

O‘ZBEKISTON RESPUBLIKASI
OLIV VA O‘RTA MAXSUS TA‘LIM VAZIRLIGI
ALISHER NAVOIY NOMIDAGI TOSHKENT DAVLAT
O‘ZBEK TILI VA ADABIYOTI UNIVERSITETI

Xolmanova Zulxumor Turdiyevna

KOMPYUTER LINGVISTIKASI

5120100 – Filologiya va tillarni o‘qitish (o‘zbek tili) yo‘nalishi
uchun

Toshkent – 2019

Mazkur o'quv qo'llanmasi bakalavriat bosqichi talabalari, magistrantlar, katta ilmiy xodim-izlanuvchilar hamda o'qituvchilarga mo'ljallangan. O'quv qo'llanmasida kompyuter lingvistikasi yo'nalishining mohiyati, asosiy masalalari yoritilgan, kompyuter lingvistikasiga doir terminlar, tayanch tushunchalar izohi keltirilgan.

Данное учебное пособие предназначено для студентов бакалавриата, магистрантов, исследователей и преподавателей. В книге обсуждены основные вопросы компьютерной лингвистики, а также приложен глоссарий терминов, основных понятий данной дисциплины.

This handbook is referred to the bacelor-students philology, masters, researchers and techers of universities. In this book basic questions of computational linguistics are analyzed and interpreted, as well as the dictionary of key terms on computational linguistics is attached at the end of the book.

Mas'ul muharrir: **f.f.d., prof. H. Dadaboyev**

Taqrizchilar: **f.f. d., prof. S. Muhamedova**
f.f.d., prof. B.Mengliyev
f.-m.f.n., dots. I. Haydarov

Ushbu qo'llanma O'zbekiston Oliy va o'rta maxsus ta'lim vazirligining 2019-yil 2-noyabrdagi 1023-sonli buyrug'iga asosan nashrga tavsiya etilgan (qayd raqami: №1023-050).

Kirish

2016-yil 13-maydagi “Alisher Navoiy nomidagi Toshkent davlat o‘zbek tili va adabiyoti universitetini tashkil qilish to‘g‘risida”gi farmon (PF-4797) hamda O‘zbekiston Respublikasi Prezidentining “O‘zbekiston Respublikasini yanada rivojlantirish bo‘yicha Harakatlar strategiyasi” to‘g‘risidagi farmoni (2017 yil, 7 feval) talablaridan kelib chiqqan holda lingvistika nazariyasi va amaliy tilshunoslik mazmunini yoritish, nazariyani amaliyotga tatbiq qilish, til birliklari, til hodisalarining amaliy qiymatini aniqlash, fan doirasida o‘rganiladigan asosiy masalalarni belgilash, nazariy tilshunoslikning muhim xususiyatlari haqida tasavvur hosil qilish, asosiy lingvistik hodisalar, jarayonlar, til birliklari, til birliklarining tizim sifatida mavjudligini asoslash, lingvistik strukturani nazariy muammolar va amaliy tilshunoslik asosida tahlil qilish kommunikatsiya tizimining asosiy vositasi bo‘lgan til mohiyatini yoritishda muhim ahamiyat kasb etadi.

Respublikamiz Prezidenti Sh.M.Mirziyoyevning ma’naviy-ma’rifiy masalalarga doir har bir chiqishlarida ajdodlar merosini qadrlash, ularning tafakkur mahsulidan unumli foydalanish, kitob mutolaasiga alohida e’tibor qaratish, shu asosda yoshlarni vatanga, millatga, tilga muhabbat ruhiyatida tarbiyalash masalalariga alohida ahamiyat berilmoqda:

“Yoshlarimizning mustaqil fikrlaydigan, yuksak intellektual va ma’naviy salohiyatga ega bo‘lib, dunyo miqyosida o‘z tengdoshlariga hech qaysi sohada bo‘sh kelmaydigan insonlar bo‘lib kamol topishi, baxtli bo‘lishi uchun davlatimiz va jamiyatimizning bor kuch imkoniyatlarini safarbar etamiz,” – degan so‘zlari biz uchun dasturulamal vazifasini o‘tamoqda¹.

Yoshlarning iqtidorli, har tomonlama yetuk mutaxassis sifatida shakllanishlari, avvalo, til haqidagi bilim hamda nutqiy ko‘nikmalarga ega bo‘lishlari bilan belgilanadi. Olib borilayotgan islohotlar samarasini yanada oshirish, davlat va jamiyatning har tomonlama va jadal rivojlanishi uchun shart-sharoitlar yaratish, mamlakatimizni modernizatsiya qilish hamda hayotning barcha sohalarini liberallashtirish bo‘yicha ustuvor yo‘nalishlarni amalga oshirish maqsadida Prezident Sh.M. Mirziyoyev tomonidan 2017-2021-yillarda O‘zbekiston Respublikasini rivojlantirishning beshta ustuvor yo‘nalishi bo‘yicha Harakatlar strategiyasi tasdiqlandi:

¹ *Mirziyoyev Sh.M.* Erkin va farovon demokratik O‘zbekiston davlatini birgalikda barpo etamiz. – T.: O‘zbekiston, 2016. –B.14.

1. Davlat va jamiyat qurilishini takomillashtirishga yo'naltirilgan demokratik islohotlarni chuqurlashtirish va mamlakatni modernizatsiya qilish.

2. Qonun ustuvorligini ta'minlash va sud-huquq tizimini yanada isloh qilish.

3. Iqtisodiyotni yanada rivojlantirish va liberallashtirish.

4. Ijtimoiy sohani rivojlantirish.

5. Xavfsizlik, millatlararo totuvlik va diniy bag'rikenglikni ta'minlash, chuqur o'ylangan, o'zaro manfaatli va amaliy ruhdagi tashqi siyosat yuritish.

Harakatlar strategiyasini o'z vaqtida va samarali amalga oshirish barcha davlat hokimiyati va boshqaruvi organlari hamda ularning mansabdor shaxslari faoliyatining birlamchi vazifasi va bosh ustuvor yo'nalishi hisoblanadi.

Harakatlar strategiyasi barcha ijtimoiy sohalar kabi tilshunoslikka doir tadqiqotlarning asosiy yo'nalishlarni belgilab berdi. Tilshunoslikning o'rganish obyekti bo'lgan til barcha davrlarda jamiyat hayotida muhim o'rin tutib kelgan. Til ijtimoiy taraqqiyot mahsuli bo'lib, shakllanishi, boyib borishi va ayrim hollarda iste'moldan chiqishi bilan tabiiylik xususiyatini namoyon etadi. Tilning asosiy ijtimoiy funksiyalaridan biri kommunikativlikdir. Tilning emotsional-ekspressiv, akkumulyativ funksiyalari kommunikativlik asosiga qurilgan. Shuni ham ta'kidlash lozimki, tilning boshqa funksiyalari ham kommunikativlik bilan teng ahamiyat kasb etib bormoqda. Emotsional-ekspressiv funksiya insonlarning ruhiy-hissiy holatlarini ifoda etib, ma'naviy ehtiyojlarini qondirishga xizmat qilsa, akkumulyativ funksiya ajdodlar merosini avlodlarga yetkazish, mental xususiyatlarni, xalq dunyoqarashi, urf-odat, an'analari, turmush tarzi, madaniyat darajasini namoyon etishda o'z ifodasini topmoqda. Til tafakkur bilan bog'liq holda insonni barcha jonzotlardan ajratib turuvchi imkoniyat sifatida baholanadi. Tilsiz tafakkur, tafakkursiz til mavjud bo'la olmaydi. Til borliqni anglash, voqelik haqida muayyan tasavvurga ega bo'lish vositasidir. Tilning shakllanishi va rivojlanishi ong, tafakkur rivoji bilan bog'liq. Kishilik jamiyatida tilning tafakkur omili sifatidagi ahamiyati katta. Til jamiyat taraqqiyotini ta'minlash, o'zaro munosabatlarni takomillashtirishda o'ziga xos o'rin tutadi. Til tafakkur mahsuli, o'z navbatida til tafakkur tadrijiga xizmat qiladi. Til fikr manbayi ekan, u barcha sohalaridagi yangiliklar, ixtirolar, kashfiyotlar zamini hamdir. Bugungi kunda tilning ijtimoiy funksiyalari yanada kengaymoqda. Til fanlararo, sohalararo munosabat

asosi sifatida ham e'tiborga molikdir. Aynan mana shu ehtiyoj tufayli kompyuter lingvistikasi fani paydo bo'ldi. Kompyuter lingvistikasi kompyuter tizimi imkoniyatlarining tilshunoslik sohasiga tatbiq qilinishi natijasida yuzaga keldi. Kompyuter lingvistikasi umumtil asoslarini shakllantirishda, tillar o'rtasidagi munosabatlarni takomillashtirishda o'ziga xos o'rin tutadi.

Tilning murakkab fiziologik, biologik, ijtimoiy hodisa sifatidagi ahamiyati ikki jihatda namoyon bo'ladi. Dastavval, tilning ichki qonuniyat asosidagi taraqqiyotini, turli fonetik, leksik, semantik, grammatik jarayonlarni, mental xususiyatlarni mujassamlantirgan holdagi takomillashuvini ta'kidlash joiz. Til insonni tafakkur jihatdan rivojlantirish omili bo'lishi bilan bir qatorda tashqi jihatdan ta'minlash, bilim berish, borliqni anglatish, ya'ni axborot uzatish vositasi sifatida ham muhim qiymatga ega. Til insonni voqelik bilan muntazam ravishda tanishtirib boradi, davrga, ijtimoiy turmush tarziga mos ma'lumotlar bilan ta'minlaydi. Tilning tashqi taraqqiyot omili, axborot, ma'lumot berish vositasi sifatidagi ahamiyati globallashuv davrida, ayniqsa, yaqqolroq namoyon bo'lmoqda.

XX asr o'rtalaridan insoniyat tarixida tubdan burilish hosil qilgan globallashuv jarayoni boshlandi. "Globallashuv" so'zi lotincha "glob" "kurra", "shar" so'zidan olingan bo'lib, qit'alaro manfaatning jadal sur'atlarda yaqinlashuvi, birlashuvini anglatadi. Tez sur'atlar bilan uyg'unlashayotgan milliy iqtisod va milliy madaniyatlar uchun umumiy muloqot vositasiga ehtiyoj seziladi. Shunday vaziyatda milliy tillar iqtisodiy, siyosiy, ma'naviy fikrning aniqroq, to'laroq ifoda etilishiga xizmat qiladi. Globallashuv ta'sirida bir tilning o'zigina ijtimoiy ehtiyojlarni qondira olmaydi.¹

Global taraqqiyotning muhim xususiyatlari sohalarga texnikaning kirib kelishi, kompyuter dasturlarining shakllanishi, integratsiya jarayoni bilan belgilanadi. Hozirgi vaqtda jahon miqyosida kechayotgan integratsiya jarayonida tabiiy til asosidagi axborot uslubini yaratish hayotiy zaruratga aylandi. Bugungi kunda o'zbek tilining mavqeyini yuksaltirish, nufuzli tillar qatoriga olib chiqish muhim zaruratlardan biri hisoblanadi. Kompyuter lingvistikasi mana shu zaruratni amalga oshirish uchun yaratilgan qulay imkoniyatlardan biridir. O'zbek tilini jahon miqyosiga olib chiqish, dunyoviy tillar sirasiga kiritish va til o'rganish va o'rgatishda kompyuter lingvistikasi fani muhim ahamiyat kasb etadi.

¹Хоназаров Қ. Глобаллашув ва тил фалсафаси. - Т., 2009. –Б.29-30.

XXI asr boshiga kelib axborot texnologiyalari taraqqiyoti tufayli globallasuv jarayoni yangi bosqichga ko'tarildi. Insoniyat tomonidan qo'lga kiritilgan fan, texnika, madaniyat, ishlab chiqarish munosabatlarini rivojlantirish, demokratiya, qonunchilik, adolatni barqarorlashtirish borasidagi yutuqlardan oqilona foydalanish O'zbekistonni jahondagi rivojlangan davlatlardan biriga aylantirishga zamin yaratadi. Globallasuv jarayoni barcha sohalarda tezkorlikni, tez sur'atlar bilan rivojlanishni talab etmoqda. Texnika taraqqiyoti mahsuli bo'lgan kompyuter tizimi barcha sohalarda qulayliklar yaratadi, ma'lumotlarning tezkor yetkazib berilishini, tarjima, tahrir jarayonlarining qisqa muddatlarda amalga oshirilishini, turli millatga mansub muloqot vakillari o'rtasida vosita sifatida xizmat qiluvchi sun'iy til, ya'ni axborot-kompyuter uslubining shakllanishini ta'minlaydi.

Texnologiya asrida taraqqiyot sur'atining tezlashgani, intellektual salohiyat, texnikaning yuqori darajada yuksalishi kishilik jamiyati oldiga qator vazifalarni qo'ymoqda. Mintaqamizdagi globallasuv jarayoni barcha sohalarda tezkorlikni, jadal taraqqiyotni talab etmoqda. Kompyuter tizimi qulayliklar, imkoniyatlar majmuyiga va tarkibiy qismiga aylandi.¹

Axborot texnologiyalari ta'lim sohasida ham asosiy o'rin egallaydi. Ta'lim samaradorligini oshirishda, zamon talablariga javob bera oladigan kadrlar tayyorlashda axborot texnologiyalari muhim ahamiyat kasb etadi. Hozirda yurtimizda ta'lim sohasiga jiddiy e'tibor qaratilishi natijasida aksariyat yoshlar axborot texnologiyalaridan unumli foydalanmoqdalar, axborotlashtirish jarayonida mustaqil ishtirok etmoqdalar. Yangi texnologiyalar kun sayin rivojlanib, axborotlashtirish jarayoni tez sur'atlar bilan o'sib borayotgan hozirgi davrda ta'lim sohasida axborot resurslarini tashkil etish va ta'limda foydalanishga mamlakatimizda ham alohida e'tibor qaratilmoqda. Jumladan, O'zbekiston Respublikasi Birinchi Prezidentining «O'zbekiston Respublikasida «Elektron ta'lim» milliy tizimini yaratish» investitsiya loyihasini amalga oshirish chora-tadbirlari to'g'risida»gi 2012 yil 16 aprelda e'lon qilingan PQ-1740-son qarori ta'lim sohasida axborotlashtirishning milliy tizimini shakllantirish, zamonaviy axborot texnologiyalarini joriy etish va undan foydalanish, jahon axborot resurslaridan bahramand bo'lishga zamin yaratadi.

Ta'lim tizimini kompyuterlashtirishga qaratilgan bir qator kurslar yuzaga kelmoqda. Jumladan, moodle – bu o'qituvchi (kurs yaratuvchi)lar tomonidan onlayn-kurslarni yaratish uchun maxsus ishlab chiqilgan sayt

¹Пулатов А.К., Мухамедова С. Компьютер лингвистикаси (ўқув қўлланма). –Т., 2008; Пулатов А. Компьютер лингвистикаси. –Т., Академнашр, 2011; Rahimov A. Kompyuter lingvistikasi asoslari. –Т., Akademnashr, 2011.

tarkibini boshqarish tizimi (Content Management System – CMS) dir. Bu tizim o‘qitishni boshqarish tizimlari (Learning Management Systems – LMS) yoki virtual o‘qitish vositasi (Virtual Learning Environments – VLE) deb ham ataladi.

Ta’lim jarayoni kompyuterlashtirilgan ekan, bevosita kompyuter tizimidagi ma’lumotlar bazasi, elektron lug‘atlar, kompyuter lug‘atlari, avtomatik tarjima, avtomatik tahrir, til o‘qitish dasturlariga ehtiyoj tug‘iladi. Bu ehtiyojni qondirish kompyuter lingvistikasi sohasida izlanishlar olib borishni taqozo etadi.

Respublikamiz Prezidenti farmonlari va hukumat qarorlari ta’lim jarayonini takomillashtirish, milliy g‘oya bilan qurollangan yangi avlodni voyaga yetkazish davlat siyosati darajasidagi eng muhim vazifalaridan biri ekanligini asoslaydi. Yosh avlodni bilimli, komil inson va global taraqqiyot, jahon sivilizatsiyasi talablariga javob bera oladigan kadrlar sifatida tayyorlash hozirgi ta’lim tizimi maqsadlaridan biridir. Bu maqsadga erishish uchun, avvalo, yoshlar chet tillarini mukammal bilishlari, kompyuter texnikasi va zamonaviy texnologiyalarni mustaqil o‘zlashtirib borishlari lozim.

Kompyuter lingvistikasi fani globallashtirish davri muammolarini hal qilishga yo‘naltirilgan, yoshlarning ham texnikani, ham xorijiy tillarni mukammal o‘rganishlarini ta’minlaydigan soha hisoblanadi. Kompyuter lingvistikasining asosiy maqsadi lingvistik masalalarni yechishning kompyuter dasturlarini ishlab chiqish hamda bu dasturlar uchun ma’lumotlar bazasini yaratishdan iborat.

Kompyuter lingvistikasining o‘rganish obyekti, tahlil doirasi keng. Kompyuter lingvistikasi tilshunoslik nazariyasi va amaliyoti, kognitiv psixologiya, matematika, informatika va falsafa bilan uzviy bog‘liqdir. Kompyuter lingvistikasi qayd etilgan sohalarga oid masalalarni ham qamrab oladi. Kompyuter lingvistikasini o‘rganish tilning o‘ziga xos tizimini, ramzlar, belgilar, modellar ko‘rinishidagi qiymatini, ijtimoiy funksiyalarini anglatishga, yangi informatsion texnologiyalar bilan tanishtirishga, tilshunoslikning nazariy masalalari, kognitiv psixologiya, mantiq, falsafa, matematika bilan bog‘liq qonuniyatlarni o‘zlashtirishga xizmat qiladi.

Kompyuter lingvistikasi fani doirasida xorijda ingliz, rus, nemis, fransuz, ispan tillarida monografiyalar, darsliklar yaratilgan, ilmiy jurnallar nashr etilgan. O‘zbek tilshunosligida kompyuter lingvistikasi bo‘yicha tadqiqotlar keyingi yillardagina yaratilmoqda. Ushbu qo‘llanmani tayyorlashda “*Новое в зарубежной лингвистике*” nashri (вып: XXIV.

Компьютерная лингвистика,1989), A.N. Baranovning “Введение в прикладную лингвистику” (2001), A.Nurmonovning “*Struktur tilshunoslik: ildizlari va yo‘nalishlari*”, A.Po‘latov, S.Muhamedovalarning “*Kompyuter lingvistikasi*” (2008), B.Yo‘ldoshevning “*Kompyuter lingvistikasi*”(2009), A.Po‘latovning “*Kompyuter lingvistikasi*” (2011), A.Rahimovning “*Kompyuter lingvistikasi asoslari*” (2011), F.Qurbonovning “*Kompyuter lug‘atlari: tezaurus*” (2012), K.Boyarskiyning “Введение в компьютерную лингвистику” (Санкт-Петербург, 2013), L.Abduhamidovaning “*Tilshunoslikning yangi yo‘nalishi: kompyuter lingvistikasi*” (2015) kitoblari va internet materiallaridan foydalanildi. Mazkur o‘quv qo‘llanmasida kompyuter lingvistikasining ayrim masalalari borasidagi ma’lumotlar kengaytirildi, o‘zbek tilidagi mavjud adabiyotlarda qayd etilmagan masalalarga munosabat bildirildi. Jumladan, kompyuter lug‘atlari, tezaurus, konkordans lug‘atlar izohlandi, Yaponiya va Xitoydagi kompyuter lingvistikasi taraqqiyoti, avtomatik tahrir, sun’iy intellekt tushunchalari, BORIS tizimi (darak gapli matnlarni avtomatik tahrir qilish tizimi) haqida ilk marta batafsil ma’lumot berildi.

I. “KOMPYUTER LINGVISTIKASI”NING RIVOJLANISH TENDENSIYASI

1-§. “KOMPYUTER LINGVISTIKASI” FANINING MAQSADI, VAZIFALARI

Reja:

1. Matematik lingvistika va kompyuter lingvistikasi.
2. “Kompyuter lingvistikasi” fani doirasida o‘rganiladigan masalalar.
3. Kompyuter lingvistikasi – tilshunoslik taraqqiyoti mahsuli.

Asosiy tushunchalar: *til algebrasi, matematik lingvistika, struktur lingvistika, hisoblash lingvistikasi kompyuter lingvistikasi, tabiiy tillarning matematik modeli, muloqotning optimal modeli, mashina tarjimasi, metatil, sun’iy til, analiz, sintez, tabiiy tilni qayta ishlash (NLP), injener lingvistikasi, generatsiya, tillarni o‘qitish, bilimlarni baholash, matnlarni tahrir qilish, nogrammatik ifodalar tahriri, o‘zbek tilining kompyuter uslubi, matnlar korpusi, korpus lingvistikasi, elektron lug‘at, tezaurus.*

XX asrning o‘rtalarida kibernetikaning rivojlanishi fan tarixida kuzatilmagan tutash sohalar kesimidagi yangi fanlarning shakllanishiga asos bo‘ldi. Jumladan, biologiya va injeneriya fanlari kesimida bionika fani yuzaga keldi. Hisoblash texnikasi va lingvistikasi kesimida bir qator nomlar bilan atalgan tutash soha vujudga keldi. Bu fan dastlab, *matematik lingvistika, struktur lingvistika, hisoblash lingvistikasi* deb nomlangan. Hozirda *kompyuter lingvistikasi* degan muqim nomga ega bo‘ldi. Bu fanning shakllanishida ikki omil qayd etiladi: 1) lingvist tadqiqotchilar zamonaviy aniq fanlar (asosan, matematika) lingvistikada yetishmayotgan aniqlikni ta’minlashda yordam berishiga umid bildirdilar. EHMning paydo bo‘lishi bu umidni mustahkamladi. Ko‘pgina tilshunoslarga avval boshdanoq kompyuter nafaqat tez ishlaydigan arifmometrilar, balki matn ustida bajariladigan amallarni avtomatlashtirishning qudratli vositasi ekanligi ma’lum bo‘ldi. Ko‘p mehnat talab qiladigan jarayonlarni avtomatlashtirish imkoniyati yuzaga keldi. Masalan, matnlarni statistik tahlil qilish, turli lug‘at va leksik kartotekalarni to‘plash avtomatik tizim asosida olib borildi; 2) kompyuter paydo bo‘lgach ular bilan aloqaga kirishishga hali tayyor bo‘lmagan foydalanuvchilarning muloqotini ta’minlash muammosi tug‘ildi. Bunday foydalanuvchilar uchun eng maqbuli kundalik faoliyatlaridagi tabiiy til hisoblanar edi. Ammo bu tarzdagi o‘zaro munosabatni tashkil etish uchun, dastavval, insonlarning

o‘zaro muloqot jarayonida ining tabiiy tildan foydalanish qonunlari va xususiyatlarini tushunish kerak bo‘lar edi. An’anaviy tilshunoslik bu masalalar bilan shug‘ullanmagani ma’lum bo‘ldi.

Kompyuter lingvistikasi – kompyuter uchun mo‘ljallangan lingvistik ta’minot (kompyuter lingvistika imkoniyatlari subyekt sifatida maydonga chiqadi);

Kompyuter lingvistikasi – kompyuterda bajariladigan lingvistika.

Kompyuter lingvistikasining muhim tushuncha kategoriyalaridan “freymlar” (vaziyatning tematik tipiklashgan bilimlar deklarativ namoyishi uchun mo‘ljallangan kotseptual strukturalar), “ssenariy” (stereotip vaziyat yoki stereotip xulq haqidagi bilimlar namoyishi uchun taqdim qilingan konseptual strukturalar), “plan” (aniq maqsadlardagi yutuqlarga olib boruvchi bilim strukturasi, imkoniyatlar) bir qator masalalarning yoritilishiga xizmat qiladi.¹

“Kompyuter lingvistikasi”ning maqsadi quyidagilarda namoyon bo‘ladi:

a) kompyuter lingvistikasi fani inson va mashina (kompyuter)ning o‘zaro aloqa-munosabatga kirishuviga asos bo‘ladi;

b) turli hududlarda yashayotgan kishilarning o‘zaro maqsadli muloqotiga imkoniyat yaratiladi;

d) kompyuter lingvistikasining asosiy ish vositasi hisoblangan axborot uslubi dunyoda yuz berayotgan siyosiy, ijtimoiy, iqtisodiy, ilmiy, madaniy yangiliklarning tezkor va aniq shaklda yetkazilishini ta’minlaydi.

Kompyuter lingvistikasining asosiy vazifalari tillarni o‘qitish, bilimlarni tekshirish, matnlarni tahrirlash, tahlil qilish, avtomatik tahrir, mashina tarjimasini dasturlarini ishlab chiqishdan iboratdir.

Kompyuter lingvistikasi kompyuterdan unumli foydalanish hamda lingvistikaga doir masalalar (axborot uslubini o‘zlashtirish, tilning funksional doirasi haqida bilimga ega bo‘lish, tillarni o‘qitish, bilimlarni baholash, matnlarni tahrir va tahlil qilish, bir tildan ikkinchi tilga tarjima qilish)ni kompyuter vositasida hal qilish yo‘llarini belgilash, kompyuter savodxonligi darajasini oshirish, mantiqiy to‘g‘ri, izchil fikrlashga o‘rgatish, nazariy bilimlarni shakllantirish va muayyan yo‘nalishlar bo‘yicha amalda qo‘llashga oid ko‘nikmalarni hosil qilishni nazarda tutadi.

Shuningdek, ushbu kurs talabalarda kompyuter lingvistikasi bo‘yicha chuqur bilimlarni shakllantirish, kompyuter lingvistikasi tamoyillarini o‘zlashtirish, tabiiy tilni matematik modellashtirish, formallashtirish, matematik aniqlikdagi fikrlashga o‘rgatish, til va adabiyotga doir

¹ Боярский К. Введение в компьютерную лингвистику. Санкт-Петербург, 2013.-С.4

kompyuter dasturlari uchun ma'lumotlar bazasini yaratish, algoritm va kompyuter dasturlari tuzish asoslarini ishlab chiqish kabi vazifalarni belgilaydi.

“Kompyuter lingvistikasi” fanining asosiy masalalari doirasida ma'lumotlar berishda:

- kompyuter lingvistikasi predmeti, vazifalari; kompyuter lingvistikasining ijtimoiy va tabiiy fanlar orasidagi o'rni va mohiyati; kompyuter lingvistikasiga doir tadqiqotlar; aksiomatik nazariyaning tilshunoslikka tatbiqi; matematik mantiq elementlari; til o'rgatish jarayonini avtomatlashtirish; tabiiy tillarni qayta ishlash, tillarning matematik modelini yaratish; avtomatik tahrir va tahlil haqida tasavvur hosil qilish;

- matematik mantiq asoslarini; matematik mantiq asoslarining tilshunoslikka tatbiqini; avtomatik 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 o'zlashtirishni; o'zbek tili grammatikasining formallashtiradigan qismi uchun model tuzishni bilish va ulardan foydalana olish;

- 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; matnni avtomatik tarjima qilish; matnni avtomatik tahrir va tahlil qilish; kompyuter lingvistikasiga oid tadqiqotlarni o'rganish, nazariy ma'lumotlarni umumlashtirish; kompyuter lingvistikasi fanining istiqboli va dolzarb masalalariga doir fikrlar bildirish va tadqiqot olib borish ko'nikmalariga ega bo'lish kabilar nazarda tutiladi.

Kompyuter lingvistikasi tabiiy tilning kompyuter imkoniyatlari asosida reallashadigan ko'rinishidir. Sotsiolingvistika, psixolingvistika, pragmalingvistika, kognitiv lingvistika tilshunoslik sohalari bo'lgani kabi kompyuter lingvistikasi ham tilshunoslikning yangi yo'nalishidir. Kompyuter lingvistikasi – informatsion texnologiyalar sohasi bilan chambarchas bog'liq yangi yo'nalish hisoblanadi. Keyingi yillarda yaratilgan dasturlar – qidiruv mashinasi, dialog tizimi, axborot boshqaruvini avtomatlashtirish, matnlarni avtomatik tahrirlashni takomillashtirish til, til tizimi imkoniyatlaridan maqsadga muvofiq foydalanishni talab qiladi. Kompyuter lingvistikasining asosiy o'rganish doirasini aynan ijtimoiy taraqqiyot asosidagi ehtiyoj – mashina tushuna

oladigan va qayta ishlashi mumkin bo'lgan tabiiy til modelini yaratish tashkil etadi.

Til – eng muhim muloqot vositasidir. Til o'ziga xos tarixiy taraqqiyot tizimiga, murakkab strukturasi ega bo'lishiga qaramasdan, ona tili bolalar tomonidan qisqa muddatda tez o'zlashtiriladi va til tafakkurning ajralmas qismiga aylanadi. Hech mubolag'asiz aytish mumkinki, til insoniyatning kishilik taraqqiyotiga muntazam ravishda ta'sir ko'rsatadigan mahsullaridan biridir. A.Rahimov Beryozinning quyidagi fikrlarini keltirgan: *“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”*.¹

Kishilik taraqqiyotining har bir davrida fan-texnika yangiliklari muhim ahamiyat kasb etib kelgan. Insoniyat uchinchi ming yillikka, axborot asriga qadam qo'ydi. Inson uchun ma'lumot olish zaruriyatga aylandi. Axborot – ma'lumot oltindan ham qimmatroq deb topildi. Axborotni topish, saqlash, qayta ishlash va boshqalarga yetkazishning qulay usullariga bo'lgan ehtiyoj kun sayin ortib bormoqda. Bu esa XX asrning buyuk kashfiyoti bo'lgan kompyuter va kompyuter texnologiyalari sohasi uchun muhim vazifalarni belgilab berdi. Kompyuter texnologiyalari xalq xo'jaligi, sport, san'at, tibbiyot, umuman, ijtimoiy hayotning barcha tarmoqlariga kirib keldi. Fan yo'nalishlarini kompyuterlashtirish ilmning asosiy shartlaridan bo'lgan *obyektivlik* va *aniqlikka* amal qilish asosida tez rivojlana boshladi va barcha sohalarda yutuqlarga erishildi.

Dastavval, kompyuter tizimi matematik hisoblash amallarida qo'llanildi, hozirgi kunda turli sohalarda kompyuter amallaridan unumli foydalanilmoqda. Til og'zaki va yozma nutq ko'rinishida kishilik jamiyatiga xizmat qilib kelmoqda. Axborotlar asrida esa til inson va mashina munosabatini ta'minlaydigan vositaga aylandi.

Kompyuter lingvistikasida tabiiy va sun'iy til tushunchalari o'zaro farqlanadi.

Tabiiy til – insonlar o'rtasida kommunikativ, emotsional-ekspressiv, akkumulyativ funksiyalar uchun xizmat qiladigan tildir.

Sun'iy til – tabiiy tilni qayta ishlash, formallashtirish asosida matematika, informatika va tilshunoslik sohalari kesimida shakllantirilgan, kompyuter va inson o'rtasidagi munosabatni ta'minlaydigan, bevosita kompyuter xotirasiga kiritiladigan va kompyuterning lingvistik doiradagi funksiyalarini amalga oshirishda foydalaniladigan tildir.

¹ Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011. –B.

Insoniyat tovush va harflarning murakkab tizimi bo'lgan, muntazam ravishda o'zgarib turadigan tabiiy tildan foydalanadi. O'z navbatida til tizimi insonlarning o'zaro muloqotini har qanday sharoitda ta'minlay oladigan qat'iy vosita hamdir.

Kompyuter tizimi tushunilishi nisbatan oson bo'lgan formal tildan, sun'iy tildan foydalanish jihati bilan chegaralangandir. Kompyuter tizimi uchun ko'p ma'nolilik va ko'chma ma'noga asoslangan nutq muammo hisoblanadi. Kompyuter lingvistikasining navbatdagi vazifasi tabiiy til xususiyatlari va tabiatini imkon qadar o'zida aks ettiradigan formal tizimni yaratishdan iborat.

Kompyuter lingvistikasi informatika va lingvistikaning shunchaki birlashuvidan shakllangan soha emas, balki o'zida kengroq doiradagi masalalarni qamrab oladi. Kompyuter lingvistikasi tilning kompyuter nuqtayi nazaridagi ilmiy talqinidir. Olimlar turli tipdagi til hodisalarining kompyuter (formal) modellarini tuzish ustida izlanmoqdalar. Bunday modellar bilimlarga, qoidalarga yoki ma'lumotlarga, faktlarga asoslanadi. Kompyuter lingvistikasi bo'yicha mutaxassislar masalaga ilmiy jihatdan yondashishlari mumkin, ya'ni ma'lumotni formal jihatdan, kompyuter tizimi nuqtayi nazaridan ifodalashlari, lingvistik yoki psixolingvistik fakt orqali izohlashlari mumkin bo'ladi. Hozirda kompyuter lingvistikasining vazifasi amaliy ahamiyatga ega bo'lgan nutqni farqlash, nutqni sintezlash, ovozli avtomatik, qidiruv mashinalari, matn muharrirlari, tillarni o'rganish uchun materiallar tizimlarini yaratish bilan belgilanadi.

Hozirgi vaqtda kompyuterlashtirish jarayoni dunyo mamlakatlarida turli darajada kechmoqda. Taraqqiyot qay darajada bo'lmasin, insoniyat jamiyatda axborot texnologiyalarining o'rni beqiyosligini tushunib yetdi. Inson faoliyatining barcha sohalarini kompyuterlashtirish bugungi kunda jamiyatning muhim vazifasi va ijtimoiy taraqqiyot omilidir. Jamiyatni kompyuterlashtirmasdan turib, ijtimoiy o'zgarishlarni amalga oshirish, munosib turmush darajasini ta'minlash qiyin. Informatika sohasi ana shu ehtiyoj natijasida vujudga keldi.

Informatika axborotga ega bo'lish, munosabat bildirish, uni saqlash va taqdim etish jarayonlarini tadqiq qilish, jamiyat hayotining barcha sohalarida axborot texnikasi va texnologiyasini yo'lga qo'yish, amaliyotga tatbiq etish, ulardan foydalanish masalalarini hal qilish bilan shug'ullanadigan fan-texnika faoliyatidir. Informatika keng ma'noda insoniyat faoliyatining barcha sohalarida, asosan, kompyuterlar va telekommunikatsiya aloqa vositalari yordamida axborotni qayta ishlash bilan bog'liq fan, texnikaning xilma-xil tarmoqlari birligini o'zida

namoyon etadi. Kompyuterlar inson faoliyatida katta qulayliklar, yengilliklar yaratadi.

Informasion texnik vositalar axborot texnologiyasining asosini tashkil etadi. Axborot kommunikativ texnologiya jismoniy mehnatni chegaralaydi, kam vaqt sarflab, ko'p natijaga erishishni ta'minlaydi, bajarilayotgan ishning samaradorligini belgilaydi. AKT (axborot kommunikasion texnologiy)da aqliy mehnatning ahamiyati oshib boradi. Axborot texnologiyasining rivojlanishi insonlar faoliyatini yengillashtirishga, bilim darajasining o'sishiga, har bir sohada yuqori malakali mutaxassislar miqdorining ko'payishiga zamin yaratadi. Mamlakat ravnaqini, dunyoning rivojlangan davlatlari darajasidagi taraqqiyotni AKTsiz tasavvur qilib bo'lmaydi. Har bir sohada o'z mavqeyini mustahkamlab borayotgan axborot texnologiyalari ta'lim sohasida ham asosiy o'rinni egallashi muqarrardir. Ta'lim samaradorligini oshirishda, yosh avlodni jahon talablari darajasidagi yetuk mutaxassis kadrlar sifatida tayyorlashda, mamlakatimizga kirib kelayotgan har qanday avtomatlashgan texnologiyani boshqarishda axborot texnologiyalarini mukammal bilish talab etiladi. AKT yordamida yoshlar, avvalo, jahon miqyosidagi bilimlar, ilmiy-texnologik jarayonlar bilan tanishadilar, bilimlarini tahlil qilish imkoniyatiga ega bo'ladilar. Kutubxonalarining kompyuter jihozlari bilan ta'minlanishi, axborot resurs markazi tizimiga birlashtirilishi, elektron darsliklar, qo'llanmalar yaratilishi, ta'lim jarayonining kompyuterlashtirilishi axborot kommunikatsiya tizimining keng yoyilganidan dalolatdir. "Kadrlar tayyorlash milliy dasturi"da keltirilgan ta'lim tizimi bosqichlarida ko'zda tutilgan "O'quv – tarbiyaviy jarayonni ilg'or pedagogik texnologiyalar bilan ta'minlash" vazifasini bajarish uchun axborot kommunikatsiya tizimi borasida yetarli ma'lumotga ega bo'lish lozim bo'ladi. Axborotning nihoyatda ko'pligi, deyarli, barcha sohalarning kompyuter texnologiyalari bilan ta'minlanganligi ish yuritishning texnologik usulda olib borilishini, axborot texnologiyalarini qo'llashni taqozo qiladi.

Bugungi kunda gumanitar sohalar bilan informatika, xususan, til va kompyuter, adabiyot va kompyuter integratsiyasi masalalariga katta e'tibor berilmoqda. XXI asrda ma'lumotni asl holicha qabul qilish, ishonchli saqlash va eng qulay tarzda boshqalarga yetkazib berish til bilan bevosita bog'liqdir. Tarjima qilish, til o'qitish jaryonini avtomatlashtirish dasturlariga bo'lgan ehtiyoj *kompyuter lingvistikasi* fanining shakllanishiga zamin yaratadi.

“Kompyuter lingvistikasi” fani XX asrning o‘rtalarida paydo bo‘lgan matematik lingvistika fani asoslaridan kelib chiqqan. Matematik lingvistika fanining asosiy maqsadi tabiiy tillarning matematik modelini ishlab chiqish edi. Matematik lingvistika umumiy lingvistikaning formal va aksiomatik nazariyalarini va aniq tillarning matematik modellarini ishlab chiqishdan iborat vazifalarni amalga oshiradi. Shu tariqa tabiiy tillar grammatikasi formal modelining tuzilishi kompyuterda tarjima qilish, til o‘rgatish, matnlarni tahrirlash kabi amaliy dasturlarni ishlab chiqishga asos bo‘ldi.

Lingvistikada til inson bilan mutanosiblikda talqin qilinadi, ya’ni inson zaruriy subyekt sifatida tavsiflanadi. Kompyuter lingvistikasi jarayonida shaxsning ishtiroki bir qadar cheklanadi. Lingvistikada kompyuter tizimining joriy etilishi asrning buyuk kashfiyoti bo‘lgan kompyuter texnologiyalarining tilshunoslikka ham kirib kelganidan dalolatdir.

Dastlabki davrlarda fan “Matematik va kompyuter lingvistikasi” deb nomlangan va ta’lim jarayonida o‘qitilgan. Kurs ikki qismdan iboratligi ta’kidlanib, ularning mohiyati quyidagicha yoritilgan:

Matematik lingvistika – 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 fanining asosiy maqsadi tabiiy tillarning matematik modelini ishlab chiqishdir. Ushbu maqsadga erishish uchun fan o‘z oldiga quyidagi vazifalarni qo‘yadi:

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

An’anaviy lingvistikada til inson bilan mutanosiblikda, aloqadorlikda talqin qilinadi, matematik lingvistika esa tavsiflash jarayonida shaxsning ishtirokini, deyarli, inobatga olmaydi, u ko‘proq kompyuterga moslashtirilgan.

Til grammatikasining matematik modeli matematik mantiqning aksiomatik nazariyasiga asoslanadi. Shu bois matematik lingvistika matematik mantiq asoslariga tayanadi.

Kompyuter lingvistikasi fanining asosiy maqsadi lingvistik masalalarni yechishning kompyuter dasturlarini ishlab chiqish, inson va

mashina (kompyuter) muloqotini optimallashtirish, tabiiy tilni qayta ishlash (Natural Language Processing)dan iborat.¹ Kompyuter lingvistikasining asosiy vazifalari quyidagilar:

avtomatik o'qitish tizimini ishlab chiqish; tillarni o'qitish; bilimlarni tekshirish; matnlarni turli jihatdan avtomatik tahrirlash va tahlil qilish; matnlarning avtomatik tarzda morfologik, sintaktik va semantik tahlilini (inglizcha *parsing*) ta'minlovchi tizimlar yaratish; mashina tarjimasini 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 shakllantirish; 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 lingvistikasi vazifalarini amalga oshirish uchun zarur dasturlarni muayyan algoritmik tilda tilshunos dasturchi yaratadi. U bu dasturlarni tuzishda maxsus sun'iy tillardan foydalanadi. Hozirda internet ma'lumotlariga ko'ra sun'iy tillarning quyidagi turlari mavjud:

a) sintez qiluvchi dasturlar uchun: *Paskal, C, Prolog, Lisp, Okkamyu* tillari;

b) yig'uvchi dasturlar uchun: *Modula-2, Ada, Eyfel, Component Paskal* tillari;

d) aniqlashtiruvchi dasturlar uchun: *C++, Smalltalk, Oberon, Java, Clarron* tillari.

NLP (tabiiy tilni qayta ishlash) 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: *Alchemy API, Expert System S.p.A., General Architecture for Text Engineering (GATE), Modular Audio Recognition Framework, Monty Lingua, Natural Language Toolkit (NLTK)*.

¹ Новое в зарубежной лингвистике. Вып. XXIV. Компьютерная лингвистика. –М.: Прогресс, 1989.- С.10

Kompyuter lingvistikasi amaliy tilshunoslikning tarkibiy qismi hisoblanadi, u nazariy tilshunoslik yutuqlaridan foydalanadi. Kompyuter lingvistikasi va nazariy tilshunoslik bir-birini to'ldiradi. Kompyuter lingvistikasining o'ziga xos xususiyatlari quyidagilarda ko'rinadi:

kompyuter lingvistikasi tavsiflash jarayonida insonni istisno qiladi, tahlillar ko'proq kompyuterga moslashtiriladi;

kompyuter lingvistikasi masalani miqdoriy xarakteristikalar va aniq parametrlar asosida hal etad, miqdoriy (kvantitativ) tavsifga, algoritmlash, modellashtirish, statistik tahlilga asoslanadi;

kompyuter lingvistikasi amaliy xarakterga ega bo'lib, til bilan bog'liq muammolarning amaliy jihatlariga e'tibor qaratadi, uni aniq maqsadga yo'naltirilgan dasturlar, metodlar, tizimlar yordamida hal etish bilan shug'ullanadi;

kompyuter lingvistikasida ko'proq sun'iy tillar (programmallashtirish tillari, algoritmik tillar)ga tayaniladi, tabiiy tillarning mavjud imkoniyatlari cheklanadi, bunda tabiiy tilga ishlov berilib, kompyuterga moslashtiriladi.

Fanlar hamisha o'zaro uzviy bog'liqlik va hamkorlikda ish ko'radi, ular bir-birisiz mavjud bo'la 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 aloqador.

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

Matematik modellashtirish, formallashtirish, kompyuter tizimining tilshunoslikka tatbiq qilinishi avtomatik tarjima muammolari bilan shug'ullanish, mashina yordamida matn tuzish, uni qisqartirish, sintezlash jarayonini amalga oshirish uchun keng imkoniyatlar yaratdi. Bugungi kunda kompyuter lingvistikasida til materiallarini formallashtirish, tabiiy tilni qayta ishlash, chastotali lug'atlar tuzish, turli qisqartmalar va simvollardan foydalanish, MT algoritmini tuzish amallari bajarilmoqda.

Kompyuter lingvistikasi fani taraqqiyoti, turli maqsadlardagi dasturlarning yaratilishi tilshunoslikning boshqa yo'nalishlari taraqqiyotini ham ta'minlashga xizmat qiladi. "Kompyuter lingvistikasi" orfografik me'yorlarning buzilish hollarini, nogrammatik ifodalarni bartaraf etishga

yordam beradi. Lingvistik ma'lumotlar statistikasida, turli mazmundagi lug'atlar yaratishda kompyuter lingvistikasining ahamiyati katta. Bu fan nafaqat tilshunoslik, balki boshqa sohalar taraqqiyotiga ham ijobiy ta'sir ko'rsatadi. Jumladan, avtomatik tahrirlash (BORIS tizimi), tarjima jarayonidan badiiy matnlarni tahlil va tarjima qilishda foydalanish mumkin.

Savol va topshiriqlar

1. Kompyuter lingvistikasi fanining amaliy ahamiyatini tushuntiring.
2. Tabiiy til va sun'iy til tushunchalarini izohlang. "Sun'iy til" termini mohiyatini yoriting.
3. Kompyuter lingvistikasining asosiy muammolarini tavsiflang.
4. Tilning aniq fanligini ko'rsatuvchi misollar keltiring.
5. Kompyuter lingvistikasi fanining yuzaga kelishi ijtimoiy ehtiyoj ekanligini asoslang.
6. O'zbek tilining rasmiy-idoraviy va kompyuter uslublarini ishlab chiqish haqida takliflaringizni bayon qiling.
7. O'zbek tili fani bo'yicha dasrliklarning (maktab, kollej, litsey va o'liy o'quv yurti uchun mo'ljallangan) elektron versiyalarini ishlab chiqish haqida mulohazalaringizni bayon eting.

2-§. KOMPYUTER LINGVISTIKASINING FAN SIFATIDA RIVOJLANISHIDA NAZARIY VA AMALIY TADQIQOTLARNING O'RNI

Kompyuter lingvistikasi bo'yicha xorijda olib borilgan tadqiqotlar

Reja:

1. Kompyuter lingvistikasining shakllanishi.
2. Kompyuter lingvistikasining xususiyatlari.
3. Kompyuter lingvistikasiga oid Yevropa mamlakatlaridagi tadqiqotlar.
4. Kompyuter lingvistikasiga oid Rossiyadagi tadqiqotlar.

Asosiy tushunchalar: *kompyuter lingvistikasi, mashina tarjimasi, avtomatik tahrir va tahlil, tillarni avtomatik o'qitish tizimi, matnning statistik tahlili, lingvostatistika, injener lingvistikasi, o'qituvchi lingvistik avtomat, lingvistik ta'min, lingvistik dasturlar, matematik mantiq, tabiiy tillarning matematik modeli, bilimlarni baholash.*

Jamiyatimizda hozirgi kunda avtomatlashgan informatsion texnologiyalar va kompyuter muhim ahamiyat kasb etmoqda. Kompyuter kundan kunga jamiyatning barcha sohalariga kirib bormoqda. Bu esa yangi yo‘nalishlarning yuzaga kelishiga sabab bo‘lmoqda. Xususan, tilshunoslikka kompyuter texnologiyalarining kirib kelishi natijasida kompyuter lingvistikasi fani shakllandi.

Kompyuter lingvistikasi algoritmlar hamda inson va EHMning tabiiy tildagi muloqotini kompyuter orqali amalga oshirish bilan shug‘ullanadi. Inson arifmetik hisoblarni yengillatuvchi elektron jihozlar o‘ylab topdi. Tez orada bu mashinalarning inson bilimi bilan bog‘liq bir qator masalalarni hal qilish imkoniga ham ega ekanligi aniq bo‘ldi. Faktik materiallar to‘plandi, kompyuter dasturlari yaratildi, bilimlarni qayta ishlovchi sun‘iy tillar yaratildi. Bu jarayonlar “sun‘iy intellekt” deb ataluvchi yangi sohaning shakllanishiga asos bo‘ldi. Bizning davrimizda sun‘iy intellekt bo‘yicha ko‘pgina nazariy tadqiqotlar amaliyotga tatbiq qilindi. Aniq mexanik operatsiyalar, obrazlar talqini amalga oshirildi. AQShdagi sun‘iy intellekt bo‘yicha nazariy tadqiqotlar amaliyotga tatbiq qilindi. Sun‘iy intellektning asosiy masalalaridan biri – tabiiy tilni kompyuter orqali tushunishdir. Bu masala kompyuter lingvistikasiga taalluqlidir. O‘z navbatida EHM boshqa mashina qurilmalarini boshqarishi mumkinligi ma‘lum bo‘ldi. Shuning uchun kompyuter lingvistikasi loyihalarini muvaffaqiyatli ravishda amalga oshirish bilan inson murakkab texnik obyektlarni ovozli buyruq va tabiiy tildagi matnlar orqali boshqarishi mumkinligi aniqlandi. Kompyuter lingvistikasi inson, so‘z, texnika olamining yaxlit birikuvidir.

XX asrning 50-yillaridan boshlab tilshunoslikda «mashina tarjimasini», «mashina tilshunosligini» atamaları qo‘llana boshladi. Kompyuter texnologiyalari tilshunoslikka shu atamalar bilan kirib keldi. Mashina tarjimasini yoki avtomatik tarjima deyilganda bir tildagi matnni ikkinchi bir tilga EHM (kompyuter) vositasida, tez sur‘atda tarjima qilish nazarda tutiladi¹.

XIX asrda tilshunoslikning mustaqil fan sifatida shakllanishiga xizmat qilgan tadqiqotlar yuzaga keldi. Fanlar integratsiyasi bir qator yangi sohalarining shakllanishiga asos bo‘ldi. Xususan, sotsiologiya va tilshunoslik kesimida sotsiolingvistika, psixologiya va tilshunoslik kesimida psixolingvistika, neyrologiya va tilshunoslik kesimida neyrolingvistika, shunga o‘xshash biokimyo, astrofizika tipidagi qo‘shma

¹ *Grishman R. Computational linguistics // Cambridge University Press. 1994.- P.4. Qarang: Rahimov A. Kompyuter lingvistikasi asoslari. –T.,2011.*

fanlar yuzaga keldi. Matematik lingvistika, kompyuter lingvistikasi fanlari ham ikki soha kesimida shakllangan fanlar hisoblanadi.

