo	O'quv fani bo'yicha aniq bilim va tasavvurga ega bo'lgan, mustaqil ijodiy fikrlay oladigan, bilimlarni amaliyotga qo'llay oladigan, YaN savollariga to'liq ilmiy asoslangan javob bera oladigan, savollar va tayanch tushunchalar mohiyatini ijodiy yondoshgan holda qisqa va lo'nda ifodalab, bayonda mantiqiy yaxlitlikka erishadigan, toshiriqlarni kamida 86% ni o'zlashtirgan talabalarga qo'yiladi.
2 2-25	Yaxshi	O'quv fani bo'yicha bilim va tasavvurga ega bo'lgan, mustaqil fikrlay oladigan, YaN savollariga javob bera oladigan, berilgan savollarning mohiyatini yoritib bera oladigan, lekin uni ilmiy asoslay olmagan, toshiriqlarni 71-85 % qismini o'zlashtirgan talabalarga qo'yiladi.
1 7-21	Qoniqar li	O'quv fani bo'yicha tasavvurga ega bo'lgan, YaN savollariga to'liq javob bera olmagan, berilgan savollarning mohiyatini yuzaki (yoki qisman) yoritib bergan, toshiriqlarni 55-70 % qismini o'zlashtirgan talabalarga qo'yiladi.
0 -16	Qoniqar siz	O'quv fani bo'yicha tasavvurga ega emas, YaN savollariga javob bera olmagan, toshiriqlarni 54 % dan kam qismini o'zlashtirgan talabalarga qo'yiladi.

**Talabalar bilimini reyting tizimi asosida baholashda
mustaqil ta'lim uchun qo'yiladigan ballar mezon**

	O'zlas htirish bahosi	B all	Talabaning bilimi va faoliyatiga qo'yiladigan talablar
0	A'lo	9 -10	Mavzu bo'yicha berilgan topshiriqlarga ijodiy yondoshgan, mustaqil fikrlay oladigan, mavzuni yoritishda ishoncli yangi ma'lumotlardan, internet saytlaridan foydalangan, fikrni aniq, tushunarli ifodalay oladigan

			talabalarga qo'yiladi.
	Yaxshi	7 -8	Mavzu bo'yicha berilgan topshiriqlarga ijodiy yondoshgan, mavzuni yoritishda yangi ma'lumotlardan foydalangan, fikrni aniq, tushunarli ifodalay oladigan talabalarga qo'yiladi.
	Qoniqarli	6	Mavzu bo'yicha berilgan topshiriqlarni bajargan, fikrni aniq, tushunarli ifodalay olmaydigan talabalarga qo'yiladi.
	Qoniqarsiz	0 -5	Mavzu bo'yicha berilgan topshiriqlarni bajarmagan, fikrni aniq, tushunarli ifodalay olmaydigan talabalarga qo'yiladi.
0	A'lo	1 8-20	Mavzu bo'yicha berilgan topshiriqlarga ijodiy yondoshgan, mustaqil fikrlay oladigan, mavzuni yoritishda ishoncli yangi ma'lumotlardan, internet saytlaridan foydalangan, fikrni aniq, tushunarli ifodalay oladigan talabalarga qo'yiladi.
	Yaxshi	1 5-17	Mavzu bo'yicha berilgan topshiriqlarga ijodiy yondoshgan, mavzuni yoritishda yangi ma'lumotlardan foydalangan, fikrni aniq, tushunarli ifodalay oladigan talabalarga qo'yiladi.
	Qoniqarli	1 1-14	Mavzu bo'yicha berilgan topshiriqlarni bajargan, fikrni aniq, tushunarli ifodalay olmaydigan talabalarga qo'yiladi.
	Qoniqarsiz	0 -10	Mavzu bo'yicha berilgan topshiriqlarni bajarmagan, fikrni aniq, tushunarli ifodalay olmaydigan talabalarga qo'yiladi.