Kompyuter lingvistikasi XX asrning 50 - yillarida 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 hodisalarini matematik bayonda tushuntiradigan fanning nomini ham taklif etgan. Olimning fikricha, bu fan «**Til algebrasi**» («**Lingvistik algebra**») deb atalishi lozim edi.¹ Amerikalik tilshunos Noam Xomskiyning formal grammatika, transformatsion grammatika haqidagi qarashlari bevosita matematik lingvistikaning alohida yo'nalish sifatida yuzaga kelishiga sabab bo'lgan.²⁴ Matematik lingvistika fani kompyuter lingvistikasi sohasining shakllanishiga asos bo'ldi. Tabiiy tillar grammatikasi formal modeli avtomatik tarjima qilish, tillarni o'rgatish, bilimlarni diagnostika qilish, matnlarni tahrir va tahlil etish kabi amaliy dasturlarni ishlab chiqishga asos bo'lgani holda kompyuter lingvistikasi yo'nalishiga zamin yaratdi.³ Kompyuter lingvistikasi inglizcha «computational linguistics» so'zining kalkasidir. Mashina tarjimasining asoschilari muhandis va matematik olimlar bo'lib, keyinchalik bu ishda tilshunoslar ham faol qatnasha boshlagan. Shu tariqa mashina tarjima g'oyalari butun dunyoda nazariy va amaliy tilshunoslikning rivojlanishida muhim ahamiyat kasb etdi. Formal grammatika nazariyasi shakllandi, til va uning alohida aspektlari modelini yaratishga e'tibor qaratildi. Tilning formal jihatlari matematik lingvistika fanida ishlab chiqilib, bu o'z navbatida, kompyuter lingvistikasi fanining yuzaga kelishi uchun poydevor bo'ldi. Shu asosda tilshunoslikning yangi yo'nalishi bo'lgan kompyuter lingvistikasi va tilshunoslikning bir qator nazariy va amaliy yo'nalishlari vujudga keldi. Kompyuter lingvistikasi matematik lingvistikaning mantiqiy davomi bo'lib, u amaliy tilshunoslikning eng muhim qismini tashkil etadi.

Kompyuter lingvistikasining nazariy va amaliy asoslarini belgilashda daniyalik olim L.Elmslev, amerikalik olimlar N.Xomskiy, K.Shenon kabilarning xizmati kattadir. Kompyuter lingvistikasining amaliy muammolari XX asrning o'rtalarida vujudga keldi. Bu narsa birinchi navbatda mashina tarjima (MT) bilan aloqadordir.

¹Нурмонов А. Структур тилшунослик: илдизлари ва йўналишлари. –Т., 2008. -Б. 145.

²Гладкий А.В., Мельчук И.А. Элементы математической лингвистики. –М.: Наука, 1969. - С.7.

³ Bu haqida qarang: Po'latov A., Muhamedova S. Kompyuter lingvistikasi. –Т., 2009; Rahimov A. Kompyuter lingvistikasi asoslari. –Т.,2011.

1954-yil AQSHda Jorjtaun universitetida mashina tarjimasi bo'yicha dunyoda birinchi tajriba o'tkazildi va IBM - 701 rusumli tarjima qurilmasi yordamida 250 ta so'zdan iborat 60 ga yaqin ruscha gaplarni ingliz tiliga tarjima qilish jarayoni amalga oshirildi.

1955-yilga kelib BRSM rusumli mashina qurilmasi yordamida inglizcha gaplarni rus tiliga tarjima qilish jarayoni amalga oshirildi (I.K.Belskaya, L.N.Korolev, S.N.Razumovskiylar loyihasi asosida). Su tajribalar asosida KL yo'nalish sifatida, 1960-yilga kelib mustaqil fan sifatida shakllandi. Angliya, Fransiya, Avstriya marniakatlarida ham mashina tarjimasi muammolariga qiziqish boshlandi.

Tilshunoslikda ma'lum bir tilni tipologik jihatdan tasniflash, til materiallarini analiz va sintez qilish, tabiiy tilni qayta ishlash, mashina tiliga o'tkazish masalalari dolzarb ahamiyat kasb eta boshladi. Lingvistik tadqiqotlarni avtomatlashtirish masalasiga doir tilshunoslarning maxsus ilmiy konferensiyalari o'tkazilgan. Masalan, 1962-yilda Kembridj shahrida jahon tilshunoslarining IX kongressi o'tkazildi. Bu anjumanda struktural tilshunoslik (ST) va matematik lingvistika ML muammolarini muhokama qilishga alohida o'rin ajratilgan edi. 1967-yilda Buxarest shahrida jahon tilshunoslarining X kongressi o'tkazilib, matematik lingvistika va mashina tarjimasi muammolari maxsus muhokama qilindi. 1960-yilda Parij shahrida o'tkazilgan Xalqaro kongressda maxsus algoritmik til – ALGOL joriy etilgan edi. Bu algoritmik til hozirgi kunga qadar EHMga matnlarni kiritish, ularni dasturlash uchun xizmat qilib kelmoqda.¹

O'tgan yetmish yilga yaqin vaqt mobaynida kompyuter lingvistikasi sohasida bir qator ilmiy va amaliy natijalarga erishildi: tabiiy tilda avtomatik tarjima tizimi yaratildi, matndagi ma'lumotlarning avtomatik qidiruv tizimi ishlab chiqildi, og'zaki nutqning avtomatik analizi va sintezi tizimi yaratildi, bir qator lingvistik muammolarni hal etuvchi kompyuter dasturlari ishlab chiqildi, inson va mashina (kompyuter) muloqoti optimallashtirildi, tabiiy tilni qayta ishlash (Natural Language Processing) tizimi shakllantirildi. Bu yo'nalishlarda sezilarli yutuqlar qolga kiritilgan bo'lsa-da, qayd etilgan masalalar hali to'liq yechimini topgani yo'q. Bu esa kompyuter lingvistikasi doirasidagi izlanishlarni davom ettirishni taqozo etadi.

Insoniyat rivojlanishining muhim xususiyatlari qatorida, avvalo, kompyuterlarning paydo bo'lishi va axborot jamiyatlarining shakllanishini qayd etish lozim. Kompyuter lingvistikasi fani shakllanishi elektron hisoblash mashinalarining paydo bo'lishi bilan bevosita bog'liqdir. XX

¹ *Йўлдошев Б.* Компьютер лингвистикаси. – Самарқанд, 2009.

asrning ikkinchi yarmidan boshlab informatika rivojlandi, o'z navbatida kompyuter lingvistikasi ham rivojlanish bosqichiga kirdi. Mazkur davrda amalga oshirilgan amerikalik olim M. Tabuening axborot izlashga qaratilgan ilk UNITERM nomi ostidagi tajribasini; F. Ruzvelt nomidagi olimlar jamiyati prezidenti Vannevr Bushning gipertekst haqidagi nazariyasini (1950); T.Nelson tomonidan rivojlantirilgan gipertekst nazariyasini (1960); chiziqli grammatika («сюженные грамматики» yoki «грамматики повествования») g'oyasi bilan fan yuksalishiga ijobiy ta'sir ko'rsatgan N. Xomskiy tadqiqotlarini; R. Shenka va V. Lenertlarning 1980-yillarda erishgan yutuqlarini kompyuter lingvistikasiga doir dastlabki izlanishlar sifatida keltirish mumkin.

Bugungi kundagi elektron ma'lumotlar bazasi, qo'l telefonlar, har xil turdagi kompyuterlarning paydo bo'lishi, ularning ishlash tizimi tadrijiy ravishdagi ilmiy tadqiqotlarning amaliy natijalari sifatida maydonga chiqdi.

Xorijda kompyuter lingvistikasi bo'yicha olib borilgan tadqiqotlarning ko'lami juda keng. Ayniqsa, AQShda olib borilgan tadqiqotlar izchilligi, qamrab olingan muammolar doirasining kengligi bilan ajralib turadi. AQShda maxsus tashkilotlar ochilgan bo'lib, kompyuter lingvistikasi bilan bog'liq masalalarni hal qilish bilan shug'ullanadi. Nashr qilinayotgan tematik to'plamlardan tashqari «Computational Linguistics» («Kompyuter tilshunosligi») nomli jurnal ham nashr etiladi. AQShda kompyuter tilshunosligi assotsiatsiyasi mavjud bo'lib, 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 10ta universitetida kompyuter lingvistikasi fakultetlari mavjud. Mazkur fakultetlarda mutaxassilar kompyuter lingvistikasining muammoli masalalari yechimiga doir izlanishlar olib BORISHmoqda. Shuningdek, Niderlandiyada «Artificial Intelligence» («Sun'iy intellekt») jurnali ham chop etiladi. Bu jurnal ham kompyuter muammolari yechimiga, yangi yaratilayotgan dasturlarga doir izlanishlarni aks ettiradi. Mexiko shahrida har yili «**Computational Linguistics and Intelligent Text Processing**» nomli xalqaro konferensiya o'tkazib kelinadi.

MDH mamlakatlarida ham kompyuter lingvistikasi bo'yicha muhim izlanishlar amalga oshirilgan. Ayniqsa, bu borada rus kompyuter tilshunosligidagi tadqiqotlar diqqatga sazovordir. Kompyuter tilshunosligining muammoli masalalari sun'iy intellekt bo'yicha o'tkazilgan turli xalqaro konferensiyalarda ham ko'rib chiqilgan. Rossiyada har yili kompyuter lingvistikasiga oid «DIALOG» xalqaro

konferensiyasi o‘tkaziladi.¹ Rus kompyuter lingvistikasi quyidagi yo‘nalishlarni o‘z ichiga oladi:

1. Mashina tarjimasi (MT).

2. Avtomatik tahrirlash (AT).

3. Til o‘rgatish jarayonini kompyuterlashtirish (avtomatik tarzda tilga o‘qitish-ATO‘).

4. 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 mos edi. L.L. Nelyubin so‘z va iboralarning avtomatik lug‘ati bilan birgalikda to‘liq tarjima qilish imkoniyatini beruvchi morfologik-sintaktik algoritm-freym yaratishga harakat qilgan; yaratilgan tarjima dasturlari asosida o‘qituvchi lingvistik avtomat (обучающий лингвистический автомат) yaratish me‘yorlarini ishlab chiqqan. Bu loyiha o‘quvchilarga ingliz tilidagi rasmiy hujjatlarni rus tiliga tarjima qilishga imkon beradi. Prof. L.L.Nelyubinning ilmiy faoliyatida mashina tarjimasi uchun (ingliz tili bo‘yicha) algoritmlar va dasturlar ishlab chiqish asosiy o‘rin tutadi. L.L.Nelyubin, tabiiyki, mashina tarjimasida ingliz tili leksikasini butunligicha qamrab ololmas edi. Shuning uchun AQSh harbiy hujjatlarini tadqiqot obyekti sifatida belgiladi va bu sohada ko‘plab ilmiy asarlar yaratdi. Y.N.Marchuk tarjimini modellashtirish metodlari va uni avtomatlashtirish yo‘llarini ishlab chiqqan va bugungi rus kompyuter lingvistikasiga asos solgan olimlardan hisoblanadi. Y.N.Marchuk tadqiqotlarida ham bir tabiiy tildan boshqa tilga bo‘lgan tarjimini modellashtirish tamoyillari, tarjima birligi, tarjima jarayonining statikasi va dinamikasi tushunchalari hamda EHM ishtirokidagi tarjimaning texnologik xaritasi tavsiflangan.

Shuningdek, ko‘plab rus, ukrain, moldovan olimlari ham mashina tarjimasi masalalari bilan shug‘ullanishgan, ularning izlanishlari kompyuter lingvistikasida muhim o‘rin tutadi. Xususan, bu sohada ko‘plab nomzodlik va doktorlik dissertatsiyalari yaratilganini alohida ta’kidlash lozim. Dissertatsiyalar sohani yanada chuqurroq o‘rganish uchun zamin yaratdi.

¹ <http://www.dialog-21.ru/>

XX asrning 60–80 yillarida mashina tarjimasida borasidagi bir qator tadqiqotlar yuzaga keldi. Mazkur tadqiqotlar dunyo kompyuter tarmoqlarida rus tilidagi ma'lumotlar bazasi yaratilishi va rus tilidagi matnlarni avtomatik tarzda boshqa tillarga (yoki boshqa tildagi matnlarni rus tiliga) tarjima qiluvchi dasturlarning vujudga kelishi uchun asos bo'ldi. 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 til o'qitish yo'nalishida ham izlanishlar olib borilgan. Bu tadqiqotlar oliy va o'rta maktabda til o'qitishni optimallashtirishning psixologik-kibernetik, semiotik, lingvostatistik, injener-lingvistik va lingvodidaktik asoslarini ishlab chiqishga yo'naltirilgan. XX asrning 80-yillaridan boshlab til o'qitishni avtomatlashtirish va kompyuterlashtirish muammolariga qiziqish kuchaydi.

Tadqiqotlar natijasida ko'plab lingvistik avtomat o'qituvchilar uchun lingvistik ta'min yaratilgan bo'lib, bu dasturlar rus tilini o'qitish jarayonini optimallashtirishga xizmat qiladi.

Rus kompyuter tilshunosligida «tillarni o'qitishni kompyuterlashtirish» deyilganda avtomatik o'qituvchi tizimiga tayanuvchi avtomatlashtirilgan o'quv kurslarini vujudga keltirish nazarda tutiladi. Avtomatik o'qituvchi tizimida o'quv jarayoni ishtirokchilariga nafaqat taqlid qilish, balki ularning intellektual faoliyatini modellashtirish ham zarur bo'ladi. Shuningdek, bu sohada yaratilgan ilmiy adabiyotlarda «lingvistika va tillarni o'qitishda model» tushunchasi asosiy o'rinni egallaydi.

Rus tilshunosligidagi tillarga o'qitish yo'nalishida lingvistik statistikadan ham keng foydalanilgan. Ma'lumki, ona tilidan boshqa tilni o'rganayotganda, avvalo, ushbu tilning lug'at boyligiga murojaat qilinadi. O'rganilayotgan tilning dastlab eng asosiy, faol qo'llanadigan qismi o'zlashtirishga kirishiladi va muntazam ravishda bosqichma-bosqich so'z boyligi orttirib boriladi. Buning uchun esa leksikostatistik manbalar – ma'lumotlar asosiy poydevor vazifasini o'taydi. L.N.Zasorinaning ta'kidlashicha, leksikostatistikaning markaziy muammosi jonli (funktional) tilning statistik qonuniyatlarini va matnning statistik strukturasi 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 zarurati tufayli kompyuter lingvistikasida statistik yo‘nalish yuzaga keldi. U rus tilshunosligida avtomatik tarzda til o‘qitish yo‘nalishi bilan mutanosib tarzda rivojlanib borgan. Statistik yo‘nalish o‘zbek, qozoq, qirg‘iz tillaridagi tadqiqotlarda ham kuzatiladi.

Kompyuterdan foydalanilgan holda ko‘plab chastotali lug‘atlar tuzildi, ular o‘z navbatida mashina tarjimasini uchun zamin yaratdi. Ko‘rinadiki, rus kompyuter lingvistikasidagi qayd etilgan to‘rt yo‘nalishning rivojlanishi rus tilining dunyoviy tillar sirasiga kirishi, ommalashuvi, uni o‘rganish va o‘rgatish jarayonining optimallasuvi uchun asos bo‘lgan. Bunda, shubhasiz, kompyuterchi- dasturchilar bilan bir qatorda tilshunoslarning ham xizmati katta. Ayniqsa, prof. R.G.Piotrovskiyning izlanishlari o‘zining serqirraligi, har to‘rt yo‘nalishni o‘zida qamrab olganligi, obyektivligi, aniq va izchilligi bilan ajralib turadi. Olim nafaqat rus tilshunosligida, balki o‘zbek, qozoq, qirg‘iz va boshqa tilshunosliklardagi injener lingvistikasi, kompyuter lingvistikasi rivojiga salmoqli hissa qo‘shgan.

Shuni ta‘kidlash lozimki, olimning rafiqasi A.A.Piotrovskaya va singlisi K.G.Piotrovskaya faoliyati matematik lingvistika va injener lingvistikasining rivojida muhim o‘rin tutadi. Olimlar injener lingvistikasining turli yo‘nalishlarida ijod qilgan holda, turkologiyada shu sohaning ravnaq topishiga o‘z hissalarini qo‘shishgan.

Rus tilshunoslari dastlab matematik, so‘ngra injener va nihoyat, kompyuter lingvistikasi bilan shug‘ullangan, mazkur sohalarning nazariy va amaliy masalalarini hal qilishga intilishgan.

XXI asr kompyuter tilshunosligida ham rus olimlarining o‘ziga xos o‘rni bo‘lib, ular doimiy ravishda COLINC konferensiyalarida qatnashib keladi, muntazam ravishda o‘z izlanishlarini davom ettirishmoqda.

Rossiyada kompyuter lingvistikasiga yo‘naltirilgan tadqiqot institutlari shakllantirildi. 2009-yilda Sankt-Peterburg universiteti Rossiyadagi 12ta universitet ichida ko‘p etapli konkursda g‘olib chiqib, “Ilmiy tadqiqot universiteti” (Национальный исследовательский университет) kategoriyasiga ega bo‘ldi. Rossiya Federatsiyasi ta‘lim va fan vazirligi tomonidan universitetning 2009-2018 yillarga mo‘ljallangan taraqqiyot dasturi ishlab chiqildi. 2011-yilda Universitet “Sankt-Peterburg informatsion texnologiyalar, mexanika va optika milliy tadqiqot universiteti” maqomini oldi.¹

¹ Боярский К. Введение в компьютерную лингвистику. Санкт-Петербург, 2013.-С.2

Xitoy kompyuter lingvistikasi

Informatsion texnologiyalar va ingliz tili ta'siri natijasida xitoy tili leksikasida o'zgarishlar sodir bo'ldi. Xitoy tilining xos xususiyatlaridan biri bo'lgan omonimiya uning butun tarixiy taraqqiyot bosqichida muhim rol o'ynadi. Hozirgi kunda xitoy ierogliflarining keng imkoniyatlarini hisobga olib, omonimlar asosida kompyuter o'yinlari yaratilmoqda. Ieroglif yozuvlar XX asrda bir qator o'yinlarga asos bo'ldi. Xitoy Xalq Respublikasi on-layn kommunikatsiyasining ommaviy xarakteri til me'yorlariga talabning pasayishi bilan belgilanadi.

Internet Xitoy xalqi hayotida muhim rol o'ynaydi. Internet tizimidan foydalanuvchilar butun Yevropa aholisidan ko'ra ko'proq miqdorni tashkil etadi. Tarmoqda barcha Xitoy ommaviy axborot vositalari, ilmiy jurnallari berib boriladi. Xitoyda favqulodda *wangluo wenxue* degan rasmiy nom bilan kiritilgan *seteratura* (сетевая литература) – tarmoq adabiyoti ommaviylashdi. Xitoy sotsial tarmoqlarida va blogosferasida sezilarli o'zgarishlar kuzatilmoqda. 90% xitoylik ovozli matn ko'rinishidagi xabarlarni uzatishda internetdan mobil aloqa vositalari yordamida foydalanadi. Bugungi kunda har kuni jahon bo'ylab millionlab xitoyliklar foydalanadigan eng faol va ko'p funksiyali *WeChat* kommunikator – ilovasi (*welxin* “kichik xabar”) mashhurdir. Til resurslaridan, grafik elementlardan, so'zdan, grammatikadan foydalanishning yangi yo'llari chatlarda va forumlarda kuzatiladi. Bugungi kunda Xitoydagi tilni va yozuvni an'anaviy standartlashda on-layn muloqot tili lingvistik fenomen sifatida dolzarb ahamiyat kasb etmoqda. Xitoyda qadim davrlardan boshlab yozma va og'zaki nutq o'rtasida katta farq kuzatiladi. Savodxonlik elitaga xos xususiyat sifatida qaralgan. Yozma nutqqa xitoyliklar hamisha hurmat va e'tibor bilan yondashgan. Xitoy tiliga hamisha e'tibor berib kelingan. Xitoyda har yili faol so'zlar va ifodalar ro'yxati tuziladi, neologizmlar lug'ati nashr etiladi. On-layn kommunikatsiya xitoy tilini og'zaki va yozma shaklida namoyon bo'lar ekan, o'zida muhim ijtimoiy ahamiyatga ega bo'lgan til vositasini namoyon qiladi. Xitoyda bir tomondan internet-muloqot yozma shaklda amalga oshmoqda. Ikkinchi tomondan, on-layn kommunikatsiya jonli muloqotni, dinamiklik va yuqori darajadagi ekspressivlik xos bo'lgan jarayonni ta'minlamoqda.

Ierogliflar bu belgilardan foydalanuvchilarga bir qadar murakkablik tug'dirmoqda. Xitoyda oddiy insonning nutq tezligi minutiga 170 bo'g'inni, ierogliflarni terish tezligi 30-50 belgini tashkil etadi. Ierogliflarni kompyuterga kiritish xitoyliklar uchun bir qadar qiyin. Ierogliflarni kiritishning ikki usulidan foydalaniladi: fonetik usul –2tadan

4ta harfgacha bo‘lgan harflar kombinatsiyasini o‘z ichiga oluvchchi pinyin transkripsiyasi va grafik usul – ieroglifni *touchscreen* texnologiyasi yordamida qo‘lda yozish. Birinchi usul o‘rtacha 9–12 belgidan iborat bo‘ladigan ierogliflarni yozishda ikkinchisiga nisbatan tezkor hisoblanadi. Pinyin transkripsiyasini qisqa vaqtda amalga oshirish uchun faol qo‘llanuvchi, so‘zlashuv uslubiga xos ifodalar va barqaror birikmalar kodlashtiriladi.

Ma’lumotni tez yetkazish maqsadida iyerogliflarning bosh harflaridan olingan qisqartmalardan foydalaniladi. Ko‘p hollarda foydalanuvchilar yaqin-atrofdagi begonalarga nisbatan ham qo‘llanuvchi qarindoshlik terminlarini qisqartiradilar:

gg – 哥哥 gēge (so‘zma-so‘z: 'aka') “yigit”, “aka”
(suhbatdoshga hurmatni ifodalaydi);

mm – 妹妹 mèimei (so‘zma-so‘z: 'singil') “singil”, “qiz”

(yoshga yoki ijtimoiy holatiga ko‘ra kichik bo‘lgan ayol suhbatdoshga nisbatan qo‘llanadi);

sg – 帅哥 shuàigē (so‘zma-so‘z: 'chiroyli erkak') “yoqimtoy (красавчик)”

tx – 同学 tóngxué “sinfdosh, guruhdosh”.

Analogiya tarzida salbiy bo‘yoqdorlikka ega bo‘lgan so‘zlar ham qisqartiriladi:

bd – 笨蛋 bèndàn “ahmoq, befahm”;

mpj – 马屁精 mǎpìjīng “xushomadgo‘y”.

Xitoy tilida omonimiyaning yuqori darajada ekanligi ko‘pgina harflarni qisqartirish usulidagi abbreviaturalarning bir xil shaklda bo‘lishiga sabab bo‘ladi. Masalan, zt qisqartmasi *zhuǎntiē* “repost”, *zhūtóu* (so‘zma-so‘z tarjimasi: “cho‘chqa bosh”) “ahmoq” so‘zlarini ifoda etadi.

Xitoyliklar kompyuterda raqamlardan faol qo‘llanuvchi leksik birliklarni ifodalashda foydalanmoqda:

5 wǔ – 我 wǒ “men”;

0 líng – 你 nǐ “sen”;

6 liù – 了 “le” suffiksi yoki yuklamasi.

88 bābā – 拜拜 báibái “Xayr! Qani (давай!)” (ingliz. *bye-bye!*).

995 jiùjiùwǔ – 救救我 jiùjiù wǒ “Meni qiyin vaziyatdan qutqar!”;

7456 qīsiwǔliù – 气死我了 Qìsǐ wǒ le “Meni qidir (moq)”;

596 wǔjiùliù – 我走了 Wǒ zǒule “Men ketyapman”;

1798 yī qī jiù bā – 一起走吧 Yīqǐ zǒu ba “Давай пойдём вместе”;

201314 èrlíngyīshānyīsi – 爱你一生一世 ài nǐ yīshēng yīshì “Meni seni umrbod sevaman”.

Xitoy tiliga xos omonimlik raqamlarda ham aks etadi:

4 sì – 是 shì “bo‘lmoq, hisoblanmoq”, 思 sī “o‘ylamoq, sog‘inmoq” 死 sǐ “o‘lasi, dahshat”;

2 èr – 愛 ài “sevmoq”, 饿 è “och”.¹

Bu shakldagi xabar namunalari o‘zbekcha matnlarda ham kuzatiladi:

«3iw 5dekan». Bu – xorijga jo‘nab ketayotgan bir yigitning global tarmoqdagi yozishmalaridan parcha. To‘g‘rirog‘i, samolyotning qachon parvoz qilishini so‘ragan do‘stlariga javobi: «Uchish 5 da ekan».

«9il 4opondaman». Internetga joylashtirilgan bir davra qizlarning surati ostidagi izoh: «To‘q qizil choponda – men».² Xitoy tilidagi ierogliflarning murakkabligi tufayli so‘zlarni ularga ohangdosh harflarga joylab berish qulaydir-ku, ammo o‘zbek tilidagi matnlarda so‘z o‘rnida raqamlardan foydalanish tilga nisbatan hurmatsizlik hisoblanadi.

Xabar matnlarida turli bgrafik elementlar kombinatsiyasidan – iyerogliflardan, inglizcha harflardan, raqamlardan, ba‘zan belgilardan foydalanilganligini ham kuzatish mumkin:

3Q sǎn-q – “Rahmat!”;

P9 p-jiǔ– 啤酒 píjiǔ “pivo”;

B4 b-sì – 鄙视 bǐshì “ko‘rish”

小 P 孩 xiǎo P hái – 年轻人 niánqīngrén “birodar”.

Harflar va iyerogliflar kombinatsiyasida so‘zlar odatda pinyan transkripsiyasida yozilgan bo‘g‘inning birinchi harfi bilan boshlanadi:

1 公 – 老公 lǎogōng “er”;

1 婆 – 老婆 lǎopó “xotin”;

Inglizcha harflar (umumga ma‘lum bo‘lgan abbreviaturalar) iyeroglif shaklidagi xitoycha elementlar bilan to‘ldiriladi:

E 化 E huà (ingl. *electronic*) “компьютеризация”.

Harflarning simvollar, simvollarning raqamlar bilan birikishidan hosil bo‘lgan til o‘yinlari uchraydi:

+U jiā U – 加油 jiāyóu “Qani! Tugmani bos”;

-7 fūqī – 夫妻 fūqī “er va xotin”.³

¹Кислов А.В., Колтачкова У.Н. Влияние интернета на современный китайский язык / Компьютерная лингвистика и вычислительные онтологии. Выпуск 1. 2017. –С.72-86.

²Niyozmetov B. Tomchida quyosh aksi. “Methodist”. Ilmiy maqolalar to‘plami. –T.,2017. –B.21.

³Кислов А.В., Колтачкова У.Н. Влияние интернета на современный китайский язык / Компьютерная лингвистика и вычислительные онтологии. Выпуск 1. 2017.-С.72-86.

Koreya va Yaponiya kompyuter lingvistikasi

Til tizimi yoki tilning ost tizimidagi o'zgarishlar aksariyat hollarda jamiyatning to'g'ridan-to'g'ri va bevosita ta'sirida emas, balki xususiy kommunikativ funksiyalarning kengayishi yoki torayishi bilan bog'liq holda, bilvosita yuz beradi. Shunga qaramay, til tizimiga insonning bevosita ta'sirini inkor qilish to'g'ri bo'lmaydi¹. Bunga Koreya Xalq Demokratik Respublikasidagi til siyosatining "Madaniylashuv davri" (1964)ni misol qilib keltirish mumkin. 1964-yilda Shimoliy Koreyada siyosiy tus olgan til islohoti boshlandi. Koreys tilini yot so'zlardan tozalash masalasi ko'tarildi. 1964-yilda o'tkazilgan Til qurultoyida mamlakat rahbari Kim Il Sung nutq so'zladi. U o'zining nutqida koreys tilidan o'zlashmalarni siqib chiqarish, ayniqsa, qiyinchilik tug'dirayotgan xitoy yozuvidagi so'zlarni xalqqa tushunarli bo'lgan koreyscha muqobillar bilan almashtirish masalasini o'rta tashladi. Muammoni hal qilish uchun quyidagilarni taklif qildi:

- o'zlashma so'z tilga o'z so'zdek singib ketgan bo'lsa, uni almashtirishga hojat bo'lmaydi;
- xitoycha so'z koreyscha sinonimiga ega bo'lsa, uni almashtirish lozim;
- xitoycha so'z va uning koreyscha sinonimi ma'no farqlariga ega bo'lsa, unda so'zlarni o'z holicha qoldirish kerak;
- agar xitoycha (yoki boshqa o'zlashma) so'zni tushunish qiyin bo'lsa, shuningdek, u koreyscha sinonimiga ega bo'lmasa, u holda, yangi so'z topish lozim.

Xullas, Shimoliy Koreyadagi til islohoti siyosiy mohiyatga ega bo'lib, u tilning so'z boyligiga katta ta'sir o'tkazgan.

Yaponiya informatika, kompyuter lingvistikasi borasida jiddiy yutuqlarga erishgan. Yaponiyada nanotexnika rivojlangan. Nano o'lchov birliklarining old qo'shimchasi bo'lib, "milliarddan bir"ni ifodalaydi. Bir nanometr metrning milliarddan biridir. *Nano* yunoncha so'zdan olingan bo'lib, "mitti" degan ma'noni bildiradi. Nano texnologiya – "mitti texnologiya" bo'lib, uning prototipi tabiatning o'zida mujassam. Ko'zga ko'rinmas bakteriyalar, chumoli, o'rgimchak va shunga o'xshash hasharotlar tabiiy nano jismlardir.

Nanotexnologiya atamasini birinchi bo'lib 1974-yilda yapon olimi Norio Naniguchi ishlatgan. Bu sohadagi dastlabki texnik vositalar Shveytsariyaning IBM laboratoriyasida ixtiro qilingan. 1982-yilda

¹ Швейцер А.Д., Никольский Л.Б. Введение в социолингвистику. –М.: Высшая школа, 1978. – С. 117.

skanerlovchi zondni ixtiro qilgan Gerd Bining va Genri Rorerlar 1986-yilda Nobel mukofotiga sazovor bo'lgan.

Yaponiyada kompyuter lingvistikasiga oid tadqiqotlar 1981-yilda e'lon qilingan 5- avlod kompyuterlarini yaratish umummilliy dasturi asosida olib borilmoqda.

Hozirda kishilik taraqqiyotining barcha jabhalarida kompyuter lingvistikasi fanining yutuqlarini ko'rish mumkin. Oddiy qo'l telefonidan tortib eng murakkab funksiyalar uchun ixtisoslashgan kompyuter va robotlar aynan mazkur fan taraqqiyoti samarasidir. Zamonaviy texnika va kompyuter lingvistikasi borasida xorijda olib borilgan tadqiqotlar natijasini quyidagicha umumlashtirish mumkin:

1) tabiiy tilni qayta ishlash asosida mashina tili sifatida xizmat qiluvchi sun'iy til shakllantirildi. Jumladan, inson va mashina munosabatini ta'minlash uchun kompyuterga ingliz tili asosidagi sun'iy til kiritildi. Sun'iy til barcha yaratilgan mashina dasturlarining asosiy va yagona kommunikativ vositasiga aylandi;

2) kompyuter lingvistikasi virtual munosabatlar shakllanishi uchun baza vazifasini o'tadi;

3) kompyuter lingvistikasi ilm-fan yutug'i bo'lgan mashina tarjimasining yuzaga kelishini ta'minladi;

4) kompyuter lingvistikasi sohasida turli xil lingvistik muammolarni hal qilishga yo'naltirilgan dasturlar yaratildi va bunday tadqiqotlar davom etmoqda. Ular, asosan, matn tahriri, avtomatik tarjima, analiz va sintez, tabiiy tilning qayta ishlanishi, kompyuter leksikografiyasiga oid dasturlarni o'z ichiga oladi;

5) kompyuter lingvistikasi nazariy asoslariga tayangan amaliy fan sifatida rivojlanib bormoqda. Uning nazariy asoslari nutqiy jarayonni, matnni o'rganishning turli xil model (gipoteza)larini yaratishda, nazariy asoslarini ishlab chiqishda ko'rinadi. Fanning amaliy xususiyati tarjima mashinalarini yaratish va uning ishlash jarayonini tashkil etish bilan beigilanadi. Xorij kompyuter lingvistikasida quyidagi uch usul va metodlardan keng foydalaniladi:

a) mantiqiy-matematik usullar: inkor amali, konyunksiya, dizyunksiya amallari kabilar;

b) nazariy -informatsion usullar;

d) ehtimollik-statistik usullar.

Turkologiyada va o‘zbek tilshunosligidagi kompyuter lingvistikasiga doir tadqiqotlar

Reja:

- 1.O‘zbek tilshunosligida kompyuter lingvistikasining tutgan o‘rni.
- 2.An’anaviy o‘zbek tilshunosligi erishgan yutuqlar – o‘zbek kompyuter lingvistikasi ma’lumotlar bazasi sifatida.
- 3.O‘zbek tilshunosligidagi kompyuter lingvistikasiga oid tadqiqotlar.

Asosiy tushunchalar: *amaliy tilshunoslik, matematik lingvistika, injener lingvistikasi, informatsion texnologiyalar, kompyuter texnologiyalari, lingvostatistika, avtomatlashtirish, turkiy tillar mashina fondi, an’anaviy tilshunoslik, o‘zbek kompyuter lingvistikasi, modellashtirish, lingvistik modellar, sodda gap modeli, qo‘shma gap modeli, o‘zbekcha matnlarning leksik -morfologik tuzilishi, statistik tahlil, avtomatik tahrir, mashina tarjimasi, avtomatik til o‘qitish, lug‘atshunoslik.*

AQSH, Yevropa mamlakatlari, Rossiyada kompyuter lingvistikasi masalalari yuzasidan olib borilgan tadqiqotlar boshqa davlatlarda, jumladan, turkiy davlatlarda, O‘zbekistonda shu sohaning rivojlanishiga zamin hozirladi.

Matematik va injener lingvistikasi, ya’ni hozirgi kompyuter lingvistikasi yo‘nalishdagi tadqiqotlar turkiy xalqlarda ham amalga oshirilgan. Turkologiyaga yangi informatsion kompyuter texnologiyalari tatbiq etilishi yaqin yillardan boshlab amalga oshirildi. Bu borada turk, qozoq, qirg‘iz tilshunosligi ma’lum yutuqlarga erishdi.

Matematik lingvistika va kompyuter lingvistikasi Qozog‘istonda izchil rivojlandi. Amalga oshirilgan dastlabki tadqiqotlar, asosan, statistik yo‘nalishga oid bo‘lib, keyinchalik qozoq tilshunoslari mashina tarjimasi, avtomatik tahrir, tillarni o‘qitish kabi sohalarda izlanishlar olib bordilar. Qozoq kompyuter lingvistikasining shakllanishida prof. Q.B.Bektayevning xizmati katta. Olim ilmiy faoliyati davomida matematik va injener lingvistikasiga oid asarlar yaratgan. Bu tadqiqotlarda til va nutq birliklarini EHM yordamida tadqiq etishning har tomonlama muhimligi ta’kidlangan, amaliy ishlar – dastur va lug‘atlar ustida tadqiqot olib borilgan¹.

Q.B.Bektayev rahbarligidagi «Lingvostatistika va avtomatlashtirish» guruhi a’zolari M.Avezovning «Abay yo‘li» romani (4 kitob) tilining

¹ Bu haqida qarang: Po‘latov A., Muhamedova S. Kompyuter lingvistikasi. – T., 2009. Rahimov A. Kompyuter lingvistikasi asoslari–T., 2011.

chastotali lug‘ati so‘zligini kompyuter yordamida tuzib chiqishgan. Mazkur chastotali lug‘atda 20000 dan ortiq leksema va 60000 so‘z shakli («glossema») mavjud bo‘lib, ular 466000 marta qo‘llangan. Ushbu lug‘at qozoq tilshunosligida kompyuter lingvistikasining rivojlanishiga asos bo‘ldi.

O‘zbek tilining Davlat tili maqomiga ega bo‘lishi, jahon miqyosidagi mavqeyining ko‘tarilishi, o‘zbek tilini o‘rganishga intilishning kuchayishi mazkur tilni o‘rgatishga mo‘ljallangan kompyuter dasturlariga bo‘lgan ehtiyojni yanada orttirmoqda. Negaki, bunday dasturlar o‘zbek tili leksikasini oson o‘zlashtirishga, u yoki bu tushunchalarni ifodalovchi so‘zning ma‘nosi, sinonimik variantlarini topishga imkoniyat yaratadi. Ma‘lumki, har qanday kompyuter dasturi uchun lingvistik ta‘min zarur bo‘ladi. Lingvistik ta‘min, ya‘ni ma‘lumotlar bazasi muayyan tilning grammatik qonuniyatlari, semantik imkoniyatlari asosida yaratiladi. O‘zbek tilidagi barcha so‘z turkumlari asosida ma‘lumotlar bazasini yaratish va uni jahon axborotlar bankiga kiritish tilning funksional doirasini kengaytirishga xizmat qiladi. Til o‘rgatish dasturlarining yaratilishi nazariy va amaliy tilshunoslik mutanosibligiga mos ravishda, nazariy tadqiqotlar natijalarini inson manfaatlari yo‘lida tatbiq qilishga imkoniyat yaratadi. Bugungi o‘zbek tilshunosligida hal qilinishi lozim bo‘lgan masalalar tadqiqotlarning xalq manfaati va mamlakat taraqqiyoti uchun yo‘naltirilishini taqozo etadi.

Hozirgi kunda o‘zbek tili AQSh, Angliya, Germaniya, Janubiy Koreya, Yaponiya, Xitoy kabi rivojlangan davlatlarda o‘qitilayotgani, o‘zbek tilini o‘rganuvchilar va bu tilda gaplashuvchilar sonining ortib borayotgani, o‘zbek tilida kompyuter dasturlarining yaratilayotgani, xullas, o‘zbek tilini dunyoviy tillar darajasiga olib chiqishga qaratilgan sa’y-harakatlar – O‘zbekistonning tashqi til siyosatini belgilaydi.

V.A.Avrarin fikriga ko‘ra, til siyosatining bevosita obyekti tilning funksional tomoni hisoblanadi. Tilning funksional tomoni tashqi ta’sirlarga sezuvchan bo‘ladi. Jamiyat tilning struktura tomonini ham ma‘lum darajada boshqarishi mumkin, biroq uni bevosita emas, tilning funksional tomoni vositasida boshqaradi¹. Odatda, tilning struktur tomoniga maktab, adabiyot, teatr, targ‘ibot va boshqa ommaviy axborot vositalari orqali,

¹Аврорин В.А. Проблемы изучения функциональной стороны языка (к вопросу о предмете социолингвистики). –Л., 1975. –С. 177.

ya'ni tilning funksional tomoni bilan bevosita aloqador bo'lgan omillar va sharoitlar vositasida ongli ta'sir ko'rsatiladi¹.

Xorijda, asosan, AQSh, Ukraina, Belorussiya, Rossiya, Xitoy, Yaponiyada kompyuter lingvistikasiga oid amalga oshirilgan tadqiqotlarning ko'lami keng. O'zbekistonda esa bu soha keyingi yillarda rivojlandi. Bajarilgan tadqiqotlar kam miqdorda bo'lsa-da, soha taraqqiyotining asosiy yo'nalishlarini belgilashda muhim o'rin tutadi. Amalga oshirilgan ishlar faqat statistik tahlil yo'nalishi bo'yicha bo'lib, ayrim ilmiy ishlar tarkibida masalaga shunchaki to'xtab o'tilgan, xolos. Bu holat o'zbek tilshunosligida kompyuter lingvistikasi bo'yicha hal qilinishi lozim bo'lgan masalalar ko'p ekanligini ko'rsatadi.

XX asr o'rtalariga kelib O'zbekiston fan-texnika sohasida ham kompyuterga oid "mashina tarjimasini", "mashina tilshunosligi" kabi bir qancha atamalar faollasha boshladi. Ammo bu atamalar, asosan, matematikaga oid izlanishlarda kuzatildi. Keyinchalik tilshunoslikda ham matematik ifodalash, kompyuter tizimi borasida izlanishlar olib borildi.

O'zbek tili materiallari bo'yicha kompyuter tilshunosligiga oid tadqiqotlar olib borgan olimlar H.Arziqulov, S.Rizayev va S.Muhamedovlardir. Mazkur olimlar, asosan, kompyuter lingvistikasining statistik tahlil yo'nalishi bo'yicha izlanish olib borishgan.

H.Arziqulov «*Информатика и переработка текста средствами вычислительной техники*» nomli monografiyasida kompyuter yordamida katta hajmdagi matnlarni qayta ishlashning avtomatik tizimi asoslarini tahlil qilgan. Professor H.Arziqulov rahbarligida Samarqand davlat chet tillari institutida injener tilshunosligi markazi tashkil etildi. Bu markaz davlat tilidagi matnlarni kompyuter yordamida qayta ishlash masalalari yuzasidan ilmiy kuzatishlar olib borgan. Matnlarni qayta ishlash kompyuter dasturlarini yaratish, o'zbek tilining barcha sathlariga oid matnlarni yig'ish, ularni muayyan bir tizimga keltirish, indeksatsiya qilish, algoritmlarini tuzish, ularga statistik jihatdan yondashish, inglizcha-o'zbekcha mashina tarjimasini yaratish sohasidagi ilmiy tadqiqot ishlarini boshqargan. Markaz vakillari o'zbekcha matnlarni qayta ishlaydigan «Uzlington» avtomatik tizimini, lotin alifbosidagi matnlarni kirill yozuviga va kirill yozuvidagi matnlarni lotin alifbosiga o'tkazuvchi «Spellchecker» dasturini, kompyuter yordamida tahrir qiladigan, imloviy xatolarni tuzatadigan kompyuter dasturlarini yaratishga birinchilardan bo'lib kirishgan.

¹ Швейцер А.Д. Современная социолингвистика: Теория. Проблемы. Методы. –М.: Наука, 1976. –С. 148.

H.Arziqulov o‘zbek mashina tilini yaratish maqsadida so‘z tarkibini tadqiq qildi. Olimning morfema va so‘z yasaliş strukturalari haqidagi fikrlari, matn statistikasi va chastotali lug‘atlar tuzish prinsiplari, o‘zbek tilidagi matnlarni kompyuter yordamida qayta ishlovchi hamda avtomatik tahrir dasturlarini yaratish borasidagi jiddiy tadqiqotlari o‘zbek tilshunosligida injener lingvistikasining shakllanishiga asos bo‘ldi, kompyuter lingvistikasining yuzaga kelishini ta‘minladi.

S.Muhamedov, R.R.Piotrovskiy bilan hammualliflikda yozgan kitobida lingvistik modellar, modellashtirish jarayoni haqida ma‘lumot keltiradi, o‘zbekcha matnlarning kvantitativ modellarini taklif qiladi.¹ Mazkur asarda turk gazeta matnlarining leksik-morfologik jihatdan amalga oshirilgan avtomatik tarjiması ham berilgan. Sun‘iy intellekt yaratish va injener lingvistikasi metodlari orqali o‘zbekcha nutqni avtomatik qayta ishlash jarayonida foydalanish mumkin bo‘lgan o‘zbekcha matnlarning leksik -morfologik tuzilishi statistik tahlili natijalari keltirilgan.

S.Rizayev tilshunoslikda aniq metodlarning qo‘llanishi va bunda EHMdan foydalanish, til va nutq hodisalariga statistik yondashuvning sabablari hamda matnlarni avtomatik qayta ishlash, mashina tarjiması muammolari, shuningdek, harflar chastotasini aniqlashda EHMdan foydalanishga doir ma‘lumotlar bergan.²

S.Rizayevning nomzodlik dissertatsiyasi ham o‘zbek adabiy tilining grammatik va fonologik sistemasini statistik tekshirishga bag‘ishlangan.³

S.Muhamedov tomonidan yaratilgan asarlar⁴ kompyuter lingvistikasiga oid keyingi tadqiqotlar uchun manba sifatida xizmat qildi.

¹ *Мухамедов С.А., Пиотровский Г.Г.* Инженерная лингвистика и опыт системно - статистического исследования узбекских текстов. –Т.: Фан , 1986.

² *Ризаев С.* Бўғин структурасини ўрганиш тажрибасидан. – Т.: Фан, 1975; *Ризаев С.* Кибернетика ва тилшунослик. –Т., Ўзбекистон, 1976, 16 б.; *Ризаев С.А., Юсунов Д.Ю.* О дистрибутивно – статистическом исследовании лексико –морфологической структуры слова современного узбекского языка // Исследования по литературоведению и языкознанию. –Т.: Фан, 1977; *Ризаев С.* Ўзбек совет болалар адабиёти тилининг частотали луғати. –Т.: Фан, 1980; *Ризаев С., Бўронов Н.Б.* А.Қахҳор «Синчалак» повести тилининг частотали луғати. –Т.: Ўқитувчи, 1986.

³ *Ризаев С.* Ҳозирги ўзбек адабий тилининг грамматик ва фонологик системасини статистик текшириш. Филол.фан.номз...дисс. –Т., 1970, 236 б.

⁴ *Мухамедов С.* Статистический анализ лексико-морфологической структуры узбекских газетных текстов. Автореф.дисс...канд.филол.наук. –Т., 1980, 25 с.; shu muallif. Ўзбек тилининг алфавитли-частотали луғати. –Т.: Фан, 1982, 110 б.; shu muallif. Лингвистика и нечетко – множественная природа естественного языка // Опыт экспериментального и структурного изучения. –Т.: Фан, 1982.

A.Ayimbetovning doktorlik dissertatsiyasida statistik metodlardan foydalanilgan holda o‘zbek, qozoq va qoraqalpoq tillarining yaqinligi isbotlangan.

1996-yil 20-21-mayda Samarqand shahrida sohibqiron Amir Temurning 660 yilligiga bag‘ishlangan "Injenerlik tilshunosligi va til o‘qitish jarayonini kompyuteriashtirish" muammolariga doir II xalqaro ilmiy anjuman o‘tkazildi. Bu anjumanni akad.R.G.Piotrovskiy "Injenerlik tilshunosligi va mustaqil davlatlarda informatsion industriyaning tikianishi" mavzusidagi ma‘ruzasi bilan boshlab berdi. Shuningdek, anjumanda "Turkiy tillar mashina fondining dolzarb masalalari", "Injenerlik tilshunosligi va lug‘atshunoslik", "Nutq sistemasi va uning unsurlarini kompyuter yordamida andozalash" singari o‘ttizdan ortiq qiziqarli ma‘ruzalar tinglandi hamda muhokama qilindi. Bu anjuman materiallari alohida to‘plam holida nashr etilgan.¹

O‘zbek kompyuter lingvistikasini shakllantirish va rivojlantirish maqsadida O‘zbekiston Milliy universitetida 2001-yilda kompyuter lingvistikasi laboratoriyasi tashkil etildi. Laboratoriyada o‘zbek tilining matematik va kompyuter modellarini yaratish, o‘zbek tilidagi matnlarni tahrir qilish, inglizcha-o‘zbekcha kompyuter tarjima dasturlarini ishlab chiqish bo‘yicha ilmiy va amaliy-tadqiqiy ishlar olib borildi. Laboratoriyada nafaqat kompyuter tilshunosligi, balki kompyuter adabiyotshunosligiga oid izlanishlar ham olib borildi. Lingvistikaga doir masalalar, ya‘ni tillarni o‘qitish, bilimlarni baholash, matnlarni tahrir qilish, bir tildan ikkinchi tilga tarjima qilish bilan bir qatorda, adabiyotshunoslikning kompyuter bilan hamkorlikdagi vazifalari: badiiy asarni referatlash, dunyo miqyosidagi kutubxona, adabiyot uchun obyektiv va universal tahlil dasturlarini yaratish, asarlarni tartibga solish, joylashtirish, qidirishga oid tadqiqotlar ham amalga oshirildi.

O‘zbek tilshunosligida kompyuter lingvistikasi yo‘nalishlariga doir ilmiy maqolalar e‘lon qilingan.²

A.Po‘latovning N.Jo‘rayeva, S.Muhamedovalar bilan hammualliflikda yozgan «Разработка формальной модели грамматики узбекского

¹ *Йўлдошев Б.* Компьютер лингвистикаси. – Самарқанд, 2009. –Б.9-17.

² *Пулатов А.К., Джураева Н.* Разработка формальной модели грамматики узбекского языка // Узбекский математический журнал. 2002, №1. –С.47-54; *Пулатов А.К., Алиходжаев Б., Джураева Н.* Разработка программы компьютерного анализа и синтеза глаголов узбекского языка // ЎзМУ хабарлари. 2002, 2-сон. –Б.17. *Пулатов А.К., Муҳамедова С.Х.* Компьютер тилшунослигида автоматик тахрир килувчи дастурнинг лингвистик таъминотини яратиш асослари // Ўзбек тилшунослигининг долзарб масалалари. –Т., ТДПУ илмий тўплами, 2003. – Б.32-36.

языка» maqolasida o‘zbek tili grammatikasi formal modeli namunasi keltirilgan. Tilning formal modeli matematik mantiqning aksiomatik nazariya konsepsiyasi asosida ishlab chiqilgan.

O‘zMU magistrantlari U.Do‘simova va G.Valiyevalarning magistrlik dissertatsiyalarida o‘zbek tilshunosligida ilk bor (rasmiy-idoraviy uslub asosida) ayrim fe’l va otlarni modellashtirish muammolari tadqiq qilindi.¹

So‘nggi yillardagi bitiruv malakaviy ishlari, magistrlik dissertatsiyalarining ma’lum qismi kompyuter lingvistikasi masalalariga bag‘ishlangan. Jumladan, 2014 yilda tezaurus lug‘at mavzusiga doir magistrlik dissertatsiyasi himoya qilindi². 2016-yilda L. Abduhamidova avtomatik tahrirga doir bitiruv ishini amalga oshirdi³. M.Abjalova o‘zbek tilidagi matnlarni kompyuter vositasida tahrir qilish dasturining lingvistik ta’mini masalalariga bag‘ishlangan magistrlik dissertatsiyasini himoya qildi.⁴ Kompyuter lingvistikasiga doir qo‘llanmalar nashr qilindi⁵.

Kompyuter lingvistikasiga oid quyidagi mavzular bo‘yicha bitiruv malakaviy ishlari va magistrlik dissertatsiyalari tayyorlanmoqda:

1. O‘zbek tilining rasmiy-idoraviy va kompyuter uslublarini ishlab chiqish.

2. O‘zbek tili grammatikasining kompyuter modelini yaratish.

3. Tabiiy tillarning optimal modelini ishlab chiqish.

4. O‘zbek-rus-ingliz tillarining qiyosiy tahlili.

5. O‘zbek kompyuter adabiyotshunosligi asoslarini aniqlash.

6. O‘zbek tilidagi internet sahifalarini (site) yaratish asoslarini ishlab chiqish.

7. “Inglizcha-o‘zbekcha tarjimon” kompyuter dasturini yaratish.

8. O‘zbek tilida matnlarni avtomatik tahrir qilish dasturlarini yaratish.

9. O‘zbekcha-ruscha-inglizcha grammatik terminlar lug‘atini yaratish.

10. O‘zbek tili fani bo‘yicha dasrliklarning (maktab, kollej, litsey va oliy o‘quv yurtlari uchun) elektron versiyalarini ishlab chiqish.

¹ *Дўсимова У.* Матндаги феълларни автоматик тахрир қилувчи дастурнинг лингвистик таъмини (расмий-идоравий услубдаги матнлар асосида). Магистрлик дисс. –Т., 2002, 56 б.; *Валиева Г.* Расмий-идоравий услубнинг лисоний бирликларини моделлаштириш. Магистрлик дисс.–Т., 2003, 60 б.

² *Курбонова Ф.* Ҳолат феъллари тезауруси. МД. –Т.,2014.

³ *Abduhamidova L.* Компьютер лингвистикасининг автоматик тахрир yo‘nalishi. ВМІ. –Т.,2015.

⁴ *Абжалова М.* Ўзбек тилидаги матнларни компьютер воситасида тахрир қилиш дастурининг лингвистик таъмини. Магистр илмий даражасини олиш учун ёзилган диссертация. –Т., 2011.

⁵ *Абдурахмонова Н.* Машина таржимаси. –Т., 2008; *F. Qurbonova.* Компьютер lug‘atlari: tezaurus. –Т.,2014; *Abduhamidova L.* Tilshunoslikning yangi yo‘nalishlari: kompyuter lingvistikasi. –Т., 2015.

11. Ijtimoiy va psixologik omillarning tilga ta'sirini o'zbek tili va jamiyati misolida tadqiq etish.

12. Inglizcha-o'zbekcha, o'zbekcha-inglizcha kompyuter lug'atlarini ishlab chiqish.

13.O'zbek tilining fundamental bazasini shakllantirish.

14.O'zbek tilining til korpusini shakllantirish.

15.O'zbek tilida mashina tarjimasini shakllantirish.

16.O'zbek tilidagi axborot uslubini shakllantirish.

17.Matnlarni "BORIS tizimi" asosida tahlil qilish.

18.Kompyuter tizimi asosida nogrammatik ifoda ko'rinishlarini aniqlash.

19. Kompyuterga xoslangan o'zbek tili tezaurusini yaratish (tezauruslar nafaqat to'g'ri yozish, balki vaqtni tejashni va har bir berilayotgan so'z haqida bir vaqtning o'zida har tomonlama, to'liq ma'lumotlarga ega bo'lishni ta'minlaydi).

Bunday masalalar olimlardan va yosh izlanuvchilardan katta bilim va mahorat talab etadi. Hozirga qadar ruslarda, inglizlarda tezaurus lug'atlarning turli ko'rinishlari mavjud.

O'zbek tilshunosligidagi kompyuter lingvistikasining bir qator yo'nalishlari bo'yicha f.-m.f.d., professor A.Po'latovning ilmiy maqolalari, "Kompyuter lingvistikasi" o'quv qo'llanmasi (2011), S.Muhamedova bilan hamkorlikda yozgan o'quv qo'llanmasi (2008,2009) muhim ahamiyatga ega. B.Yo'ldoshevning "Kompyuter lingvistikasi", A.Rahimovning "Kompyuter lingvistikasi asoslari" o'quv qo'llanmalari mazkur fanni o'zbek tilidagi ta'lim yo'nalishlarida o'qitishda asosiy adabiyotlar sifatida xizmat qilib keladi.

Mutaxassislar va tilshunoslar o'zbek tilining matematik va kompyuter modellarini yaratish, o'zbek tilidagi matnlarni tahrir qilish, inglizcha-o'zbekcha kompyuter tarjima dasturlarini ishlab chiqish bo'yicha ilmiy va amaliy-tadqiqiy ishlar olib borishmoqda. Lingvistikaga doir masalalar, ya'ni tillarni o'qitish, bilimlarni baholash, matnlarni tahrir qilish, bir tildan ikkinchi tilga tarjima qilish kabi masalalar o'rganilmoqda.

Ko'rinadiki, so'nggi o'n yillikda o'zbek tilshunosligida kompyuter lingvistikasining avtomatik tahrir (AT), mashina tarjimasini (MT), avtomatik tarzda til o'qitish (ATO') kabi yo'nalishlarida bir qator tadqiqotlar, amaliy ishlar yuzaga keldi. Mavjud tadqiqotlar tavsifiy xarakterda bo'lib, o'zbek kompyuter lingvistikasining shu yo'nalishdagi tadqiqotlari uchun manba vazifasini o'taydi. Axborot va kompyuter texnologiyalari rivojlangan bugungi kunda o'zbek kompyuter lingvistikasi taraqqiyot yo'lini tutdi.

Kompyuter lingvistikasiga bag‘ishlangan ilmiy ishlar hajmi, ko‘lami kengaydi. Mashina tarjimasiga bag‘ishlangan doktorlik dissertatsiyasilari himoya qilindi.¹ O‘zbek tilidagi matnlarni tahrir va tahlil qiluvchi dasturning lingvistik modullarini ishlab chiqish maqsadi qo‘yilgan tadqiqotlar amalga oshirildi². O‘zbek tili mualliflik korpusini tuzishning lingvistik asoslari monografik planda ishlab chiqildi³. O‘zbek tili milliy korpusining sinonim so‘zlar bazasini shakllantirishga doir ilmiy tadqiqot yaratildi⁴.

Bugungi kunda o‘zbek tilshunosligida kompyuter lingvistikasi bilan bog‘liq hal etilishi lozim bo‘lgan quyidagi vazifalarni ta’kidlash mumkin:

- o‘zbek tilining kompyuter uslubini yaratish;
- axborot matnlaridagi qoliplilik, qisqalik mezonlarini belgilash;
- kompyuter izohli va tarjima lug‘atlarini yaratish;
- o‘zbek tili va adabiyot darsliklarining elektron versiyasini ishlab chiqish;
- kompyuterda inglizcha-o‘zbekcha tarjima dasturlari uchun lingvistik ta’min yaratish;
- kompyuterdagi matnlarni avtomatik tahrirlash dasturlarini yaratish.

Bu vazifalarni amalda bajarish uchun tilshunoslar va kompyuter mutaxassislari birgalikda ish olib borishlari lozim.

O‘zbek kompyuter lingvistikasida ilk marta bajarilgan amaliy ishlar sifatida kirill va lotin yozuviga asoslangan yangi o‘zbek alifbosining kompyuter dasturlariga ilova qilinishini, o‘zbek tilidagi ba’zi o‘yin dasturlari, ba’zi darslik va qo‘llanmalar, “orfoslovar”lar, kirill yozuvidan lotin yozuviga, lotin yozuvidan kirill yozuviga tabdil qiluvchi “TransEdit” tizimi hamda antivirus dasturlarini keltirish mumkin.

O‘zbekistonda kompyuter lingvistikasi sohasini rivojlantirish maqsadida Xalqaro konferensiyalar o‘tkazilmoqda, davlat ahamiyatiga molik hamkorlik aloqalari yo‘lga qo‘yilmoqda. 2018-yil 18-20 oktyabrda Alisher Navoiy nomidagi Toshkent davlat o‘zbek tili va adabiyoti

¹ *Абдурахмонова Н.* Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти. Филол. фан. фалс. д-ри. (PhD) diss. –Т., 2018.

² *Абжалова М.* Ўзбек тилидаги матнларни таҳрир ва таҳлил қилувчи дастурнинг лингвистик модуллари (расмий ва илмий услубдаги матнлар таҳрири дастури учун). Филол. фан...(PhD) дисс. –Т., 2019.

³ *Захаров В.П., Богданова С.Ю.* Корпусная лингвистика. –Иркутск: ИГЛУ, 2011. – С.5.

⁴ *Эшмўминов А.* Ўзбек тили миллий корпусининг синоним сўзлар базаси: Филол. фан. фалс. д-ри (PhD) дисс. автореф. – Қарши, 2019. –Б.21.

universitetida kompyuter lingvistikasi masalalariga bag‘ishlangan “Turklang.2018” Xalqaro konferensiyasi o‘tkazildi¹.

Kompyuter lingvistikasi yirik sohadir. Uni qo‘shimcha fan sifatida chuqur o‘zlashtirib bo‘lmaydi. Dunyoning rivojlangan mamlakatlarida bu sohada yuzlab olimlar, o‘nlab ilmiy tekshirish institutlari ish olib borayotgan, yo‘nalishlar, fakultetlar, xalqaro miqyosdagi uyushmalar, yig‘inlar tashkil etilayotgan bir paytda o‘zbek kompyuter lingvistikasini rivojlantirish masalalari imkon doirasida hal etilmoqda. Birinchi vazifa sohaning har tomonlama yetuk mutaxassisini shakllantirishdan iborat. Shu maqsadda universitetlarda alohida kompyuter lingvistikasi yo‘nalishlarini ochish, fakultetlar tashkil etish lozim bo‘ladi. Mazkur yo‘nalishda qiyosiy tilshunoslik, matematika, mantiq, kompyuter dasturlash tili, o‘zbek tili, xorijiy til kabi fanlardan olingan bilimlar kompyuter lingvistikasi bo‘yicha malakali mutaxassis shakllantirishni ta‘minlaydi. Kompyuter lingvistikasi rivojlangan xorijdagi yirik ilmiy tadqiqot markazlari bilan muttasil ilmiy aloqalar yo‘lga qo‘yiladi.

Dunyo kompyuter tarmoqlarida o‘zbek tilidagi ma‘lumotlar bazasining yaratilishi va o‘zbek tilidagi matnlarni avtomatik tarzda boshqa tillarga (yoki boshqa tildagi matnlarni o‘zbek tiliga) tarjima qiluvchi dasturlarning yaratilishi dolzarb ahamiyat kasb etadi.

Tilshunoslikning yangi yo‘nalishi bo‘lgan kompyuter lingvistikasi bir qator nazariy va amaliy tilshunoslik yo‘nalishlarining vujudga kelishiga zamin hozirlaydi.

Savol va topshiriqlar:

1. Internet orqali kompyuter lingvistikasi bo‘yicha xalqaro konferensiya materiallarini o‘rganing.
2. Kompyuter lingvistikasiga oid xorijdagi tadqiqotlarni tasniflang.
3. Avtomatik tarjima muammosiga doir izlanishlar haqida ma‘lumot bering.
4. O‘zbek tilshunosligidagi kompyuter lingvistikasiga oid tadqiqotlarga tavsif bering.
5. Kompyuter lingvistikasi fanini rivojlantirish uchun xizmat qiladigan takliflaringizni bayon eting.
6. Kompyuter lingvistikasining shakllanish tarixini yoriting.
7. Kompyuter lingvistikasining milliy til taraqqiyotidagi o‘rnini klaster usulida tasvirlang.

¹ VI International Conference on Computer Processing of Turkic Languages. –Tashkent, 2018. www.turklang.uz, www.navoiy-uni.uz, www.complinglab.uz.

3-§. “KOMPYUTER LINGVISTIKASI”NING MATEMATIK MANTIQ VA BOSHQA SOHALAR BILAN ALOQADORLIGI

Reja:

1. Kompyuter lingvistikasida mantiqiy birliklar va tushunchalarning o‘rni.
2. Grammatikaga mantiqiy yondashuv.
3. Semiotika va informatsiya nazariyasida tilning kod sifatidagi talqini.
4. Algoritm – dasturlash asosi.

Asosiy tushunchalar: *mantiq, mantiq funksiyalari, grammatika, semiotika, algoritm, tilshunoslikda matematik metodlardan foydalanish, matematik mantiq, mulohazalarni hisoblashning formal nazariyasi, “uchinchisi mustasno” qonuni, oraliq uchunchi, dialektik mantiq, konyunksiya, dizyunksiya, implikatsiya, ekvivalensiya, lingvistik protsessor, tabiiy tilli interfeys, semantika, pragmatika, kod, kodlash, dekodlash, denotat, signifikat, algoritmik tillar, translyator dasturlar (kompilyatorlar).*

Tilni algoritmik jihatdan tasvirlash jarayoni matnni analiz va sintez qilishni taqozo etadi. Bu esa “mazmun – matn” muammosi faqat sof lingvistik muammo bo‘lib qolmay, balki matnni bir qator logik-matematik interpretatsiya qilish (sharhlash) jarayonini ham taqozo etuvchi murakkab muammo ekanligini ko‘rsatadi. Bu jarayonlarning barchasi matematik mantiq (logika) elementlari bilan aloqadordir.¹

Mantiq, grammatika, semiotika va algoritm kompyuter lingvistikasining muhim omillari hisoblanadi. Ular til va nutqning funkcionallashuv jarayonini anglashda hamda lingvistik muammolarni avtomatik hal etish tizimini yaratishda asos bo‘lib xizmat qiladi. Mantiq, grammatika, semiotika va algoritm bir-biri bilan chambarchas bog‘liq tushunchalar bo‘lib, ular kompyuter lingvistikasi masalalari doirasida bir-birini to‘ldiradi, taqozo etadi.

Matematik mantiq, dastlab, tabiiy fanlarda qo‘llanib keldi, keyinchalik u gumanitar sohalarga ham tatbiq etildi. Mantiq – muhokama yuritishning qonun-qoidalari, usullari va formalari haqidagi fan bo‘lib, unda tushuncha, hukm, xulosa chiqarish kabi mantiqiy jarayonlar

¹ *Йўлдошев Б.* Компьютер лингвистикаси. – Самарқанд, 2009. – Б.17

o‘rganiladi. Demakki, har bir fan mohiyatida u yoki bu ma’noda mantiqiy tushunchalar amal qiladi.

Mantiq dastlab Qadimgi Xitoy va Hindistonda miloddan avvalgi V asrda grammatika bilan uzviy bog‘liq holda shakllana boshladi. Miloddan avvalgi IV asrda Qadimgi Yunonistonda mantiq mustaqil fan maqomiga ega bo‘ldi. Bu hodisa yunon faylasufi Aristotel nomi bilan bog‘liq, u mantiq fani asoschisidir. Olim “Organon” (bu so‘z “qurol”, “vosita” ma’nolarini anglatadi) asarini yozib, tafakkur qonuniyatlarini birinchi bo‘lib tushuntirdi, deduksiya nazariyasini – mantiqiy xulosa chiqarishning formal xarakterga ega ekanligini ko‘rsatdi, mantiqni fikrlash quroli sifatida tavsifladi. Aristotel nominativ (otli) ifodalarning mantiqini ishlab chiqdi. Keyinroq Aristotel izdoshlari va stoiklar maktabi vakillari jumla mantiqini yaratdilar.¹ Aristotel ta’limoti Farobiy, Beruniy, Gegel, Frege, Leybnis kabi buyuk olimlar tomonidan davom ettirildi. Mantiq fani turli yo‘nalishlarda rivojlana boshladi. Dastlab formal mantiq yuzaga keldi, unda hodisalar faqat formal asosda tushuntiriladi. Bu yo‘nalishda *uchinchisi mustasno* qonuni amal qiladi. Unga ko‘ra, bayon qilingan ikki zid fikrdan biri to‘g‘ri, boshqasi noto‘g‘ri, lekin uchinchi holat bo‘lishi mumkin emas. Bu qonun «A –V dir yoki V emasdir» formulasi orqali beriladi. Ko‘p yillar davomida ushbu qarash to‘g‘ri, deb baholab kelindi.

XVIII – XIX asrlarda falsafada hodisalarni doimo rivojlanish va o‘zaro ta’sirda o‘rganuvchi dialektik ta’limot mantiqqa ham ta’sir ko‘rsata boshladi. Natijada dialektik mantiq yo‘nalishi yuzaga keldi. Bu yo‘nalishga nemis faylasufi Georg Vilgelm Fridrix Gegel asos soldi. U o‘z qarashlarini 4 jildli “Mantiq fani” nomli monografiyasida bayon qildi. Olim mantiqqa *oraliq uchinchi* qoidasini olib kirdi. Unga ko‘ra, hodisalarni baholashda ikkita bir-birini inkor etuvchi hodisa o‘rtasida neytral – “oraliq uchinchi” holat ham ajratiladi. Masalan:

issiq	Iliq	Sovuq
A	V	S

Dialektik mantiqqa asosan hodisalar ko‘p parametrlarda baholanadi. Keyinchalik Leybnis va Gilbertlar mantiq faniga matematik simvollarni tatbiq etdi. Nemis faylasufi va matematigi G.Leybnis (1646-1716) birinchilardan bo‘lib mantiqiy fikrlashning hisob xarakteriga ega ekanligini ko‘rsatdi. Uning fikricha, barcha ilmiy tushuncha va

¹ Белонозов Г.Г. Компьютерная лингвистика и перспективные информационные технологии. - М.: Русский мир, 2004. - С. 15.

mulohazalarning asosini mantiqiy elementlar tashkil qiladi. Mana shu mantiqiy elementlarni muayyan simvollar bilan belgilash mumkin bo‘ladi.

Leybnis g‘oyalari XIX asrdagina amalga oshdi. Nemis olimi G.Frege, ingliz olimi J.Bul, rus olimi V.Poretskiy ishlarida simvolik mantiq asoslari yaratildi.⁹ Xususan, J.Bul mantiq operatsiyalari uchun quyidagi simvollarni:

-predmetlar mavjudligini belgilash uchun: **x, y, z** harflarini;

-predmetlar sifatini belgilash uchun: **X, Y, Z** harflarini;

-bir-biriga akslantirilgan mulohazalar uchun **perpendikulyar** \perp belgisini;

- bildirilayotgan mulohazalarda yo‘qlikni bildiruvchi «**0**» belgisini;

- mulohazalarning mantiqiy qo‘shilishini bildiruvchi belgi sifatida «**+**» ni;

- 2- mulohazaning 1-mulohazaga zidligini ifodalovchi «**-**» belgisini;

- mulohazada odatiy so‘z birikuvlari mavjudligini bildirish uchun «**U**» belgisini;

- mulohazadagi noodatiy birikuvlarni ko‘rsatish uchun «**∩**» belgisini taklif qildi.

Antik davrda ham, o‘rta asrlarda ham grammatika va mantiq yonma-yon o‘qitilgan. Ayniqsa, o‘rta asrlarda grammatika va mantiq klassik ta‘lim poydevori bo‘lgan “yetti erkin san’at” (*artesliberales*) tarkibiga kirgan. Bunda grammatika, ritorika (notiqlik san’ati) va mantiq “formal san’atlar” (*artes-formales*) sifatida *trivia* (uchlik)ga birlashgan. Arifmetika, geometriya, musiqa va astronomiya “real san’atlar” (*artesreales*) sifatida *kvadrivie* (to‘rtlik)ga birlashgan. Ushbu fanlarni o‘zlashtirgan insongina yetuk va barkamol hisoblangan.

Tabiiy tilning strukturasi o‘rganishda mantiq, grammatika va semiotika uzviy aloqadorlikda ish ko‘radi. Mantiq va grammatika orasidagi aloqadorlik til va tafakkur, so‘z va tushuncha, gap va hukm, ega va subyekt, kesim va predikat kabi kategoriyalarning dialektik munosabatida namoyon bo‘ladi.

Mantiqiy kategoriyalar	Grammatik tushunchalar
Subyekt	Ega
Predikat	Kesim
Obyekt	To‘ldiruvchi
Atribut	Aniqlovchi
Predmet, hodisa	Ot so‘z turkumi
Jarayon (harakat, holat)	Fe‘l so‘z turkumi
Belgi (sifat)	Sifat so‘z turkumi

Miqdor	Son so‘z turkumi; son kategoriyasi
Vositalar; munosabatlar	Kelishik, egalik, ko‘makchilar, bog‘lovchilar. ¹

Til tafakkur bilan chambarchas bog‘liq bo‘lib, fikrni moddiylashtiradigan, kishilarning fikr almashishini ta‘minlab beradigan quroldir. Til bilan tafakkur bir-birini taqozo etadigan ajralmas hodisadir. Tilsiz tafakkur bo‘lmaganidek, til ham tafakkursiz mavjud bo‘lmaydi. Tafakkur inson oliy nerv sistemasining, ya‘ni bosh miyaning faoliyati bo‘lib, nutqda o‘z ifodasini topadi, har qanday so‘zlardan tarkib topgan ifoda qolipi orqali reallashadi.

Tafakkurning reprezentatsiyasi til hisoblanadi. Ong amalda til formasida mavjud. Til bilan tafakkur murakkab, qarama-qarshiliklarga boy bo‘lgan dialektik birlikni tashkil etadi. Til fonetika, grammatika qonunlari asosida tashkil topgan material hodisa bo‘lsa, tafakkur obyektiv borliqning inson miyasida aks etgan oliy formasi bo‘lgan ideal hodisadir. Til qonunlari bilan tafakkur qonuniyatlarining tuzilishi va rivojlanishi bir xil emas. Tilning tizim sifatidagi qonunlarini grammatika o‘rganadi. Tafakkur qonunlari va birliklari esa mantiqda o‘rganiladi. Grammatik kategoriya va mantiqiy kategoriya orasidagi munosabat ham juda murakkabdir. Grammatik kategoriyalarni mantiq, mantiqiy kategoriyalarni grammatik kategoriyalarga bo‘ysundirish yoki ularni qorishtirish mumkin emas. Bu kategoriyalar ba‘zida o‘zaro mos kelishi, ba‘zida mos kelmasligi ham mumkin.

Kompyuter lingvistikasida inson tafakkuri va tilini modellashtirish masalasi markaziy muammo hisoblanadi. Sun‘iy intellekt, tabiiy tilni qayta ishlash (NLP), lingvistik protsessorlar, kompyuterga mo‘ljallangan tabiiy tilli interfeys bo‘yicha olib borilayotgan tadqiqotlar mantiq, grammatika, semiotika, kompyuter texnologiyalari sohalari mutanosibligida amalga oshirilmoqda. Bu sohada juda ko‘p yutuqlarga erishildi, lekin hali-hanuz tabiiy til, inson tafakkuri to‘la modellashtirilgani yo‘q. Ayni shu masala – til va tafakkurni modellashtirish kompyuter lingvistikasi fanining istiqbolli yo‘nalishi sanaladi.

Semiotika yoki semiologiya (yunoncha *semiotike* “belgilash”, “tasvirlash”, “ifodalash”) belgilar sistemasi haqidagi fan bo‘lib, axborotni saqlovchi, uzatuvchi turli xil belgi va belgilar tizimlarining yaratilishi, vazifalari, belgilar sistemasiga kiruvchi hodisalar, ularning tabiati, mohiyati, turlari, ko‘rinishlarini o‘rganadi. Semiotika tilshunoslikda,

¹ Лазута О.Н. Логика и лингвистика. - Новосибирск, 2000. - С.7.

xususan, kompyuter lingvistikasida ham muhim metodologik ahamiyat kasb etadi. Negaki, til ham belgi sanaladi, ma'lum ma'noda ramzlar tizimidir. Semiotikada tabiiy tillar ham, sun'iy tillar ham, yo'l belgilari, audio va videosignallar ham tadqiq etiladi.

Semiotika fani XIX asr oxirlarida vujudga keldi, uning asoschisi amerikalik olim Charlz Sanders Pirs (1839-1914) hisoblanadi. Strukturalizm asoschisi Ferdinand de Sossyur "Umumiy tilshunoslik kursi" asarida bu fanni "semiologiya" deb nomlagan. Ushbu fan yutuqlaridan hozirgi zamon kompyuter, avtomatik boshqaruv, telekommunikatsiya, uyali telefon tizimlari, sun'iy yo'ldoshlar va boshqa sohalarda keng foydalaniladi. Semiotikada belgilar tizimi to'rt aspekt asosida o'rganiladi: 1) *sintaktika* – belgilar strukturasi 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, foydalilik darajasi, amaliy jihatlarini tadqiq etadi.

Tilning semiotik tabiati, tushunchalarni ifodalashda turli ramzlardan, belgilardan foydalanish kompyuter lingvistikasi uchun dastlabki omillardan sanaladi.

Semiotika belgilar haqidagi keng qamrovli nazariy-falsafiy fan bo'lib, tilni belgilardan iborat butunlik sifatida o'rganadi. Tilshunoslik va semiotikaning o'zaro maqsadli aloqaga va munosabatga kirishuvidan lingvosemiotika sohasi shakllandi. Lingvosemiotika tilning belgi tabiatiga egaligini, informatsiya uzatuvchi til birliklarining xususiyatlarini tadqiq qiladi. Til birliklari, tilning tarkibini tashkil qiluvchi struktural elementlar – morfema, so'z, so'z birikmasi, gap, matn va boshqalar muayyan informatsiyani berish, yetkazish – ijtimoiy vazifani bajarish uchun xizmat qiluvchi belgilar sistemasi sifatida tadqiq qilinadi. Shunga ko'ra til jamiyatda mavjud bo'lgan boshqa belgilar bilan umumiylikka, o'xshashlikka ega bo'ladi. Shunga ko'ra, til semiotika fanining – lingvosemiotikaning o'rganish obyektiga aylanadi. Tilning tabiatini, mohiyatini, jamiyatdagi o'rnini yoritish uchun uni faqat qaysidir til bilan qiyoslashning o'ziga yetarli bo'lmaydi.

Tilni – til sistemasi tarkibidagi birliklarni, belgi deb olingan elementlarni boshqa belgilar – nolisoniy belgilar sistemasi bilan ham, jamiyatga faol xizmat qilayotgan ramzlar bilan ham, sun'iy tillar bilan ham qiyoslash lozim bo'ladi.

Ma'lumki, nutq birliklari (so'z, so'z birikmasi, gap va boshqalar) muayyan belgilar sifatida ma'no, tushuncha, fikr ifodalaydi. Til birliklari

ma'no, tushuncha, fikr kabilarning simvoli (ramzi) yoki belgisi hisoblanadi. Shunga ko'ra, Ferdinand de Sossyur tilni fikr ifoda etuvchi belgilar tizimi sifatida e'tirof etadi.

Til birliklari o'zaro zaruriy, mantiqiy aloqaga kirishib, til tizimini shakllantiradi, uning obyektivligini, bir butunligini ta'minlaydi, tilning ijtimoiy vazifa bajarishi uchun imkoniyat beradi, shart-sharoit yaratadi. Ushbu jarayonda til birliklari o'ziga xos xususiyatlarga egaligi, nisbiy mustaqilligi, umumiy va farqli jihatlari bilan ajralib turadi.

Til tizimining asosiy birliklari bo'lgan fonema, morfema va so'zlarning ijtimoiy vazifa bajarishdagi eng muhim jihati ularning ikki xil xususiyatida ko'rinadi. Ushbu birliklar ifoda va mazmun jihatiga egaligi, shu nuqtayi nazardan umumiy hamda farqli belgilari bilan o'zaro ajralib turadi.

So'zning mazmun tomonini to'g'ridan-to'g'ri ma'lum qilish mumkin bo'lmaganligi uchun uning tovush jihati, ramzi talaffuz qilinadi. Bu ramz (belgi) tinglovchi ongida tegishli tushuncha, tasavvur uyg'otadi. Demak, til sistemasida ramziy munosabat tovush va ma'no orasidagi shartli bog'lanish tarzida yuzaga chiqadi. Ma'noda aks etadigan aniq narsa yoki hodisa tasavvuri (denotat) ikkinchi planda bo'ladi.

So'z mazmun-mundarijasida ikki hodisa farqlanadi: leksik ma'no (signifikat) va denotat.

Denotat so'z ma'nosi emas, balki so'z yordamida ko'rsatiladigan, nomlanadigan, ataladigan hodisadir. Ma'no esa, qayd etilganidek, obyektiv borliq elementlariga xos bo'lgan xususiyatlarning inson ongida mavhumlik kasb etishi, umumlashuvi natijasida shakllanishidir. Nutqda so'z denotatning (obyektiv borliq elementining) lingvistik ifodasi hisoblanadi. So'zda tovush ifodalovchi (ramz, belgi) bo'lsa, ma'no ifodalanuvchi bo'ladi.

So'zdan yirik so'z birikmalari va gaplar ham ifoda va mazmun jihatlariga ega, ularning mazkur ikki jihati ko'proq grammatika va semantikaning aloqasi doirasida qaraladi. Nutq birliklari ham, aslida, ifoda va mazmun jihatiga ega so'zlarning birikuvidan, o'zaro bog'lanishidan yuzaga keladi. Shu sababli ularni o'rganish tarkibidagi so'zlarni tahlil qilishdan boshlanadi.

Kompyuter tizimiga oid informatsiya nazariyasida ham tilning semiotik tabiati namoyon bo'ladi. Informatsiya nazariyasi informatsiyani uzatish va saqlashning mazmun tomonini emas, balki uning statistik qurilishi bilan shug'ullanadi, ya'ni informatsiya nazariyasi axborot matnini statistik struktura sifatida o'rganadi. Statistik struktura axborotdagi signal

(simvol) va signallar birlashuvining voqelanish chastotasi soni, miqdoridir. Signalning ko‘p marta takrorlanishida signal orqali uzatiladigan axborotning miqdorini aniqlash juda muhim. Demak, informatsiya nazariyasi muayyan axborotdagi ma’lumotlarni ushbu ma’lumotlarning tabiatiga - mazmuniga bog‘liq bo‘lmagan holda o‘lchash (tekshirish) yo‘llarini, usullarini ishlab chiqadi. Shunga ko‘ra, informatsiya nazariyasining muhim va asosiy xulosalari tilshunoslikda ham o‘z ifodasini topadi. Informatsiya nazariyasining asoschisi amerikalik olim K.Shennon informatsiyaning quyidagi aloqa kanallari bo‘yicha uzatilishini sxema asosida tushuntiradi:

*Manba —► uzatuvchi—p-kanal—&■ priyomnik—► axborotni oluvchi.*¹

Ushbu holat til vositasida insonlarning kommunikatsiyaga kirishuv jarayoniga ham to‘la mos keladi. Aslida, ushbu model til, tafakkur va nutq munosabatidan kelib chiqqan. Informatsiya tabiiy yoki sun‘iy tillar orqali uzatilar ekan, til sistemasi kodlash va dekodlash jarayonlariga bog‘liq holda talqin etiladi.

Informatsiya nazariyasi ma’lumotlarini o‘rganish jarayonida maxsus terminlardan – tushunchalardan foydalaniladi. Bular quyidagilar:

1. **Kod** – axborotni yozib olish vositasi (usuli).
2. **Matn** (tekst) – muayyan axborotni berish, uzatish jarayonidagi belgilarning (til birliklarining, simvollarning) izchilligi, ketma-ket kelishi.
3. **Kanal** – axborot beriluvchi, uzatiluvchi muhit, sharoit.
4. **Informatsiya** asosiy birlik bo‘lib, u **bit** (yoki **binit** inglizcha “binary unit” – (“ikkilik birligi”) deb ataladi.²
5. **Kodlash** va **dekodlash** – informatsiyani mashina (kompyuter) tushunadigan simvolga o‘tkazish va, aksincha, simvoldan tabiiy tilga aylantirish jarayoni.

Tashqi olam haqidagi informatsiya insonning bosh miya qobig‘ida model va strukturalar ko‘rinishida saqlanadi, aks etadi va qayta ishlanadi. Ushbu holatning fiziologik asosi nerv-sensor hujayralar – aksonlar hisoblanadi. Bosh miya qobig‘ida so‘z tizimlari hamda semantik maydon shakllanadi. Funkzionallashuv jarayoni (xabarning bir individumdan boshqasiga yetkazilishi va qabul qilinishi) da til diskret shaklda – nutq ko‘rinishida namoyon bo‘ladi. Nutqning idrok qilinish jarayoni nutq tovushlaridan informatsiyaning ajralishi, tarqalishidir. Eshitish organi havoning tebranishini qabul qiladi, signalni bosh miya qobig‘iga jo‘natish

¹ Эко У. Отсутствующая структура. Введение в семиологию. –М.: Петрополис, 1998. –С.45.

² Qarang: *Rahimov A.* Kompyuter lingvistikasi asoslari. –T., 2011

uchun uzluksiz impulslarni kodlashtiradi. Birinchi bosqichda bosh miyada tovushlar zaxirasidan “tovushlar modeli” (N.I.Jinkin buni “*fonemalarning barqaror katakchalari*” deb atagan) shakllanadi. Inson miyasidagi obrazlar va tushunchalar tovushlar kombinatsiyasi ko‘rinishida moddiy qobiqqa ega bo‘ladi. Inson xotirasida so‘zlarning modellari aniq hujayralar kompleksiga muvofiq ravishda saqlanadi. Informatsiyani solishtirish natijasida miyada yangi kod – so‘z kodi shakllanadi va uni qobiqning keyingi bosqichiga uzatadi. U yerda informatsiyaning yangi modeli shakllanadi, bunda oldingiga nisbatan kodlashtirish oson kechadi, chunki miya o‘zida mavjud bo‘lmagan ramzlarnigina kodlashtiradi, keyingi so‘zni kodlashtirish oldingisidan ko‘ra ham oson kechadi, keyingi bosqichda – birikmalar kodlashtiriladi hamda informatsiya keyingi bosqichga uzatiladi. Bunda yuqoridagi jarayon qaytariladi: informatsiya ajratiladi va mazmun modellashtiriladi.¹

Tilni informatsiya nazariyasiga ko‘ra kod bilan tenglashtirish, kod bilan parallel qo‘yish mumkin.² Chunki kod, aytilganidek, axborotni yozib olish usuli, vositalaridan biridir. Demak, til, til birliklari kod sifatida muayyan axborot tashiydi, uzatadi, yetkazadi. Binobarin, ayni vaqtda, ushbu axborotni hisoblab chiqish, statistik ma’lumot to‘plash mumkin. Xullas, til semiotika va informatsiya nazariyasiga ko‘ra kod sifatida namoyon bo‘ladi.

4-§. MULOHAZALARNI HISOBLASHNING FORMAL NAZARIYASI

Reja:

- 1.Funksiyalar ekvivalentligi.
- 2.Matematik mantiqda fikrlarning rost yoki yolg‘onligi.
- 3.“Algoritm” tushunchasi.

Asosiy tushunchalar: *mulohazalar algebrasi (MA), mulohaza, rost yoki yolg‘on mulohaza, gap, darak gap, mantiqiy amallar, o‘zgaruvchi mulohaza, inkor amali, rostlik jadvali, konyunksiya amali, mantiqiy ko‘paytma amali, dizyunksiya amali, mantiqiy yig‘indi amali, implikasiya amali, ekvivalentlik amali, teng kuchlilik amali, Shiffer amali yoki Shiffer shtrixi, asosiy chinlik jadvallari.*

¹Qarang: *Амосов Н.М.* Моделирование мышления и психики. –Киев,1965.

²*Шемакин Ю.И.* Начало компьютерной лингвистики. –М.: Высшая школа, 1992. -С.7.

Mantiq algebrasi funksiyalari quyidagi nomlanishlarga ega.

1. 0 – konstanta (o‘zgarmas) 0, ya’ni mutlaqo xato (yolg‘on) gap: *Bolalar sho‘x bo‘lmaydi.*

2. 1-konstanta 1, ya’ni mutlaqo to‘g‘ri gap: *Bolalar, odatda, sho‘x bo‘ladilar.*

3. X – bir-biriga aynan o‘xshash funksiya: *Bolalar sho‘x bo‘ladilar. Bolalar bor joyda shovqin bor.*

4. X – X ni rad etish, yoki «X emas»: *Kattalar bolalardek sho‘x emas. Kattalar mo‘min – qobildir.*

5. (X1 & X2) – konyunksiyasi X1 va X2 o‘rniga X1 & X2 belgisi ishlatiladi, «&» belgisi «-u» yuklama-bog‘lovchisi yoki «va» bog‘lovchisini modellashtiradi: *Bolalar sho‘x bo‘ladilar va bolalar bor joyda shovqin bor.*

6. (X1 v X2) – X1 va X2 dizyunksiyasi. X1 v X2 operatsiyasi «yoki» bog‘lovchisini modellashtiradi: *Kattalar bolalardek sho‘x emas yoki kattalar mo‘min-qobildir.*

7. (X1 ⇒ X2) X1 va X2 implikatsiyasi. X1 ⇒ X2 operatsiyasi «agar, ... u holda» bog‘lovchi nisbatini modellashtiradi: *Agar bolalar sho‘x bo‘lsa, u holda kattalar mo‘min-qobildir.*

8. (X1 | X2) – bu belgi «Sheffir tayoqchasi» deb nomlanadi, u «va ... emas» bog‘lovchisiga to‘g‘ri keladi: *Bolalar sho‘x va kattalar sho‘x emas.*

9. X1FX2 – bu belgi «Pirs strelkasi» deb nomlanadi, u «yoki ... emas» bog‘lovchisiga to‘g‘ri keladi: *Kattalar mo‘min-qobil yoki kattalar sho‘x emas.*

Funksiyalar ekvivalentligi. Elementar funksiyalar xususiyatlari.

Ta’rif: N va D formulalari, agar ularga mutanosib bo‘lgan f_n va f_D funksiyalar teng bo‘lsa, ekvivalent deb ataladi. N&D modeli N va D formulalari ekvivalent ekanligini bildiradi. A = V – bu belgi mantiqiy teng kuchlilik (ekvivalensiya) uchun ishlatiladi.

Masalan, 1.1. $0 + (x \& x)$

2.2. $X_1 \& X_2 + X_2 \& X_1$

Elementar funksiyalar xususiyatlarini xarakterlovchi quyidagi ekvivalentliklar (ayniliklar) ro‘yxatini keltirish mumkin. Har qanday (X1 & X2) funksiyalardan birini $X_1 \circ X_2$ bilan belgilaymiz,

$(X_1 \vee X_2), (X_1 \oplus X_2)$

1. $(X_1 \circ X_2)$ funksiyasi assotsiativlik xususiyatiga ega.

$((X_1 \circ X_2) \circ X_3) + (X_1 \circ (X_2 \circ X_3))$

2. $(X_1 \circ X_2)$ funksiyasi kommutativlik xususiyatiga ega.

3. Dizyunksiya va konyunksiyani rad qilish orasida o‘zaro munosabat mavjud.

Bu ayniliklar osonlikcha tekshirilishi mumkin. Formulani yozishni soddalashtirish maqsadida quyidagicha tartib belgilanadi: «&» operatsiyasi «V» operatsiyasidan kuchlidir, agar qavslar bo‘lmasa, unda avval «&» operatsiyasi, so‘ngra esa «V» operatsiyasi bajariladi. Bundan tashqari, assotsiativlik qonuniga binoan $(X_1 \circ X_2)$ uchun $((X_1 \circ X_2) \cdot X$ va $(X_1 \circ (X_2 \circ X_3))$ formulalari o‘rnida $(X_1 \circ X_2 \circ X_3)$ ifodalaridan foydalanish mumkin.

Matematik mantiqda fikrlarning rost (to‘g‘ri) yoki yolg‘onligi matematik yo‘l bilan isbotlanadi. Bunda xi argument (o‘zgaruvchi) deb belgilanadi. Agar $x = 1$ bo‘lsa, mutlaq rost hukm, $x = 0$ bo‘lsa, mutlaq yolg‘on hukm deb olinadi. Matematik mantiq turli simvollar bilan ish ko‘radi. Ular mantiqiy bog‘lovchilar deb ham yuritiladi. Ushbu simvollar sun‘iy programmalashtirish tillarida ishlatiladi.

∃ = inglizcha <i>exists</i> so‘zining qisqartmasi bo‘lib, mavjudlik kvantori sifatida ishlatiladi.
∑ = summa, barcha narsaning jami.
∀ = inglizcha <i>all</i> so‘zining qisqartmasi bo‘lib, hammasi uchun e kvantori sifatida ishlatiladi.
x, y, z – predmetlarning borligini belgilash uchun ishlatiladi.
X, Y, Z – predmetlarning sifatini (atributlarini) belgilash uchun ishlatiladi.
S-subyekt
P-predikat
A = 0 – mutlaqo noto‘g‘ri
A = 1 – mutlaqo to‘g‘ri
A – emas (inkor). Bir mulohaza ikkinchi bir mulohazaga zid kelganda ishlatiladi.

Matematik mantiqda mulohazalarning rost yoki yolg‘onligini formal nazariya asosida tekshirib ko‘rish mumkin. Bunda mantiqiy hukmlar shartli ravishda simbolik belgilar, mantiqiy bog‘lovchilar yordamida belgilab olinadi. U gapning propozitsional ko‘rinishi deb nomlanadi. So‘ng jumla formal nazariya asosida hisoblanadi. Bunda hukmlarning qiymatini ko‘rsatuvchi jadval yordam beradi.

X1,X2	X1&X2	X1VX2	X1=>X2	X1 = X2
00	0	0	1	1
01	0	1	1	0
10	0	1	0	0
11	1	1	1	1

Formal mantiq va simvolik mantiq til strukturasi o'rganishda amaliy ahamiyat kasb etadi, grammatikani formallashtirishda, sun'iy dasturlash tillarini yaratishda asos bo'lib xizmat qiladi, kompyuter lingvistikasining o'ziga xos simvolik tili sifatida namoyon bo'ladi. Mantiq kompyuter lingvistikasining murakkab va ko'p aspektli muammosi hisoblangan tabiiy tilni qayta ishlash (NLP) uchun formallashtirish, analogiya, modellashtirish kabi optimal metodlarni taklif etadi¹.

Algoritm. Algoritm (bu so'zning etimologiyasi buyuk alloma Al-Xorazmiy nomi bilan bog'liq. Lotin tilida "ibtido" ma'nosidagi algoritmus so'zi ham mavjud) – ma'lum bir tipga oid masalalarni yechishda ishlatiladigan jarayonlar tizimining muayyan tartibda bajarilishi haqidagi aniq qoida yoki qoidalar tizimi.²

O'rta asrlarda sanoqning o'nli sanoq sistemasi bo'yicha to'rt arifmetik amal bajariladigan qoida *algoritm* deb atalgan. Bu qoidalarni fanga IX asrda qomusiy olim Muhammad al-Xorazmiy kiritgan. XX asrda algoritmlar kibernetikaning nazariy va mantiqiy asosi sifatida tan olindi.

Qo'shuv, ayiruv, ko'paytiruv, bo'luv kabi arifmetik qoidalar, kvadrat ildiz chiqarish qoidalari, har qanday ikki natural son uchun eng umumiy bo'luvchini topish usuli eng oddiy arifmetik misollardir. Aslida, biror vazifani umumiy tarzda hal qilish vositalariga ega bo'lish algoritm asosida ish korishni taqozo qiladi. Har bir hayotiy hodisa muayyan algoritm asosida amalga oshiriladi. Masalan, ta'lim jarayoni, ishlab chiqarish tizimi, mehnat faoliyatini amalga oshirish o'ziga xos algoritmik tartibga ega.

Kompyuter texnologiyalari taraqqiyoti bilan bog'liq holda odam bilan mashina orasida o'ziga xos ko'prik vazifasini bajaruvchi sun'iy tillar yaratildi, ular dasturlash jarayonida qo'llanadi. Sun'iy tillar buyruqlarning tizimli algoritmiga tayanib yaratilgani uchun ular algoritmik tillar (programmallashtirish tillari) deb ham ataladi. Dastlabki algoritmik til 1958-yilda Parijda yaratilgan va "Algol" (to'liq nomi inglizcha – "Algorithmic language") deb atalgan. Hozir jahonda 500 dan ortiq algoritmik til mavjud.

Algoritmik tillardagi matnlar xuddi tabiiy tildagidek aniq interpretatorga - axborotni qabul qilish va tushuish uchun insonga yoki avtomatik qurilmaga yo'naltirilgan bo'lishi lozim. Agar bu jarayon (qabul qilish va tushunish) amalga oshmasa, u holda algoritmik tillar va

¹Bu haqida qarang: *Po'latov A., Muhamedova S. Kompyuter lingvistikasi. – T., 2009. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.*

² Фалсафа. Қисқача изоҳли луғат. – Т.: Шарқ, 2004. - Б. 18.

kompyuter orasida vositachi sifatida translyator dasturlar (yoki kompilyatorlar)dan foydalaniladi.¹

Algoritmik tillar uch asosiy komponentdan iborat bo‘ladi: 1) alfavit – asosiy simvollar majmuyi; 2) sintaksis –simvollar kombinatsiyasi orqali jumla hosil qilish qoidalari; 3) semantika – algoritmik tillardagi birliklarning mazmun tomoni. Algoritmik tillar, asosan, imperativ xarakterga ega bo‘ladi, ya’ni asosan buyruqlar tizimiga tayanadi (*begin* – boshla, *end* – tugalla kabi. Bu jihatdan o‘zbek tili asosidagi algoritmik tilni shakllantirishda turkiy tillarga xos tub fe‘l-asoslarning buyruq maylida kelishi qulay imkoniyatlardan biri ekanligini ta’kidlash joiz). V.A.Uspenskiy va A.L.Semyonovlar algoritmnini shunday ta’riflaydilar: *“Algoritm nafaqat algoritmlar nazariyasi, matematika va informatikaning, balki umuman zamonaviy fanlarning asosiy tushunchalaridan biridir. Axborot asrining kirib kelishi bilan algoritmlar sivilizatsiyaning eng muhim faktorlaridan biri bo‘lib qoldi”*. Algoritmik tillarning buyruqlar tizimi asosidagi imperativ xarakterda bo‘lishi haqida shunday yozadilar: *Algoritm “buyruqlashtirilgan”, “ijro etilishi, bajarilishi shart” mazmunini ifoda etadi. Aytish mumkinki, algoritmlarni o‘rganuvchi nazariya buyruq gaplar lingvistikasi sifatida talqin etilishi mumkin»*.²

Algoritmlar, asosan, quyidagi shakllarda tuziladi: a) so‘z ko‘rinishida – tabiiy til yordamida; b) grafik ko‘rinishida – tasvir, chizma va grafikli simvollar yordamida; c) algoritmlarning yarim formallashtirilgan tavsifi hisoblangan psevdokod ko‘rinishida (dasturlashning dastlabki bosqichlarida qo‘llanadi). Bunda dasturlash tillari unsurlaridan ham, tabiiy tildagi so‘z va birikmalardan ham foydalaniladi; d) programma ko‘rinishida –dasturlash tillaridagi belgilar yordamida³.

Algoritm buyruqlar tizimi, qat’iy bajariladigan amallar ketma-ketligi hamda dasturlashning tayanch nuqtasi sanalgani bois u orqali kompyuter dasturlarini ishlab chiqish va muammoning avtomatik tarzda (mashina orqali) hal etilishini amalga oshirish mumkin. Algoritmnini keng miqyosdagi vazifalarni hal etishga yo‘naltirish, algoritm nazariyasini mukammallashtirish, ayniqsa, kompyuter texnologiyalari, shuningdek, kompyuter lingvistikasi rivojlanayotgan davrda dolzarb ahamiyat kasb

¹ Кревский И.Г., Селиверстов М.Н., Григорьева К.В. Формальные языки, грамматики и основы построения трансляторов. - Пенза, 2003. - С.7. Qarang: Rahimov A. Kompyuter lingvistikasi asoslari. –T.: Akademnashr, 2011.

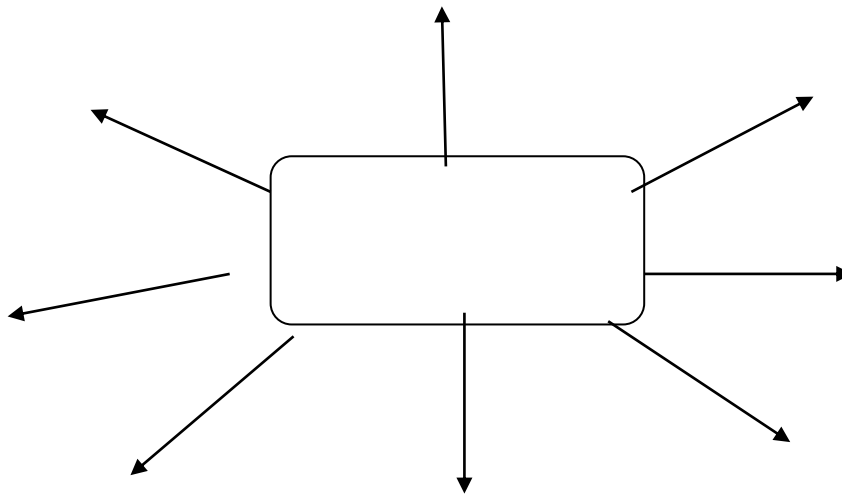
² Успенский В.А., Семенов А.Л. Теория алгоритмов: основные открытия и приложения. – М.: Наука, 1987. –С.13.

³ Bu haqida qarang: Po‘latov A., Muhamedova S. Kompyuter lingvistikasi. – T., 2009. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

etadi. Shu o‘rinda ta’kidlash joizki, kompyuter lingvistikasida shunday masalalar ham borki, ularni algoritmlar asosida hal etish mumkin emas. Jumladan, sun’iy intellekt tizimi doirasidagi neyron tarmoqlar, sifatli avtomatik tarjima kabilarni hal etishda «algoritmgaga asoslangan yondashuv» («rule based approach»)dan tashqari, analogiya prinsipiga asoslanadigan «namunalarga asoslangan yondashuv» («example based approach»)ga ham ehtiyoj seziladi.

Savol va topshiriqlar

- 1.Mantiq va semiotikaning kompyuter lingvistikasidagi ahamiyatli o‘rnini yoriting.
- 2.Konyunktiv, dizyunktiv, implikativ hukmlarga o‘zbek tilidan misollar keltiring.
- 3.Matematik mantiq fani masalalariga doir tadqiqotlarni tavsiflang.
- 4.Algoritmik tillarni ayting
- 5.Algoritmgaga hayotiy misollar keltiring.
6. Mantiqiy tushunchalar algoritmini tavsiflang.
7. Quyidagi klaster shaklini to‘ldiring.



II. KOMPYUTER LINGVISTIKASIDA TILSHUNOSLIK MASALALARI

5-§. LINGVISTIK BILIMLAR BAZASI

Reja:

1. “Model” tushunchasining mohiyati.
2. Tilshunoslikda modellashtirish metodidan foydalanish.
3. Modellashtirishning pragmatik jihatlari va cheklangan tomonlari.
4. So‘z turkumlarini modellashtirish.
5. Gap bo‘laklarini modellashtirish.

Asosiy tushunchalar: *model, tabiiy model, matematik model, mantiqiy-matematik model, metamodel, induksiya, deduksiya, lingvistik model, sintaktik modellashtirish, so‘z turkumlarini modellashtirish, so‘z turkumlari simvollari, gap bo‘laklarini modellashtirish, gap bo‘laklari simvollari, asosiy gap konstruksiyalari, analogiya metodi.*

Modellashtirish, garchi bugungi kunda yangi yo‘nalish, zamonaviy metod sifatida qaralayotgan bo‘lsa-da, uning ildizlari qadim davrlarga borib taqaladi. Stenogrammalarni modellashtirishning dastlabki ko‘rinishlari sifatida qayd etish mumkin. Stenografiya grekcha “tor, zich yozish” ma’nosidagi so‘z bo‘lib, yozuv tezligining 4-7 baravar oshirilishiga asoslanadi. Stenografiyaga asos bo‘lgan turli yozuv ko‘rinishlari mavjud. Geometral yozuv –geometrik elementlardan foydalanilgan yozuv: nuqta, to‘g‘ri chiziq, aylana, aylana qismi, geometrik shakldagi harflar. Morfologik yozuvda morfemalardan, fonetik yozuvda tovushlar ifodasi bo‘lgan harflardan foydalanilgan.

N.N.Sokolov stenogramma sifatida oddiy grafik elementlar tizimidan – “qo‘shib yoziladigan belgi”lardan foydalangan: *st, sn, str., pr.* Stenogramma sifatida maxsus belgilar qo‘llangan: so‘zning boshlanish qismlari (*pac-, za-, nep-*), so‘zning oxirgi qismlari (*-ение, -ства*), so‘zning o‘zagi (*здрав-, держ-*). *JD-железная дорога.* So‘zlarning boshlanish qismi asosidagi qisqartirish ko‘p uchraydi: *результат=рез., время=вр.*

Stenografiyada 10ga yaqin ideogrammadan foydalaniladi:

“*frazeogramma*” – so‘z o‘rtasida tushirish:

a, ya unlilari tushirib qoldiriladi;

sifatlardagi qo‘shimchlar yozilmaydi;

satr usti va satr osti belgilari unlilarni ifodalashda qo‘llangan (arab yozuvi).

Stenografiyadan qadimgi misrda fir'avn nutqini yozib olish maqsadida foydalanilgan. Keyinchalik bu yozuv shakllari yunonlar va rimliklarga o'tgan. Stenografiya eramizdan avvalgi 63-yilda shakllangan. Rimda stenografistlarni tez yozuvchilar, "notariylar" deb atashgan. Rim imperiyasi davrida stenografiya maxsus o'qitilgan. 1887 yilda birinchi stenografiya kongressi bo'lib o'tgan¹.

Model (lotincha «modelus» so'zidan olingan bo'lib, «nusxa», «andaza», «o'lchov», «me'yor» ma'nolarini anglatadi) tabiiy fanlar yoki, umuman, fanda muayyan original – obyekt haqidagi ma'lumotlar majmuyi sifatida yuzaga kelgan hosila – obyekt, moddiy qurilma, grafik, sxema, umuman, bilish vositasi. Boshqacharoq tushuntirganda, model tabiiy obyektlarning imitatsiyasidir (o'xshashi, taqlidiy ko'rinishi), u o'zbek tilidagi "qolip", "andaza" so'zlariga mos keladi. Model hodisalarning yuzaga kelishi uchun asos vazifasini o'taydi, bunda aniq yoki mavhum obyektlar kichraytirilgan obyektlar, sxemalar, chizmalar, fizikaviy konstruksiyalarda ifoda etiladi. Model, dastavval, *moddiy* va *fikriy* modelga bo'linadi. Masalan, olmani xarakterlovchi belgilar, atributlar – uning dumaloqligi, mevaligi, shirinligi o'sha tushunchaning fikriy modeli hisoblanadi. Agar olma loydan yoki sun'iy bir materialdan yasalsa, bu uning moddiy modeli hisoblanadi. Yoki globus yerning modeli, o'yinchoq mashina haqiqiy mashinaning modeli sifatida baholanishi mumkin.² Model quyidagi asoslarga ko'ra bilishda muhim hisoblanadi:

- o'rganish obyektini soddalashtiradi;
- o'rganish obyektini boshqa obyektlar ta'siridan ajratadi;
- obyektini ta'riflashni osonlashtiradi.

Model dastlab amaliy sohalarida ishlatilgan, matematika, kimyo fanlariga keng tatbiq qilingan. Keyinchalik ilm-fanning ijtimoiy sohalariga ham kirib keldi.

¹ Internet materiallari asosida tayyorlandi.

² <http://www.wikipedia.org/wiki/model.html>

Lingvistik model tushunchasi struktur tilshunoslikning E. Sepir, L.Blumfild, R. Yakobson, N. Xomskiy, Z. Harris faoliyatida qo'llangan.



Modellarni shartli ravishda quyidagicha tasniflash mumkin:

1. *Tabiiy modellar* – o'rganilayotgan obyekt bilan bir turda bo'ladigan va undan faqat o'lchamlari, jarayonlarining tezligi va ba'zi hollarda yasalgan materiali bilan farq qiladigan modellar.

2. *Matematik modellar* – prototipdan (asl nusxadan) jismoniy tuzilishi bilan farq qiladigan, lekin prototip bilan bir xil matematik tasvirga ega bo'lgan modellar.

3. *Mantiqiy-matematik modellar* – belgilardan iborat bo'lgan, tafakkur jarayonini o'rganishda qo'llanadigan abstrakt modellar.

4. *Kompyuter modellari* – matematik, mantiqiy modellashtirish metodlariga asoslanib, kompyuterda algoritm va dasturlardan foydalangan holda yaratiladigan modellar.

Modellashtirishning quyidagi turlarini ham ajratib ko'rsatish mumkin: informatsion modellashtirish, lingvistik modellashtirish, matematik-kartografik modellashtirish, iqtisodiy-matematik modellashtirish, molekulyar modellashtirish, pedagogik modellashtirish, psixologik modellashtirish, statistik modellashtirish, struktur modellashtirish, fizik modellashtirish, imitatsion modellashtirish, evolutsion modellashtirish, kognitiv modellashtirish¹.

XX asr II yarmidan boshlab tilshunoslikda matematik metodlarni qo'llash ommalashdi. Tilshunoslikda qo'llanadigan matematik metodlar, asosan, uch xil: to'plam nazariyasi, mantiq algebrasiga asoslangan transformatsion metod, statistik metod.

¹ Po'latov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2009; Rahimov A. Kompyuter lingvistikasi asoslari. –T.,2011.

1.To‘plam nazariyasi. Har qanday elementlarning yig‘indisi **to‘plam** deyiladi. To‘plamning ayrim elementlari, ba’zi hollarda kesishadi.

$N(a, b, c) \cap M(a, e, d)$. Bu yerda a kesishgan nuqta hisoblanadi.

Nazariy to‘plam metodidan tilshunoslikda ko‘p ma’noli so‘zlar tahlilida foydalanish mumkin. Polisemantik so‘zlar ma’nolari orasida integral semalar kuzatiladi, ular o‘rtasida mantiqiy aloqadorlik bo‘ladi. Leksemalarning polisemantik ekanligi to‘plamlik belgilari asosida tahlil qilish, to‘plam elementlari va belgilari orasidagi kesishuvni topish yordamida aniqlanadi. Masalan, *ko‘z*:

A.Odamning *ko‘zi*.

V. *Yog‘ochning ko‘zi*.

D. *Buloqning ko‘zi*.

E. *Uzukning ko‘zi*.

F. *Ishning ko‘zi*.

G. *Qalb ko‘zi*.

Ko‘z leksemasi bir qator semalardan iborat butunlik. *Ko‘z* leksemasida quyidagi sema (belgi)lar mavjud: 1) “ko‘rish (sezish) organi”; 2) “ovalsimon shaklga egalik”; 3) “juftlik”; 4) “bo‘rtiqlik”; 5) “cheti botiqlik”; 6) “asosiylik; 7)“noziklik”; 8) shaffoflik.

To‘plam	1	2	3	4	5	6	7	8
A to‘plam	+	+	+	+	+	-	+	+
V to‘plam	-	+	+	+	-	-	-	-
D to‘plam	-	-	-	+	-	+	-	-
E to‘plam	-	+	-	+	+	-	-	-
F to‘plam	-	-	-	-	-	+	+	-
G to‘plam	-	-	+	+	-	-	-	-

Sinonimik qatorda dominanta (bosh, asosiy) so‘zni aniqlash ham yuqorida berilgan to‘plam metodi asosida amalga oshirilishi mumkin. Bunda turli nutq uslublaridagi matnlar tahlilga tortiladi va barcha uslublarda qo‘llanishi kuzatilgan so‘z shu sinonimik qatorda dominanta hisoblanadi.

Bunda *ko‘z* leksemasining A, V, D, E, G to‘plamdagi ma’nolari 4-belgi (“bo‘rtiqlik” semasi), A, E to‘plamdagi ma’nolari 5-belgi (“cheti botiqlik” semasi), D, F to‘plamdagi ma’nolari 6- belgi (“asosiylik” semasi), A, F to‘plamdagi ma’nolari 7- belgi (“noziklik” semasi) asosida kesishadi.

2. Mantiq algebrasiga asoslangan transformatsion metodda gapning shakli o‘zgartiriladi, lekin bu uning mazmuniga ta’sir ko‘rsatmaydi.

Masalan, *Biz vatan posbonlaridirmiz. Biz Vatan posbonlari bo‘lamiz. Bizmiz Vatan posbonlari.* Bu metod asoslari Amerika tilshunosligida Ch.Friz va Xomskiylar tomonidan ishlab chiqilgan.

3. **Statistik metod.** Matematik statistika metodidan tilshunoslikda, ayniqsa, til o‘rgatish jarayonida keng foydalaniladi. Buning uchun muayyan tilda qo‘llanuvchi fonemalar, leksemalar, so‘z shakllarining ishlatilish chastotasi tekshiriladi. Masalan, o‘zbek tilida 6ta unli, 23ta undosh tovush, 8ta jarangli undoshning jarangsiz jufti mavjudligi, sinonimik qatorlarning 2ta va undan ortiq (10 tagacha) so‘zdan iborat bo‘lishi, so‘zlarning qo‘llanish miqdori, chastotasi statistika bilan bog‘liqdir.

«Model» tushunchasi fan va texnikada turli ma’nolarda ishlatilgani bois modellashtirish turlarining yagona tasnifi mavjud emas. Tasnif modelning xarakteriga ko‘ra, modellashtirilayotgan obyektning tabiatiga ko‘ra, modellashtirish tatbiq qilinayotgan soha yoki yo‘nalishga qarab amalga oshirilishi mumkin.

Modellashtirish har bir fan obyektini soddalashtiruvchi metoddir. Lingvistik birliklarni modellashtirish bu belgilar tarkibidagi elementlarning barqaror munosabatlariga asoslanadi. Shuning uchun ham butunlik tarkibidagi elementlar o‘rtasida munosabatlarning barqaror va beqaror turlarga ajratilishi lingvistik modellashtirish uchun muhim ahamiyatga ega. Modellashtirish barcha fanlar uchun xos bolgan fanlararo metod hisoblanadi va u quyidagi tamoyillarga amal qiladi:

- deduktivlik – mantiqiy xulosa chiqarishga asoslanish, xususiylikdan umumiylikka borish;

- tafakkur eksperimentidan foydalanish;

- evristik funksiyaga ega bo‘lish, yangi g‘oyalar bera olish va uni amaliyotda sinab ko‘rish imkoniyatiga ega bo‘lish;

- modelni ideallashtirilgan obyekt sifatida talqin qilish.

-eksplanatorlik xususiyatiga, ya’ni tushuntirish kuchiga ega bo‘lish. Shu tamoyillarga asoslansagina model an’anaviy nazariyalar tushuntirib bera olmagan muammoni hal qiladi, obyektning ilgari kuzatilmagan, ammo kelajakda amalga oshishi mumkin bo‘lgan tomonini kashf etadi.¹

Modellashtirish jarayoni uch asosiy tushunchani o‘z ichiga oladi:

- subyekt (tadqiqotchi);

- tadqiqot obyekti;

- o‘rganuvchi subyekt va o‘rganiluvchi obyekt munosabatini aks ettiruvchi model.

¹ Po‘latov A., Muhamedova S. Kompyuter lingvistikasi. - T., 2007. -B.23

Kompyuter asosida modellashtirish bugungi kunda barcha fanlarda, xususan, kompyuter lingvistikasida ham samarali qo‘llanayotgan metod hisoblanadi.¹

Kompyuter asosida modellashtirish quyidagi asosiy bosqichlardan iborat:

- masalaning qo‘yilishi, modellashtirish obyektining aniqlashtirilishi;
- konseptual (tushunchaviy, fikriy) modelning ishlab chiqilishi, asosiy tushunchalarning ajratib olinishi;
- formalizatsiya, ya’ni matematik model bosqichi; algoritm yaratilishi hamda dastur tuzilishi;
- kompyuter eksperimentlarini o‘tkazish;
- natijalar tahlili va talqini.

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

Modellashtirishning obyektini umumlashtirish darajasiga ko‘ra turlari quyidagicha:

1.Lingvistik fakti tavsiflashga qaratilgan analitik model.

2.Oraliq model yoki to‘ldiruvchi model.

3.Maksimal umumlashtirishga asoslangan sintezlovchi model.

Lingvistik model tushunchasi struktur tilshunoslikning E.Sepir, L.Blumfeld, R. Yakobson, N.Xomskiy, Z.Harris, Ch.Hokket kabi namoyandalari qarashlari asosida shakllandi. Lingvistik model taraqqiyoti XX asrning 60 – 70-yillariga (matematik va kibernetik lingvistika rivojlana boshlagan davrga) to‘g‘ri keladi. Lingvistik modelni quyidagi turlarga ajratish mumkin:

1.Inson nutqiy faoliyati modellari. Bu modellar konkret nutq jarayonini va hodisalarini aks ettiradi. Masalan, aniq bir tovushning talaffuz modeli yoki nutqning yuzaga chiqish modeli.

2.Lingvistik tadqiqot modellari. Bunda muayyan til hodisalari asosida olib borilgan tadqiqot jarayoni aks ettiriladi. Masalan, o‘zbek tilida morfologik usul asosida so‘z yasalişining umumiy modeli: **asos + so‘z yasovchi qo‘shimcha**; xususiy modellari: **asos + -chi; asos + -dosh; asos + -do‘z** kabi.

3.Metamodellar – bunda lingvistik modellar saralanadi, u gipotetik-deduktiv xarakterga ega, abstraktlashgan va ratsionallashgan bo‘ladi.

¹ *Grishman R. Computational linguistics // Cambridge University Press. 1994. -P.-7-8. Bu haqida qarang: Rahimov A. Kompyuter lingvistikasi asoslari.-T.,2011.*

Modellashtirish metodi ayrim tillarga, jumladan, ingliz tiliga faol tatbiq qilingan.

Ko'p yillar davomida tilshunoslikda kuzatish metodi yagona metod hisoblanib keldi. Ammo bu metod yordamida lingvistik hodisalarning ichki mohiyati ochilmaydi. Kuzatish metodi orqali so'z shakllari, gap tuzilishi va boshqa struktur xususiyatlarni o'rganish mumkin bo'ladi. Til va nutqning tuzilishi murakkab bo'lganligi sababli kuzatish metodi yordamida ularni to'liq o'rganib bo'lmaydi. Negaki, til va nutq o'ziga xos murakkabliklarga ega.

Keyingi davrlarga kelib til hodisalarini yoritishda modellashtirishdan foydalanila boshladi. Modellashtirish metodida tadqiqotchi obyektning o'zini emas, balki uning modelini o'rganadi. Original bilan model o'rtasida o'zaro bog'liqlik va mutanosiblik mavjud bo'ladi. Tilshunoslikda modelning uch turi farqlanadi:

1. Original modellar – obyektning tuzilishini o'rganadi.

2.Funksional modellar – original modellarning ishlash tartibini o'rganadi.

3.Struktur modellar – har ikkalasi haqida ma'lumot beradi.

Original modellarga quyidagi talablar qo'yiladi:

1.Model tabiiy obyektning aynan nusxasi bo'lishi kerak.

2.Original model o'zida elementlarning murakkab tuzilishini namoyon eta olishi kerak.

3.Tabiiy obyektning hamma xususiyatlari original modelga to'g'ri kelishi lozim.

Gipotetik (nazariy) modellarga xos xususiyatlar sifatida o'z-o'ziga zid emaslik, tavsifning to'liqligi, modelning soddaligini keltirish mumkin.

Yaratuvchi modellar tilshunoslik rivojlanishi natijasida yuzaga keldi. Bunda lingvistik modellar mavjud tushunchani tekshirishdan o'tib, yaratuvchanlik xususiyatini kasb etadi. Yaratuvchi modellar invariantligi asosida variantlar haqida xulosa chiqariladi.

Olimlar modellashtirish metodining bir qator afzalliklari va pragmatik jihatlarni nazarda tutadilar. Buni quyidagicha izohlash mumkin:

- modellashtirish tavsifiy emas, amaliy metod hisoblanadi;

- modellashtirish metodi har qanday sharoitda optimal («eng qulay», «eng maqbul») hisoblanadi;

- modellashtirish metodi tejash prinsipiga tayanadi. Bunda uzun ta'rif va tavsiflarga ehtiyoj bo'lmaydi;

- obyektning tushuntirish va izohlashni osonlashtiradi va soddalashtiradi.

Modellashtirish metodining pozitiv va pragmatik tomonlari bilan birga cheklangan jihatlari ham bor. Ular quyidagilarda namoyon bo‘ladi:

- modellashtirishda faqat struktur va formal belgilarga asoslaniladi. Bunda mazmuniy tomon, semantik qirralar e’tibordan chetda qoladi. Vaholanki, har qanday hodisaning, xususan, lisoniy hodisaning mohiyati shakl va mazmun birligida o‘z ifodasini topadi;

-obyekt haqidagi bilimlar rivojlanib borgan sari modellar eskiradi. Fan, ma’lum ma’noda, avvalgi bilimlarni inkor qilish hisobiga rivojlanadi. Mavjud modellar bilimlar chegarasini nisbatan cheklaydi va xato tasavvurlarga olib kelishi mumkin;

-tabiat va jamiyatdagi hech bir narsa boshqa obyektlar ta’siridan xoli bo‘lmaydi. Ya’ni modellar nisbiylik xarakteriga ega, ularni mutlaq haqiqat sifatida qabul qilib bo‘lmaydi.

Modellashtirish tilshunoslikda strukturalizm yo‘nalishi ta’sirida faol tatbiq qilina boshlandi. Gap strukturasi modellashtirish g‘oyasi XX asrning 50-yillarida amerikalik tilshunos Charlz Friz tomonidan ilgari surildi. Olim o‘z qarashlarini umumlashtirib 1952-yilda «The Structure of English» nomli tadqiqotini yaratdi. Ch.Friz o‘z konsepsiyasini *distributiv model* deb nomlagan. Unga ko‘ra, gap muayyan turkumlarga oid bo‘lgan so‘zlar zanjiri hisoblanadi va tahlilda morfologiya bazasiga tayaniladi. Masalan, «*The young man painted the door yesterday*» jumlasini distributiv model asosida quyidagicha tahlil etiladi: D 3 P 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, mazkur modelda turli turkumlarga mansub so‘z shakllarining nutq zanjiridagi distributsiyasi (tarqalishi, qurshovi) gap strukturasi modellashtirishning asosiy mezonini sanaladi¹.

So‘z turkumlarini modellashtirishda quyidagi shartli belgilar e’tirof etilgan.

So‘z turkumlari modellari:

Ot turkumiga oid quyidagi simvollardan foydalaniladi:

N– birlik shaklidagi ot: *olma*;

Ns – ko‘plik shaklidagi ot: *olmalar*;

N com -turdosh otlar: *qalamlar*;

N – atoqli otlar uchun: *O‘tkir*;

NP– otli birikma uchun: *olma yemoq*.

¹ Bu haqida qarang: Po‘latov A., Muhamedova S. Kompyuter lingvistikasi. – T., 2009. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

Sifat uchun **Adj** yoki **A**; sifatli birikma uchun **AP** simvollaridan foydalaniladi.

Son predmetning miqdori va narsa-buyum, voqea-hodisalar tartibini bildiradi. Son soʻz turkumini ifodalash uchun **Q** simvolidan foydalaniladi.

Feʼl turkumi simvollari:

V – feʼl: *yurmoq*;

Vo – oʻtimli feʼl: *chizmoq*;

Vno – oʻtimsiz feʼl: *kulmoq*;

Vf – feʼlning predikativ (shaxsli shakli – *finite form*) shakli: *koʻrdim*;

Vnf – feʼlning nopredikativ shakli (shaxssiz shakli – *non-finite form*): *koʻrmoq*;

Vn – hozirgi zamon sifatdoshi yoki gerundiy shaklidagi feʼl: *koʻrayotgan*;

Vcn – oʻygan zamon sifatdoshi yoki majhul nisbatdagi feʼl: *koʻrgan*, *koʻrilgan*;

VP – feʼlli birikma: *otda yurmoq*;

aux – yordamchi feʼllar;

mod – modal feʼllar.

Ravish asosan feʼlga bogʻlanib, keladi. Ravish uchun **Adv** yoki **D** simvolidan foydalaniladi.

Olmosh gapda ot, sifat yoki ravish oʻrnida qoʻllanuvchi soʻz turkumidir. Olmosh uchun **Pron** simvolidan foydalaniladi.

Artikl yordamchi soʻz boʻlib, otning maxsus belgilovchisi hisoblanadi. Artikl barcha tillarda ham uchramaydi, artikl uchun **Art** simvoli qoʻllanadi.

Predlog (old koʻmakchi) yordamchi soʻz boʻlib, otning gapdagi boshqa soʻzlarga nisbatan boʻlgan munosabatini (yaʼni maʼkon, vaqt va sabab kabi munosabatlarni) ifodalaydi. Predlog uchun **Prep** yoki **p** simvolidan foydalaniladi.

Bogʻlovchi yordamchi soʻz bolib, gap boʻlaklari yoki gaplarni bogʻlash uchun xizmat qiladi. Bogʻlovchi uchun **Com** simvolidan, yuklama uchun **Part** simvolidan, undov soʻzlar uchun **Interj** simvolidan, taqlid soʻzlar uchun **Mim** simvolidan foydalaniladi.¹

Har bir til lugʻatlarda koʻrsatilmagan soʻz shakllariga ega boʻladi: yasama soʻz + qoʻshimcha

Ingliz tilidagi ayrim qoʻshimchalar jadvali

¹ *Rahimov A. Kompyuter lingvistikasi asoslari. –Toskent: Akademnashr, 2011.*

Qo'shimcha:	Misollar
-s	Words
-is	
-ed	Played
-ing-	Reading ¹
-th	

So'z + boshlang'ich so'z (lug'atdagi so'z) – yasama so'z (shakli o'zgargan so'z)

1-izoh: ta'kidlash zarurki, so'z+so'z+affiks so'z yasash formulasi orqali aniqlanuvchi so'zlar ham sodda so'zlar hisoblanadi, ular lug'atda shu ko'rinishda keltirilgan. Masalan, *boxer, dislike*.

2-izoh: Shuningdek, **so'z+so'z** (**so'z+so'z+so'z** qolipi asosida yasalgan qo'shma so'zlar sifatida *gultojixo'roz, nomozshomgul, gulhamishabahor* kabi birliklar keltiriladi. Aslida bu qo'shma so'zlar ham ikki komponentdan iborat: *gul+tojixo'roz, nomozshom+gul, gul+hamishabahor*) modeli orqali aniqlanuvchi qo'shma so'zlar ham lug'aviy birlik hisoblanadi, chunki ular ham shundayligicha lug'atda ko'rsatilgan bo'ladi. Masalan,

Qo'shimcha:	Misollar
at once	At
Mother in fon	In
The rich	The
A comic	A
on Sunday	On

3-izoh: lug'atda barcha atoqli otlar (insonlar ismi, geografik nomlar va h.k) ham ko'rsatiladi.

So'z turkumlarini formallashtirish: ot

Birlikdagi otdan ko'plikdagi ot hosil bo'lishi:	<i>boy – o'g'il bola</i>
Ko'plikdagi otning birlikdagi otdan ot hosil bo'lishi:	<i>boys – o'g'il bolalar</i>
Kelishikdagi otdan lug'atdagi ot hosil bo'lishi:	<i>boy – o'g'il bola</i>
Kelishikdagi otdan –s ot hosil bo'lishi	<i>boys – o'g'il bolalar</i>

¹Bond.uz.

Bu qoida otlarga nisbatan tatbiq qilinadi. Bunda ko‘plik shakli yasalishining o‘ziga xos formalari ham mavjud bo‘ladi. Ularni maxsus jadvalda ko‘rsatish mumkin: (noto‘g‘ri otlar)

Birlikdagi ot	Ko‘plikdagi ot
Child	Children
Man	-
Goose	geese
Cuctus	Cucti
Bacterium	Bacteria
News	faqat birlik shaklida qo‘llanadi
faqat ko‘plik shaklida qo‘llanadi	Glasses
Fish	Fish

4-izoh: so‘z+so‘z+ -s formulasi nafaqat otning ko‘plik shaklini hosil qilish, balki so‘z yasalishi uchun ham xizmat qiladi:

forse – kuch	forses – armiya
ich – boy	ciches – boylik
colour – rang	colours – bayroq

1.1. Sifat

Ingliz tilidagi sifat 3 ta darajaga ega: oddiy, qiyosiy va orttirma.

Oddiy daraja	Qiyosiy daraja	Orttirma daraja
Lug‘atdagi sifat	lug‘atdagi sifat+-er	(the) lug‘atdagi sifat+-est
Big	bigger	the biggest

Ko‘p bo‘g‘inli so‘zlar:

Oddiy daraja	Qiyosiy daraja	Orttirma daraja
Lug‘atdagi sifat	More-oddiy daraja	(the) most+oddiy daraja
Big	bigger	the biggest
Big	More useful	the most useful

Ushbu qoidaga bo‘ysunmaydigan noto‘g‘ri sifatlar ro‘yxati ham mavjud:

Oddiy daraja	Qiyosiy daraja	Orttirma daraja
Good	Better	(the) most+oddiy daraja
Far	farther	the farthest
	further	the furthest

Son. Sonlar ma‘no jihatidan ikki guruhga bo‘linadi: miqdor sonlar va tartib sonlar:

1) 0 dan 9 gacha bo‘lgan miqdor sonlar lug‘atda keltiriladi. Tarkibli sonlar matematika qoidalariga binoan o‘qiladi. Masalan: 7.687.564;

2) tartib sonlar: miqdor sonlar + th.

Noto'g'ri sonlar bundan mustasno:

One	First
Two	Second
Three	Third

Ravish. Ravish ham ikki guruhga bo'linadi:

1. Tub ravishlar (lug'at bo'yicha).

2. Yasama ravishlar (sifatdan yasalgan ravishlar).

Yasama ravish sifatdan +ly.

Sifat	Ravish
Hard	Hard
Good	Well

Fe'l. Ma'lumki, fe'ning tuslovchi, munosabat shakllari shaxs-son, zamon, mayl qo'shimchalaridir. Fe'llar ham ikki guruhga bo'linadi:

1) tub fe'llar (lug'atlarda keltiriladigan shakli);

2) yasama fe'llar (tub fe'llardan yasalgan fe'llar).

tub fe'l 1) yetakchi fe'l To be, to have, to do, shall, will can

2) modal fe'l

3) fe'l bog'lama

4) yordamchi fe'l.

Yasama fe'l tub fe'ldan qo'shimcha orqali hosil qilinadi:

Qo'shimcha	
-fs	Received
-fing	Reading

Fe'ning dastlabki qo'shimcha formulasi faqatgina to'g'ri fe'l uchun taalluqlidir.

Noto'g'ri fe'llar uchun yasama shaklning yasalishi quyidagi tablitsada keltirilgan:

Tub	Yasama	Yasama
See	Saw	seen
Do	Did	does

Yuqorida fe'ning shaxsli shakllari haqida fikr yuritildi. Fe'ning shaxssiz shakllari ham mavjud bo'lib, ular *infinitiv* va *gerundiy* deb nomlanadi.

Infinitiv to + dastlabki fe'l: to walk

Gerundiy dastlabki fe'l+ ing: reading

Gap bo'laklari

Ixtiyoriy tildagi yozma matn yoki shaxsning nutqi "jumla" deb ataluvchi alohida segmentlardan iborat bo'ladi. Ular sodda va qo'shma gaplarga bo'linadi. Har qanday tilning grammatikasini o'rganish sodda gaplarni tadqiq etishdan boshlanadi. Sodda gaplar tarkibida ega, kesim, to'ldiruvchi, aniqlovchi, hol kabi gap bo'laklari mavjud bo'ladi.

Kesim – gapning ham semantik, ham grammatik markaziy bo'lagi bo'lib, gapning yadrosi hisoblanadi, kesimsiz gap shakllanmaydi. Kesim harakat,

holat, xususiyat kabilarni anglatadi. Kesim uchun **P** (ba'zan **V**) simvolidan foydalaniladi.

Ega – gapning zaruriy bo'lagi bo'lib, boshqa bolaklar bilan grammatik bog'lanadi, u predmet yoki shaxsni bildiradi. Egani ifodalash uchun **S** simvolidan foydalaniladi.

To'ldiruvchi – predmetni bildiruvchi va ma'no jihatidan kesimni to'ldiruvchi gap bo'lagi. To'ldiruvchi uchun **O** simvolidan foydalaniladi.

Aniqlovchi o'tning belgilarini tushuntiruvchi, izohlovchi gap bo'lagidir. Aniqlovchi kompyuter lingvistikasida gap bo'lagi sifatida ajratilmaydi, chunki u gapda boshqa bo'laklarga tobelanib keladi, u gap uchun nozaruriy (fakultativ, obligator emas) bo'lak hisoblanadi.

Hol kesimda ifodalanayotgan harakat yoki holatning o'rni, payti, maqsadi, sababi va bajarilish sharti kabilarni bildiruvchi gap bo'lagidir. Hol uchun **M** simvolidan foydalaniladi.

Ingliz tili grammatikasining butun modeli juda katta hajmga ega. Bunda o'quv materiallarining hajmidan kelib chiqqan holda nisbatan qisqartirilgan model keltirildi. Gaplarni tadqiq qilishning uch bosqichli ko'rinishi umum tomonidan tan olingan:

- 1) so'zlar (so'z turkumlari);
- 2) frazalar (gap bo'laklari);
- 3) gaplar (gap konstruksiyalarining sxemalari).

Ingliz tilidagi gaplarning asosiy konstruksiyalari:

1. *Ega+kes+to'ld.+hol* (kengaytirilgan darak gap): Students bought books yesterday.

2. *Ega+kesim* (kengaytirilmagan gap): It is good.

3. *Ega+gram.f.+hol +k+to'ld.+hol* (inkor gap): Students did buy books yesterday.

4. *Aniq+ega+kes+to'ld+aniq+hol* (aniqlovchili gap): The blue car stopped at the big gate.

5. *Gram.f+ega+k'+to'ld+hol* (umumiy so'roq gap): Has he been working since morning?

6. *So'roq so'z+gram.f.+ega+k+to'ld+hol* (maxsus so'roq gap)¹: Why are you sitting here?

Xuddi shu qolipda rus va o'zbek tillaridagi gap konstruksiyalarini tuzish mumkin.

Ingliz, rus tillarida sodda gap qurilishi quyidagicha belgilangan:

S + V + O: S = ega, O = to'ldiruvchi, V = kesim.

Я пишу диссертацию. Он читает книги.

I have done my work. I am collecting apples.

O'zbek tilida sodda gap qurilishi:

S + O + V: Sen she'r bitding. Men she'r yod oldim.

S = ega, O = to'ldiruvchi, V = kesim.

O'zbek tilida qo'shma gapning eng kichik modeli quyidagicha bo'ladi:

S1+ V1 S2 + V2: Bahor keldi, qalblar huzurga to'ldi.

T –H so'z birikmasi modeli (tobe – hokim).

WPm minimal sodda gap modeli. M: *O'qidim. Yozding.*

Analogiya. Kompyuter lingvistikasida eng ko'p qo'llanilayotgan metodlardan biri analogiya metodi hisoblanadi. **Analogiya** (yunoncha «moslik», «o'xshashlik» ma'nosini anglatadi) – munosabatlarning tengligi, muvofiqdigi; hodisa va jarayonlarning muayyan xossalarini qiyoslash yo'li bilan anglashda o'xshash jihatlari. Qiyoslanayotgan obyektlar o'rtasida o'xshashlik va farqlanish kuzatiladi, bu qiyoslashning asosi sanaladi. Bunda ma'lum obyektga xos bo'lgan aniq belgining qiyoslanayotgan obyektlarda kuzatilishiga tayanib ularning identifikatsiyasi asosida mushtarak jihati aniqlanadi. Masalan, sifat va ravish so'z turkumiga mansub so'zlar berildi. Ular bir qarashda bir-biriga juda yaqin, so'roqlari o'xshash, shuningdek, farqli tomonlari ham mavjud. Analogiyaga tayanib berilgan so'zlarning mushtarak jihatlari ajratish mumkin. Analogiya *induksiya* (umumiylikdan xususiylikka; qoidadan misolga borish) va *deduksiya* (xususiylikdan umumiylikka; misoldan qoidaga borish) metodlari bilan uzviy aloqadorlikda bo'ladi. Bilish jarayonida ular o'rtasiga qat'iy chegara qo'yib bo'lmaydi. Analogiya deduksiya va induksiya o'rtasida bog'lovchi «halqa» vazifasini o'taydi.

Bilishda analogiya metodining ahamiyatini mashhur matematik Dyord Poya shunday izohlaydi: «*Analogiyaga asoslanmasdan na elementar, na*

¹ Modellarning o'zbek tili grammatikasi asosidagi muqobillari A.Rahimov tomonidan ishlab chiqilgan. Qarang: *Rahimov A. Kompyuter lingvistikasi asoslari.* –T., 2011.

oliy matematikada, na fanning boshqa sohalarida hech qanday kashfiyot qilib bo‘lmaydi». Stefan Banax esa shunday yozadi: «Matematik isbotlar o‘rtasidagi analogiyani topa olgan odam, eng yaxshi matematik; isbotlar o‘rtasida analogiya o‘rnata olgan odam, bundan ham ko‘ra kuchli matematik – nazariyalar analogiyasini sezgan odam; biroq odam o‘zini eng kuchli hisoblashi uchun analogiyalar o‘rtasidagi analogiyani ko‘ra olishi kerak»¹. Ushbu mulohazalardan kelib chiqqan holda aytish mumkinki, analogiya lingvistikadagi, xususan, kompyuter lingvistikasidagi tadqiqotlar uchun ham o‘ziga xos kompas vazifasini bajaradi.

Akademik L.V.Sherba bir ma‘ruzasida yozuv taxtasiga shunday jumlani yozadi: «Глокая кудра итеко будланула бокра и кудрячить бокренка». Talabalardan uning grammatik tahlilini talab qiladi. Mazkur jumla aslida hech qanday mazmunga ega emas, lekin u rus tilidagi mazmunga ega bolgan qaysidir jumlani eslatadi. Bu jumlaning grammatik tahlilini amalga oshirish mumkin. Masalan, «kuzdra» so‘zi – birlik, bosh kelishik, jenskiy rod shaklidagi ot bilan ifodalangan ega, «budlanula» so‘zi – o‘tgan zamon shaklidagi fe‘l bilan ifodalangan kesimlardan biri, «kudryachit» so‘zi – noaniq shakldagi fe‘l bilan ifodalangan boshqa bir kesim. Masalaga bunday yondashish analogiya metodi asosida amalga oshadi. Mazkur yondashuv, ayniqsa, kompyuter lingvistikasidagi matnning avtomatik analizi yo‘nalishida muhim ahamiyat kasb etadi. Bunda so‘zlarning grammatik tavsifi va so‘zlarning oxiridagi harf tarkibi o‘rtasidagi kuchli korrelyatsion aloqaga tayaniladi.²

Oxirgi harfi bir xil bo‘lgan so‘zlarning derivatsion va relyatsion modellari (so‘z yasalishi va shakl yasalishi) hamda grammatik ma‘lumotlari (qaysi so‘z turkumiga mansubligi, qanday grammatik kategoriyalarga egaligi) bir xil bo‘lishi tabiiy. Masalan, ingliz tilida *-er, -ness, -dom* kabi, rus tilida *-nost, -ujik, -nik* kabi, o‘zbek tilida *-shunos, -chi, -lik* qo‘shimchalariga ega bo‘lgan so‘zlar ot (noun) so‘z turkumi hisoblanadi. Demak, tezaurusda (lug‘at bazasida) mavjud bo‘lmagan so‘zlarning avtomatik analizini mavjud bo‘lganlarining harf tarkibidagi moslikdan, so‘zlar analogiyasidan kelib chiqqan holda amalga oshirish mumkin. Bu jarayonda tizimning lingvistik ta‘minoti chappa (ters) lug‘atlar bilan ham to‘ldirilgan bo‘lishi lozim.

¹Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

² Белоногов Г.Г. Об использовании принципа аналогии при автоматической обработке текстовой информации // Проблемы кибернетики. 1974. №28. -С.33. Bu haqida qarang: Po‘latov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2009. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

Modellashtirish lingvistik qoidalarni tushuntirishda eng qulay vositalardan biridir.

Savol va topshiriqlar

1. Lisoniy model va matematik modelning uzviy bog‘liqligi nimalarda namoyon bo‘lishini tushuntiring.
2. Tilshunoslik bo‘limlarini modellashtirish darajasi haqida bilganlaringizni so‘zlang.
3. So‘z yasash modellariga turli tillardan misollar keltiring.
4. Gaplarni modellashtirishda e‘tiborga olinishi lozim bo‘lgan asosiy jihatlarni bayon eting.
5. Gap bo‘laklari modellari haqida ma’lumot bering.
6. Fikrni davom ettiring: Kompyuter modellari – bu ...
7. Ingliz, rus tillarida sodda gap qurilishiga doir quyidagi modellarga misol keltiring: **S + V + O**: S = ega, O = to‘ldiruvchi, V=kesim.

6-§. LINGVISTIK PROTSESSORNING GRAMMATIK XUSUSIYATLARI

Reja:

1. Kompyuter analizi va ingliz tilidagi gaplarni o‘zbek tiliga tarjima qilish algoritmi.
2. Dastur algoritmi.
3. Dasturning ishlash tamoyili.

Asosiy tushunchalar: *dastur algoritmi, muayyan ketma-ketlik, dasturning ishlash prinsipi, tarjima protsedurasi, tarjimon dasturlar, ma’lumotlar bazasi, kompyuter analizi va sintezi, avtomatik tarjima, lingvistik ta’minot, dasturiy ta’minot, dastur algoritmi, tarjimon dasturlar, formal model, morfologik, semantik-sintaktik va konseptual analiz, lingvoanalizatorlar, parser dasturlar.*

Ingliz tilidagi matnlarni o‘zbek tiliga va, aksincha, o‘zbek tilidagi matnlarni ingliz tiliga o‘girishning ommaviy kompyuter modellari qo‘llash, tillarni kompyuter yordamida o‘qitish, bilimlarni baholash, matnlarni tahrirlash eng dolzarb muammolar hisoblanadi.

Inglizcha-o‘zbekcha va o‘zbekcha-inglizcha kompyuter tarjimasi dasturlari katta ahamiyatga ega. Ma’lumki, ingliz va o‘zbek tillari leksik-

grammatik xususiyatlariga ko'ra bir-biridan tubdan farq qiladi. Shuning uchun ingliz tilidan o'zbekchaga va o'zbekchadan inglizchaga kompyuter tarjimasini yaratish o'ziga xos qiyinchliklarni keltirib chiqaradi. Bugungi kunda rus tili vositasida ingliz tilidan o'zbek tiliga avtomatik tarjima qiluvchi dasturlarning versiyalari e'lon qilingan.

Ayrim dasturlarda ingliz tilidan o'zbek tiliga tarjima qilish vositachi tilsiz amalga oshiriladi. Ushbu jarayon gaplarning sintaktik analiz qilish algoritmlari va formal modellari bazasi asosida yuzaga keladi.

Ingliz tilidan o'zbek tiliga kompyuterda tarjima qilish algoritmi

Algoritm quyidagi vazifalarni hal qilish uchun mo'ljallangan:

I.I. **Analiz** – bunda ingliz tilidagi gap soddalashtirilgan model doirasida sintaktik tahlil qilinadi.

1.1.Ushbu model faqat sodda gaplarni qamrab oladi.

1.2.Gapning har bir bo'lagi bitta so'zdan iborat bo'ladi.

1.3.Gaplarda aniqlovchilar bo'lmaydi.

1.4.Gaplarning standart tiplari ko'rib chiqiladi (darak gap: ega + kesim+ to'ldiruvchi + hol, so'roq, inkor va so'roq-inkor gaplar).

1.5.Fe'lning quyidagi tuslanish shakllarini qamrab oluvchi kesimli gaplar ko'rib chiqiladi:

a) shaxs (I, II, III shaxs);

b) son (birlik va ko'plik);

d) zamon (o'tgan, hozirgi, kelasi zamon);

e) harakat tipiga ko'ra (simple)

f) maylga ko'ra (in)

g) nisbatga ko'ra (active)

II.Tarjimada gaplar ingliz tilidan o'zbek tiliga o'giriladi. Algoritm quyidagi etaplardan tashkil topadi:

- gap kiritiladi;

- gapning har bir so'zi α_i massivining elementlariga qo'shiladi;

- α_i massivining elementlari yordamida lug'at elementlari bilan taqqoslanadi, bu lug'atda olmosh, ko'makchilar, ko'makchi hamda modal fe'llar, artikllar va noto'g'ri fe'llar ro'yxati mavjud bo'ladi;

- agar so'zlar yordamchi lug'atda topilmasa, unda taqqoslash maxsus lug'at yordamida davom ettiriladi;

- topilgan so'zlar yordamchi lug'atda beriladi, bu yerda so'zga ushbu so'zni va uning tarjimasini saqlovchi kod keltiriladi;

- bunday so'z lug'atlarda mavjud bo'lmasa, so'z shakl yasovchi affikslardan ajratib olinadi va 5-ish bajariladi;

- agar soʻzlar yordamchi va maxsus lugʻatlardan topilmasa, ushbu soʻzning yoʻqligi haqida maʼlumot kiritiladi;
- gap ikki guruhga boʻlinadi: kesimgacha boʻlgan soʻzlar ega guruhiga kiradi (ega guruhi: ega);
- kesimdan boshlanib gapning oxirgacha boʻlgan soʻzlar kesim guruhi hisoblanadi (kesim guruhi: kesim, toʻldiruvchi, hol);
- kesim guruhidan kesim ajratib olinadi;
- toʻldiruvchi ajratiladi;
- gapning qolgan qismi hol hisoblanadi;
- gapning har bir boʻlagi shakl yasovchi qoʻshimchalarsiz tarjima qilinadi;
- tarjima qilingan gap boʻlaklaridan oʻzbek tilidagi gap tuziladi, u, albatta, ingliz tilidagi gap konstruksiyasiga mutanosib boʻladi;
- oʻzbek tilidagi soʻzlarga ingliz tilidagi soʻzlar mazmuniga mutanosib ravishda affiks va qoʻshimchalar qoʻyib chiqiladi;
- tarjima chiqarib beriladi («tarjima» rejimida);
- analiz chiqarib beriladi («analiz» rejimida).

Igliz tilidagi gaplarni oʻzbek tiliga tarjima qilish va kompyuter analizi dasturi

(1-versiya)

Boshlanish.

Kirish.

Dastur ishlashi prinsipining nazorat misoli yordamidagi tahlili:

We received a letter from school.

I. Morfologik tahlil.

- 1) We – kishilik olmoshi, koʻplik birinchi shaxs, tarjimasi: “biz”;
- 2) Received – receive+ed, feʼl, tarjimasi: “qabul qilmoq”;
- 3) a – noaniq artikl;
- 4) letter – birlikdagi ot, tarjimasi: “xat”;
- 5) from – koʻmakchi, tarjimasi: “-dan”;
- 6) school – birlikdagi ot, tarjimasi: “maktab”.

II. Gap boʻlaklarini ajratish.

We received a letter from school – ega + kesim + toʻldiruvchi + hol

III. Gap tahlili.

1. Ijro mayli.
2. Aniq nisbat.
3. Simple (harakatlar).
4. Oʻtgan zamon.

5. Darak gap.

VI. Tarjima: *Biz maktabdan xatni qabul qilgandik.*

Taklif qilinayotgan dastur inglizcha –o‘zbekcha kompyuter lug‘atini yaratishning asosi (Computer Dased Dictionay) va undan effektiv va har tomonlama foydalanish uchun kalit hisoblanadi. Avvalo, shuni ta’kidlash zarurki, bu dastur keng doiradagi mutaxassislar bilan bir qatorda tillarni o‘rganish va tarjimada har kuni muammolarga duch kelayotgan har qanday insonlarga mo‘ljallangandir. Mazkur lug‘at foydalanuvchiga vaqtni tejash imkoniyatini beradi.

Dasturning ishlash algoritmi o‘zida quyidagi bosqichlarni qamrab olgan:

1.1. Boshlanish.

2.2. Rejimlarni tanlash.

3.3. So‘zni kiritish va uning kodini xotiradan qidirish.

4.4. Tanlangan rejimlarning maxsus dasturlari bilan topilgan kodni qayta ishlash va talab qilinayotgan ma’lumotlarni chiqarish.

5.5. Joriy rejimda ishni davom ettirish haqida so‘rash.

6.6. Ishni yakunlash haqida so‘rash.

7.7. Tugatish¹.

Yaratilgan dastur versiyasi 10 000 ta umumiste’ moldagi inglizcha so‘z bazasiga asoslanadi va Turbo Pascal 7.0 dasturlash tilida ishlab chiqiladi. U Windows, Norton Commander, Far tizimida va MS-DOS operatsiyasi sistemasida eksluatatsiya qilinadi.

Kelajakda dasturning “Delphi”ga asoslangan versiyasini ishlab chiqish unga ovoz effektlarini qo‘shish rejalashtirilgan.

Savol va topshiriqlar

1. Til o‘rgatuvchi o‘quv dasturlariga namunalar keltiring.

2. Til o‘rgatuvchi o‘quv dasturlarining samaradorligi haqida mulohazalaringizni bayon eting.

3. Boshqa turdagi o‘quv dasturlari haqida ma’lumot bering.

4. Internet tizimi yordamida (on-line) til o‘qitish kurslaridan foydalaning.

5. GALL tizimining rivojlanishi va istiqbollari haqida fikr bildiring.

6. Matn tarjimasida grammatik qurilishning o‘rnini yoriting.

7. Tarjimon dasturlarining ishlash prinsiplarini o‘zlashtirib oling.

¹ Qarang: *Po‘latov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2009. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.*

7-§.MORFOLOGIK TAHLILNING AVTOMATIK TIZIMI

Reja:

- 1.Morfologik analiz jarayoni.
2. So‘z yasalish tizimini modellashtirish.
3. Morfologik tahlil va sintaktik tahlilning uzviyligi.

Asosiy tushunchalar: *morfema, analiz, so‘z tarkibi, sintez, affiks, prefiks, suffiks, kompyuter analizi va sintezi, avtomatik tarjima, tarjima dasturi, lingvistik ta‘minot, morfologik, sintaktik analiz, kodlash, fragment, lingvistik ta‘min, semantik freym, konseptual analiz, semantik va konseptual analiz, protsedura,*

Matnning formal ajraladigan eng kichik birligi so‘z hisoblanadi, lekin u ma‘no anglatuvchi eng kichik birlik morfemalardan tashkil topadi. So‘z tarkibi o‘zak morfemalar, prefikslar (old qo‘shimchalar) va suffikslar (o‘zakdan keyin keladigan qo‘shimchalar)dan iborat bo‘ladi. Masalan, *bilimdonlikdan* so‘zida 5ta morfema mavjud. So‘zlarning morfologik analizi uchun tildagi prefiks va suffikslar bazasi hamda so‘zlarning grammatik shakllarda o‘zgarishi (qanday qo‘shimchalarni qabul qilishi)ni ko‘rsatuvchi lingvistik ta‘min yaratilishi lozim. Bunday baza avtomatik tarjima tizimi mavjud bo‘lgan rus, ingliz, fransuz tillarida yaratilgan. Shulardan birini (rus tili grammatikasi asosidagi) tarjima dasturini keltirib o‘tamiz.¹

Слово-представитель	Окончания: 1) им.п., ед.ч. 2) тв.п., ед.ч. 3) им.п., мн.ч. 4) род.п., мн.ч.	Примечание
Существительные мужского рода неодушевленные		
Телефон	+*), ом, ы, ов	
Тираж	+, ом, и, ей	
огонь	ь, ем, и, ей	См. п. 012
перебой	й, ем, и, ев	См. п. 005
санаторий	й, ем, и, ев	ОСНОВЫ НА БУКВУ <И>
бланк	+, ом, и, ов	
сапог	+, ом, и, +	

¹ Белоногов Г.Г. Компьютерная лингвистика и перспективные информационные технологии. - М.: Русский мир, 2004. - С.68 - 90.

Лес	+, ом, а, ов	
колодец	+, ем, ы, ев	
путь (класс состоит из одного слова)		
край	й, ем, я, ев	
брус	+, ом, я, ев	
глаз	+, ом, а, +	
зародыш	+, ем, и, ей	
волос	+, ом, ы, +	
лагерь	ь, ем, я, ей	
Существительные мужского рода одушевленные		
кузнец	+, ом, ы, ов	
солдат	+, ом, ы, +	
сосед	+, ом, и, ей	См. п. 024
врач	+, ом, и, ей	Основы на ж, ч, ш, щ
пролетарий	й, ем, и, ев	
воробей	ей, ем, и, ев	
конь	ь, ем, и, ей	
учитель	ь, ем, я, ей	
сапожник	+, ом, и, ов	
испанец	+, ем, ы, ев	
юноша	а, ей, и, ей	
мужчина	а, ой, ы, +	
судья	я, ей, и, ей	
товарищ	+, ем, и, ей	
гражданин	+, ом, е, +	
профессор	+, ом, а, ов	
Существительные женского рода одушевленные		
женщина	а, ой, ы, +	
переводчица	а, ей, ы, +	
нутрия	я, ей, и, й	Основы на букву «и»
швея	я, ей, и, й	см. п. 046
цапля	я, ей, и, ь	
санитарка	а, ой, и, +	
мышь	ь, ю, и, ей	
Иванова	а, ой, ы, ых	
Существительные женского рода неодушевленные		
речь	ь, ю, и, ей	Основы на ж, ч, ш, щ
грань	ь, ю, и, ей	см. п. 054
колба	а, ой, ы, +	

задача	а, ей, и, +	
заготовка	а, ой, и, +	
линия	я, ей, и, й	Основы на букву «и»
галерея	я, ей, и, й	см. п. 061
земля	я, ей, и, ь	
эскадрилья	я, ей, и, ий	
статья	я, ей, и, ей	
башня	я, ей, и, +	
улица	а, ей, ы, +	
Существительные среднего рода		
место	о, ом, а, +	
облако	о, ом, а, ов	
поле	е, ем, я, ей	
сомнение	е, ем, я, и	
жилище	е, ем, а, +	
перо	о, ом, я, ев	
время	я, ем, а, +	
побережье	е, ем, я, ий	
колени	о, ом, и, ей	
очко	о, ом, и, ов	
ружье	е, ем, я, ей	

*) Символ + обозначает нулевое окончание.

Прилагательные

Слово-представитель	Окончания: 1) им. п., муж. р., ед.ч., 3) род. п., муж. р., ед.ч., 4) им. п.,	Примечание
главный	ый, ая, ого, ые	
передний	ий, яя, его, ие	
хороший	ий, ая, его, ие	
Легкий	ий, ая, ого, ие	
Годовой	ой, ая, ого, ые	
Плохой	ой, ая, ого, ие	
Третий	ий, я, его, и	
этот, сам	+, а, ого, и	
мой, твой, свой	и, я, его, и	
наш, ваш	+, а, его, и	
Весь	ь, я, его, е	

Глаголы в личной форме

Слово-представитель	Окончания 1,2 и 3-го лица ед.ч. и 3-го лица мн. ч.	Примечание
Делать	ю, ешь, ет, ют	
Строить	ю, ишь, ит, ят	
Писать	у, ешь, ет, ут	
Стучать	у, ишь, ит, ат	
Бежать	у, ишь, ит, ут	
Хотеть	у, ешь, ет, ят	
Зависеть	у, ишь, ит, ят ¹	

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 hosil qilish, analiz natijalarini umumlashtirishdan iborat. Morfologik analiz va sintez jarayoni kompyuterda kodlash-dekodlash tamoyili asosida ishlaydi. Bunda so'z turkumlari maxsus kodlar yordamida belgilanadi. RETRANS (inglizcha-ruscha, ruscha-inglizcha frazeologik avtomatik tarjima tizimi) avtomatik tarjima tizimida morfologik analiz jarayonida foydalaniladigan inglizcha so'zlarning grammatik tasnifida so'z turkumlarini quyidagicha kodlash mumkin:

ot – «1», sifat – «2», fe'l – «3», ravish – «4», predlog – «5», simvoli bog'lovchi – «6», artikl – «7», yuklama – «8» simvollar bilan ifodalanadi. Morfologik analiz jarayonida ushbu simvollar orqali so'zlar kodlanadi.

O'zbek tilidagi so'z yasalihi o'zining tarixiy qonuniyatiga ega. So'zshakllar tarkibi **asos+so'z yasovchi+lug'aviy shakl yasovchi+sintaktik shakl yasovchi** qolipida bo'ladi. So'z yasalihsida esa **asos+so'z yasovchi** qolipdan foydalaniladi. O'zbek tilidagi so'z yasalihsini quyidagicha umumlashtirish mumkin:

Shaxs otlarining yasalihs		
terim	-chi	terimchi
paxta	-kor	paxtakor
til	-shunos	tilshunos

¹Po'latov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2009. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

mulk	-dor	mulkdor
osh	-xo‘r	oshxo‘r
kabob	-paz	kabobpaz
dor	-boz	dorboz
bog‘	-bon	bog‘bon

Narsa-buyum otlarining yasalishi

kura	-κ	kurak
tara	-q	taroq
o‘r	-oq	o‘roq
to‘pla	-m	to‘plam
o‘yin	-choq	o‘yinchoq
sachrat	-qi	sachratqi
supur	-gi	supurgi
o‘rin	-diq	o‘rindiq

Sifat yasalishi

gul	-dor, -li	guldor, gulli
gap	-don	gapdon
jahl	-dor	jahldor
o‘yna	-qi	o‘ynoqi
kuz	-gi	kuzgi
qo‘rq	-qoq	qo‘rqoq
maqtan	-choq	maqtanchoq
tarqa	-q	tarqoq

Sifat yasovchi qo‘shimchalarning ayrim ot yasovchi qo‘shimchalar bilan omonimligi kompyuter lingvistikasida o‘ziga xos murakkabliklar tug‘diradi. Bu muammolar kontekst mazmuni orqali bartaraf qilinadi.

Inglizcha matn fragmentining morfologik analiz namunasi

Language 0/11 engineering 0/11 is 0/31 the 0/71 core 0/11 of 0/51 information 0/11 technology 1/11 and 0/61 information 0/11 technology 1/1! will 0/31 be 0/31 the 0/71 key 0/1121 industry 1/11 of 0/51 the 0/71 next 0/21 decades 1/11 .

Bu yerda ayrim so‘zlar grammatik omonimiyaning hisobiga dastlab («tahlil natijalari»da) ikki xil raqam bilan berilgan, so‘nggi bosqichda kontekstdan kelib chiqib, aniq tasnif keltiriladi. Yuqoridagi matn fragmentida shunday so‘zlar sirasida «state» so‘zini misol qilish mumkin. «State» so‘zi dastlabki tahlilda 0/1132 kodi orqali beriladi, bunda «0» raqami so‘zda qo‘shimchaning yo‘qligini (nol morfemani), «11» raqami ot so‘z turkumini, «32» raqami esa hozirgi zamon shaklidagi fe‘l yoki infinitivni bildiradi. Bunga o‘xshash yana bir so‘z «wished» so‘zining dastlabki tahlilida bir qiymatli bo‘lmagan ma’lumotlarga ega bo‘lamiz. «Wished» so‘zi 0/2433 kodi bilan beriladi, bunda «0» raqami

yuqoridagidek nol morfemani, «24» fe'l asosidan hosil bo'lgan sifat turkumini, «33» o'tgan zamon shaklidagi fe'lni anglatadi.

So'zlarning qaysi turkumga mansubligini aniqlashda chappa (ters) lug'atlardan foydalaniladi. Bunga turli xil turkumga mansub so'zlarning morfologik analizini ko'rsatuvchi inglizcha so'z shakllarining grammatik chappa lug'ati yaqqol misol bo'la oladi.

Inglizcha so'z shakllarining grammatik chappa lug'ati fragmenti (ushbu lug'at 54 000 leksik birlikni o'z chiga oladi)

So'z Grammatikaga oid ma'lumot		
Ot (Noun)	Sifat (Adjective)	Fe'l (Verb)
America 0/11	morphologic 0/21	tightened 0/2433
Africa 0/11	epidemiologic 0/21	shortened 0/2433
Corsica 0/11	physiologic 0/21	fastened 0/2433
Arctica 0/11	immunologic 0/21	moistened 0/2433
Antarctica 0/11	hydrologic 0/21	flattened 0/2433
Casablanca 0/11	geohydrologic 0/21	aligned 0/2433
Canada 0/11	paleohydrologic 0/21	standardize 0/32
transcanada 0/11	micrometeorologic 0/21	focalize 0/32
Granada 0/11	metrologic 0/21	realize 0/32
Nevada 0/11	neurologic 0/21	specialize 0/32
Andromeda 0/11	paleontologic 0/21	commercialize 0/32
Florida 0/11	allergic 0/21	rationalize 0/32
Propaganda 0/11	readable 0/21	neutralize 0/32
addenda 1/11	achine-readable 0/21	visualize 0/32
agenda 0/11	degradable 0/21	conceptualize 0/32
thermobalance 0/11	biodegradable 0/21	stabilize 0/32
microbalance 0/11	upgradable 0/21	utilize 0/32
heatbalance 0/11	decidable 0/21	straightening 0/142444
resemblance 0/11	undecidable 0/21	heightening 0/142444

Morfologik analizning mantiqiy davomi sintaktik analiz hisoblanadi. Sintaktik analiz morfologik analizga asoslanadi. Bunda sintaktik analiz turli xil modellar asosida amalga oshirilishi mumkin. Formal grammatika asoschisi bo‘lgan Noam Xomskiy belgilagan algoritm o‘zining optimalligi bilan boshqa sintaktik modellardan ajralib turadi¹. Matnning avtomatik sintaktik analizi masalasi XX asrning 50-yillarida bir tildan boshqa tilga kompyuter yordamida tarjimini amalga oshiruvchi tizimlarga oid tadqiqotlar bilan uzviylikda o‘rganilgan. Zero, matn tarjimasida gaplarning sintaktik strukturasi, sintaktik aloqasi hal qiluvchi ahamiyat kasb etadi. Dastlabki sintaktik analiz protseduralari ancha cheklangan bo‘lib, unda faqat gapning strukturasi analizi e‘tiborga olingan, bunda har qanday gap so‘zlarning muayyan ketma-ketligi asosida tashkil topgan qurilma sifatida baholangan. Gap strukturasi formallashtirilgan modeli sifatida, odatda, tobelik daraxti (sentence tree)dan foydalanilgan.² Shuni alohida ta’kidlash zarurki, sintaktik analizning samaradorligi hozirgi kunda juda yuqori saviyada, ular faqat avtomatik tarjima sistemasi ishlab chiqilgan tillarda maksimal darajada ishlatilmoqda. Gapning semantik analizi kompyuter lingvistikasidagi inson tafakkurini modellashtirish, sun‘iy intellekt tizimi, semantik freym, Word-net masalalari bilan bog‘liq. Agar sanab o‘tilgan muammolar o‘z yechimini topsa, semantik va konseptual analizning avtomatik tizimi ham takomillashadi.

Savol va topshiriqlar

- 1.Morfologik analiz va sintez qanday protseduralarni o‘z ichiga oladi?
- 2.Sintaktik analiz va sintezning mohiyati nimadan iborat?
- 3.Ingliz tilida gaplarning dastur orqali tarjimasini qanday amalga oshiriladi?
- 4.Algoritmni hayotiy misollar bilan tushuntiring.
- 5.Tarjima algoritmi nima?
- 6.Morfem tahlilga oid analizga misol keltiring.
- 7.Quyidagi jadvalni to‘ldiring.

O‘zbekcha shaxs	Ruscha muqobili	Inglizcha

¹ Po‘latov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2009. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

² Jurafsky D., Martin J.H. Speech and Language Processing. - New Jersey, 2000.-P.353.

8-§. SINTAKTIK TAHLILNING AVTOMATLASHGAN TIZIMI

Reja:

1. Grammatika haqidagi lingvistik qarashlar.
2. Grammatikaning turlari (tavsifiy, formal, transformatsion grammatika).
3. Formal grammatika nazariyasi.

Asosiy tushunchalar: *morfologik tizim, grammatik shakllar, grammatik kategoriyalar, sintaktik tizim, sintaktik kategoriyalar, konstruksiyalar, soʻzlarning oʻzaro birikish qonuniyatlari, «formal grammatika nazariyasi», kategorial grammatika, tobelik grammatikasi, leksik-funksional grammatika, konstruktiv grammatika, transformatsion grammatika, kontekst.*

Grammatika tabiiy tillar strukturasi haqidagi fan boʻlib, u morfologik kategoriya va shakllar, sintaktik kategoriya, konstruksiyalar hamda soʻz yasash usullari tizimini tashkil etadi. Til qurilishi birliklarining paradigmatic va sintagmatic munosabatga kirishuvi (vertikal va gorizontol yoʻnalishda) muayyan qoidalar asosida yuz beradi. Ushbu qoidalar yigʻindisi tilning grammatik qurilishi va grammatik tizimi deb yuritiladi. Grammatika ikki yirik tizimni oʻz ichiga oladi:

1. Morfologik tizim – soʻz formalari, grammatik shakllar, grammatik kategoriyalar.

2. Sintaktik tizim – sintaktik kategoriyalar va konstruksiyalar, soʻzlarning oʻzaro birikish qonuniyatlari.

Grammatika tushunchasi, dastlab, xat-savod, imloni oʻrgatuvchi fan sifatida ishlatilgan. Keyinroq grammatika tilning fonologik, morfologik, sintaktik va semantik belgilarini tavsiflovchi til qoidalari majmuyini ifodalagan. Grammatika haqidagi ilk qarashlar falsafa, mantiq, germenetika (diniy matnlarni sharhlash, talqin etish) doirasidagi izlanishlarda aks etgan. Bu manbalarda tilning paydo boʻlishi, til va tafakkur munosabati, lingvistik belgi motivatsiyasi, diniy matnlar mazmunini toʻgʻri tushunish bilan bogʻliq masalalar oʻrganilgan, til strukturasi esa yetarli darajada tadqiq etilmagan. Miloddan avvalgi II – I asrlarda shakllangan Aleksandriya grammatika maktabining eng yirik vakili Dionisiy Traks (miloddan avvalgi 170 – 90-yillarda yashagan) oʻzigacha boʻlgan tadqiqotchilarning ishlari va tajribalarini oʻrganib, rimliklar uchun «Grammatika sanʼati» («Grammatike techne») nomli

sistemalashtirilgan dastlabki yunon grammatikasini yaratdi. Shundan buyon grammatika falsafadan mustaqil soha sifatida ajralib chiqdi.

Antik davrda grammatika preskriptiv (to'g'ri gapirish, to'g'ri jumla tuzishni ko'rsatuvchi yo'riqnomaga) 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'rganish emas, balki sinxron (muayyan vaqtdagi holatini) tadqiq qilish lozim, deb hisobladi. XX asrning 50-yillarida Noam Xomskiy grammatikani tavsiflashning yangi metodini tavsiya etdi. Bu tilshunoslikda «formal grammatika nazariyasi» deb ataladi. Olimning tadqiqot natijalari kompyuter lingvistikasining shakllanishiga asos bo'ldi. N.Xomskiyning formal grammatika haqidagi qarashlari «*Syntactic Structures*» hamda «*Aspects of the Theory of Syntax*» nomli kitoblarida yoritilgan.

Grammatikaga bo'lgan yondashuvlar turlicha bo'lganligi bois tilshunoslik tarixida grammatikaning bir qator ko'rinishlari yuzaga kelgan. 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.

Grammatik tavsifning eng oddiy formal modeli gap bo'laklariga ajratish modeli sanaladi. Tabiiy tillar grammatikasida sodda gap ikki qismdan – ega va kesimdan tashkil topgan sintaktik qurilma sifatida tavsiflanadi. Ega nutqning predmetini, kesim esa uning xususiyatini ko'rsatadi. Ega va kesim alohida so'zlar bilan ham, so'zlar guruhi bilan ham ifodalanishi mumkin. Agar ega va kesim so'zlar guruhi bilan ifodalansa, sodda gap ikki struktur elementga ajratiladi: ega guruhi (noun phrase) va kesim guruhi (verb phrase).

Ega guruhida asosiy mazmun ifodalovchi unsur ot so'z turkumi bilan ifodalangan ega yoki otlashgan so'zlar (substantivatsiya), ega guruhidagi boshqa so'zlar esa eganing aniqlovchilari vazifasida keladi. Eganing aniqlovchilari ot, sifat, son, olmosh, kvantifikatorlar (miqdor bildiruvchi so'zlar), determinatorlar (artikllar) orqali ifodalangan gap bo'lagidir.

Kesim guruhida asosiy mazmun ifodalovchi unsur fe'l-kesim hamda ot-kesim hisoblanadi, shuningdek, turli so'z turkumlari bilan ifodalangan to'ldiruvchi va hol kesimning aniqlashtiruvchilari bo'lib keladi. Mazkur modelda ega va kesim gapning bosh bo'laklari, gapning semantik-sintaktik yadrosi, aniqlovchi, to'ldiruvchi, hol esa ikkinchi darajali bo'laklar deb nomlanadi. Ushbu talqin mumtoz sintaksis tamoyillariga muvofiq keladi.

Gap bo'laklari iyerarxiyasida egani markazga qo'yish tendensiyasi mantiq fanidan kelib chiqqan. Chunki mantiqda ism (subyekt) substansiyani (mohiyatni), fe'l (predikat) aksidensiyani (hodisani) ifodalaydi, degan qarash mavjud. Bunda ega gapning tub mohiyatini aks ettiruvchi birlamchi bo'lak, kesim esa egani to'ldiruvchi, eganing turli belgilarini ko'rsatuvchi ikkilamchi bolak sifatida talqin etiladi.

XX asrning ikkinchi yarmida struktur sintaksis, semantik sintaksis yo'nalishlari paydo bo'lishi bilan gapning sintaktik tuzilishiga, gap bo'laklariga nisbatan yondashuv ham o'zgardi. Formal grammatika talqiniga ko'ra faqat kesimgina gapning grammatik yadrosi sifatida e'tirof etildi. Gapning markaziy bo'lagini belgilashga doir qarashlar dunyo tilshunosligida turlicha bo'lib kelgan. Fransuz tilshunosi Lyusen Tenyerning **verbosentrizm** (lotincha «verbum» – «fe'l», «kesim» ma'nolarini anglatadi, «centrum» - «markaz») konsepsiyasi fe'lning gap markazi bo'lgan kesim vazifasida kelishini ifoda etadi.¹

L.Tenyer g'oyalari ta'sirida shakllangan tobelik grammatikasi modellari kompyuter lingvistikasining avtomatik analiz dasturlarida ishlatiladi. Bu model tobelik daraxti (shajara) deb ataladi. Bunda gap 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.

Tilshunoslikda an'anaviy gap bo'laklaridan tashqari, gapni kommunikativ nuqtayi nazardan qismlarga ajratish ham mavjud, u gapning aktual bo'linishi deb ataladi. Bunda gap tema va rema qismlariga ajratiladi. Tema nutqiy muloqotning ma'lum qismidir, rema esa muayyan narsa, voqea-hodisa, jarayon kabilar haqidagi yangi informatsiyani anglatadi. Ko'pincha tema ega guruhi orqali, rema esa kesim guruhi orqali ifodalanadi. Ba'zan buning aksi bolishi ham mumkin. Masalan, «*Botirjon*

¹ *Теньер Л. Основы структурного синтаксиса. - М.: Прогресс, 1988. Bu haqida qarang: Po'latov A., Muhamedova S. Kompyuter lingvistikasi. – T., 2009. Rahimov A. Kompyuter lingvistikasi asoslari.–T., 2011.*

universitetda ishlayapti» gapida ega vazifasida kelgan *Botirjon* tema hisoblanadi, chunki nutq vaziyati uchun *Botirjon* aniq va avvaldan tanish bo'lgan subyektdir. Gapda kesim guruhini tashkil etgan *universitetda ishlayapti* remadir. Demak, bunda *Botirjonning* universitetda ishlayotganligi fakti nutq vaziyati uchun yangi xabardir, shuning uchun u rema hisoblanadi.

Tobelik daraxti (shajara) modeliga eng yaqin model bevosita ishtirokchilar (immediate constituents) sintaktik modeli hisoblanadi. Bevosita ishtirokchilar metodiga ko'ra, tayanch nuqta konstruksiya hisoblanadi. Konstruksiya tarkibidan ishtirokchilar va bevosita ishtirokchilar ajratiladi. Konstruksiya atamasi ostida ma'noli qismlarning ketma-ket munosabatidan tashkil topgan butunlik tushuniladi. Ishtirokchilar konstruksiya tarkibiga kirgan so'z yoki birikmalardir. 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 bog'langan. Demak, bevosita ishtirokchilar sintaktik modelida dastlab so'z birikmalari ajratiladi, so'ngra ular gap ichida birlashtiriladi.

Professor A.Nurmonov fikricha, mazkur modelda nutqning har bir mustaqil parchasi ikki qismdan tashkil topadi. O'z navbatida, har qaysi qism yana o'z ichida ana shunday qismlardan iborat bo'ladi. Ko'rinadiki, har bir konstruksiya binar tamoyilga ko'ra qismlarga bo'linadi va bu bo'linish konstruksiya doirasida graduallik (darajalanish) xususiyatiga ega bo'ladi.¹

Bevosita ishtirokchilarga ajratish modeli asosida birmuncha mukammalroq va formallashtirishning negizi hisoblangan generativ va transformatsion grammatika (*generative and transformational grammar*) maydonga keldi. Transformatsion grammatika 1950-yillarda dastlab Z.Harrisning til strukturasiidagi transformatsiyaga oid qarashlari ta'sirida yuzaga kela boshladi, bu yo'nalish Z.Harrisning shogirdi N.Xomskiy tomonidan rivojlantirildi. Unga ko'ra, dastlabki yadro strukturalar, birikma strukturalar qoidalar tizimi orqali yuzaga keladi. Ya'ni har qanday til grammatikasi yadro strukturalar (kernel structures) yig'indisidir. Yadro

¹ Nurmonov A. Struktur tilshunoslik: ildizlari va yo'nalishlari. - T., 2008. -B. 124.

gaplar deganda sodda, yig‘iq, darak gaplar tushuniladi.¹ Ingliz tilida yadro gaplar sifatida quyidagi konstruksiyalar ajratib ko‘rsatiladi.²

1	NV	Noun Verb	Ot- fe’l	John came
2	VpN	Noun Verb Preposition Noun	Ot – fe’l predlogli ot	John looked at Mary
3	NVN	Noun Verb Noun	Ot-fe’lot	John saw Mary
4.	NisN	Noun is Noun	Ot –is- ot	John is a teacher
5.	N.ispN	Noun is Preposition Noun	Ot -is -predlogli ot	John is in bed
6.	NisD	Noun is Adverb	Ot- is -ravish	John is out
7.	NisA	Noun is Adjective	Ot- is- sifat	John is angry ³

Mazkur konstruksiyalarda N (ot) turli xil determinatorlar (masalan, artikl, olmosh) hisobiga, V (fe’l) ravishlar, adverbial (ravishli) birikmalar hisobiga kengayishi mumkin. Ushbu yadro konstruksiyalarning bir-biri bilan yoki turli transformatsiyalar asosida birikishi natijasida ingliz tilida boshqa turdagi gaplar hosil qilinishi mumkin. Masalan, *The old man saw a black dog there* jumlasida uchta yadro strukturaning yig‘indisi sanaladi: 1) the man saw a dog there; 2) the man was old; 3) the dog was black.

Generativ grammatikada yadro strukturalar (operandlar) turli shakliy o‘zgarishlar asosida asl ma’noni saqlagan holda ikkilamchi sintaktik qurilmalarning hosil qilinishi (transformatsiyalar = transformalar), bu jarayondagi vositalar (operatorlar) va transformatsiya hosil qilish modellari tadqiq etiladi.⁴ Masalan, *Bobur o‘qidi – Boburning o‘qishi – Bobur o‘qigach – Bobur o‘qiganda – Bobur o‘qishi bilan...; Kamola oyna artdi – Oyna Kamola tomonidan artildi*. Ushbu misollarda transformatsiya uchun asos bo‘lgan gap (*Bobur o‘qidi* yoki *Kamola oyna artdi* gaplari) operand, transformatsiya natijasi bo‘lgan hosila gap (*Boburning o‘qishi – Bobur o‘qigach* yoki *Oyna Kamola tomonidan artildi* kabi sintaktik strukturalar) transformalar yoki transformand, operanddan transformandni hosil qiluvchi vosita (ravishdosh, sifatdosh, harakat nomi yoki majhul nisbat shakllari) transformatsiya operatori deb ataladi.

Transformatsion grammatika matematik lingvistik va kompyuter lingvistikasi fanining rivojlanishida muhim ahamiyat kasb etadi.

¹ Nurmonov A., Iskandarova Sh. Tilshunoslik nazariyasi. - T.: Fan, 2008. -B. 140 -141.

² Херпус З.С. Совместная встречаемость и трансформация в языковой структуре // Новое в лингвистике. Вып. II. - М., 1962. - С.627 - 628.

³ Bu haqida qarang: Po‘latov A., Muhamedova S. Kompyuter lingvistikasi. – T., 2009. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

⁴ Qarang: Хомский Н. Синтаксические структуры // Новое в лингвистике. Вып. II.-М., 1962.- С.412-527.

Birinchidan, transformatsion grammatika tabiiy tillarning grammatikasini formallashtirish nazariyasi negizlarini yaratishga asos bo'ldi. Ikkinchidan, transformatsion grammatika avtomatik tarjima tizimining nazariy asosi sanaladi. Transformatsiya va derivatsiya jarayonlarining tabiiy tillarga faol tatbiq etilishi avtomatik tarjima samaradorligini oshiradi. Uchinchidan, transformatsion grammatika til o'qitishni modellashtirishda ham muhim ahamiyatga ega.¹

To'rtinchidan, transformatsion grammatika tabiiy tillardagi paradigmatic va sintagmatic munosabatlarning me'yoriyligi, grammatik to'g'ri va noto'g'ri jumalarni ajratishning aniq mezonlarini belgilaydi. N.Xomskiy tavsiya etgan grammatikada gaplarni hosil qilish uchun o'rin almashtirish qoidalaridan foydalaniladi. Chap qismda gapning struktur komponenti simvollari joylashadi, o'ng qism esa ularning simvollari o'rniga almashuvchilar yoki simvollar zanjiri (birikuvi)dan iborat bo'ladi. Quyidagi jadvalda ingliz tili uchun ayrim gapning struktur komponenti simvollari ko'rsatilgan:

No	Simvollar	Simvollar birikuvi
1	S – boshlang'ich simvol, inglizcha «Sentence» («gap») so'zining bosh harfidan olingan.	S = NP + VP
2	VP – kesim guruhining simvoli, inglizcha «Verb Phrase» birikmasidan olingan.	VP = Verb + NP
3	NP – ega guruhining simvoli, inglizcha «Noun Phrase» birikmasidan olingan.	NP = Det + N
4	Verb – kesim gap bo'lagining simvoli, inglizcha «Verb» so'zidan olingan.	V = Aux + V
5	Det – artikl simvoli, inglizcha «Determiner» («aniqlagich») so'zidan olingan.	Det = the, a,...
6	N – ot so'z turkumi simvoli, inglizcha «Noun» so'zidan olingan.	N = book, computer...
7	Aux – yordamchi (ko'makchi) fe'llar simvoli, inglizcha «Auxiliary» so'zidan olingan.	Aux = will, can...
8	V – fe'l so'z turkumining simvoli, inglizcha «Verb» so'zidan olingan.	V = see, know...

¹ Шаумян С.К. Теоретические основы трансформационной грамматики // Новое в лингвистике. Вып. I. - С.408.

Jadvaldagi 5 – 8- qatorlarda ko‘rsatilgan simvollar (Det, N, Aux, V) *terminal simvollar* deyiladi, chunki ular lug‘atdagi aniq so‘zlardir. 1 – 4- qatordagi simvollar esa *noterminal simvollar* deb ataladi.¹

Formallashtirish – bu bilimlar majmuyini, mazmunini muayyan shakllar yordamida sun‘iy til belgilari orqali ifodalashdir. Grammatikani formallashtirish tendensiyasi XX asrning o‘rtalarida kuchaydi. Bu bevosita til strukturasi yanada aniq parametrlarda ifodalashga, uni kompyuterga moslashtirishga urinish bilan bog‘liqdir. Formal grammatika nazariyasi matematik mantiq, matematik lingvistikaning rivojlanishi natijasida yuzaga kelgan. Grammatikani formallashtirish konsepsiyasining asoschisi amerikalik tilshunos Noam Xomskiydir. Formal grammatika abstrakt simvollar orqali bayon etiladigan grammatika hisoblanib, unda til qoidalari, so‘z yasalihi, 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.

N.Xomskiy qoidalarining ifodalanish turiga qarab formal grammatikani 4 tipga ajratgan va bu tasnif Xomskiy iyerarxiyasi deb nomlanadi:

1.1 grammatika (unrestricted phrase-structure grammar) – juda murakkab generativ grammatika, bunda $G=(N, \sigma, P, S)$ belgilari yordamida ish ko‘riladi. N, – alifbo (N – noterminal simvollar, u grammatikadagi simvollarini o‘z ichiga oladi; σ – metatilning terminal simvollarini, 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 ($V=TKN$). Bu grammatika o‘ta abstraktlashgan matematik modellarga, informatika nazariyasiga tatbiq etilgan.

1.2 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.

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

¹ Bu haqida qarang: Po‘latov A., Muhamedova S. Kompyuter lingvistikasi. - T., 2007. Rahimov A. Kompyuter lingvistikasi asoslari. -T.,2011.

1.4. grammatika (regular grammar) – “regulyar grammatika” deb ataladi, u sodda va cheklangan grammatika bo‘lib, tilning sodda unsurlari uchun ishlatiladi (miqdor, konstantalar, o‘zgaruvchilar uchun). Formal grammatikada ikki tushuncha asosiy: a) generation – tug‘dirish, hosil qilish; b) recognition – tanish, bilish. Shunga ko‘ra, grammatika ikkiga bo‘linadi: 1) yuzaga keltiruvchi grammatika (generative grammar) – tog‘ri gaplarni qabul qilish protsedurasi; 2) tushunuvchi, tanuvchi grammatika (recognizing grammar) – to‘g‘ri gaplarni tushunish protsedurasi.

Demak, grammatika tilning asosida yotuvchi shunday ichki strukturaki, uni har bir til tashuvchisi intuitiv ravishda his qiladi va undan g‘ayrihuuriy tarzda foydalanadi. Har bir inson o‘z ona tili grammatikasini bolaligidanoq o‘zlashtiradi. Bu jihatlar esa grammatikani formallashtirish masalasi juda murakkab ekanligini tasdiqlaydi.

Yuqoridagilardan kelib chiqib aytish mumkinki, shu kungacha olimlar tomonidan ishlab chiqilgan barcha grammatik modellar shartli va nisbiydir. Keltirilgan barcha modellar – gap bo‘laklari modeli ham, bevosita ishtirokchilar modeli ham, tobelik daraxti modeli ham, transformatsion va generativ grammatika modeli ham mutlaq va benuqson emas. Ular va nutq strukturasi qaysidir aspektlarini aks ettiradi, xolos. Murakkab sistema hisoblangan tilning barcha qirralarini formallashtirishning imkoni yo‘q, bunda muayyan cheklovlar mavjud (ayniqsa, tilning semantik aspektida).

Demak, har qanday formallashtirish model til grammatikasining butun qirralarini qamrab ololmaydi.

Savol va topshiriqlar

1. “Grammatika” termini mohiyatini yoriting.
2. Tilshunoslik tarixidagi grammatikaga nisbatan yondashuvlar haqida ma'lumot bering.
3. An'anaviy sintaktik qarashlar haqida ma'lumot bering.
4. Transformatsion grammatikaning formal tavsif uchun ahamiyatli jihatlari va cheklangan tomonlari haqida mulohaza yuriting.
5. Formal grammatika iyerarxiyasining mohiyatini tushuntiring.
6. Grammatik qonuniyatlarning kompyuter lingvistikasidagi ifodasini yoriting.
7. “Grammatika” tushunchasini klaster usulida ifoda eting.

III. KOMPYUTER LINGVISTIKASINING MA'LUMOTLAR BAZASI

9-§. KORPUS LINGVISTIKASI

Reja:

1. Matnlar korpusi – elektron holda saqlanadigan ma'lum hajmdagi til birliklari.
2. Korpus lingvistikasi masalalari.
3. Miiliy korpus va uning ahamiyati.

Asosiy tushunchalar: *kompyuter leksikografiyasi, elektron matnlar korpusi, parallel matnlar korpusi, korpus lingvistikasi, empirik yondashuv, korpuslarning nazariy va amaliy ahamiyati, korpuslar yaratish, lingvistik tadqiqotlar, tilning leksik qatlami, fanlar integratsiyasi, umumiy tilshunoslik, leksikologiya, stilistika, lingvokulturologiya, madaniyatlararo kommunikatsiya, korpuslar tizimi va asosiy tavsifi, Milliy korpus, maxsus korpuslar.*

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 maqsadida va turli yo'nalishdagi tadqiqotlar uchun zaruriyatga qarab turli shakllarda tuziladi. Bular fonema, grafema, morfemalardan tortib undan kattaroq birliklar – leksema, gap va matnlar (badiiy yoki ilmiy asar, gazeta va jurnal matnlari)dan tashkil topishi mumkin. Ularning qay tarzda saqlanishiga qarab maxsus dasturlar yordamida har bir kerakli so'z yoki so'z birikmasining misollari, imlo bo'yicha variantlari, sinonimik qatorlari topilishi mumkin. Matnlar korpusiga oid ilmiy tadqiqotlar salmog'ining ko'payishi natijasida tilshunoslikda korpus lingvistikasi yo'nalishi shakllandi.¹

Korpus lingvistikasining maqsadi til o'rganishda empirik yondashuv negizidagi korpus lingvistikasi bilan bir qatorda korpus texnologiyalari asoslari bilan ham tanishtirishdan iborat. Shuningdek, korpuslar bilan ishlash malakalarini hosil qilish, tom ma'noda ilmiy lingvistik tadqiqot

¹Баранов А.Н. Введение в прикладную лингвистику. - М.: Эдиториал УРСС, 2001; Po'latov A., Muhamedova S. Kompyuter lingvistikasi. - T., 2007. Rahimov A. Kompyuter lingvistikasi asoslari.-T., 2011.

olib borishda korpuslarning nazariy va amaliy ahamiyatini ko'rsatish, lingvistik sohalarga oid fanlar tizimida korpus lingvistikasining o'rnini aniqlash vazifalari belgilanadi. Korpus lingvistikasi korpuslarni yaratish va ular asosida lingvistik tadqiqotlarni amalga oshirish, lisoniy tizimlarni obyektiv va lingvistik yo'nalishlar: leksikografik tadqiqotlar, tilning leksik qatlami tavsifi, u yoki bu so'zlarning qo'llanish nisbati, tildagi leksik-semantik, struktur o'zgarishlar, tabiiy tillar grammatikasining o'rganilishi, til tizimi mohiyati va uning qo'llanish tavsifi bilan shug'ullanadi.

Fanlar integratsiyasida korpus lingvistikasining o'rnini katta va u kompleks xarakterga ega. Korpus lingvistikasi fanini o'rganishda fundamental lingvistik bilimga ega bo'lish, umumiy tilshunoslik, leksikologiya, stilistika, lingvokulturologiya, madaniyatlararo kommunikatsiya asoslari, til tarixi, ingliz tili grammatikasi, tarjima nazariyasi va amaliyoti kabi gumanitar sohalarni ham o'zlashtirish talab etiladi.

Korpus lingvistikasi doirasida quyidagi masalalar ham qamrab olinadi:

- korpuslarning tizimi va asosiy tavsifi;
- korpuslarning asosiy turlarini bilish;
- lingvistik korpuslar tarixi: kartotekadan korpusgacha;
- korpus texnologiyalari asoslarini bilish;
- asosiy terminlarga ta'rif berish;
- korpuslar tasnifi. Vazifasiga ko'ra korpuslar turi. Formal omillarga ko'ra korpuslar turi. Birinchi tipdagi korpuslar. Ikkinchi tipdagi korpuslar;
- korpuslar texnologiyasi;
- turli xil turdagi korpuslar sharhi. Milliy korpus. Maxsus korpuslar. Korpus va korpus menedjerlarining qiyosiy tahlili. Braun korpusi. Britaniya milliy korpusi. Hozirgi amerika ingliz tili korpusi. S. Sharovning ruscha-inglizcha korpusi. Korpuslar yaratish texnologiyasi. Turli xil asoslarga ko'ra korpuslar tasnifi tipologiyasi;
- qidiruv vositalaridan (konkordanserlar va korpusli menedjerlardan) foydalana bilish;
- korpuslar asosida qidiruvni amalga oshirish va tadqiq qilish;
- korpus lingvistikasi shakllanish tarixi va zamonaviy holati. Rossiyadagi korpus lingvistikasi. Elektron lingvistik korpuslar tarixi. Korpus lingvistikasi maktablari. Korpuslarning birinchi avlodi: manbalar. Korpuslarning ikkinchi avlodi. Xomskiy lingvistikasiga empirik yondashuv;

- korpus lingvistikasi shakllanishiga asos bo'lgan lingvistik yo'nalishlar;

- korpus lingvistikasining asosiy tushunchalari;

- korpus – qidiruv tizimi sifatida;

- korpuslar kompilyatsiyasi metodologiyasi;

-internet qidiruv tizimi va lingvist orasidagi korpussifat interfeyslar;

- korpuslarga asoslanadigan lingvistik tadqiqotlar;

- korpuslarning reprezentativ masalalari;

- granulyarlik tushunchasi;

- tasniflash (korpuslar tipologiyasi);

- belgi qo'yish tushunchasi; belgi qo'yish jarayoni mohiyati va asosiy standartli korpusli belgi qo'yish. Belgi qo'yish turlari. Belgi qo'yilgan korpuslar, belgilash vositalari. Avtomatik morfem va sintaktik tahlil. Metalingvistik belgi qo'yish. Parallel korpuslar. Tekislash masalasi. Belgilangan matnlarni taqdim etishdagi lisoniy vositalar. Ekstralingvistik belgilangan korpuslarni yaratish metodlari, vositalari;

-lingvistik razmetka; ekstralingvistik razmetka;

- korpus lingvistikasidagi standartizatsiya;

-korpusdan axborotni toppish – chiqarish metodlari. Topilgan axborot turlari. Chastota (qo'llanish miqdori) bo'yicha axborot. Metaaxborot;

- konkordans. Konkordans korpus bilan ishlash dasturlari: konkordanslar va parallel korpuslar. Boshqa fanlarda korpuslarning qo'llanishi. Sotsiologiya va tarixda korpuslardan foydalanish;

-korpuslarga asoslangan lingvistik va nolingvistik tadqiqotlar. Korpus asosidagi lingvistik tadqiqotlar: leksikaning, grammatikaning o'rganilishi. Belgilangan matnlarni qayd qilishdagi lisoniy vositalar. Grammatik kategoriyalar bo'linishi. Korpuslarga asoslangan leksikografik tadqiqotlar;

- tavsiya qilingan adabiyotlardagi ma'lumotlarni tahlil qilish, qiyoslash va baho berish ko'nikmalariga, korpus lingvistikasining dasturiy vositalari va axborot resurslari bilan ishlash malakasiga ega bo'lish.

Lingvistlar, dastavval, 1960 yillarda kompyuterlashtirilgan matnlar korpusini to'pladilar. "Kompyuterda yaratilgan birinchi matnlar korpusi Braun korpusi (BK, inglizcha *Brown corpus*, VS) hisoblanadi, u 1961 yilda Braun universitetida yaratilgan, har biri 2000 so'zli 500 ta matn fragmentini o'z ichiga oladi. Braun korpusi (The Brown Corpus) matnlari AQShda nashr etilgan jurnallardan, amerika kitoblari, gazetalaridan olingan. Korpus mualliflari U.Frensis va G.Kucherlar uni dastlabki statistik ishlov berilgan katta hajmdagi materiallar miqdori: har xil statistik

taqsimotga asoslangan chastotali va alfavit-chastotali lug‘at sifatida shakllantirdilar.

Braun korpusini yaratishdan maqsad – ingliz tilidagi yozma janrlarini o‘rganish va qiyoslashdan iborat. Tuzuvchilar, birinchi navbatda, masalani hal qilish nuqtai nazaridan yondashib, matnni saralash va tuzish prinsipiga amal qildilar. Bir tomondan: korpus statistik protsedura asosiga qurildi, ikkinchi tomondan, statistika korpus mualliflarining professional intuitsiya bazasiga asoslangan erkin qarorlari bilan belgilandi. Mazkur murakkab jarayonda maksimal darajada ob'ektivlikka erishish uchun maksimal formallashgan, shaffof protsedurani tekshirish va nazorat talab qilinar edi.

Keyinchalik yevropalik tadqiqotchilar shu prinsip asosida 1961 yilda Buyuk Britaniyada birinchi marta e‘lon qilingan matnlar korpusini tuzishdi: 15 xil janr (registr), 2000 so‘zli (so‘zshaklli) 5000ta matn dan iborat. U ingliz tilining britaniya variantiga oid 1 mln. so‘zni qamrab oldi, uni britaniya va ikki norvegiya universiteti nomi bilan “Lankaster-Oslo-Bergen” yoki qisqacha LOB korpusi (The Lancaster-Oslo-Bergen Corpus) deb atashdi.

Braun tipida yaratilgan boshqa korpuslar ham tadqiqotchilar uchun muhim ahamiyat kasb etdi. 1963- yilda AQShda Braun universiteti (Brown University)da The Brown Standard Corpus of American English korpusi yaratildi. Mazkur korpus lingvistika sohasida yaratilgan bo‘lib, lingvistik tavsif va tahlillarga xizmat qiladi.

Birinchildan bo‘lib yaratilgan ikki katta korpus ingliz tilining yozma amerika va britaniya variantida yaratilgan bo‘lib, bu korpuslar hozirgi kunda ham ahamiyatini yo‘qotgani yo‘q, ingliz tili bo‘yicha bir qator tadqiqotlarga asos bo‘lib kelmoqda.

Bu korpuslar yaratilganidan keyingi o‘n yillikda kompyuter arzonlashdi, qudratliroq kompyuter sinflari yuzaga keldi, matnlarni jamlashning klaviatura usulidan tashqari skaner variantlari shakllandi. Bu imkoniyatlar korpus yaratish jarayonini yengillatdi, korpuslarning milliard so‘z hajmidagi tiplari yuzaga keldi¹.

1970-yillarda 1 mln so‘zni o‘z ichiga olgan matnlar korpusi asosida rus tilining chastotali lug‘ati yaratildi. 1980-yillarda Shvetsiyaning Upsala universitetida ham rus tilida matnlar korpusi yaratildi. Keyinchalik kompyuter leksikografiyasining rivojlanishi natijasida katta hajmli matnlar korpusiga ehtiyoj tug‘ildi. Ya‘ni 1 mlnta so‘z elektron lug‘atlar bazasi uchun yetarli emas. Shu asosda yirik hajmli matnlar korpusi yaratila boshladi. Ko‘pgina mamlakatlarda bunday korpuslar XX asrning 80-

¹Zaxarov V.П., Богданова С.Ю. Корпусная лингвистика. -Иркутск: ИГЛУ, 2011. –С.16-17.

yillaridan boshlab tuzila boshladi. 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.¹ Masalan, Rus tilining milliy korpusi hajmi hozirgi kunda 149 mln soʻzdan iborat. Keyingi yillarda Internet tizimining rivojlanishi virtual matnlar korpusining yuzaga kelishiga olib keldi. Internetdagi qidiriv saytlari, elektron kutubxonalar, virtual ensiklopediyalar korpus vazifasini bajarmoqda. Korpusning janri va tematik rang-barangligi internetdan foydalanuvchining qiziqishlariga bogʻliq. Masalan, ilm-fan doirasida “Wikipedia” katta hajmdagi matnlar korpusi sifatida foydalanilmoqda.²

Korpus lingvistikasida parallel matnlar korpusi ham muhim ahamiyat kasb etadi. Parallel matnlar korpusi esa, oʻz navbatida, badiiy asar, qoʻllanma, ommaviy axborot vositalari, turli xil hujjatlarning ikki yoki undan koʻp tillardagi elektron holdagi koʻrinishlaridir. Masalan, Yevropa Ittifoqi oʻzining barcha qonun va hujjatlarini ingliz, fransuz, nemis, ispan va italyan tillarida nashr qiladi hamda ular internet tizimiga ochiq arxiv sifatida qoʻyiladi. Bunday korpuslarning afzalligi shundaki, ular yordamida nafaqat biron bir soʻz 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 lugʻatlari tuzish imkoniyati tugʻiladi. Ushbu korpuslar kompyuter leksikografiyasi uchun amaliy hamiyat kasb etadi.

“Kompyuter lingvistikasi” masalalariga bagʻishlangan maqolalar seriyasida korpus lingvistikasiga oid maʼlumotlar ham keltirilgan³. Rus tilshunosligida korpus, uning turlari, oʻziga xos xususiyati, korpusning ijtimoiy ahamiyati, korpus tuzish tamoyillari borasida tadqiqot olib borilgan.

Oʻzbek tilshunosligida kompyuter lingvistikasi, tabiiy tilni qayta ishlash, statistik tahlil masalalariga doir izlanishlarda korpus lingvistikasiga ham toʻxtalib oʻtilgan⁴. Kompyuter lingvistikasi

¹ <http://www.corpus.leeds.ac.uk/list.html>

² http://www.wikipedia.org/wiki/corpus_linguistics

³ Новое в зарубежной лингвистике. Вып. XXIV. Компьютерная лингвистика. –М., 1989.

⁴ A.Poʻlatov, S.Muhamedova. Kompyuter lingvistikasi. –T.,2008; B.Yoʻldoshev. Kompyuter lingvistikasi. –T.,2009; A.Poʻlatov. Kompyuter lingvistikasi. –T., 2011; A.Rahimov. Kompyuter lingvistikasi asoslari. –T.,2011; F.Qurbonova. Kompyuter lugʻatlari: tezaurus. –T.,2014; L.Abduhamidova “Tilshunoslikning yangi yoʻnalishi: kompyuter lingvistikasi. –T.,2015.

yoʻnalishlari monografik tadqiqot obyekti sifatida oʻrganila boshladi¹. Kompyuter lingvistikasining yoʻnalishi sifatida shakllangan, hozirda oʻz taraqqiyot yoʻnalishiga ega boʻlgan alohida soha hisoblanuvchi korpus lingvistikasi masalalari keyingi yillarda monografik planda tadqiq etilmoqda². Tezauruslarning maʼlumotlar bazasi sifatidagi qiymatlari yoritilgan tadqiqotlar ham korpuslar haqida muayyan tasavvur berishga xizmat qiladi³. Tezauruslarning tuzilishi, ishlash tamoyillari, kompyuter bazasi sifatidagi imkoniyatlari, WordNet tezaurus bazasi haqidagi maʼlumotlar ham ilmiy-amaliy ahamiyatga ega⁴.

M.V.Kopotev, A.Mustayokilar hozirgi tilshunoslikda korpus asosidagi yondashuvning keng qoʻllanishi ahamiyatini bir qator misollar orqali koʻrsatib berishgan: 1.Korpuslarning grammatika va leksikaga oid tadqiqotlarda keng qoʻllanishi hozirgi tadqiqot amaliyotida odatiy holga aylandi. 2.Chastotaga asoslangan roʻyxatlar va kalit soʻzlar hozirgi elektron korpuslar tuzilishidan ancha muddat oldin yaratilgan va faol qoʻllangan. 3. Kollokatsiyalar (yaʼni leksemalar birikmasi) tadqiqi hozirda korpus tadqiqotlarining eng ommabop mavzularidan biri boʻlib qolmoqda. 4. Meʼyor /uzus tadqiqi. Meʼyor tadqiqi odatda korpus tilshunoslarining vazifalariga kirmasa-da, keng jamoatchilikni qiynaydigan koʻpgina keskin masalalar subʼekti baho asosida emas, balki statistik jihatdan koʻp miqdorni tashkil qiladigan materiallar mazmuni bilan hal qilinadi. 5.Korpus metodlari shakllanish davridanoq sotsiolingvistik tadqiqotlarda faol qoʻllanib kelmoqda. 6. Ogʻzaki nutq korpusini yaratish va oʻrganish korpus lingvistikasining alohida va faol ishlab chiqilgan sohasiga aylandi. 7.Korpus lingvistikasi shakllanish davridanoq xorijiy auditoriyada til oʻqitish bilan chambarchas bogʻliq boʻlib keladi. 8. Xato turlarini tasniflaydigan va taʼlim jarayonida ularga eʼtibor qaratadigan oʻquvchilar matni korpusini yaratish yangi soha hisoblanadi. 9.Turli pedagogik vazifalar bilan chambarchas bogʻliq va aynan lingvistik ahamiyatga ega boʻlgan vazifa koʻp tilli parallel korpuslarni yaratishdir. 10. Bir avtorga

¹ Абдурахмонова Н.З. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти (Содда гаплар мисолида): филол. фан. бўйича фалсафа доктори (PhD)... дисс. автореф. –Т., 2018; Абжалова М. Ўзбек тилидаги матнларни тахрир ва таҳлил қилувчи дастурнинг лингвистик модуллари (расмий ва илмий услубдаги матнлар тахрири дастури учун). Филол. фан. фалс. д-ри (PhD) дисс. –Т., 2019.

² Ҳамроева Ш.М. Корпус лингвистикаси атамаларининг қисқача изоҳли луғати. – Т.: «Камалак» нашриёти. – 47 б.; А.Эшмуминов.Ўзбек тили миллий корпусининг синоним сўзлар базаси. –Қарши, 2019.

³ Poʻlatov A., Muhamedova S. Kompyuter lingvistikasi. – Т., 2009. Rahimov A. Kompyuter lingvistikasi asoslari.–Т., 2011.

⁴Thye Global WordNet Assoiation <http://global.wordnet.org/> wordnets-in-thye –world.

tegishli bo'lgan elektron matnlar an'anaviy tarzda hal qilingan uslub va muallif stilemetriyasi masalalari doirasini kengaytiradi. 11. Korpus metodlari yordamida muvaffaqiyatli tarzda amalga oshirib kelinayotgan vazifalardan biri plagiat va yashirin ko'chirma olishni aniqlashdir. 12. Korpus metodlari sud-lingvistik ekspertizasi masalalarini hal qilishda qo'llanmoqda¹.

Savol va topshiriqlar:

1. Kompyuter leksikografiyasi haqida ma'lumot bering.
2. O'quv jarayonida qo'llanadigan kompyuter lug'atlari haqida so'zlang.
3. Kitobiy va elektron lug'atlar orasidagi mushtarak va farqli jihatlarni tavsiflang.
4. Kompyuter lug'atlarining xususiyatlarini bayon eting.
5. Kompyuter lug'atlarining ma'lumotlar bazasi sifatidagi ahamiyatini yoriting.
6. Kitobiy lug'atlarning turlarini ayting.
7. Kompyuter lug'atlarining turlarini tavsiflang.

10-§. 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.

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

Leksikografiya tilshunoslik fanining amaliy sohalaridan biri hisoblanadi. U yunoncha «lexikos» – «so'z», «lug'at», «grapho» – «yozaman» degan ma'nolarni anglatadi. Leksikografiya bo'limida

¹ Копотев М.В., Мустайоки А. Современная корпусная русистика// Slavica Helsingiensia 34.Инструментарий русистики: корпусные подходы. Хельсинки, 2008. –С.10-13.

lugʻatlar, ularni tuzish yoʻllari va tamoyillari oʻrganiladi. Avtomatizatsiya ijtimoiy hayotning barcha jabhalariga kirib kelishi natijasida lugʻatlarni elektron holatda yaratish imkoniyati yuzaga keldi. Kompyuter yordamida lugʻatlar bilan ishlashning optimallasuvi natijasida kompyuter leksikografiyasi yoʻnalishi shakllandi. Kompyuter leksikografiyasi amaliy tilshunoslikning muhim tarkibiy qismi boʻlib, unda lingvistik va dasturiy taʼminot asosida ishlaydigan, kodlash va dekodlash prinsipi asosida yaratilgan kompyuter lugʻatlari, ularni tuzish dasturlari, algoritmlari oʻrganiladi. Mazkur sohaning rivojlanishi natijasida turli nomlarda elektron lugʻatlar yaratilmoqda. Eng mashhur elektron lugʻatlar sirasiga CONTEXT, ABBY LINGVO, MULTI-TRAN, POLYGLOSSUM, MULTILEKS kabilar kiradi.

Leksikografiya sohasiga kompyuter texnologiyalarining tatbiq etilishi lugʻat tuzish ishlarini ancha yengillashtiradi. Axborot manbalarini yigʻish boʻyicha maxsus dasturlar (Database Software) lugʻat asosini tashkil etuvchi barcha maʼlumot va misollarni jamlash va sistematik tarzda ishlov berishga qulaylik tugʻdiradi. Bundan tashqari, boshqa maxsus dasturlar lugʻatni tahrir qilish va chop etish borasidagi ancha mashaqqatli mehnatni bir qadar osonlashtiradi. Elektron lugʻatlar ham, oʻz navbatida, anʼanaviy lugʻatlarga qaraganda birmuncha afzalliklarga ega. Bugungi kunda, masalan, *ABBY LINGVO* kompaniyasining lugʻat va tarjima dasturlari ulardan foydalanuvchi har bir xaridorga oʻz lugʻatini tuzish yoki mavjud lugʻat soʻz boyligini toʻldirib BORISH imkonini beradi. Ochiq turdagi Internet lugʻatlari (on-line lugʻatlar) ham koʻp hollarda barcha foydalanuvchilarga ushbu lugʻatlarni boyitish imkonini beradi. Shuningdek, hajm jihatidan ixchamligi, kompaktligi, boshqa manbalarga (internet tarmogʻi orqali, giper-murojaatlar yordamida) ulanish imkoniyati, mavjud matnlar korpuslari yordamida illyustrativ misollar tuza olish imkoniyati, multimedia tasvirlari bilan boyitilganligi, soʻzlarning sinonimik variantlari, omonimlik xususiyatlari, grammatik maʼlumotlar bilan taʼminlanganligi va ularning tezlik bilan foydalanuvchiga havola etilishi kabi jihatlari elektron lugʻatlarning optimalligini taʼminlaydi.

Kompyuter leksikografiyasiga semantik maydon, semantik tarmoq, semantik toʻr hamda freym semantikasining faol tatbiq etilishi natijasida kompyuter leksikografiya resurslari yaratildi. Shunday yirik leksikografik resurslardan biri FRAMENET boʻlib, u internet tizimida on-line rejimida ishlaydi.¹ Mazkur tizim Ch.Fillmorning «Tools for Lexicon Building» loyihasi asosida Kaliforniya shtati, Berkli shahridagi Xalqaro informatika

¹ <http://framenet.icsi.berkeley.edu/>

institutida ishlab chiqilgan. Framenet resursining ma'lumotlar bazasida 10 000 ta leksik birlik mavjud, shundan 6000 dan ortig'i to'liq annotatsiyaga ega. Bundan tashqari, ma'lumotlar bazasida 800 ta semantik freym ko'rsatilgan, 135 000 annotatsiyali gaplar keltirilgan.

Ingliz va rus tilshunosligida kompyuter lug'atlari borasida izlanishlar amalga oshirilgan.

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

1. Kitobiy lug'atlar kartotekalar asosida tuziladi. Elektron lug'atlar korpus (elektron kartotekalar) asosida yaratiladi.

2. Kitobiy lug'atlar qo'lyozma holatda, kompyuter lug'atlari elektron variantda tuziladi.

3. Kitobiy lug'atlar qo'lyozma holida tahrir qilinadi; elektron lug'atlar nisbatan tayyor lingvistik baza asosida yaratiladi.

4. Kitobiy lug'atlar bosma holida nashr etiladi; elektron lug'atlarda lingvistik ta'minot dasturiy ta'minot bilan uyg'unlashtiriladi.

5. Kitobiy lug'atlar tuzilishi sahifalar ketma-ketligiga tayanadigan chiziqlik 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'atlarda har bir so'zga mutanosib keluvchi kod ishlab chiqiladi va qo'llanadi, kodni qayta ishlash jarayonida zaruriy bo'lgan ma'lumotlar, tarjimalar, sinonim, antonim va sharhlarga ega bo'lish mumkin.

So'zlarni kodlashtirish quyidagicha amalga oshiriladi. Ma'lumotlar 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 tarjimasi, sinonimlari va antonimlarini ko'rsatish ularning birinchi harflarini kodlashdagi tartib raqamlari bilan birgalikda keltirish bilan amalga oshiriladi.

«So'z» – alfavit harflari bilan yozilgan oddiy so'z.

«Kod» – raqam va harflar ketma-ketligi bo'lib, unda so'zning barcha zaruriy morfologik, sintaktik, ieksik xususiyatlari haqidagi ma'lumotlar jamlangan bo'ladi: grammatik ma'lumot, adreslar, tarjima, sinonim, antonim, mutanosib sharhlar.

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

So'z kiritiladi.

1. Grammatik ma'lumotlar yaratiladi – kompyuter so'z haqida morfologik va sintaktik ma'lumotlarni so'raydi va ular darhol aniq nisbat asosida shifrlanadi.

2. Tayyor shifr «Grammatik ma'lumotlar» bo'limiga yozib qo'yiladi.

3. Tarjimalarda sinonimlar, antonimlar adreslari yaratiladi:

a) tarjimasi bo'lishi mumkin bo'lgan so'z so'raladi;

b) kiritilgan so'z kodning mutanosib bo'limiga yozib qo'yiladi; v) kodlashning oxirida tarjimalar (sinonimlar, antonimlar) kodi topiladi va kiritilgan so'zlar o'rniga ularning birinchi harflari va tartib raqamlari yoziladi (masalan, «katta» so'zi o'rniga K0083), agar tarjimalar (sinonim, antonimlar) kodi topilmasa, u holda shu so'zga nisbatan kodlash operatsiyasi amalga oshiriladi va bu bilan bosqich tugallanadi.

4. Mazkur so'zga sharh kiritiladi:

a) matn kiritiladi;

b) maxsus dastur asosida so'z va uning sharhi orasidagi moslik belgilanadi, so'ng kiritilgan matn xotiraga yoziladi, ular orasidagi moslik esa so'z kodida o'z ifodasini topadi.

Dastur ishlay boshlashi bilan ekranda ma'lumotlar paydo bo'ladi. Ular bilan tanishilgandan so'ng klavish bosiladi va dastur quyidagi bosqichlarda o'z ishini davom ettiradi:

I. Ekranga 6 rejimga ega oyna chiqadi.

1-rejim. Only translation (faqat tarjima) – muqobil tarjimalarni va mazkur so'zning qaysi sohaga tegishliligi haqidagi ma'lumotlarni beradi. Mazkur rejimning asosiy funksiyasini “Only-tran” maxsus protsedurasi amalga oshiradi. Ushbu protsedura, avvalo, mazkur so'z kodining adreslar bo'limidan tarjimalar adresini izlab topadi. So'ng uni qayta ishlashni boshlaydi: tarjimaning birinchi harfini olib, shu harf bilan nomlangan tekst faylini ochadi va berilgan tartib raqamiga ko'ra tarjimini topib, uni ekranga chiqaradi.

2-rejim. **Grammatical information (grammatik ma'lumotlar)** – so'zning barcha morfologik va sintaktik xususiyatlari va undan foydalanishdagi ayrim noaniqliklar haqidagi ma'lumotlarni beradi. “Gram info” maxsus protsedurasi ishlaydi. Grammatik ma'lumotlar kodi bo'limida jamlangan ma'lumotlarni tahlil qiladi.

3-rejim. **List of synonyms (sinonimlar ro'yxati).**

4-rejim. **List of antonyms (antonimlar ro'yxati).**

Ushbu rejimlar Syn-List va Ant-list maxsus protseduralari yordamida sinonimlar va antonimlar ro'yxatini beradi. Ular quyidagi tartibda ishlaydi:

sinonim va antonimlar kiritilgan soʻz kodning «sinonimlar adresi» va «antonimlar adresi» boʻlimlaridan topiladi hamda ekranga chiqariladi.

5-rejim. **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 soʻzlar ekranga chiqariladi.

6-rejim. **Comments (sharhlar)** – oʻrganilayotgan soʻz tez esda qolishi uchun «aytib berish» (подсказка) 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 izlash-

dan iborat. Agar u topilsa, maxsus protseduralar uzatiladi, aks holda 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.¹

Chastotali lugʻatlar

Reja:

- 1.Chastotali lugʻat va uning ahamiyati.
2. Tezaurus lugʻatlarning yaratilishi.
- 3.Tezaurus yuzasidan olib borilgan tadqiqotlar.
- 4.Tezurus lugʻatlarning ahamiyati.

Asosiy tushunchalar: *kompyuter leksikografiyasi, tezaurus lugʻat, deskriptor, maʼlumotlar bazasi, chastotali lugʻatlar, tematik bogʻlanishlar, leksik-semantik bogʻlanishlar, semantik-grammatik bogʻlanishlar,*

¹ Poʻlatov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2007. –B.58-61.

tillararo bog'lanishlar, so'zlarning faollik darajasi, ideografik (ideologik) tezaurus, analogik (qiyoslash) tezaurusi, assotsiativ (muvofiqlashgan) tezaurus.

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 yuzaga kelmoqda. Shuningdek, keyingi paytlarda muayyan sohalarga ixtisoslashgan elektron lug'atlar ham yaratilmoqda. Ya'ni, dastlab, kompyuter lug'atlari faqat umumiy leksikon bilangina cheklangan edi, hozirda fanga oid bo'lgan, qurilish, aviatsiya, avtomobilsozlik, harbiy, diniy, yuridik sohalarga oid so'zlar bazasi ham e'tiborga olinmoqda. Hatto so'zlarning matnda qo'llanish imkoniyatlari, distributiv holatlari, birikma holdagi ko'rinishlari ham nazarda tutilgan holda lug'atlar yaratilmoqda. Bu ko'rsatkichlar kompyuter leksikografiyasi sohasining taraqqiy etayotganidan darak beradi.

Kompyuter lug'atlarining bir necha turlari bor.

Chastotali lug'atlar matnda so'zlarning qo'llanish chastotasi (qanday miqdorda ishlatilishi)dan kelib chiqqan holda tayyorlanadigan so'zliklar hisoblanadi. 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 lingvostatistik tahlillar yordamida amalga oshiriladi. Chastotali lug'atlar kompyuterning lingvistik informatsion bazasi hisoblanadi. Masalan, chastotali lug'atdagi birinchi 10 va 50 eng ko'p uchraydigan so'z formalarining miqdori umumiy so'z birliklariga nisbatan agglutinativ tillarda flektiv tillarga qaraganda kamroq foizni tashkil etadi. Agglutinativ tillarda dastlabki 1000 so'z formasi umumiy so'z birliklarining 50 – 60 foizini tashkil etsa, bu ko'rsatkich flektiv tillarda 62 – 87 foiz atrofida bo'ladi. Flektiv tillarda chastotali lug'atning dastlabki 1000 so'zi umumiy har xil so'zlar sonining 69 – 89 foizini tashkil etsa, bu ko'rsatkich agglyutinativ tillarda 64 – 89 foiz atrofidadir. Flektiv tillarda dastlabki ko'p uchraydigan 100 – 200 so'z shakli (ko'pchiligi yordamchi so'zlar) jami har xil so'z shakllarining 50 foizini tashkil etsa, bu ko'rsatkich uchun agglyutinativ konstruksiyali turkiy tillarda 600 – 1000 ko'p uchraydigan so'z formalari zarur bo'lar ekan, buning asosiy sababi agglyutinativ va flektiv tillar morfologik strukturasi, so'z shakllarining tuzilishidadir.

Alfavit chastotali lug‘at yordamida muayyan old qo‘shimchalarning mahsuldorligini so‘z turkumlari bo‘yicha aniqlash mumkin. Masalan, rus tili chastotali lug‘atida B3- old qo‘shimchasi 74 xil fe‘l, 27 sifat, 14 ot va faqat 2 ravishda uchragan. Bez- old qo‘shimchasi asosan ot, sifat va ravish yasashda ishlatilar ekan, fe‘lda esa atigi 4 martagina ishlatilgan, xolos. Alfavit-chastotali lug‘atlar tilda yoki qardosh tillarda u yoki bu so‘z paradigmasining amalga oshishini kuzatish imkonini beradi.

Chappa (ters) chastotali lug‘atlar tilda qo‘shimchalarning qanday qo‘llanishini aniqlab beradi. Chastotali va ters chastotali lug‘atlarning ma‘lumotlariga ko‘ra, ot, olmosh, sifat, ravish va son tillardagi jami so‘zlarning 70%ini tashkil qilgani holda, fe‘l ulushi ularning 1/4 ga teng. Tillarda so‘z turkumlarining ishlatilishi ham tadqiqotchilar e‘tiborini o‘ziga

jalb etadi. Hind-yevropa tillarida yordamchi so‘z turkumlarining ko‘proq ishlatilishi kuzatiladi (28 – 34%), o‘zbek tilida esa kam o‘rin egallaydi (atigi 8%).

Jahon tilshunosligida sanoq usuli – chastotalarni hisoblash ishlari o‘tgan asrdan boshlangan. Xususan, Y.Arnold hindshunoslikka doir bir qator ishlarida, jumladan, «Vedalarining vazni haqida» (1905) nomli asarida chastotalarni hisoblash orqali Rigvedadagi turli fragmentlarning arxaikligini aniqlagan. Hindu dinining muqaddas kitoblaridan biri «Gita»ga tuzilgan konkordans «Gita lug‘ati» deb atalgan. Unda yordamchi so‘zlar va olmoshlar, (she‘r)lar keltirilgan. Hindu dinining boshqa muqaddas kitoblari Upanishadlar va Puranalarga ham konkordanslar tuzilgan. Muqaddas kitoblardan Qur‘oni karim va Bibliyaga ham shunday ro‘yxatlar tuzilganligi haqida ayrim ma‘lumotlar mavjud.

Ayrim dalillarga murojaat qilamiz: *M.Y.Lermontov* tilining chastotali lug‘ati, bir milliondan ortiq so‘zshaklga asoslangan rus tilining chastotali lug‘ati, *F.M.Dostoyevskiyning* 1971 – 1990 yillarda chop etilgan 30 jildlik asarlari asosida yozuvchi publitsistikasining chastotali lug‘at – konkordansi nashr etilgan. Bulardan tashqari, «*A.S.Pushkin asarlari tili lug‘ati*» (1956), «*Gyote lug‘ati*» (1966), «*T.Shevchenko tili lug‘ati*» (1964), «*Adam Miskevich tili lug‘ati*», «*Abay tili lug‘ati*», shuningdek, *M.Gorkiy, N.A.Nekrasov, A.V.Kolsov, M.A.Sholoxov, M.Avezov* va boshqa bir qator jahonga tanilgan yozuvchi, shoirlar asarlari matni yuzasidan tuzilgan lingvostatistik lug‘atlar shular jumlasidandir.

Hozirga qadar quyidagi chastotali lug‘atlar yaratilgan:

1. 1861-yilda bosmaxonada harf terish ishini osonlashtirish maqsadida dastlabki chastotali lug‘at tuzildi.

2. 1963-yilda eng faol 2500 leksema kiritilgan chastotali lug‘at Shteyn-feld tomonidan yaratildi.

3. 1970-yilda Leypsig universitetida prof. L.Gofman guruhi tomonidan tibbiyot, fizika, kimyo, qurilish ishi, matematika, veterinariya sohalarida eng ko‘p qo‘llanuvchi so‘zlarni jamlagan 3 tilli chastotali lug‘at yaratildi.

4. 1972-yilda Kissen tomonidan «Словарь наиболее употребительных слов современного узбекского языка» («Hozirgi o‘zbek adabiy tilining ko‘p ishlatiladigan so‘zlar lug‘ati») chastotali lug‘ati yaratildi.

5. 1974-yilda harbiy chastotali lug‘at-minimum yaratildi.

6. 1981-yilda R.Qo‘ng‘urov va S.Karimovlar tomonidan «Zulfiya poeziyasining lug‘ati. Konkordans»i tuzildi.

7. 1982-yilda matematika fani yuzasidan inglizcha-ruscha chastotali lug‘at ishlab chiqildi.

8. 1984-yilda Leningrad davlat universitetida tarix, iqtisod fani bo‘yicha inglizcha-ruscha chastotali lug‘at tuzildi.

9. 1986-yilda S.Rizayev va N.Bo‘ronovlar tomonidan «A.Qahhor «Sinchalak» povesti tilining chastotali lug‘ati» yaratildi.

10. 1987-yilda falsafa fanidan chastotali lug‘at tuzildi.

11. 2006-yilda D.O‘rinboyeva tomonidan «O‘zbek xalq dostonlari tilining chastotali lug‘ati» ishlab chiqildi. Bu lug‘at uchun «Alpomish», «Ravshan», «Rustamxon» dostonlari matni statistik tadqiqot obyekti qilib olingan, unda 28 499 so‘z shakli statistik tahlil qilingan¹.

Konkordans lug‘atlar

Konkordans lotincha so‘z bo‘lib, lug‘aviy ma‘nosi “*muvofiglashtirish*”, “*uyg‘unlashtirish*”dir. Atama sifatida so‘z, morfema, so‘z birikmasi yoki frazeologizm kabi birliklarning ma‘lum hajmdagi matn yoki muayyan ijodkorning asarlari asosida tuzilgan so‘zligi yig‘indisi, birliklarning necha marta uchrashi (qo‘llanish chastotasi) hamda asarning qaysi o‘rinlarida kelishini anglatadi.

O‘zbek tilshunosligida konkordans lug‘atlarning yaratilayotgani e‘tiborga molik. Ayniqsa, mumtoz manbalar konkordanslarining yaratilayotgani kompyuter leksikografiyasini takomillashtirish bilan bir qatorda ajdodlarimiz ma‘naviyati, ma‘rifiy-axloqiy qarashlarining bugungi avlodga yetkazilishida ham muhim ahamiyat kasb etadi.

¹ Po‘latov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2007. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

Alisher Navoiyning “Xamsa” asari turkiy adabiyotning eng sara namunasi. Unda mumtoz o‘zbek tilining noyob so‘z boyligi aks etgan. Toshkent davlat sharqshunoslik institutining bir guruh olim va tadqiqotchilari “Xamsa” dostonlari konkordanslarini yaratishdi. 2012-yilda boshlangan ish 2016-yilda nihoyasiga yetkazildi va besh dostonning konkordansi har bir bayti raqamlangan xalqaro transkripsiyadagi matni bilan nashrdan chiqdi. Professorlar konkordans tuzishning dolzarbligini shunday izohlaganlar: “Afsuski, turkiy tilli xalqlarning orzu-armonlari, ma’naviy va ma’rifiy qadriyatlarining buyuk kuychisi bo‘lgan, turkiy xalqlar sharafini butun olamga tarannum qilgan Navoiydek bir ulug‘ zot badiiy merosining ilmiy tanqidiy matni hanuz yaratilmadi va shoirning biror asarining konkordansi (jome’ so‘zligi) va mukammal izohli lug‘ati ishlab chiqilmadi. Shu bois olimlar ishni “Xamsa” dostoni konkordansini tuzishdan boshladilar”.¹

Chastotali lug‘atlarning kompyuter lingvistikasi uchun ahamiyati quyidagilarda ko‘rinadi:

- chastotali lug‘atlar avtomatik qidiruv tizimi asosida ishlaydigan elektron lug‘atlar yaratish uchun tayyor lingvistik ta’minot vazifasini o‘taydi;

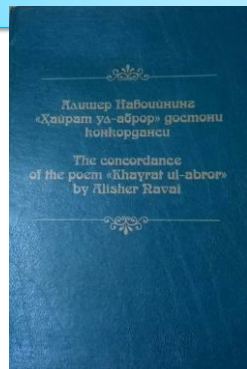
- chastotali lug‘atlar til o‘qitishni avtomatlashtirish jarayonida, kompyuter lingvodidaktikasi sohasida metodologik ahamiyat kasb etadi;

- chastotali lug‘atlar chet tillarini o‘qitishning samaradorligini oshirish uchun muhim omil bo‘lib, ular xorijiy tillarning leksik va leksik-morfologik minimumlarini yaratishga imkon beradi.

- chastotali lug‘atlar so‘zlarning faollik darajasini o‘rganishda, leksik-semantik xususiyatlarini yoritishda o‘ziga xos o‘rin tutadi.

¹ Умаров Э. Алишер Навоийнинг “Хамса” дostonлари конкорданслари. –Т., 2011.-Б.4.
Алишер Навоийнинг “Хайрат ул-аброр” дostonи конкорданси. –Т., ТДШИ,2012.

CHASTOTALI LUG'ATLAR MATNDA SO'ZLARNING ISHLATILISH CHASTOTASI (QAY DARAJADA ISHLATILISHI, FAOL QO'LLANISHI) DAN KELIB CHIQQAN HOLDA TAYYORLANADIGAN SO'ZLIKLAR HISOBLANADI. 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 LINGVOSTATISTIK TAHLILLAR YORDAMIDA AMALGA OSHIRILADI. CHASTOTALI LUG'ATLAR KOMPYUTERNING LINGVISTIK INFORMATSION BAZASIHISOBLANADI KEYINCHALIK CHASTOTALI LUG'ATLAR KONKORDANS DEB YURITILA BOSHLADI.



№	Словосочетание / Sary phrase / Obyekt / Group	№	№
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	7	7
8	8	8	8
9	9	9	9
10	10	10	10
11	11	11	11
12	12	12	12
13	13	13	13
14	14	14	14
15	15	15	15
16	16	16	16
17	17	17	17
18	18	18	18
19	19	19	19
20	20	20	20
21	21	21	21
22	22	22	22
23	23	23	23
24	24	24	24
25	25	25	25
26	26	26	26
27	27	27	27
28	28	28	28
29	29	29	29
30	30	30	30

Tezaurus lug'atlar

Tezaurus (thesauros) grek tilidan olingan bo'lib, «xazina» ma'nosini anglatadi. *Tezaurus* termini ikki xil ma'noda ishlatiladi:

1) lingvistik tahlillarga asoslangan to'liq ma'lumot yoritilgan lug'atlarni;

2) inson yoki hisoblash mashinalarida bilimlarning muayyan sohasiga taalluqli tizimlashtirilgan axborotlar majmuyini ifoda etadi. Tezaurus tarkibidagi so'zlar alfavit bo'yicha emas, balki mavzu va mavzular majmuyi sifatida shakllangan lingvistik lug'at hisoblanadi. Tezaurus – so'zlar va so'z birikmalari o'rtasidagi turli bog'lanishlar (munosabatlar) haqidagi axborotlarni saqlashga xizmat qiluvchi ma'lumotlar bazasidir.

Tezaurus aslida nafaqat voqelik haqidagi axborotni, balki yangi ma'lumotlarni qabul qilib olish imkoniyatini yaratuvchi metainformatsiyani (axborot haqidagi ma'lumotlarni) ham qamrab oladi.

Tezaurus “maxsus tashkil qilingan axborot”, tildagi leksik birliklarning normaga keltirilgan holatdagi yig'indisidir. Tezaurus til birliklarini, ular o'rtasidagi munosabatlarni ifodalaydi. Tezaurus milliy til, ma'lum fan tili yoki avtomatlashtirilgan boshqaruv tizimi uchun rasmiy qabul qilingan til semantikasini belgilaydi. Dastavval, tezaurus so'zning matnlardagi ma'nosi to'liq yoritilgan lug'atlar sifatida ham qo'llangan va u semantikaga asosan bo'limlarga ajratilgan so'zlar guruhlari bilan belgilanadigan bir tilli lug'at sifatida baholangan. Ayni vaqtda esa bir tilli lug'at va tezaurusning o'xshash va farqli jihatlari mavjud. Tezaurus

lugʻatning bir tilli lugʻatdan bir qator afzalliklari mavjud, tezauruslarda muayyan soʻz har tomonlama keng yoritiladi.

Tezaurus, ayniqsa, elektron formatdagi tezaurus har bir sohaga doir soʻzlarni toʻliq ochib beradigan lugʻatlardan biri sanaladi. Oddiy lugʻatlardan farqli oʻlaroq, har bir soʻz mohiyatining toʻliq yoritib berilishini taʼminlaydi. Tezaurus lugʻatda bir qancha lugʻat belgilarining mujassamlangan holatini koʻrish mumkin. Bu lugʻatning elektron formatdagi koʻrinishi insonning vaqtini tejash, bir soʻz haqida toʻliq maʼlumot olish imkonini beradi.

Tezaurus lugʻatda har bir soʻzning sinonimi, antonimi, etimologiyasi, omonimi va u bilan birika oladigan soʻzlar qatori ochib beriladi. Tezaurus birinchi navbatda soʻzlar va ularning matndagi qoʻllanishiga doir misollar izohlangan maksimal darajada toʻliq lugʻat hisoblanadi. Hozirgi tilshunoslikda tezaurus leksik birliklar oʻrtasidagi semantik munosabatlar (sinonimlar, antonimlar, paronimlar, giponimlar, giperonimlar) koʻrsatilgan umumiy yoki maxsus leksika lugʻatidir. Tezauruslar, asosan, elektron koʻrinishlarda yaratiladi va muayyan predmetlar sohasini atroflicha tavsiflash vositasi hisoblanadi hamda muayyan predmet haqida toʻliq holatda axborot beradi.

Izohli lugʻatdan farqli ravishda tezaurus tushunchani izohlabgina qolmay, uning sunʼiy intellekt doirasida faol qoʻllanadigan boshqa tushunchalar va shu tushunchalar guruhiga munosabatini ham aks ettiradi. Shu bilan bir qatorda tezaurus informatsiya (axborot) nazariyasida subyekt egallagan barcha maʼlumotlarning majmuasi sifatida ham qoʻllanadi.

Tezaurus korporativ kommunikatsiyaga ham yordam beradi. Korporativ kommunikatsiya – kompyuter lingvistikasiga oʻxshash, soddaroq qilib aytganda bitta kasb yoki bir xil tartib-intizom bilan bogʻliq insonlar oʻrtasidagi muloqotlarning, munosabatlarning asosi sanaladi.

Tilshunoslikda tezaurus lugʻatlarning quyidagi koʻrinishlari mavjud:

Amara – kosha (Amara lugʻati) – dastlabki sanskrit lugʻati (III – IV asrlar);

inglizcha *Tezaurus Roje* (XIX asr) – 15 000 ga yaqin leksik birlikni oʻz ichiga oladi;

Vikislovar – koʻp tilli, erkin toʻldirilib boriluvchi lugʻat va tezaurus (“*Rusvikislovar*”da 250 000 tushuncha, 67 000 munosabat mavjud);

RuTez – тезаурус русского языка – (informatsion tadqiqotlar markazi tomonidan) avtomatik indekslash maqsadida yaratilgan. 1997-yildan hozirga qadar 45 000 tushuncha, 107 000 soʻz va ifoda, 177 000 munosabatni aks ettiradi.

1962- yilda nashr qilingan (1-nashr, 1852) ingliz tezaurusi (muallif P. M. Rodjet), tarkibida 240 000 soʻz mavjud; bu lugʻat 1040 boʻlimga ega. Ushbu tezaurusning koʻrsatkichi (kaliti), boʻlim va boʻlim tarkibiy qismlari koʻrsatilgan soʻzlarning alfavit boʻyicha joylashtirilgan roʻyxatini oʻz ichiga oladi.

“Тезаурус” в словаре Д.Н. Ушаков “Большой толковый словарь современного русского языка” (Zamonaviy rus tilining katta izohli lugʻati);

“Тезаурус” в словаре “Этимологический словарь русского языка”. Фасмер Макс (Rus tilining etimologik lugʻati);

“Тезаурус” на португальский язык в словаре “Русско-португальский словарь” (Portugal tili tezaurusi);

“Тезаурус” на итальянский язык в словаре “Большой русско-итальянский словарь” (Italyan tili tezaurusi , katta rus-italyan tili lugʻati);

Перевод “тезаурус” на английский язык в словаре “Русско-английский машиностроительный словарь”;

Перевод тезаурус на английский язык в словаре “Русско-английский машиностроительный словарь”;

Перевод тезаурус на английский язык в словаре “Русско-английский юридический словарь”;

Перевод “тезаурус” на английский язык в словаре “Русско-английский медицинский словарь”;

Перевод тезаурус на английский язык в словаре “Русско-английский медицинский словарь”;

Перевод “тезаурус” на английский язык в словаре “Русско-английский физический словарь”;

Перевод тезаурус на английский язык в словаре “Русско-английский физический словарь”;

Перевод “тезаурус” на английский язык в словаре “Русско-английский политический словарь”;

Перевод тезаурус на английский язык в словаре “Русско-английский политический словарь”;

Перевод “тезаурус” на английский язык в словаре “Русско-английский словарь по электронике”;

Перевод “тезаурус” на английский язык в словаре “Русско-английский индекс к англо-русскому биологическому словарю”¹.

¹ Poʻlatov A., Muhamedova S. Kompyuter lingvistikasi. – T., 2009. Rahimov A. Kompyuter lingvistikasi asoslari.–T., 2011.

Ingliz, fransuz, ispan tillari uchun umumiy an'anaviy tezauruslar (alohida tillarning semantik tizimlari ta'riflari) mavjud.

Tezaurus yaratish borasida xorijda muhim tadqiqotlar olib borilgan. Ingliz tilidagi lug'at- tezaurus, V.Dal tomonidan yaratilgan izohli lug'at – tezaurus, T.F. Yefremovning rus tilidagi yangi so'zlarning lug'at tezaurusi, S.I. Ojegov, N.Y. Shvedova tomonidan yaratilgan rus tilining izohli lug'ati – tezaurusi, Fasmer Maksning rus tilining etimologik lug'ati – tezaurusi, yuridik terminlar tezaurusi yaratilgan.

XX asrning 70- yillarida axborot qidiruvchi tezauruslar keng tarqaldi. Bu tezauruslarda hujjatli axborot qidirilishini avtomat ravishda amalga oshirish imkonini beruvchi maxsus leksik birliklar - *deskriptorlar* ajratilgan. Axborot qidiruvchi tildagi leksik birliklar – "*deskriptorlar*" deyiladi. Deskriptor ma'lum bir soha matnlaridan ajratib olinadi va shu sohaga tegishli bo'lgan, nutqda boshqa so'zlarga nisbatan ko'proq qo'llaniladigan so'zlar – kalit so'zlar sifatida tanlanadi. Masalan, deskriptor sifatida har qanday kalit so'z (ko'p ishlatiladigan yoki qisqa so'zlar), yoki so'z birikmasi, yoki sonli kod tanlanishi mumkin. Bir necha so'zdan tashkil topgan sinonimik qatorlarga birdan ortiq deskriptorlar mos keladi va, aksincha, bir necha sinonimik qator so'zlariga va iboralarga bitta deskriptor mos kelishi mumkin. Avtomatlashtirilgan axborot qidiruvida nafaqat savol deskriptorlari indeksiga ega bo'lgan hujjatlar, balki ular bilan ma'lum semantik munosabatda bo'lgan deskriptorga ega hujjatlar ham qidiriladi. Ayrim vaqtlarda ushbu tematik sohaga xos bo'lgan tezauruslarda ma'lum uyushgan munosabatlarni belgilash to'liq ma'lumot olishga asos bo'ladi: *kasallik – qo'zg'atuvchi, asbob – vazifasi* (yoki o'lchanayotgan kattalik) va boshqalar. Leksik birlik (so'z, til birligi)ning tezaurusdagi o'rnini uning tildagi ma'nosini ifodalaydi: ushbu so'z (shu jumladan, u mansub bo'lgan bo'limlar) ishtirok etayotgan semantik munosabatlar tizimini bilish, shu so'z ma'nosini tahlil qilish imkonini beradi. Axborot qidiruvchi tezauruslarda har bir so'zga sinonim bo'lgan deskriptor taqqoslanadi va deskriptorlar uchun semantik munosabatlar ko'rsatiladi: *jins – ko'rinish, qism – butun, maqsad – vosita* va boshqalar. Tezaurus lug'atlarda birliklar iyerarxik va uyushgan munosabatlarga ajratiladi. Deskriptorlar o'rtasidagi uyushgan munosabatlarda tezaurusga tematik guruhlarning semantik kartalari ham kiritiladi. 1973- yilda chop etilgan "Informatika bo'yicha axborot-qidiruv tezaurusi"da har bir deskriptorga sinonim so'zlar, jinsga oid, ko'rinishga oid va uyushgan deskriptorlar alohida ko'rsatilgan. Tezaurus lug'atlarda so'zlar avtomat ravishda tartibga solinadi, saqlanadi va qidirish uchun ma'lum bir tizimga

joylashtiriladi; tezaurusda, raqamlash uchun tavsiya qilinadigan kalit soʻzlar (deskriptorlar) ajratib olinadi, hujjat mazmunini ochib beruvchi muhim (kalit) soʻz olinadi, shuningdek, ular orasidagi semantik bogʻlanishlar ifoda etiladi.

Tezaurus, keng maʼnoda, alohida tashuvchi yoki tashuvchilar guruhi ega boʻlgan voqelik haqidagi bilimlar tizimini tasvirlash vositasi sifatida namoyon boʻladi. Bu tashuvchi qoʻshimcha maʼlumot qabul qiluvchi vazifasini ham bajarishi mumkin, natijada uning tezaurusi ham oʻzgarib ketadi. Bunda boshlangʻich tezaurus qabul qiluvchining u tomonidan semantik axborotni qabul qilish imkoniyatlarini belgilaydi. Psixologiyada sunʼiy intellektga ega tizimlarni oʻrganishda, individumlarning axborotni qabul qilishi va tushunishi jarayonida namoyon boʻluvchi tezaurus xususiyatlari oʻrganiladi. Sotsiologiya va muloqot nazariyalarida esa tezaurusning umumiyliigi asosida oʻzaro hamfikrlilik imkoniyatlarini taʼminlovchi individum va jamoalar tezauruslarining xususiyatlari talqin etiladi. Bunday vaziyatlarda tezaurusga murakkab tizim tarkibidagi maʼlumotlar zaxirasini belgilovchi murakkab mazmunlar va ularning semantik bogʻlanishlarini kiritishga toʻgʻri keladi.

Tezaurusda bilimlar qaysi sohada yaratilgan boʻlsa, shu sohaga tegishli boʻlgan soʻzlar mavzusiga koʻra joylashtirilgan va leksik birliklar oʻrtasidagi semantik munosabatlar koʻrsatilgan boʻladi. Axborot-qidiruv tezauruslarida matnning leksik birliklari deskriptorlar bilan almashtiriladi.

Tezaurus kompyuter lingvistikasi algoritmlarida saqlanayotgan maʼlumotdan foydalanish, jumladan, tarjima jarayonini maksimal darajada osonlashtirish maqsadida ishlab chiqilgan.

Tezaurusdagi bogʻlanishlar quyidagicha boʻladi:

1. Soʻzlar bilan ifodalangan predmet va jarayonlar vazifalarining umumiyliigi va oʻxshashliigi natijasida soʻzlarning bir guruhga jamlaydigan predmetli (koʻgazmali) yoki tematik bogʻlanishlari: uy-roʻzgʻor predmetlari, tana aʼzolari, kiyim turlari, qurilishlar va b. Ideografik tezaurus – bu koʻrgazmali (tematik) guruh tarkibidagi leksik birliklarni ifodalovchi va ularni dunyo haqidagi konseptuallashtirilgan bilimlarni reprezentlash (qayta tasvirlash) uchun moʻljallangan iyerarxik tuzilmaga tashkillashtiruvchi leksikografik hosiladir.

2. Leksik-semantik bogʻlanishlar; bunday turdagi bogʻlanishli guruhlariga umumlashtirish soʻzlar uchun asosiy boʻlgan alomat – leksik qiymatga koʻra amalga oshiriladi. Bunda shaklida soʻzlarning alohida qiymatlari namoyon boʻladigan leksik-grammatik bogʻlanishlar ham inobatga olinadi. Analogik tezaurus – bu makrotuzilmaning asosiy birligi

leksik-semantik guruh bo'lgan leksikografik ma'lumotnomadir; guruhlarda ma'noviy dominantlar alfavit ketma-ketligi tartibida sistemalashtirilgan. Semantik bog'lanishlar giperonimlar va giponimlar (*xrom –metall*), sinonimlar (*chaqqon – epchil*), gender antonimlar (*moda – nar*), kichraytiruvchi-erkalovchi (*singil – singilcha*), kuchaytiruvchi /kattalashtiruvchi (*qo'l – katta qo'l*) semalarni ifoda etadi.

3. Semantik-grammatik bog'lanishlar – harakat: *yozuvchi – yozmoq*, xususiyat: *jasur – jasurlik*, instrument: *gitarachi – gitara*, manzil: *kon – konchi*, *xitoylik ayol-xitoy* bog'lanishlari. Bu bog'lanishlarning o'ziga xosligi shundaki, ular, asosan, bir o'zakli so'zlardan iborat va bu xususiyati ularni *quruqlik – quruq* tipidagi grammatik bog'lanishlar bilan yaqinlashtiradi, *it – itlik qilmoq*, *ilon – ilonlik qilmoq* kabi juftliklardan ajratib turuvchi birliklar orasidagi ma'lum munosabatlarni ifodalaydi.

Semantik-sintaktik bog'lanishlar asosida kelib chiqishiga va faoliyatiga ko'ra ikki taraflama bog'lanishlar uchun mo'ljallangan so'zlar guruhlariga yoki ma'noviy hamda sintaktik juftliklarga birlashtiriladi. So'zlarning sintaktik bog'lanishlari, asosan, gapda predikativ vazifani bajaruvchi fe'llar, sifatlar va otlar doirasida bog'lanadi, masalan:

- harakat va uning yordamida harakat sodir etuvchi organ (instrument) o'rtasidagi bog'lanish: *ushlamoq – qo'l*, *ko'rmoq – ko'z*, *suzmoq – qayiq* va b.

- bir subyektni talab etuvchi harakat fe'li va subyekt o'rtasidagi bog'lanish: *hurmoq – kuchuk*; *kishnamoq – ot* va b.

- fe'llar va talab etiluvchi ma'lum grammatik to'ldiruvchilar o'rtasidagi bog'lanishlar: *chopmoq – daraxt*, *yemoq – ovqat* va b. Bu tarzdagi bog'lanishlar assotsiativ tezaurusga zamin yaratadi. Leksik birliklarni ular orasidagi ma'noviy va sintaktik bog'lanishlar asosida va guruhlarini ularning so'z markazi shakllariga qarab joylashtirishni ta'minlaydigan lug'at-tezaurusdir.

4. Tillararo bog'lanishlar: *dog – sobaka – chien*. Bu turdagi bog'lanishlar mashina tarjimasini qurilishining asosi hisoblanadi.

Alohida ilovalar uchun esa boshqa turdagi bog'lanishlar zarur bo'lishi mumkin, masalan, izohli lug'atlar uchun izohlash bog'lanishlari, tarjimalar uchun transkripsiya bog'lanishlari talab qilinadi. Grammatik qo'zg'atuvchi bog'lanishlar turlarini cheklamaydi, zaruriyat tug'ilganda lug'atning dastlabki matnlariga yangi turdagi bog'lanishlar haqidagi e'lonlarni qo'shib BORISH mumkin. Bog'lanishlarning ayrim turlari to'liq sharhi alohida sahifalarga kiritib boriladi. Tezaurus aslida til (milliy til, aniq fan yoki avtomatlashtirilgan boshqaruv tizimi uchun rasmiylashtirilgan til)

semantikasini belgilaydi. Avvaliga tezaurus tarkibidagi semantik munosabatlar tematik bo‘limlarga joylashtirilgan so‘z guruhlari bilan belgilanadigan bir tilli lug‘at sifatida tadqiq qilindi (masalan, P.M.Rodjet lug‘ati). Ingliz, fransuz, ispan tillarining an’anaviy umumtilli tezauruslari tillarning alohida semantik tizimlariga asoslangan. Tezaurus har bir so‘zning asosiy parametrlari hosilalarini belgilovchi bir tilli lug‘atlarga juda yaqin turadi. S.I.Ojegovning rus tili lug‘ati bunga misol bo‘la oladi. Tezaurusda so‘zlar o‘rtasidagi semantik-ma’no jihatdan bog‘lanishlar antonimlar, sinonimlar, giponimlar, giperonimlar, guruhlar orqali ifodalanadi:

Sinonimlar – yozilishi har xil, ma’nosi bir-biriga yaqin so‘zlar (so‘z birikmalari): masalan: *chiroyli = go‘zal*.

Antonimlar – qarama-qarshi ma’noli so‘zlar: *yaxshi - yomon*.

Giponim – boshqa umumiy tushunchaning xususiy hodisasi bo‘lgan atama.

Zobit = giponim (*harbiy*); *inson* = giperonim (*harbiy*)

Giperonim – boshqa qator shaxsiy tushunchalar uchun umumiy bo‘lgan atama. Giponim (*ovqatni shirin pishiradi*) = giponim (*uyini toza tutadi*), giponim (*tikishni biladi*) = bu xususiyatlar umumlashtirganda *yaxshi uy bekasi* giperonimi ostida birlashadi.

Davlat standartlariga javob beradigan, ko‘pgina davlatlarda qo‘llanadigan “Bir tildagi axborot-qidiruv tezaurusi”da bog‘lanishlarning quyidagi ko‘rinishlari belgilangan:

– jins-tur: *turar joy inshootlari – uy, chodir, o‘tov*; bu misollarda, zot sifatida *turar joy inshootlari* olinsa, uning turlari sifatida – *uy, chodir, o‘tov* va b. olinadi;

– qism-butun: tana *so‘zi* butun sifatida olinsa, *qo‘l, oyoq, yuz* uning qismlarini ifodalaydi;

– sabab-oqibat: *ko‘p siqildi – xafaqonga chalindi* gapida *ko‘p siqildi* sabab bo‘lsa, *xafaqonga chalindi* uning oqibati sifatida yuzaga chiqadi;

– xom-ashyo – mahsulot: *kumush soat, ya’ni kumushdan tayyorlangan soat*;

– ma’muriy Iyerarxiya: *hokim – o‘rinbosar – xodim*;

– jarayon – subyekt: *tahsil olmoq – talaba*;

– jarayon – obyekt: *tahsil olmoq – kurs, bosqich*;

– xususiyat – xususiyat tashuvchisi: *yorituvchi – oy*;

– antonimiya;

– sinonimiya.

Uyushgan munosabatlar boshqa, iyerarxik munosabatlarga yoki sinonimik (ya'ni, so'zlar orasidagi bog'lanishlarning har qanday ko'rinishlari, hattoki, g'oyat spetsifik, faqatgina ma'lum bir fan sohasida mavjud bo'ladigan) munosabatlarga kirmaydigan tizim umumlashmasidir.

Tezurusdagi lug'at maqolasi (norasmiy darajada) quyidagi ko'rinishda bo'ladi:

YAXSHI

sinonimi – durust, ma'qul

antonimi – yomon

giponimi – samimiy, mehribon, oq ko'ngil, ko'ngilchan.

tur – fe'l-atvor ko'rsatkichi (neytral darajada).

Tezaurus va grammatika axborot-qidiruv uslubini tashkil etadi. Grammatika til hodisalarining sodir bo'lish qoidalarini (semantik kodlarni, sintagmalarni, gaplarni) qamrab oladi va sintaktik munosabatlarni belgilash vositalaridan, masalan, bog'lanishlar ko'rsatkichlaridan foydalanishni belgilaydi.

Yuqorida ko'rib chiqilgan axborot-qidiruv xizmatidan tashqari axborot tanishuv xizmati ham mavjud. Axborot-tanishuv xizmatida tezaurus mijozlar xatlarida uchraydigan barcha sifat va tavsiflarni, ularning darajalanishini ifodalashi lozim. Grammatika va tezaurus shunday mutanosiblikda, aniq holatda tashkil etilishi lozimki, tizim nima haqida topshiriq berilayotganini tushunishi kerak. Masalan, so'rovda ko'rsatilgan sonni: bo'y, yoshni (albatta, bularda kalit so'zlar, mijozning o'zi to'g'risida bergan ma'lumotlarini) to'liq holatda yetkazib berilishi talab qilinadi.

Tezaurus hamda grammatika qoidalari asosida hujjat va so'rovlarning qidiruv ko'rinishi (qidiruv ko'rsatmalari) shakllanadi. *Qidiruv ko'rsatmalari* – bu foydalanuvchi tomonidan talab qilingan, hujjatlar belgilarini o'z ichiga olgan axborot qidiruv tilidagi matndir.

Ma'lumotning qidiruv ko'rinishi – ma'lumot bilan bir xil mutanosiblikka qo'yilgan va uning sifatlarini ifodalovchi, uni so'rov talabiga ko'ra izlashi lozim bo'lgan, axborot-qidiruv tilidagi matndir. Ma'lumotning mazmunini ochib berish uning mavzusini belgilovchi qidiruv sifatlaridan tashqari, ma'lumotning qidiruv ko'rinishi aynan tenglashtirilgan hamda ayrim qo'shimcha ma'lumotlarni o'z ichiga oladi (nashrga oid ma'lumotlar, hujjat turi, uning tili va boshqalar).

Qidiruv ko'rsatmalari so'rovlar kelib tushishi natijasida shakllanadi, ma'lumotlarning qidiruv obrazi esa tizimning yangi hujjatlar bilan boyitilishi orqali to'ldirilgani kabi so'rovnomalarga javoblar izlash

jarayoni natijasida ham to'ldirilishi mumkin. Axborot oqimi katta va tez-tez yangilanib turadigan tizimlar uchun resurslarni kodlashga zarurat yo'q va qidiruv so'zi sifatida ko'pincha hujjatning o'zi yoki uning nomi qabul qilinadi, chunki yangi ma'lumotlar bilan to'ldirilgan sari asos sifatida olinadigan so'z ikkinchi darajali bo'lishi yoki umuman matnda ko'rinmay qolishi mumkin. Agar uning nomi yoki hujjatning o'zi deskriptor so'z sifatida qabul qilinsa, bu kodlar inson xotirasida uzoq saqlanadi, yo'qotilganda ham uni tiklash nisbatan oson kechadi. Har qanday holatda ham kompyuter lingvistikasi birinchi navbatda inson manfaatlariga xizmat qiladi.

Ta'kidlanganidek, bir qancha davlatlarda tezaurus tizimi samarali yo'lga qo'yilgan. Xususan, rus tilshunosligida yaratilgan tezauruslar quyidalarni o'z ichiga oladi:

1. So'zning yoki til birligining morfologiya va sintaksisga aloqadorligi.
2. Talaffuz qilinishi.
3. Semantik tarkibi.
 - A) so'zning asl ma'nosi;
 - B) sinonimi;
 - D) antonimi;
 - E) giperonimi;
 - F) giponimi.
4. Yaqin ma'nodagi so'zlar.
5. Etimologiyasi.
6. Frazologizmlarda qo'llanishi.
7. Boshqa tillarda qo'llanishi.

Masalan, *tezaurus* terminining o'zini olib ko'radigan bo'lsak,

1. Morfologiya va sintaksisda

Asos *tezaurus* – *tezaurusning*, *tezaurusga*, *tezaurusdan*, *tezaurusni*

Ko'plikda: *tezauruslar*

Voqelik nomi, ot so'z turkumiga mansub, turlanadi.

2. Talaffuz qilinishi: birlikda [tʲɪ 'zavrus], **ko'plikda** [tʲɪ 'zavruslar]

3. Ma'no tarkibi: *tezaurus* so'zning hamma ma'nolarini birlashtirib, lingvistik jihatdan ta'rif beradigan, so'zlar orasidagi semantik munosabatni yoritadigan lug'at. Hozirgi paytda semantikaga va lingvistikaga oid matnlar yetarli darajada tahlil qilinmagan.

- A) sinonimlari: ideografik lug'at;
- B) antonimi: -
- D) giperonimi: lug'at;
- E) giponimi: -

F) yaqin ma'nodagi so'zlar:

4. Etimologiyasi: grekchadan olingan bo'lib, (θησαυρός) "xazina" ma'nosini anglatadi.

5. Frazeologizmlarda yoki turg'un birikma, birikma shaklida qo'llanishi:

Lingvistikaga oid tezaurus

Statiskaga oid tezaurus

Logika (mantiq)ga oid tezaurus

6. Boshqa tillar (xalqlar)da qo'llanishi.

Angliyada: *thesaurus* Germaniyada: *thesaurus*

Niderlandiyada: *thesaurus* Bolgariyada: *tezaurus*

Portugaliyada: *tesauro*, Fransiyada: *thésaurus, thesaurus*

Chexiyada: *tezaurus* Ispaniyada: *tesauro, tesoro*

Italiyada: *thesaurus* Malayziyada: *θησαυρομ*

Kompyuter intellektual sistemasidagi o'zbek tilining holat fe'llari asosida tezaurus lug'atlarini yaratish jarayoni quyidagi tartibda amalga oshiriladi:

TINCHIMOQ

1. Morfologiya va sintaksisda: O'zak *tinmoq* = - *tin* + *ch* - so'z yasovchi qo'shimcha.

Fe'l so'z turkumiga mansub, tuslanadi.

2. Talaffuz qilinishi: birlikda [tinchidi], ko'plikda [tinchidilar].

3. Ma'no tarkibi: *tinchimoq* - "tinmoq-", "sokin bo'lmoq" ma'nolarini anglatadi.

A) sinonimlari: *tinchimoq, nafasni rostlamoq, to'xtamoq;*

B) antonimi: -**no** old qo'shimchasi orqali *tinch* so'zining antonimi *notinch* so'zi yasaladi.

D) giperonimi:

E) giponimi:

E) qarindoshlik so'zlari: *tashvishdan, xatardan xoli bo'lmoq, osoyishta bo'lmoq;*

F) eng yaqin qarindoshlik so'zlari.

4. **Etimologiyasi:** *тын-* fe'lining "harakat-faoliyatdan to'xta" ma'nosidan, eski o'zbek tilida *-(ы)* *ch* qoshimchasi bilan yasalgan; keyinroq ikkinchi bo'g'indagi tor unli talaffuz qilinmay qo'ygan, *ы* unlisining qattqlik belgisi yo'qolgan. Qadimgi turkiy tilda ham xuddi shunday ma'noni anglatgan bu fe'l asli *тынch-* shaklida talaffuz qilingan. Bu fe'l dastlab "tashvishdan", "xatardan xoli", "osoyishta bo'lmoq"

ma'nosini ifodalagan: *тын-+- (ы) ch = тыныч > тынч > tinch* ko'rinishida shakllangan.

5. Frazeologizmlarda yoki turg'un birikma shaklida qo'llanishi: ko'ngli tinchidi; yuragi tinchidi.

6. Boshqa tillarda qo'llanishi.

Quvonmoq fe'li:

1. Morfologiya va sintaksisda: o'zak - *quvon*; ko'plikda - *quvondilar*.

Fe'l so'z turkumiga mansub, tuslanadi.

2. Talaffuz qilinishi: birlikda [quvondi], ko'plikda [quvondilar]

3. Ma'no tarkibi: *quvonmoq* – “xursand bo'lmoq” ma'nosidagi *quvon*- asosidan shakllangan fe'lining noaniq shakli.

A) sinonimlari: *xursand bo'lmoq, sevinmoq*;

B) antonimi: *xafa bo'lmoq*;

V) giperonimi: –

D) giponimi: –

E) qarindoshlik so'zlari: –

F) eng yaqin qarindoshlik so'zlari: –

4. Etimologiyasi: qadimgi turkiy tilda “baxt” ma'nosini anglatgan *qыv* otining quv shaklidan eski o'zbek tilida *-a* qo'shimchasi bilan yasalgan fe'lga “o'zlik” ma'nosini ifodalovchi *-n* qo'shimchasini qo'shib hosil qilingan; keyinchalik ikkinchi bo'g'indagi *-a* unlisi *-o* unlisiga almashgan: *quv+a=quva+n=quvan->quvon*.¹

5. Frazeologizmlarda yoki turg'un birikma shaklida qo'llanishi: Ko'zlari quvondi; ko'ngli yayradi; dili yayradi.

6. Boshqa tillarda qo'llanishi.

Asramoq fe'li:

1. Morfologiya va sintaksisda: o'zak – *asra*,

Fe'l so'z turkumiga mansub, tuslanadi.

2. Talaffuz qilinishi: birlikda [asradi], ko'plikda [asradilar]

3. Ma'no tarkibi: *asra-* – “ehtiyot qilmoq, saqlamoq” ma'nolarini bildiruvchi *asra-* fe'lining noaniq shakli.

A) sinonimlari: *avaylamoq, ehtiyot qilmoq*

B) antonimi: –

D) giperonimi: –

E) giponimi: –

F) qarindoshlik so'zlari: –

G) eng yaqin qarindoshlik so'zlari –

¹ *Rahmatullayev Sh. O'zbek tilining etimologik lug'ati.-T.,2000.*

4.Etimologiyasi: bu soʻz asli qadimgi turkiy tildagi “avaylab boqilgan, parvarish qilingan” maʼnosini anglatuvchi *asbi* sifatidan *-ra* qoʻshimchasi bilan yasalgan, keyinchalik ikkinchi boʻgʻindagi tor unli talaffuz qilinmay qoʻygan:

asbi+ *ra*= *asbira*->*asra*. Bu feʼl dastlab “oʻz boqimiga olmoq” maʼnosini anglatgan.¹

5.Frazeologizmlarda yoki turgʻun birikma shaklida qoʻllanishi:
Alloh asradi. Xudo asrasin.

6.Boshqa tillarda qoʻllanishi.

Oʻzbek tilida **tezaurus** lugʻat yaratilishi tilshunoslikda globallashuv jarayoniga mos ravishda erishilgan yutuqlardan hisoblanadi.

Oʻzbek tilshunosligida ham kompyuter leksikografiyasi yuzasidan tadqiqotlar olib borilmoqda. Oʻzbek tilidagi har bir soʻzning tezaurus koʻrinishini ochib berish qiyin vazifa boʻlganligi sababli oʻzbek tilshunosligida fikrning mohiyatini ifodalaydigan, gapning markazini tashkil etuvchi feʼl soʻz turkumi yuzasidan soʻzlar jamlanib tadqiq etilgan.²

“Bir qator davlatlar tomonidan ishlab chiqilgan WordNet tipidagi tezauruslar tabiiy tillarni qayta ishlashda muvaffaqiyatli ravishda foydalanib kelmoqda. Leksik informatsiyalarning tezaurus tipida qurilishi va tashkil qilinishi V.N.Lukashevich tadqiqotida batafsil bayon etilgan”³. Tezauruslarning tuzilishi, ishlash tamoyillari, kompyuter bazasi sifatidagi imkoniyatlari, WordNet tezaurus bazasi haqidagi maʼlumotlar ham ilmiy-amaliy ahamiyatga ega.⁴ Kompyuter lingvistikasi yoʻnalishida *tezaurus* lugʻatlar nafaqat filologlar, balki axborot texnologiyalari sohasi xodimlari uchun ham zarur manba hisoblanadi.

Kompyuter lingvistikasi yoʻnalishida *tezaurus* lugʻatlar nafaqat filologlar, balki axborot texnologiyalari sohasi xodimlari uchun ham zarur manba hisoblanadi.

Tezaurus lugʻatlar tarkibi

Tezaurus bir necha darajalarga boʻlinadi. Tezauruslarda darajalash muammosi esa yangilik emas va bir necha yillar davomida mamlakatimiz va xorijning qator lingvistlari (V. V. Morkovkin, L. P. Stupin, K.Marello va b.) eʼtiborini tortgan. Bu sohada olib borilgan tadqiqotlar natijasida

¹ *Rahmatullayev Sh.* Oʻzbek tilining etimologik lugʻati. –T.,2000.

² *Kurbonova F.* Kompyuter lugʻatlari: tezaurus. –T.,2014; *Kurbonova F.* Holat feʼllari tezaurusi. MD. –T., 2014. Tezaurus lugʻat haqidagi maʼlumotlar shu manbadan olindi.

³ *Лукашевич В.Н.* Тезаурусы в задачах информационного поиска. –М.:Изд-во Моск. ун.-та,2011.61-97.

⁴Thye Global WordNet Assoiation <http://global.wordnet.org/> wordnets-in-thye –world.

ko'rsatilgan leksikografik asarlarning alternativ darajalari yaratildi. Ammo mualliflar tezaurus turlarini tanlashda, ayrim vaqtlarda darajalanishning umumiy prinsiplarini nazarga olmagan holda har doim ham bir xil usullarni qo'llamaydilar va natijada mantiqiy yondashuv buziladi. Bundan tashqari, keltirilgan darajalanishlar yetarli ravishda tizimli bo'lmaydi va lug'at-tezauruslarning barcha turlarini aks ettirmaydi.

Darajalanish, asosan, ingliz tilining zamonaviy tezauruslar materiallari asosida, bosma va elektron variantda ifodalangan qayta paydo bo'lgan leksikografik hosilalarni inobatga olgan holda amalga oshiriladi. Ushbu darajalashda quyidagi mezon asos qilib olingan:

1. So'zlik birliklari o'rtasidagi ma'noviy bog'lanishlar turi;
2. So'zlik hajmi;
3. So'zlikning yaxlitligi;
4. Leksemalar qiymatini ishlab chiqish;
5. Leksemalarning grammatik-stilistik xususiyati;
6. Leksemalar faoliyatining namoyishi;
7. Berilgan tillar soni;
8. Leksemlarni sistemaga solish uchun ishlatiladigan semiotik vositalar turi.

Darajalashda O.M.Karpova va I.Burxanovlar tomonidan avvalroq yaratilgan mezonlarga, shuningdek, inglizlar tomonidan tadqiq qilingan 30 dan ortiq lug'at-tezauruslar ko'rinishlaridan olingan natijalarga asoslanilgan. Darajalashda foydalaniladigan terminologiya V.V.Morkovkin, Y. N.Karaulov, I.Burxanov, K.Marellolar tomonidan leksikografik apparatga kiritilgan. Mavjud lug'at-tezauruslar turlarini yanada to'liqroq namoyon qilish uchun ko'p bosqichli darajalanishni yaratish taklif qilingan.

So'zlik birliklari o'rtasidagi mazmuniy bog'lanishlar turi, so'z ma'nolari o'rtasidagi aloqalar, so'zlarni birlashtirish tizimiga asosan tezauruslar uch guruhga bo'linadi:

1. Ideografik (ideologik) tezaurus.
2. Analogik (qiyoslash) tezaurusi (V.V.Morkovkin termini).
3. Assotsiativ (muvofiqlashgan) tezaurus (Y.N.Karaulov termini).

Ideografik tezaurus quyidagi ko'rinishlarda bo'ladi:

1. Shaxsiy ideografik tezaurus;
2. Tematik lug'at tezaurus;
3. Sistematik lug'at tezaurus;
4. Tematik-sistematik lug'at (I.Burxanov termini).

Shaxsiy ideografik tezaurus – katta strukturasi tilning leksik tarkibiga bog‘langanligiga muvofiq ravishda shakllantirilgan ideografik turlarning alohida ko‘rinishidir. Ideografik lug‘atning boshqa turlaridan farqli ravishda, shaxsiy ideografik tezaurus ilmiy taksonomiya asosida yaratilgan mantiqiy va qat‘iy ravishda tartibli darajalanish tuzilmasi bilan ifodalanadi.

Tematik tezaurus – ideografik tezaurusning alohida ko‘rinishi bo‘lib, makrostrukturaning asosiy birligi, leksemalarning denotatlari darajalanishi asosida birlashtirilgan va ma‘lum mavzuga mos ravishdagi nuqtayi nazardan kelib chiqib muhokama qilinuvchi leksemalarni o‘z ichiga oladigan majmuadir.

Tematik tezaurusda sinoptik kartaning mavjudligi ixtiyoriydir. Tematik lug‘atlar, odatda, boshqa tilda gapiruvchi foydalanuvchilar ehtiyojini qondirish maqsadida yaratiladi, shuning uchun bunday lug‘atlardagi mavzular ro‘yxati ularning foydalanuvchi muloqot kengligidagi qiymati darajasiga qarab belgilanadi.

Sistematik tezaurus – bu ideografik tezaurusning alohida ko‘rinishi bo‘lib, uning sinfiy tuzilmasi tilning leksik birliklari o‘rtasida mavjud bo‘lgan haqiqiy semantik munosabatlarni ifodalashga mo‘ljallangan. Darajalash tuzilmasi o‘z mohiyatiga ko‘ra vokabulyarning leksik-grammatik darajalanishini ifodalaydi, boshqacha qilib aytganda, uning bo‘ysunish nuqtayi nazaridan tasvirlangan paradigmatic tuzilmasini aks ettiradi.

Tematik-sistematik tezaurus – ideografik lug‘atning sistematik va tematik lug‘atlar birikishini ifodalaydigan alohida ko‘rinishidir.

Analogik (o‘xshatishga asoslangan) tezaurus, o‘z navbatida ikki turga bo‘linadi.

1. Shaxsiy analogik tezaurus;
2. Sinonimlar / antonimlar lug‘ati.

Leksik-semantik guruhlarining cheklanishi berilgan so‘zning butun ma‘noviy hajmi yoki uning ayrim qiymatlari bo‘yicha amalga oshirilishi mumkin. Bir ma‘noda kesishuvchi va bir turdagi ma‘noviy munosabatda bo‘lgan so‘zlar, hajm jihatidan leksik-semantik guruhlardan doimiy ravishda kichikroq bo‘luvchi sinonimik qatorlarni tashkil etadi.

Shunday qilib, so‘zlarning faqatgina sinonimik/antonimik munosabatlarini aks ettiruvchi lug‘atni analogik tezaurus (sinonimlar, antonimlar lug‘ati)ning alohida turi, deb hisoblash mumkin. Y.N.Karaulov fikriga asosan, assotsiativ tezaurusni ham kichik turlarga bo‘lish mumkin:

1. Bir ko‘rinishli assotsiativ tezaurus.

2. Shaxsiy assotsiativ tezaurus (Y.N. Karaulov termini).

Assotsiativ tezauruslarning oddiy assotsiativ lug‘atlardan prinsipial farqi quyidagilardan iborat:

a) assotsiativ lug‘atlarda rag‘batlantiruvchi omillar soni 100 - 200; assotsiativ tezauruslarda esa – 1000ga yaqin.

b) oddiy assotsiativ lug‘at rag‘batlantiruvchi omillarni bir marotaba taqdim qilish asosida quriladi; assotsiativ tezaurusni tuzish dasturi 3 bosqichni ko‘zda tutadi: birinchi bosqich an’anaviy assotsiativ tajribalardan farq qilmaydi (uning yakunida oddiy assotsiativ lug‘at hosil bo‘ladi); ikkinchi bosqichda rag‘batlantiruvchi omil sifatida birinchi bosqich jarayonlari asos bo‘ladi; uchinchi bosqich esa ikkinchi bosqich jarayonlaridan foydalangan holda ikkinchi bosqichni qayta takrorlash orqali amalga oshiriladi. Uchinchi bosqichdan so‘ng assotsiativ semantik kenglik berkitiladi, chunki jarayonlarda yangi so‘zlar va mazmunlarning paydo bo‘lishi kuzatilmaydi.

Assotsiativ tezaurus shaxs (sotsium)ning tilga oid hayotini boshqaruvchi nisbatan yopiq leksik-semantik va grammatik kenglikdir.

So‘zlik hajmi nuqtayi nazaridan qaraganda barcha mavjud tezauruslar kichik lug‘atlar sirasiga kiradi, vaholanki, u yoki bu ma’lumotnomaga kiritilgan leksik birliklar soni turlicha (150 ming – 800 ming atrofida). Shuni alohida ta’kidlash kerakki, tematik, tematik-sistematik lug‘atlar, sinonimlar, antonimlar lug‘ati va assotsiativ tezauruslar, asosan, kichik hajmda chop etiladi.

O‘ziga xosligiga nisbatan qaraganda tezauruslar umumiy va maxsus tezauruslarga bo‘linadi. Umumiy tezauruslar umumadabiy til leksikasini ifodalaydi. Maxsus tezauruslar esa asosan, nutq uslublari leksikasini aks ettiradi:

1. Adabiy til tezauruslari yozuvchi tili yoki alohida asar lug‘atlari, evfemizmlar lug‘ati va boshqalar bilan ifodalangan bo‘lishi mumkin.

2. So‘zlashuv uslubi slenglar lug‘atida ifodalanadi.

3. Hududiy tilning o‘ziga xosligi shevaga oid ideografik lug‘atlarda ifodalanadi.

4. Ilmiy til esa ilmiy-texnik tezauruslar (atamalar)da ifodalanadi.

Leksemalarni ishlab chiqish nuqtayi nazaridan qaraganda mavjud tezauruslar uchta asosiy darajaga bo‘linadi:

1. Kumulyativ – ya’ni leksemalar guruhini, ularning qiymatlarini ta’riflamasdan turib ifodalaydi.

2. Definitiv – guruhdagi har bir leksik birlikning bir tildagi izohlarini o‘z ichiga oladi (K. Marelloni termini);

3. Tarjimaga oid darajalanish – leksik birlikning mazmunini boshqa til yordamida ochib beradi.

Hozircha kumulyativ tezauruslar ustunlikni saqlab turibdi. Bu, avvalo, ideografik tezauruslarga taalluqli, chunki lug‘atning bu turi tashuvchilar uchun mo‘ljallanganligi sababli tilning butun leksik ko‘rinishlarini o‘z ichiga olishi lozim, ammo har bir leksik birlikni ta‘riflash uni foydalanuvchi uchun noqulay qilib qo‘yar edi. Yangi avlod lug‘atlari ko‘pincha ta‘riflarni o‘z ichiga oladi va definitiv tezauruslar guruhini to‘ldiradi (asosan, analogik – o‘xshash tezauruslar).

Grammatik-stilistik malakaga, shuningdek, leksemalar faoliyatining namoyishiga nisbatan mavjud tezauruslarni to‘liq va tabaqalashtirilgan (differenziatsiyalangan) tezauruslarga ajratish mumkin.

To‘liq tezauruslar mukammal qo‘shimcha semantik-funksional tabiatga ega va suratlar bilan bezatilgan materiallarni o‘z ichiga oladi; differenziatsiyalangan tezaurusda belgilar tizimi va tasviriy misollar nisbatan kamroqdir.

Taqdim qilingan tillar soniga qarab tezauruslar bir va ikki tilli tezauruslarga bo‘linadi.

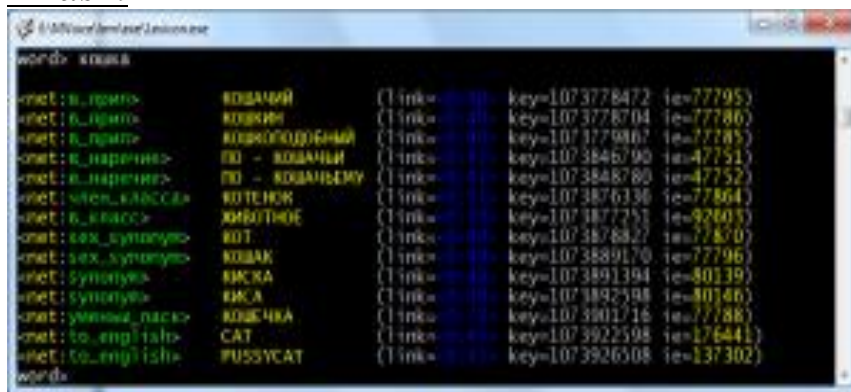
Leksemalarni semantizatsiyalash uchun ishlatiladigan semiotik turdagi vositalar nuqtayi nazaridan qaraganda, bosma nashrdan chiqarilgan an’anaviy tezauruslar, illyustrativ (rasmli) lug‘atlar va aralash semiotik vositali lug‘atlar to‘g‘risida gapirish mumkin. Dunyo leksikografiyasida kombinatsiyalangan lug‘at yangilikdir. Uning asosiy xususiyati – bu leksikani tashkil qilishning bir necha turlaridan foydalanilganidir. So‘zlikdagi leksemalarning alfavitli tartibi, odatda, lug‘atda keltirilgan leksik birliklar ularning ishlatilish sohasi yoki ular izohlaydigan tushunchalarga qarab tezaurus-indekslar bilan muvofiqlashtiriladi. Shu yo‘l bilan yaratilgan lug‘atlar sifatida izohli-ideografik (kombinatsiyalangan) lug‘atlarni, izohli-analogik (o‘xshash) (kombinatsiyalashtirilgan) lug‘atlarni va (potensial) izohli-assotsiativ lug‘atlarni ko‘rsatish mumkin. Bir vaqtning o‘zida izohli, analogik lug‘atlarni va bilvosita bog‘lanishlar lug‘atini o‘z ichida mujassamlantiruvchi elektron kombinatsiyalashtirilgan lug‘atlar soni ortib bormoqda. O‘z ichida ham lingvistik, ham ensiklopedik axborotlarni mujassamlantiruvchi bir va ikki tilli maxsus lug‘atlar ham ommalashib bormoqda; shu turdagi lug‘atlar, grafik ko‘rinishning umumiyligi jihatidan, ya’ni alfavit bo‘yicha, shuningdek, tematik hosilalar bilan

ta'minlangan leksikani sistemalashtiradi. Bundan tashqari, ilovalarning giperizohli sistemasi tufayli foydalanuvchi denotativ bog'langan leksemalar, leksema-giponimlar (giperonimlar)dan foydalanishi mumkin. Bunday lug'at-tezauruslarni shartli ravishda lingvistik-ensiklopedik kombinatsiyalashtirilgan lug'atlar, deb atash mumkin. Bundan tashqari, lug'atlarning texnik imkoniyatlari tufayli leksemalarning qo'shimcha semantik-funksional doirasi kengaymoqda; bunga, masalan, ovozashtirilgan lug'atlar, leksemalarning audioko'rinishlarini beruvchi lug'aviy maqolalar misol bo'la oladi.

Tezaurus – so'zlar va ularning matndagi qo'llanishi keltirilgan maksimal darajada to'liq lug'at hisoblanar ekan, unda bilimlar qaysi sohada yaratilgan bo'lsa, shu sohaga tegishli bo'lgan so'zlar mavzusiga ko'ra joylashtiriladi va leksik birliklar o'rtasidagi semantik munosabatlar ko'rsatiladi. Axborot-qidiruv tezauruslarida matnning leksik birliklari deskriptorlar bilan almashtiriladi.

Tezaurusning lug'at tarkibidagi muhim vazifasi tezaurus – so'zlar va so'z birikmalari o'rtasidagi turli bog'lanishlar (munosabatlar) haqidagi axborotlarni saqlovchi ma'lumotlar bazasi ekanligidadir.

1-rasm



Rus tilshunosligida tuzilgan tezaurus ko'rinishi (1-2 rasm)

2-rasm

Word1	Link type	Word2
кошка	to_japanese	ねこ
кошка	to_japanese	猫
кошка	в_прил	кошачий
кошка	в_прил	кошкин
кошка	в_прил	кошкоподобный
кошка	в_наречие	по-кошачьи
кошка	в_наречие	по-кошачьему
кошка	член_класса	котенок
кошка	в_класс	животное
кошка	sex_synonym	кот
кошка	sex_synonym	кошак
кошка	synonym	киска
кошка	synonym	киса
кошка	уменьш_ласк	кошечка

Tezaurusning elektron ko‘rinishi (versiyasi) amalga oshirilishining o‘ziga xos xususiyatlaridan biri tasvirlashdir. Bunda asosiy e‘tibor tezaurusda qo‘llanadigan atamalar haqidagi ma‘lumotlar tizimining mantiqiy bir butunligini va undagi ma‘lumotlardan foydalanish qulayligini ta‘minlashga qaratiladi. Hozirgi kunda, o‘z ichiga matnlardan axborotni ajratib olish usullari haqidagi bilimlarni, hujjatlarni mazmuniga qarab qidirish va indekslash, tilga oidlarini, jumladan nutqning rivojlanishi haqidagi bilimlarni o‘z ichiga olgan ilmiy-amaliy fan sifatida kompyuter lingvistikasiga bo‘lgan e‘tibor ortmoqda. Kompyuter lingvistikasi atamalarini o‘zlashtirishga, ta‘lim samaradorligini oshirishga, shu sohaga oid nashrlarni indekslashga, kompyuter lingvistikasi bo‘yicha bilimlardan foydalanishni osonlashtirishga, soddaroq holatda avtomatlashtirilgan, avtomatik ravishda ishlaydigan, tizimlashtirilgan bilimlarga katta ehtiyoj sezilmoqda.

Kompyuter lingvistikasi – o‘zbek tilshunosligidagi yangi yo‘nalishlardan biridir. Hozirgi kunda o‘zbek tilshunosligida kompyuter lingvistikasiga oid bir qator tadqiqotlar yuzaga kelmoqda. O‘z navbatida, soha oldida turgan muammolar, hal qilinishi lozim bo‘lgan masalalar anchagina. Shulardan biri kompyuter lingvistikasi atamalarini jamlash, izohlash, kompyuter leksikografiyasini shakllantirish, o‘zbek tilidagi kompyuter lug‘atlarini yaratish, ma‘lumotlar bazasi sifatida ahamiyatli bo‘lgan tezauruslar tuzish bilan bog‘liqdir.

Zamonaviy kompyuter lingvistikasi atamalarining to‘liq ixcham to‘plami nafaqat O‘zbekistonda, balki Rossiyada va hatto chet elda ham mavjud emas. Aslida lingvistik terminlar ham bir necha asosiy manbalarda, jumladan, “Lingvistik ensiklopedik lug‘at”da (LES), O.S.Axmanova lug‘atida va tarkibida lingvistikaning yangi tushunchalari bo‘lgan “Крыгосбет” internet-ensiklopediyasida aks ettirilgan. 2007- yilda RFAda yaratilgan tilshunoslik tezaurusi ushbu fan bo‘limlariga tegishli bo‘lgan 3000 ga yaqin atamani o‘z ichiga olgan. O‘zining barcha afzalliklariga qaramasdan bu tezaurus, avvalambor bibliografik qidiruvga mo‘ljallangan, shuning uchun lug‘aviy atamalar ta‘rifi (defenitsiyalari)ga ega emas. Bundan tashqari, RFAda yaratilgan tezaurus bir tilli va aynan kompyuter lingvistikasi atamalari ham keltirilgan (kompyuter lingvistikasi sohasiga tegishli bo‘lgan umumiy atamalar tezaurusda keltirilgan atamalar umumiy hajmining 4% iga teng).

Kompyuter lingvistikasi atamalari ta‘riflari Rossiyada nashr etilgan “Sun‘iy intellekt” nomli lug‘atda ham mavjud. Biroq u 1980 yillar oxiridagi atamalarni aks ettiradi. Bundan tashqari, kompyuter

lingvistikasiga oid atamalarning kam qismini qamrab oladi va unda keltirilgan atamalar kompyuter lingvistikasi nuqtayi nazaridan emas, balki sun'iy intellekt nuqtayi nazaridan izohlangan.

Kompyuter lingvistikasi fanlararo bog'lanishga ega bo'lganligi sababli shu sohaga oid terminlar izohini umumiy ensiklopedik lug'atlar (masalan, "Katta ensiklopedik lug'at" –BES)da ham uchratish mumkin Hozirgi kunda kompyuter lingvistikasi bilimlari bo'yicha eng mashhur manba – bu kompyuter lingvistikasi tushunchalarining ko'p qismiga izohlar, ko'rsatmalar topish mumkin bo'lgan "Vikipediya"dir, biroq bu ma'lumotlar ko'pincha bir tomonlama, to'liq emas va ularning ko'pi xomakidir.

Hozirgi kunda kompyuter lingvistikasining barcha terminlarini bir tizimga birlashtirish uchun ingliz – rus yoki rus – o'zbek tillaridagi kompyuter lingvistikasiga oid terminlar izohini qamrab olgan ikki tilli tezauruslar yaratish lozim bo'ladi. Bunday tezaurus bilimlarning shu sohasiga tegishli bo'lgan terminlar to'g'risidagi ma'lumotlarni to'plash va bir tizimga solish imkonini yaratadi, ma'lumot manbalari, shu tizimdagi axborotlarni mazmuniga qarab qidirish imkoni mavjud bo'ladi. Bundan tashqari, ushbu tezaurus nafaqat kompyuter lingvistikasi sohasining bo'lg'usi mutaxassislarini, balki, umuman axborot texnologiyalari sohasi mutaxassislari faoliyatini ham yengillashtirib, bilim darajalarining oshirilishiga yordam beradi.

O'zbek tilshunosligida tezauruslarning yaratilishida o'zbek tili xususiyatlarini hisobga olish lozim bo'ladi. Har bir til birligiga oid tushunchalar tezaurusini berishda ularning nutqdagi ahamiyatlilik darajasidan kelib chiqib yondashiladi. Turkiy tillarda, jumladan o'zbek tilida fe'llar fikr ifodalashda muhim o'rin tutadi. Shu bois o'zbek tilidagi fe'llar tezaurusini boshqa tillarga qiyoslagan holda yaratish tilning ijtimoiy funksiyasini oshirishda, axborot uslubini shakllantirishda va takomillashtirishda muhim o'rin tutadi.

Tezaurus tuzishda, dastavval, kompyuter lingvistikasi bo'yicha tezaurus yaratishga yondashuvlar o'rganiladi. Tezaurusning umumiy tuzilmasi, tezaurus manbalar strukturasi va terminlar orasidagi bog'liqliklar to'plamini izohlash lozim bo'ladi. Atamalar o'rtasidagi munosabatlarni tanlash va aks ettirishdagi muammolarga, shuningdek, tezaurus terminologik tizimining mantiqiy bir butunligini ta'minlashga alohida e'tibor qaratish kerak.

Tezaurusning mukammal variantini yaratishda Rossiya va Xalqaro standartlar asosida tezaurusga kiritilishi mumkin bo'lgan asosiy birliklarni

va ular o'rtasidagi munosabatlar to'plamini, leksik birliklar turlarini yig'ish qoidalarini o'rnatish, so'zlarni shakllantirish, so'z manbalarini tuzish va axborot-qidiruv tezauruslarini rasmiylashtirish lozim.

Tezaurusning asosiy atamalari fanda *deskriptor va askriptor*larga bo'linadi. Davlat standartlariga asosan axborot-qidiruv tezaurusiga leksik birliklarning quyidagi turlari kiritiladi: yakka so'zlar (otlar, sifatlar, fe'llar, ravishlar), nom (ot)ga oid so'z birikmalari, murakkab so'zlarning leksik qiymatga ega komponentlari, so'z va so'z birikmalarining qisqartmalari. Leksik birliklar axborot-qidiruv tezauruslarida, agarda bir leksik birlikning ikkinchisiga almashtirilishi, qidirilayotgan axborot matni ma'nosining o'zgarib ketishiga olib kelmasa, ekvivalent hisoblanadi. Ekvivalentlik sinfining leksik birliklaridan biri ushbu sinf vakili sifatida tanlab olinadi va deskriptor maqomini, qolgan leksik birliklar esa askriptor maqomini oladi. Shuningdek, abbreviatura (qisqartma so'z) yoki yozuvning boshqa ko'rinishlari (chiziqcha, oraliq joy) bilan tasvirlangan atamalar ham askriptor maqomini oladi.

Axborot-qidiruv tezauruslari o'z tarkibiga ko'ra, barcha birliklari deskriptor (maqbul bo'lgan) va askriptor (maqbul bo'lmagan) atamalarni aks ettiruvchi tezauruslarga ajratiladi. Bunda deskriptorlar hujjatlarni indekslashda va qidiruv so'rovnomalarida ishlatilishi, askriptorlar esa (matnga kirish sifatida) bir yoki bir necha deskriptorlar bilan almashtirilishi lozim bo'ladi.

Tezaurusning lug'at manbasi tarkibiga o'z maqomidan qat'iy nazar yana quyidagi elementlar ham kiritiladi:

-so'z, so'z birikmasi yoki tildagi murakkab so'zlarning leksik ahamiyatga ega bo'lgan tarkibiy qismini aks ettiruvchi fan sohasidagi *atama nomi*.

- atamalar nomi berilgan *til*.

- atamalardan foydalanish qoidalari va tavsiyalarini, shuningdek, lug'at muallifining fikr-mulohazalarini va izohlarini o'z ichiga oluvchi *sharhlar*.

- *lug'at muallifi* – tezaurusning jamoaviy ishlab chiqish jarayonini nazorat qilish maqsadida beriladi.

Deskriptor atamalar yuqorida keltirilgan alomatlardan tashqari quyidagi qo'shimcha alomatlar bilan ham ifodalanadi:

- *atamani ifodalash* – atamaning ushbu tildagi ma'nosi yoki mazmunini izohlash. Tezaurusda atamalar ifodasining mavjudligi uni nafaqat qo'l vositasida yoki avtomatlashtirilgan holda indekslash, balki

ushbu fan sohasidagi bilimlarning tizimlashtirilgan manbasi sifatida foydalanish imkonini ham yaratadi;

- *relyator (ma'lumotnoma)* – ta'riflanayotgan fan sohasi doirasida omonim terminlarning ma'nolarini farqlash maqsadida kiritilgan belgilarni ifodalaydi. U atamaning bir qismi bo'lib, uning muayyan bir tushuncha kategoriyasiga yoki fan sohasiga taalluqliligini ifodalagan holda, mazmunini ochib beradi;

- deskriptor tushunchalar Iyerarxiyasining eng yuqori cho'qqisida turganligini ko'rsatuvchi *negiz* atamaning belgisidir.

Tezaurus atamaları har bir atamaning tanlangan fan sohasi tizimidagi o'rnini aks ettiruvchi turli semantik munosabatlar bilan bog'lanadi.

Deskriptorlarni askriptorlar bilan bog'lash uchun sinonimlar munosabatlaridan foydalaniladi. Deskriptor har qanday matnda, ma'nodosh sifatida qaysidir askriptor o'rnini bosa olsa, bunda u "sinonim" munosabati bilan bog'lanadi va sinonimik munosabatni yuzaga chiqaradi; bunda askriptordan deskriptorga tomon teskari munosabat yuzaga keladi. Agar deskriptor va askriptorlar o'rtasida o'zaro munosabat mavjud bo'lmasa, u holda askriptorlar va ular bilan almashtiriladigan deskriptorlar o'rtasida o'zaro munosabatni ta'minlovchi "Alternativ ravishda foydalan" yoki "kombinatsiyadan foydalan" kabi munosabatlar istifoda etiladi; bunda "Alternativ tanlovni taqqoslab ko'r" va "Kombinatsiyani taqqoslab ko'r" kabi ularga qaytuvchi munosabatlar kiritiladi. Masalan, **partitsipant** askriptorini **semantik valentlik** va **vaziyat ishtirokchisi** deskriptorlari bilan "Alternativ ravishda foydalan" munosabati orqali bog'lash mumkin. Xuddi shunday, **statistik mashina yordamidagi tarjima tizimi** askriptorini "Kombinatsiyadan foydalan" bog'lanishi yordamida ikki deskriptorlar kombinatsiyasi (birikmasi) sifatida quyidagicha ifodalash mumkin – **mashina yordamidagi tarjima tizimi va mashina yordamida statistik tarjima**.

Deskriptorlar o'rtasiga quyidagi iyerarxik munosabatlar kiritiladi: quyi deskriptordan yuqoriga yo'naltirilgan "Yuqori" darajalanmagan iyerarxik bog'lanish, quyi deskriptor tushunchasi hajmi yuqori deskriptor tushunchasi hajmiga kirgandagi "Yuqori tur" o'zaro bog'lanishi, quyi deskriptor yuqori deskriptor obyekt komponentlaridan birini ifodalaganda ikki deskriptor o'rtasida "Yuqori butun" partonimik bog'lanishi, shuningdek ularning aksi bo'lgan "Quyi (past)", "Quyi ko'rinish", "Quyi qism" munosabatlari ham kiritiladi.

Tushunchalar sinfini va shu sinf namunasini ifodalovchi deskriptorlar orasidagi munosabatlarni belgilash maqsadida “Yuqori namuna sinfi” va “Quyi namuna” bog‘lanishlari tanlab olinadi.

Ayrim deskriptorlarga oid iyerarxik munosabatlarni o‘rnatish uchun “Iyerarxiyaning bo‘linish jihatlari” belgisini ko‘rsatish mumkin. Masalan, “Quyi ko‘rinish” munosabati asosida qurilgan iyerarxiyada mashina yordamida tarjima “yondashuv” belgisiga ko‘ra mashina yordamida statistik tarjima, Qoidalarga asoslangan holda Mashina yordamidagi tarjima va pretsendentlarga asoslangan holda mashina yordamidagi tarjimalarga bo‘linadi, “insonning ishtiroki darajasi” belgisiga ko‘ra – to‘liq avtomatik tarjima va inson-mashina yordamidagi tarjimaga ajraladi. Ko‘rinadiki, aynan bir deskriptor bir vaqtning o‘zida turli bog‘lanishlar asosida qurilgan (“Yuqori”, “Yuqori-tur”, “Yuqori-butun”) va iyerarxik bo‘linishning turli jihatlarga qarab bir nechta iyerarxik tushunchalarga mansub bo‘lishi mumkin.

Deskriptorlar yordamida ifodalanayotgan tushunchalararo semantik bog‘lanishlarning bir tilli ko‘rinishi (versiya)larida iyerarxik va birikmali munosabatlar o‘rnatiladi.

Deskriptorlar o‘rtasidagi, iyerarxik va sinonimik munosabatlardan farq qiluvchi bog‘lanishlar, “bog‘liqlik” munosabati tomonidan belgilanadi. Bunday munosabat deskriptorlar orasida erkin assotsiativ bog‘lanishlar, masalan, “jarayon-obyekt”, “sabab-oqibat” va boshqa ko‘rinishlardagi tobeliklarni aks ettiruvchi munosabatlarni hosil qilish imkonini beradi.

Turli tildagi versiyalar deskriptorlarining teng ma‘noga ega ekanligini ko‘rsatish uchun ular o‘rtasida “boshqa tildagi ekvivalent” munosabati o‘rnatiladi. Agar tushunchani boshqa tilda aynan bir deskriptor yordamida ifodalash imkoni bo‘lmasa, unda standartga muvofiq ekvivalent sifatida bir necha deskriptorlar kombinatsiyasi ko‘rsatiladi.

Tezaurusga kiritish uchun atamalar tanlash muammolari, shuningdek, bir qancha sinonimik atamalar orasidan asosiy deskriptor-atamani tanlab olish masalalari ushbu mavzu ustida ko‘proq ish olib BORISHni talab qiladi. Tezaurusning elektron ko‘rinishi (versiyasi) amalga oshirilishining o‘ziga xos xususiyatlarini tasvirlashda asosiy e‘tibor tezaurusda qo‘llanadigan atamalar haqidagi ma‘lumotlar tizimining mantiqiy bir butunligini va ma‘lumotlardan foydalanishning qulayligini ta‘minlashga qaratiladi.

O‘zbek tilshunosligida kompyuter lingvistikasini shakllantirish va rivojlantirishda kompyuter tizimidagi ma‘lumotlar bazasi, terminlar

lug'ati, boshqa lug'at turlari o'ziga xos o'rin tutadi. Bir sohadagi muammoni hal etish boshqa yo'nalishdagi muammoning yechimi bilan bog'liq. Shunday ekan, o'zbek tilshunosligining kompyuter lingvistikasiga doir muammolari qisqa davrda hal etilishi mushkul bo'lgan, yaxlit tizim holatidagi ilmiy tadqiqotlarni talab etadigan masalalar ekanligi ma'lum bo'ladi.

Kompyuter lingvistikasi terminlarini izohlash, aniq va umum tomonidan qabul qilingan ilmiy terminlar glossariylarini yaratish, sohaning asosiy tushunchalari haqidagi ma'lumotlar bazasini shakllantirish, bilimlarni tizimlashtirish dolzarb hisoblanadi.

Tezaurus bilimlarni to'plash va bir tizimga solish imkonini yaratar ekan, ma'lumot manbalari va shu tizimdagi axborotlarni ma'noga ko'ra qidirish, nafaqat kompyuter lingvistikasi sohasi mutaxassislarini, balki, umuman axborot texnologiyalari sohasi mutaxassislari faoliyatini yengillashtirishi, bilim darajalarini oshirishi mumkin bo'ladi. Buning uchun, dastavval, tezaurusning umumiy tuzilmasi, tezaurus manbalar strukturasi va atamalar orasidagi bog'liqliklar to'plamini izohlash, terminlar o'rtasidagi munosabatlarni tanlash va aks ettirishdagi muammolarga e'tibor qaratish, shuningdek, tezaurus terminologik tizimining mantiqiy bir butunligini ta'minlash lozim. Bu borada ingliz, rus kompyuter lingvistikasida yaratilgan tadqiqotlar yo'naltiruvchi manba sifatida xizmat qiladi.¹

Tilga yoki tillarga yo'naltirilganligiga qarab tezauruslar bir tilli va ko'p tilli bo'lishi mumkin. Ikki til asosida tezaurus yaratish tadqiqotchilarga keng imkoniyat yaratadi. Kompyuter lingvistikasi bo'yicha rus-ingliz tezaurusi ko'p tilli axborot-qidiruv tezaurusining bir tilli to'plami sifatida ishlab chiqilganida bir vaqtning o'zida tezaurusning bir-biriga mos keluvchi terminlari ko'rinishlari – ingliz va rus tilidagi ko'rinish(versiya)larigina tahlil qilinadi.

Bir tilli versiyalarning har biri ISO 2788-1986 Xalqaro standarti, GOST 7.25-2001 Davlatlararo standarti asosida ishlab chiqiladi.

Ko'p tilli axborot-qidiruv tezaurusi bir necha tillar terminlariga ega bo'ladi va ularning har birida ma'no jihatidan o'zaro teng bo'lgan tushunchalar aks etadi. Ko'p tilli axborot-qidiruv tezaurusining asosiy tuzilish birligi sifatida ularning doimo teng qiymatda ekanligini aks ettirish uchun foydalaniladigan vositalar bilan bog'liq bo'lgan bir tilli, o'zaro teng deskriptorlardan yig'ilgan va tashkil etuvchi deskriptorning o'zini qabul qilish mumkin.

¹ Kurbonova F. Holat fe'llari tezaurusi. -MD.-T., 2014.

Kompyuter lingvistikasida rus va ingliz tillari bo'yicha ikki tili tezaurusni yaratish ISO 5964-1985 Xalqaro standartining asosiy me'yoriy nizomlarini inobatga olgan holda ishlab chiqilgan GOST 7.24-2007 Davlatlararo standarti talablariga asosan bajariladi va u ko'p tili axborot-qidiruv tezaurusini qurishning asosiy talablari va tuzilmalarini, tarkibini belgilab beradi.

Tezaurus atamalari manbasini aks ettirish

Tezaurusga kiritilgan atamalarning muhimligini tasdiqlash va tezaurusdan foydalanish tajribasi bilan tanishtirish maqsadida har bir termin uchun uning manbalari, ya'ni ushbu termin uchraydigan yoki ifodalanadigan matnli hujjatlar yoki matnli hujjatlar to'plami bilan bog'lanish belgilanadi. Buning uchun ikki munosabat xizmat qiladi: **manba** – matnlar to'plami bo'lsa, terminning manbada takrorlanish darajasini ko'rsatishi mumkin bo'lgan "...da uchraydi" bog'lanishi va ushbu atama manbaning predmetga oid yo'riqnomasida yoki glossariy (biror tekstdagi notanish, tushunilishi qiyin so'zlar lug'ati)da uchrashi bu atamaning muhimligini ko'rsatuvchi unga bo'lgan ishonchning oshishini belgilovchi "hujjat qismlarida uchraydi" bog'lanishidir. Izoh-ta'riflar bilan ta'minlangan deskriptor-atamalar manba bilan "...da ta'rif beriladi" munosabati yordamida bog'lanadi.

Tezaurusda manbalar nomlanishi, bibliografik izoh, til, tur (kitob, monografiya, ilmiy maqola, hujjatlar to'plami, darslik, lug'at, tezaurus, internet-manbalar, matnlar to'plami va b.), internet tarmog'idagi manzil va qisqa ta'rif, matnlar to'plami uchun esa qo'shimcha ravishda matnlar soni va so'zlar qo'llanishi namoyon bo'ladi.

Tezaurusni elektron manba sifatida ifodalash maqsadida ma'lumotlarning ikki darajali ombori (saqlash joyi) ishlab chiqiladi, uni ishlab chiqish va ilova qilish sifatida esa – tezaurus muharriri yaratiladi.

Birinchi darajada terminlararo belgilanadigan tezaurus maqolalari sxemasini, turlarini va munosabatlarining xususiyatlarini, shuningdek, atamalar manbalarining ta'riflari va ularning izohlarini belgilovchi tezaurus tuzilmasi turadi. Tezaurus tuzilmasining yaratilishi tezaurus muharririning maxsus bo'limida amalga oshiriladi. Jumladan, bunda atamalar sinfi va atamalar manbalarining turlari, shuningdek, munosabatlar to'plami, ularning xususiyatlari belgilanadi. Vaholanki, munosabatlarning, nafaqat ularning argumentlari turini va munosabatlarning sonini cheklash vazifalarini ko'rsatish yo'li bilan amalga oshiriladigan tuzilmali xususiyatlari, balki rasmiy xususiyatlarini ham belgilash mumkin.

Atamalarini, ularning izohlarini va manbalarini yuklash, shuningdek, ular o'rtasida aloqa o'rnatishda tezaurus muharriri tilshunos-ekspertlar uchun qulay interfeys yaratib beradi. Shuni ta'kidlash kerakki, atamalar ta'riflarini, manbalar va ular o'rtasidagi bog'lanishlarni kiritish yoki tahrir qilish tugallanishi bilan tezaurus foydalanuvchilariga veb-interfeysi orqali yangi axborotlarni ko'rish imkoni yaratiladi.

Tezaurus muharriri veb-ilova sifatida amalga oshiriladi va ro'yxatga olingan iste'molchilarda internet orqali foydalanish imkoni mavjud bo'ladi. Taqsimlangan jamoa bo'lgan holda ishlab chiqishni ta'minlash maqsadida tezaurus muharririda turli darajadagi ekspertlar huquqini himoya qiluvchi vakillar saqlanadi. Bu mexanizmga ko'ra faqatgina eng yuqori daraja ekspertlari tezaurus tuzilmasini tahrir qilishlari mumkin, boshqa darajadagi ekspertlar esa faqatgina uning tarkibi - mazmunini (atamalarning ta'rifi va manbalarini) tahrir qilishlari mumkin. Bunda quyidagi cheklovlar mavjud: ikki ekspert (foydalanuvchi) bir vaqtning o'zida aynan bir maqolani (yoki manba ta'rifini, yoki atamani) tahrir qila olmaydi. Bundan tashqari, lug'at manbasini faqatgina muallifning o'zi tahrir qilishi mumkin, degan yondashuv ham mavjud. Agar ekspertlardan biri boshqa bironing atamasiga qo'shimcha yoki umuman o'zgartirish kiritmoqchi bo'lsa, muallif bilan kelishib olishi mumkin va buni elektron tezaurusda ilova sifatida keltirilgan maxsus forum orqali amalga oshirishi mumkin.

Tezaurus matnli hujjatlarni indekslash (kodlash) va qidirishda foydalanish uchun fan sohasining yaxlit va bir-biriga zid bo'lmagan tushunchalar tizimini ifodalashi lozim. Bu atamalar muharririga kiritilgan, tezaurus tushunchalari tizimining mantiqiy yaxlitligini ta'minlovchi va chiqarib beruvchi mexanizmlar tomonidan amalga oshiriladi. Ularning vazifasi tezaurus muharririda aksioma va cheklovlar sifatida ifodalangan tezaurus munosabatlari ta'riflariga asoslanadi.

Jumladan, ushbu xususiyatlar asosida tezaurus atamalari o'rtasidagi maqbul bog'lanishlar o'rnatish hodisasi yuz beradi, zaruriyat tug'ilganda esa ular avtomat tarzda qo'shimcha qilinishi yoki o'chirib tashlanishi mumkin. Bundan tashqari, ularning u yoki bu sinflarga mansubligiga ko'ra tezaurus atamalari o'rtasidagi u yoki bu bog'lanishlar mavjud bo'ladi va ularning soniga nisbatan bo'lgan cheklovlar tartibga solinadi. Masalan, agar sal avvalroqda ko'rilgan "Qara" munosabatiga qaytuvchi munosabat (sinonim) keltirilgan bo'lsa va mavjud munosabatlar cheklangan bo'lsa (ushbu turda har bir askriptor atama uchun faqat birgina bog'lanish bo'ladi), u holda **avtomatik tarjima** askriptori va **mashina yordamidagi**

tarjima deskriptori bilan bog‘langanda sinonim qaytuvchi bog‘lanishining yaratilishi sodir bo‘ladi (**mashina yordamida tarjima, avtomatik tarjima**), shuningdek, **avtomatik tarjima** askriptoriga boshqa deskriptorlar bilan “Qara” va “Sinonim” bog‘lanishlarini yuzaga keltirishda cheklanganlik kuzatiladi.

Elektron tezaurusdan foydalanishni osonlashtirish maqsadida foydalanuvchiga tezaurus tarkibidagi ma’lumotlarni tarmoq ichida o‘zaro bog‘langan obyektlar, ya’ni tezaurus elementlari – atamalar (deskriptorlar va askriptorlar), atamalar manbalarining ta’riflari va ularning tavsiflari sifatida aks ettiruvchi foydalanuvchi veb-interfeysi yaratildi. Atamalarning o‘ziga xos xususiyatlari to‘plami va ular o‘rtasida o‘rnatilgan bog‘lanishlari yuqorida ta’riflangan tezaurus tuzilmasiga muvofiqdir.

Tezaurusda qidiruv jarayoni foydalanuvchi uchun kerakli bo‘lgan atamalarni tanlash, ularning ta’riflarini (tezaurus maqolalarini) barcha tafsilotlari bilan birga ko‘rib chiqish, shuningdek, atama yoki uning ta’rifi uchrashi mumkin bo‘lgan manbalar (nashrlar va matnlar to‘plami) bilan tanishish imkoni mavjud.

Foydalanuvchi uni axborotning qaysi turi qiziqtirayotganini, ya’ni barcha atamalarni, deskriptorlarni, askriptorlarni yoki atamalar manbalarini ko‘rsatishi mumkin. Bunda unga tanlangan obyekt turi bo‘yicha tezaurusda mavjud bo‘lgan axborot to‘liq holda, ushbu obyektga yo‘llanmalar to‘plamiga ega bo‘lgan *html*-sahifa sifatida aks ettiriladi.

Muayyan obyekt va uning bog‘lanishlari haqidagi axborot ham *html* – sahifa sifatida ifodalanadi. Bunda ushbu obyekt bilan bog‘langan obyektlar sahifasida giperilova (giperssilka) ko‘rinishida aks ettiriladi va ular yordamida obyektlarning mayda tafsilotlarini ko‘rishga o‘tish mumkin.

Tezaurus bo‘ylab keyingi yo‘naltirish tezaurusning bir obyektidan ular o‘rtasida mavjud bo‘lgan tezaurus (atamalar o‘rtasida) yoki bibliografik (atamalar va manbalar o‘rtasidagi) munosabatlarni aks ettiruvchi bog‘lanishlarni inobatga olgan holda boshqa obyektlarga o‘tish jarayonini aks ettiradi.

Tezaurusni tuzishning muhim jihati atamalar, ya’ni tezaurusga kiritiladigan ma’lumotlarni saralash, shuningdek, sinonimik atamalar orasidan atama-deskriptorlarni tanlash uslubidir.

Kompyuter lingvistikasi bo‘yicha o‘zbek tiliga xos bo‘lgan tezaurusga kiritish uchun atamalarni tanlash kompyuter lingvistikasining yangi yo‘nalish sifatidagi o‘ziga xos xususiyatlari va uning Rossiyada rivojlanish holatiga aloqador masalalar bilan chambarchas bog‘langan. Bunda

umuman kompyuter lingvistikasini va rus kompyuter lingvistikasini tasvirlovchi quyidagi faktorlarni alohida ta'kidlash zarur:

- kompyuter lingvistikasining fanlararo tabiati;
- kompyuter lingvistikasining turli munosabatlar tizimi sifatidagi tavsifi;
- kompyuter lingvistikasi yo'nalishlari taraqqiyotining xilma-xilligi;
- rus kompyuter lingvistikasining ingliz kompyuter lingvistikasidan farqi (jumladan, rus kompyuter lingvistikasi yo'nalishlari yuzasidan bajarilishi lozim bo'lgan masalalarning mavjudligi).

Kompyuter lingvistikasi "sun'iy intellekt" tadqiqot yo'nalishining bir qismi hisoblanar edi. "Sun'iy intellekt va intellektual tizimning maxsus atamaları XX asrning 60 yillaridan boshlab shakllana boshladi. Soha shakllanishining birinchi bosqichi turli maktablar va mutaxassislar guruhining mavjudligi, sinonimik atamalar ko'pligi bilan ajralib turadi. 70-yillar o'rtasiga kelib sun'iy intellekt sohasida terminlar majmuyi qat'iylasha boshladi. Aksariyat mutaxassislar tomonidan tan olingan atamalar yuzaga keldi. Bu terminlarning aksariyati ingliz tilidan kirgan. Kompyuter lingvistikasiga oid asosiy atamalar majmuyi 80- yillarning birinchi yarmida to'liq shakllandi. Kompyuter lingvistikasining atamalar majmuyi alohida bo'limlarda birinchi bosqich xususiyatlarini saqlab qolgan (sinonimlarning ko'pligi, masalan, semantik munosabatlar bo'limida).

Sun'iy intellekt – bu oxirgi o'n yillikda turli fan sohalarida, jumladan, kompyuter lingvistikasida faol bo'lgan sohalardan biridir. Sun'iy intellekt ham fanlararo soha hisoblanadi, ammo bu parametr nuqtayi nazaridan qaraganda sun'iy intellekt va kompyuter lingvistikasi o'zaro farqlanadi: sun'iy intellekt fanlararo sohadir, chunki uning usullaridan turli fanlarda foydalaniladi, kompyuter lingvistikasi esa lingvistika (nutqni va matnlarni tadqiq qilish bilan bog'liq bo'limlarni), psixologiya, mantiq kabi turli fanlarni, sun'iy intellektning ayrim bo'limlarinigina qamrab oladi.

Rus tilidagi terminlarning asosiy manbasi sifatida 2000– 2010 yillarda bo'lib o'tgan "Dialog (Suhbat)" Xalqaro anjumanida taqdim etilgan ma'ruzalar matnlari to'plamini qayd etish mumkin. Bu to'plamga lingvistik modellar (morfologik va sintaktik tahlil) va statistik ko'rsatkichlar asosida ushbu fan sohasidagi so'zlar va so'z birikmalari, ya'ni fan sohasining terminlari sifatida qo'llanishi mumkin bo'lgan atamalarni tanlab olishga ko'mak beruvchi yo'l-yo'riqlar tatbiq qilingan. Bu ro'yxat kompyuter lingvistikasi sohasi ekspertlari tomonidan tuzilgan. Bunda ekspertlar nafaqat kompyuter lingvistikasi sohasidagi bilimlarga,

balki atamashunoslik haqidagi umumlingvistik tushunchalar va terminologik soʻzlarni shakllantirishga asoslanadi. Fan sohasining yuzaga kelish jaryonini eʼtiborga olgan holda yondashish darajalanuvchi sxemalar asosida shakllanishning umumiy metodikasi bilan muvofiqlashtiriladi. Agar bu lugʻat ingliz tili uchun tuziladigan boʻlsa, yaratilayotgan tezaurusning rus-ingliz tilida ekanligini eʼtiborga olib, kompyuter lingvistikasining ingliz tilidagi manbalaridan tarjimaga oid ekvivalentlari tanlab olinadi. Boshqa tarafdin esa, rus tilidagi kompyuter lingvistikasining ayrim boʻshliqlarini toʻldirish uchun bu qismlarga atamalar yigʻish va ularni qoʻllash, lugʻatga kiritish jarayonida yana ingliz tilidagi manbalarga tayanishga toʻgʻri keladi. Chunki ingliz tilidagi atamalar va ularning izohi yoritilgan manbalar doirasi anchagina kengdir. Oʻquv-tanishtiruv ixtisosligiga tegishli boʻlgan zamonaviy hamda nufuzli ingliz tilidagi kitob manbalarining fan koʻrsatkichlari va taniqli tovush analizatorlari hujjatlariga kiruvchi glossariylar shular jumlasidandir. Ushbu terminologik asosda parallel terminlarning ingliz-rus tilidagi lugʻatini yaratish imkoniyati tugʻilgan.

Koʻp sonli sinonimik qatorlardan asosiy atama – deskriptorni tanlab olish muammosi ham qiyin kechmoqda, chunki kundan kunga yangi tushunchalar va ularga mos keluvchi atamalar paydo boʻlib BORISHi va ularni oʻzlashtirish asosiy muammo boʻlib kelmoqda. Inglizlarda avtomatik tarjima sohasida *translation memory* tizimlarining paydo boʻlishi, ilmiy jamiyat tomonidan tan olinmagan edi va *tarjimaviy xotira* atamasi bunga qarama-qarshi qoʻyilgan, ammo tarjimon-amaliyotchilar tomonidan *tarjimalar xotirasi* tushunchasi keng qoʻllanildi va asosiy isteʼmoldagi tushuncha sifatida shakllandi (sinonimik qatorlari: *tarjimaviy xotira* – 8, *tarjimalar xotirasi* – 0, *tarjimalar arxivi* – 1, *tarjimalar toʻplovchisi* – 0, *tarjimalar toʻplami* – 0).

Kompyuter lingvistikasi ayrim yoʻnalishlarining rivojlanishi (masalan, *onlayn* tizimidagi avtomatik tarjima) natijasida atamalarning eski koʻrinishlari bilan ziddiyat yuz beradi. **Inion** va **LES**da *avtomatik tarjima* va *mashina yordamidagi tarjima* juftligida *avtomatik tarjima* asosiy hisoblanadi va unga deskriptor maqomi beriladi. Biroq “Dialog” toʻplamida tez-tez uchrashiga koʻra *mashina yordamidagi tarjima* ustun turadi: *mashina yordamidagi tarjima* –318 vs., *avtomatik tarjima* –58. “Vikipediya” va “Krugosvet (dunyo boʻylab)” internet-ensiklopediyasi, shuningdek, darsliklar ham mana shu anʼanaga asoslanadi. Yevropa assotsiatsiyasi (uyushmasi)ning mashina yordamidagi tarjima saytida ham qayd qilinganidek, mashina tarjimasi (*machine translation*) atamasi

eskirgan bo'lsa ham, butun soha uchun asosiy umumiy atama bo'lib qolmoqda.

Deskriptor atamalarni tanlab olishda nafaqat statistikaga, balki tilshunoslar ilmiy jamiyatida hozirgi kunga kelib yuzaga kelgan so'zlardan foydalanish an'alariga ham tayanmoq lozim. Kompyuter lingvistikasi bo'yicha rus-ingliz tezaurusida qo'llanadigan so'zlarning semantik xususiyatlarini ta'minlovchi vositalar mavjudligi tufayli ulardan har qanday til va fan sohalari uchun ko'p tilli tezauruslar tuzishda foydalanish mumkin.

Hozirgi kunda rus va ingliz tilshunosligida 1000 dan ortiq atamalarni, atamalar o'rtasidagi 3500 atrofidagi bog'lanishlarni, 120 dan ortiq atamalar manbalarini va ularning ta'riflarini qamrab oluvchi elektron tezaurus tarkibini to'ldirish bo'yicha faol ishlar olib borilmoqda. Bu boradagi tadqiqotlarni yaratish o'zbek tilshunosligi uchun ham dolzarb hisoblanadi¹.

Savol va topshiriqlar

1. Tezaurus lug'atlar tarkibi haqida fikr bildiring.
2. Tezaurus lug'atlarning turlarini ayting.
3. Ideografik, analogik (qiyoslash), assotsiativ (muvofiqlashgan) tezaurus lug'at turlarining xususiyatlarini yoriting.
4. Tezaurusdagi bog'lanish turlarini tushuntiring.
5. Tezaurus lug'atga namuna keltiring.
6. Tezaurus lug'atlarning ma'lumotlar bazasi sifatidagi ahamiyatini yoriting.
7. Tezaurus lug'at tarixi haqida bilganlaringizni ayting.

¹ *Po'latov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2007. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.*

IV.KOMPYUTER LINGVISTIKASI YO‘NALISHLARI VA METODLARI

11-§. KOMPYUTER LINGVISTIKASI YO‘NALISHLARI TAVSIFI

Reja:

1. Kompyuter lingvistikasining asosiy yo‘nalishlari.
2. Avtomatik tarjima yo‘nalishi tarixi. Tarjimon dasturlar.
3. Avtomatik tarjima yo‘nalishining zamonaviy holati va istiqboli.
4. Avtomatik tarjima uchun ma’lumotlar bazasini yaratish.

Asosiy tushunchalar: *matn tahriri, avtomatik tarjima, til o‘qitishni avtomatlashtirish, kompyuter lingvodidaktikasi, on-line tarjima tizimi, ekvivalent moslik, transformatsion moslik, tarjima protsedurasi, generatsiya, translyatsiya, bilingv baza, example-based translation (namunalarga asoslanadigan tarjima), rule-based translation (algoritmga asoslangan tarjima), translation memory (tarjimon xotirasi).*

Kompyuter lingvistikasining asosiy yo‘nalishlari:

1.Tabiiy tildagi matn analizi. 1954-yilda Jojtaun loyihasi doirasida rus tilidagi 60ta gapning ingliz tiliga tarjimasini amalga oshirildi. Olib borilgan izlanishlar natijasida har bir matn ortida bitta emas, uch darajaga ajratish mumkin bo‘lgan bir necha formal struktura yashiringan bo‘ladi:

1. Birinchi daraja – bu yuzaki sintaktik struktura. Bu strukturada matndagi har bir gap alohida ko‘rib chiqiladi va har bir gap alohida gap bo‘laklariga ajratiladi. Ega, kesim, aniqlovchi, to‘ldiruvchi, hol aniqlanadi.

2. Ikkinchi daraja – chuqur sintaktik struktura. Har bir matnda u yoki bu olamning manzarasi tasvirlanadi. Masalan, *Bola gulni uzdi. Bola tomonidan uzilgan gul.* Birinchi gapda subyekt – *bola*. Uning bosh kelishikda ekanligi ega sifatidagi morfologik ko‘rsatkichidir. Ikkinchi gapda ega vazifasini *gul* so‘zi bajarmoqda. Ammo subyekt *bolaligicha* qoladi, chunki harakat *bola* tomonidan sodir etilmoqda. *Gul* har ikki gapda ham harakat obyektini ko‘rsatadi.

3. Semantik daraja. Yuzaki semantik daraja chuqur semantik daraja bilan mustahkam bog‘liq. Gaplarning chuqur tahlili pragmatik hisoblagnadi.

Bormoq – Subyekt _____ Qayerga? _____ Qayerdan? Maqsad _____
Ushbu modelga mos tushuvchi fraza:

Bormoq – Subyekt: *Petr*. Qayerga ? *Do‘konga*. Maqsad: *sut uchun*.
Pyotr do‘konga sut uchun bordi.

2. Tabiiy tildagi matn sintezi. Matn sintezi faqatgina avtomatik tarjima tizimida ishlatilmaydi. XX asrning 70- yillarida sobiq ittifoqda yartailgan TALE (ertak) sehrli ertak matnlarining avtomatik tizimi ham sintezga asoslanadi. *Bir X bor edi. X da Y yo‘q edi. X xudodan so‘ray boshladi. Xudo unga rahm qildi. Y tug‘ildi. Y o‘sdi. Ketdi, X unga Zga BORISHga izn bermadi. Ammo Y Zga ketdi. X qaytdi. Y yo‘q. X tushundiki, Y Zga ketgan. X Yni izlab ketdi.* Tizim xotirasi bu o‘rinlarni to‘ldiruvchilarni joylashtiradi: *Bir podsho bor edi. Prodshoda voris yo‘q edi. Podsho xudodan so‘ray boshladi. Xudo unga rahm qildi. Voris tug‘ildi. Voris o‘sdi. Ketdi, podsho unga o‘zga yurtga BORISHga izn bermadi. Ammo voris o‘zga yurtga ketdi. Podsho qaytdi. Voris – yo‘q. Podsho tushundiki, voris o‘zga yurtga ketgan. Podsho vorisni izlab ketdi.*

3. Matnni tushunish. Tabiiy tillarda matnni tushunish muammosi nafaqat lingvistik aspektni qamrab oladi, balki psixologiya, falsafa, semiotika doirasida halq ilinuvchi masalalarni ham qamrab oladi. Matn analizi matn mazmunini tushunish vositasi hisoblanadi.

4. Matn mazmunini tasavvurda jonlantirish – matn voqeligini ongda qayta tiklashdir. Inson uchun ikki modallik – simvulli (matniy) va ko‘rish, anglash muhim sanaladi. Nimaningdir nomi yoki qandaydir vaziyatning matndagi tavsifi bu obyekt va vaziyatlar haqida tasavvur uyg‘otadi. O‘z navbatida biz ko‘rganlarimizni tabiiy tilimiz yordamida tasvirlab bera olamiz. Matn va unga muvofiq keluvchi tasavvur kartinasi bizning ongimizda birlashgan holda namoyon bo‘ladi.

5. Kommunikatsiya modellari. Inson nutqini qabul qiluvchi va tushunuvchi sun‘iy tizim inson va kompyuter muloqotiga zamin yaratdi.¹ Axborot uslubi, kompyuter modellari, turli simvollar, kodlash va dekodlash kommunikatsiya vositalariga aylandi.²

Kompyuter lingvistikasi amaliy tilshunoslikning tarkibiy qismi bo‘lib, u hozirda turli aspektlarda jadal rivojlanmoqda. A.Po‘latov, S.Muhamedova, A.Rahimov o‘quv qo‘llanmalarida kompyuter lingvistikasining quyidagi yo‘nalishlari ham qayd etilgan:

1) avtomatik tarjima – berilgan matnni bir tildan boshqa tilga qisqa vaqt ichida tezkor tarjima qilish tizimi;

2) matn tahriri – kompyuter yordamida muayyan matnni tahrir qilish, xatolarini aniqlash va tuzatish;

3) CALL (Computer Assisted Language Learning) – til o‘qitishni avtomatlashtirish – xorijiy tillarni kompyuter yordamida o‘qitish tizimini

¹ Боярский К.К. Введение в компьютерную лингвистику. -Санкт-Петербург, 2013. -С.5-9.

² Боярский К.К. Введение в компьютерную лингвистику. -Санкт-Петербург, 2013. -С.5-9.

ta'minlovchi amaliy dasturlar hamda bilimlarni avtomatik baholash mexanizmlari;

4) kompyuter leksikografiyasi (elektron lug'atlar tuzish) – muayyan til leksikonidagi so'zlarning ma'lumotlar bazasi va maxsus dasturlash tillari yordamida ishlash tizimiga ega bo'lgan dasturlar (kompyuter lug'atlari) yaratish bilan shug'ullanuvchi soha;

5) **sun'iy intellekt tizimini yaratish;**

6) **NLP** – tabiiy tilni qayta ishlash;

7) **gipertekst texnologiyasi;**

8) **on-line o'qitish tizimi** (virtual sinf va laboratoriyalarda) – kompyuter lingvodidaktikasi;

9) **informatsion qidiruv tizimi.**

Informatsiyaning geometrik progressiya asosida keskin ko'payishi ularning 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 boshladi. Mashina tarjimasiga oid izlanishlar kompyuter lingvistikasi fanining yuzaga kelishida asosiy manba hisoblanadi.

D.O'rinboyev kompyuter lingvistikasi yo'nalishlarini quyidagicha guruhlagan: "Kompyuter lingvistikasi fani yangi soha bo'lsa-da, hozirgi kunga qadar, qisqa vaqt ichida, turli yo'nalishlari vujudga keldi. Tabiiy tillar negizida mashina tili yaratish, foydalanuvchining talab va ehtiyojlari kengayishi ushbu yo'nalishlarni ko'paytirdi. Kompyuter lingvistikasi fanining yo'nalishlari shartli ravishda quyidagicha guruhlanadi:

1. Qoliplash va dasturlash ishlariga qaratilgan yo'nalishlar:

- muloqotni kompyuter yordamida qoliplash;
- syujet strukturasi kompyuter yordamida qoliplash;
- matnni katta hajmdagi matn (gipertekst) darajasida tasavvur etish;
- axborot – izlanish sistemasi.

2. Til o'rgatishga qaratilgan yo'nalishlar:

- til o'rgatish jarayonlarini kompyuterlashtirish;
- mashina tarjimasi.

3. Matnni tahlil etishga qaratilgan yo'nalishlar:

- avtomatik tahrirlash;
- statistik tadqiqotlar;
- kompyuter leksikografiyasi.

Kompyuter lingvistikasining ushbu yo'nalishlarini amaliy va nazariy ahamiyati jihatidan quyidagicha guruhlash mumkin:

1. Amaliy yo'nalishlar:

- * Mashina tarjimasi.
- * Avtomatik tahrir.
- * Til o‘rgatish tizimini kompyuterlashtirish.
- * Statistika tadqiqotlar.
- * Kompyuter leksikografiyasi.

2. Nazariy yo‘nalishlar:

- * Tabiiy tilni qayta ishlash.
- * Syujet strukturasi modelashtirish.
- * Muloqotni kompyuter yordamida qoliplash.
- * Matnni gipertekst texnologiyalarida berish.

Kompyuter lingvistikasi quyidagi yo‘nalishlarni ham o‘z ichiga oladi:

Tabiiy tillarni qayta ishlash (inglizcha. natural language processing; matnning semantik, morfologik, sintaktik tahlili). Bunga quyidagilar kiradi:

1. Lingvistik korpus, ya’ni matnning elektron korpusini yaratish va foydalanish.

2. Elektron, tezaurus lug‘atlar, ontologiyalar yaratish. Masalan, Lingvo lug‘atlardan avtomatik tarjima va orfografik tekshirish uchun foydalaniladi.

3. Matnning avtomatik tarjimasi. Rus tarjimon dasturlari orasida eng mashhuri “Promt” hisoblanadi.

4. Matndan statistikani avtomatik ravishda ajratish (inglizcha, fact extraction, text mining).

5. Matnni avtomatik jamlash (inglizcha, automatic text summarization). Bu vazifa Microsoft Wordga kiritilgan.

6. Bilimlarni boshqarish tizimini barpo qilish.

7. Savol-javob tizimlarini yaratish (inglizcha, question answering systems).

- Belgilarni optik aniqlash.
- Nutqni avtomatik aniqlash.
- Nutqning avtomatik sintezi.

Shuningdek, kompyuter lingvistikasi markazida quyidagi masalalar turadi:

- axborot yetkazish uchun umumiy til yaratish;
- axborot yetkazish uchun tillarni takomillashtirish;
- inson va mashina aloqasi uchun tabiiy til interfeysini yaratish (sun’iy intellektning asosiy masalalaridan biri);

- axborot tizimini yaratish va modellashtirish”.¹

12-§. MASHINA TARJIMASI

Reja:

- 1.Mashina tarjimasi – kompyuter lingvistikasi yo‘nalishi sifatida.
- 2.Avtomatik tarjima tizimlari.
- 3.Mashina tarjimasiga doir tadqiqotlar tavsifi.

Asosiy tushunchalar: *avtomatik tarjima turlari, avtomatik tarjima rejimi, tarjima xotirasi, lingvistik tahlilga asoslangan mashina tarjimasi, avtomatik tarjimaning struktural tizimi, avtomatik tarjimaning lingvistik muammolari, SOKRAT sistemasi, YEUROTRA sitemasi, LINGVO sistemasi, DIANA, mashina tarjimasi strategiyasi.*

Mashina tarjimasi yoki avtomatik tarjima – muayyan so‘z yoki matnlarning kompyuter yordamida bir tildan boshqa tilga qilingan tarjimasi. Avtomatik tarjima (mashina tarjimasi) kompyuter lingvistikasining asosiy yo‘nalishlaridan bo‘lib, avtomatik tarjima tizimi prinsip va metodlari haqidagi bilimlarni hosil qilish, lingvistik ta‘minot yaratish, mavjud tizimlar va lug‘at yaratish tarixi, til haqidagi bilimlar, ko‘nikmalarni shakllantirish, mashina tarjimasi tizimining lingvistik protsessorini shakllantirishda lingvistik tahlilni algoritmlash tizimini amalga oshiradi. Avtomatik tarjima oldiga quyidagi vazifalar qo‘yiladi:

- avtomatik tarjima tizimini yaratish va undan foydalanishning asosiy prinsiplarini o‘rgatish;

- tilning nazariy asoslari (formalizatsiya) metodlari va vositalari bilan tanishtirish;

- avtomatik tarjima tizimi lingvistik protsessorni yaratishda so‘z va uning ma‘nosi, gap strukturasi, semantikasi va matn tahlilida lingvistik algoritm bosqichlarini o‘rgatish.

Avtomatik tarjima quyidagi masalalarni qamrab oladi:

1.Avtomatik tarjima – sun‘iy intellekt yo‘nalishi.

2.Avtomatik tarjima turlari qo‘l tarjimasi, avtomatik tarjima.. Qo‘l tarjimasining turlari, shakllari, tiplari. Avtomatik tarjima strategiyasi: so‘zma-so‘z tarjima, transfer yordamidagi tarjima, interlingvo yordamidagi, lingvistik tahlilga asoslangan batafsil tarjima, tarjima xotirasi.

¹Ўринбоев Д.Б. Компьютер лингвистикаси (Маърузалар матни). –Самарқанд, 2009, 5-бет; L.Abdusamidova. Компьютер лингвистикасининг автоматик таҳрир yo‘nalishi. BMI.-Т., 2016.

3.Lingvistik tahlilga asoslangan mashina tarjimasini algoritmi. Matnning morfologik, sintaktik, semantik tahlili. Tarjima qilayotgan tilda matn sintezi: matnning semantik ko‘rinishi, sintaktik strukturalar qurilishi, morfologik sintez.

4.Avtomatik tarjimaning struktural tizimi. Mashina tarjimasining lingvistik bilimlar bazasi strukturali: lingvistik ma’lumotlar bazasi va lingvistik protsessor. Lingvistik ma’lumotlar bazasi komponentlari. Lingvistik protsessor strukturali (tuzilishi).

5.Tarjima xotirasi. Misollarga asoslangan mashina tarjimasini tizimlari. Tarjima xotirasi tizimi uchun matnlar korpusi (to‘plami) qurilishini shakllantirish prinsiplari. Tizim strukturali (tuzilishi) va ishlash algoritmi. Tarjima xotirasining qo‘llanish sohalari. Mashina tarjimasini tizimlarida tarjima xotirasining qo‘llanish sohalari .

6.Avtomatik tarjimaning boshqa bilim sohalari bilan bog‘liqligi. Sun’iy intellekt tadqiqida mashina tarjimasini va kompyuter lingvistikasi. Mashina tarjimasining matematik lingvistik va psixolingvistik bilan bog‘liqligi.

7.Avtomatik tarjimaning lingvistik muammolari. Morfologik va semantik doiradagi muammolar. O‘zakdosh va o‘zakdosh bo‘lmagan tillarning formal va mazmuniy nomutanosibliigi. Muammolar yechimi yo‘llari. Mashina tarjimasining semantik va pragmatik muammolari. Polisemiya masalasini hal qilishda kontekst va ekstralingvistik faktlarni hisobga olish. Avtomatik tarjima sifatini oshirishda va semantik tahlil borasida qiyinchiliklarni hal etish yo‘llari.

Avtomatik tarjima “Kompyuter lingvistikasi”, “Kvantitativ lingvistikasi”, “Bilim injeneriyasi”, tabiiy tilning avtomatik ishlab chiqilishi, sun’iy intellekt muammolari bilan bog‘liq. Bu sohalarning kompleks rivoji sun’iy intellekt taraqqiyoti haqida to‘liq tasavvur hosil qilishga yordam beradi.

Avtomatik tarjima til o‘rgatishning zamonaviy yo‘nalishlari va lingvistik tahlilni informatsion texnologiyalar yordamida algoritmlash, bilimlarni tatbiq qilish, bilimlar bazasi, lingvistik bilimlarni yaratish prinsiplari bilan chuqurroq tanishishga imkon yaratadi.

Avtomatik tarjima Yevropada alohida fan sifatida o‘qitiladi.

Avtomatik tarjima bo‘yicha dastlabki g‘oyalar ingliz olimi Charlz Bebbidj tomonidan 1836-1848 yillarda ilgari surilgan. Uning fikricha, mexanik-elektron mashinalar kodlashtirilgan avtomatik tarjimaning amalga oshirishi mumkin bo‘ladi. Uning loyihasi bo‘yicha 50mingta so‘z xotira kuchiga ega bo‘lgan elektron mashinalar 100000 ta so‘zni avtomatik

ravishda tarjima qilishi zarur bo‘ladi.¹ Bebbidj g‘oyalari 100 yildan keyin amalga oshdi.

1949-yilda Ch.Bebbidjing g‘oyalariidan ta’sirlangan amerikalik olim Uorren Uiver mashina tarjimasini yaratishning nazariy asoslarini ishlab chiqdi hamda lingvistika, deshifrovka, programmashtirish nazariyasi bo‘yicha mutaxassislarni hamkorlikka chorladi.²

Shundan so‘ng AQSHda MT bo‘yicha laboratoriyalar, hamkorlik loyihalari yaratildi. Mashina tarjimasini bo‘yicha birinchi xalqaro konferensiya 1952-yilda Massachusetts texnologiya institutida o‘tkazildi.

1954-yilda AQSHdagi Jorjtaun universitetida dunyoda birinchi marta avtomatik tarjima amalga oshirildi. U GAT tizimi (inglizcha «*Georgetown Automatic Translation*») birikmasining qisqartmasi) deb ataldi. Bu tajribalar IBM-701 kompyuterida o‘tkazildi. Bu tajribaning bazasi algoritmlardan iborat bo‘lib, 50 000 ta so‘z (60 ta gap)ni rus tilidan ingliz tiliga tarjima qildi.³

XX asrning 70-yillarida LIFER (Language Interface Facility with Elipsis and Recursion) ingliz tilidagi dastlabki ma’lumotlar bazasi yaratildi. Undan keyin kompyuter bozorida qayta ishlangan tabiiy tilga asoslanuvchi tizimlar yuzaga keldi. XX asrning 80-yillarida AQSHda tabiiy til interfeysiga asoslangan dasturlarni ishlab chiqish va sotish bilan shug‘ullanuvchi bir qator kompaniyalar shakllandi. 1985-yilda «Semantek» korporatsiyasi, «Karnegi Gruppya» paketlari yaratildi. Avtomatik tarjima tizimini yaratish bo‘yicha faol ish olib borildi.

Yevropada kompyuter tarjima tizimini yaratish Европейский информационный сеть (EURONET DIANA)ning yuzaga kelishi bilan me’yoriy tarzda amalga oshirila boshladi.

Hozirgi kunda dunyoda keng tarqalgan jahon tillari uchun avtomatik tarjima dasturlari ishlab chiqilgan. Jumladan, ruschadan inglizchaga, inglizchadan ruschaga, ruschadan fransuzchaga yoki aksincha tarzdagi dasturlar yaratilgan.

Bugungi kunda avtomatik tarjimaning quyidagi sistemalari mavjud:

SYSTRAN avtomatik tarjima sistemasi– AQSH harbiy havo kuchlari buyurtmasi bo‘yicha D.Tom rahbarligida ishlab chiqilgan. 1974–1975 yillarda bu tizimdan NASA aerokosmik assotsiatsiyasi tomonidan

¹ Po‘latov A., Muhamedova S. Kompyuter lingvistikasi. – T., 2007. -B.52.

² Баранов А.Н. Введение в прикладную лингвистику. - М.: Эдиториал УРСС, 2001.-С.89. Qarang: Po‘latov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2007. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

³ Убин И.И. Современные средства автоматизации перевода: надежды, разочарования и реальность // Перевод в современном мире. - М.: ВЦП, 2001. -С.60-69.

Apollo – Soyuz loyihasi hujjatlarini tarjima qilishda foydalanildi. 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.

YEUROTRA sistemasi – Yevropa olimlari tomonidan ishlab chiqilgan. 1982-yilda “Yevropa iqtisodiy hamkorligi” (Европейское экономическое сообщество) barcha Yevropa tillari uchun mo‘ljallangan EUROTRA avtomatik tarjima dasturi yaratilganligini ma’lum qildi. Dastavval, loyiha 12 mln dollarga baholandi, 1987 -yilda mutaxassislar bu loyihaga 160 mln dollardan ko‘p mablag‘ sarflash kerakligini aniqladilar. Sistema 40 mlnta so‘zdan iborat xotira kuchiga ega. Bu sistema barcha Yevropa tillari bo‘yicha avtomatik tarjimani amalga oshira oladi.

SOKRAT sistemasi – 40 mln ta so‘zdan iborat xotira kuchiga ega. Dunyodagi 40 ta tilda avtomatik tarjimani amalga oshiradi. Windows operatsion tizimi versiyasi mavjud bo‘lgan tillarda shu tizim asosida avtomatik tarjimani amalga oshirish mumkin bo‘ladi

LINGVO sistemasi (lingvo kompyuter lug‘ati) – bu lug‘atga ham dunyoning 40 tilidagi ma’lumotlar jamlangan.

DIANA – avtomatik tarjima bo‘yicha dunyodagi eng yirik loyiha. 1991-yilda AQShda NASA mutaxassislari tomonidan yaratilgan. Uning hajmi 10 mlnta so‘zni tashkil etadi. U dunyoning 6 ta yirik tiliga – ingliz, rus, fransuz, nemis, arab, ispan tillariga asoslanadi:

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.

Tarjimon dasturlar ikki xil ta’minotga asoslanib ishlaydi: 1) lingvistik ta’minot – bunda mavjud tillarning umumiy leksikoni kiritiladi, uning grammatikasi joylashtiriladi, bir so‘z bilan aytganda, ma’lumotlar bazasi yaratiladi; 2) dasturiy ta’minot – bunda bir tildan ikkinchi tilga tarjima qilish algoritmi va kompyuter analizi ishlab chiqiladi (maxsus programmalashtirilgan tillar yordamida).

Avtomatik tarjimaning ishlash algoritmi quyidagicha bo‘ladi:

1. Boshlanish – dasturlar paneli va menyu satri ko‘rinadi.
2. Ishlash rejimi tanlanadi (tarjimaning yo‘nalishi – inglizchadan ruschaga yoki ruschadan inglizchaga).

Jumla kiritiladi va xotiradagi kodi qidirib topiladi.

3. Topilgan kod maxsus dasturlar yordamida qayta ishlanadi va tarjima oynada paydo bo‘ladi.

4. Joriy rejimda ishlashni davom ettirish haqida so‘raladi.

5. Dasturni yopish haqida savol so‘raladi.

6. Dastur yopiladi.

Avtomatik tarjimada samaradorlikning yuqori bo‘lishiga tabiiy tillarning ayrim jihatlari – polisemiya (ko‘p ma’nodlilik), omonimlik, frazeologizmlar, idiomalar, paremlar, troplar to‘sqinlik qiladi. Shularni istisno qilganda, hozirda avtomatik tarjima yo‘nalishida yuqori samaradorlikka erishilgan. Hatto hozirda faqat avtomatik tarjima yo‘nalishiga bag‘ishlangan saytlar ham yaratilgan. Bu avtomatik tarjimaning keng qamrovda taraqqiy etayotgan yo‘nalishlardan ekanligini ko‘rsatib turibdi.

Hozirgi kunda inglizcha-o‘zbekcha va o‘zbekcha-inglizcha kompyuter tarjimasini dasturlari juda katta ahamiyatga ega. Ma’lumki, ingliz va o‘zbek tillari leksik-grammatik xususiyatlariga ko‘ra bir-biridan tubdan farq qiladi. O‘zbek tilida gap qurilishi SOV (ega+to‘ld+kes) qolipida bo‘lsa, ingliz tilida SVO (ega+kes+to‘ld) qolipidadir; o‘zbek tili agglutinativ tillar sirasiga kiradi, ingliz tili esa flektiv til hisoblanadi. Shuning uchun ingliz tilidan o‘zbekchaga va o‘zbekchadan inglizchaga kompyuter tarjimasini yaratish o‘ziga xos qiyinchiliklarni keltirib chiqaradi. Bugungi kunda rus tili vositasida ingliz tilidan o‘zbek tiliga avtomatik tarjima qiluvchi dasturlarning versiyalari e’lon qilingan.

Hozirda tadqiqotchilar tomonidan taklif qilinayotgan ingliz tilidan o‘zbek tiliga kompyuterda tarjima qilish dasturi vositachi tilsiz amalga oshiriladi. Ta’kidlash lozimki, mazkur ish gaplarning sintaktik analiz qilish algoritmlari va formal modellari bazasi asosida amalga oshiriladi.

Ingliz tilidan o‘zbek tiliga kompyuterda tarjima qilish algoritmi quyidagi vazifalarni hal etish uchun mo‘ljallangan:

I.I. Analiz – bunda ingliz tilidagi gap soddalashtirilgan model ramkasida sintaktik tahlil qilinadi.

1.1. Ushbu model faqat sodda gaplarni qamrab oladi.

2. Gapning har bir bo‘lagi bitta so‘zdan iborat bo‘ladi.

3. Gaplarda aniqlovchilar bo‘lmaydi.

4. Gaplarning standart tiplari ko‘rib chiqiladi (darak gap (ega+kes+to‘ld+hol), so‘roq, inkor va so‘roq-inkor gaplar).

5. Fe‘lning quyidagi tuslanishli shakllarini qamrab oluvchi kesimli gaplar ko‘rib chiqiladi:

a) shaxs (I, II, III shaxs);

- b) son (birlik va ko'plik);
- v) zamon (o'tgan, hozirgi, kelasi);
- g) harakat tipiga ko'ra (Simple)
- d) maylga ko'ra (Indicative)
- e) nisbatga ko'ra (Active)

II. Tarjimada gaplar ingliz tilidan o'zbek tiliga o'giriladi. Algoritm quyidagi bosqichlardan tashkil topadi:

- 1) gap kiritiladi;
- 2) gapning har bir so'zi massivining elementlariga qo'shiladi;
- 3) massivining elementlari yordamida lug'at elementlari bilan taqqoslanadi, bu lug'atda olmosh, ko'makchilar, ko'makchi va modal fe'llar, artikllar va noto'g'ri fe'llar ro'yxati mavjud bo'ladi;
- 4) agar so'zlar yordamchi lug'atda topilmasa, unda taqqoslash maxsus lug'at yordamida davom ettiriladi;
- 5) topilgan so'zlar yordamchi lug'atga beriladi, bu yerda so'zga ushbu so'zni va uning tarjimasini saqlovchi kod beriladi;
- 6) bunday so'z lug'atlarda mavjud bo'lmasa, so'z shakl yasovchi affiks-lardan ajratib olinadi va 5-ish bajariladi;
- 7) agar so'zlar yordamchi va maxsus lug'atlardan topilmasa, ushbu so'zning yo'qligi haqida ma'lumot kiritiladi;
- 8) gap ikki guruhga bo'linadi: kesimgacha bo'lgan so'zlar ega guruhiga kiradi (ega guruhi);
- 9) kesim va unga bog'langan so'zlar (kesim, toidiruvchi, hol) kesim guruhi hisoblanadi;
- 10) kesim guruhidan kesim ajratib olinadi;
- 11) so'ng to'ldiruvchi ajratiladi;
- 12) gapning qolgan qismi hol hisoblanadi;
- 13) gapning har bir bo'lagi shakl yasovchi qo'shimchalarsiz tarjima qilinadi;
- 14) tarjima qilingan gap bo'aklaridan o'zbek tilidagi gap tuziladi, u, albatta, ingliz tilidagi gap konstruksiyasiga mutanosib bo'ladi;
- 15) o'zbek tilidagi so'zlarga ingliz tilidagi so'zlarga mutanosib ravishda affiks va qo'shimchalar qo'yib chiqiladi;
- 16) tarjima chiqarib beriladi («tarjima» rejimida);
- 17) analiz chiqarib beriladi («analiz» rejimida).

Avtomatik tarjima texnika taraqqiyoti mahsuli, qo'l mehnatini osonlashtirishga qaratilgan, vaqtni tejaydigan tezkor tizimdir. Avtomatik tarjimaning fan-texnika, ishlab chiqarishdagi ahamiyatli o'rnini ta'kidlash

joiz. Shuning barobarida, avtomatik tarjimada ayrim nuqsonlar ham kuzatiladi.

Avtomatik tarjimaning cheklangan tomonlari:

1. Soʻzlar *avtomatik tarjima* jarayonida toʻgʻridan-toʻgʻri tarjima qilinadi. Shuning uchun tarjima grammatik qoidalarga, deyarli, rioya qilinmagan holda amalga oshiriladi

2. *Avtomatik tarjima* qilingan matnni uslub jihatdan qayta koʻrib chiqish zarur boʻladi.

3. Hozirgi mavjud avtomatik tarjima qiluvchi tizimlar ayrim hollarda soʻzlar birikuvini ham, deyarli, notoʻgʻri tarjima qiladi.

4. Avtomatik tarjimada polisemiya, omonimiya, troplar, frazeologizmlar, idiomalarni toʻgʻri oʻgirish imkoniyati yoʻq.

Masalan, *The bat looks like a mouse with wings* gapini rus va oʻzbek tiliga tarjima qilishda har bir soʻzning polisemantik va omonimlik kabi semantik qirralarini eʼtiborga olish zarur. Aks holda tarjima muvaffaqiyatli chiqmaydi. Jumladagi *bat* soʻzi «koʻrshapalak», «belkurak», «toʻqmoq», «kichik ziyofat» kabi maʼnolarni, *look* soʻzi esa «qaramoq», «nigoh», «koʻrinish» maʼnolarini anglatadi. Bunda mazkur soʻzlar rus yoki oʻzbek tiliga tarjima qilinayotganda mos ekvivalentini tanlash uchun kontekst va uning distributsiyasidan (qurshovidan) kelib chiqish zarur. Ushbu holat, ayniqsa, idiomatik ifodalar qatnashgan matnlar tarjimasida murakkablik tugʻdiradi. Masalan, *You are an apple of my eyes* jumlasida «Siz koʻzlarimning olmasisiz» tarzida soʻzma-soʻz tarjima qilinadi, chunki «apple» soʻzining denotativ maʼnosi «olma»dir. Lekin tarjimaga kontekstdan kelib chiqib yondashilsagina, mazmun toʻgʻri ifodalanadi. Bunda jumlaning «an apple of my eyes» qismi idioma ekanligi nazarda tutiladi va «Siz koʻzimning oq-u qorasisiz» tarzidagi gap hosil boʻladi.

Tarjimada turli til birliklari oʻrtasidagi moslik muhimdir. Agar leksik birliklar semantik jihatdan mos kelmasa, tarjimadan koʻzlangan maqsad amalga oshmaydi. Tarjimadagi moslik 3 xil boʻladi:

1) ekvivalent moslik – bunda bir tildagi soʻz boshqa bir tildagi soʻzga tarjimada toʻla muvofiq keladi. Masalan, rus tilidagi «книга» soʻziga inglizcha «book» soʻzi toʻla mos keladi;

2) variantli moslik – bunda sinonimiya va polisemiya asosida soʻzning maʼnosiga mos keluvchi variantlar ham chiqariladi;

3) transformatsion moslik – bu asosan birikmali ifodalar tarjimasida ishlatiladi, bunda mazmunga taʼsir koʻrsatmagan holda shakliy oʻzgarishlar 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 har xil, 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 tizimini qat’iy belgilash asosida mexanik tarjimadagi muammolarni bartaraf etish mumkin. Bugungi kunda dunyo olimlari tomonidan mavjud kamchiliklarni bartaraf qilish borasidagi takliflarni yorituvchi maxsus jurnal nashr qilingan.¹

Eng sifatli tarjima insonning aralashuvi bilan yuzaga chiqadi. Chunki tarjimon inson tarjimaga tabiiy tilining nozik qirralarini, semantik, grammatik va pragmatik jihatlarini nazarda tutgan holda yondashadi. Bunda yondashuv quyidagicha bo‘lishi lozim: mexanik tarjima yordamida tanlangan matnning tarjima qilingan dastlabki versiyasi tahrir etiladi, 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 parametrlar tekshiriladi.

Ixtisoslik bo‘yicha to‘plangan bilimlar maktab va oliy ta’limda chet tilini yangi informatsion texnologiyalarni qo‘llagan holda o‘qitish hamda mashina tarjimasi tizimini yaratish, ma’lumotlarni lingvistik banerlarga joylashtirish va ilmiy ishlarda shu ma’lumotlar banerini qo‘llashga zamin yaratadi.

Rus olimi Y.I.Shemakin «Начало компьютерной лингвистики» kitobida avtomatik tarjimaning ikki 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 u tarjima mosligiga asoslangan.

Birinchi yondashuvda tilning semantikasiga maksimal darajada e’tibor bergan holda avtomatik sifatli tarjimaga erishishga harakat qilingan. Tizimning lingvistik asosi «*matndan uning mazmuniga qarab borish*» («*from surface structure to deep structure*») hisoblanadi. Bunda tarjima

¹Qarang: Po‘latov A., Muhamedova S. Kompyuter lingvistikasi. – T., 2007; Rahimov A. Kompyuter lingvistikasi asoslari. –Toshkent, 2011.

tilning morfologik, sintaktik va semantik tahlili, lugʻatlar, grammatik qoidalar, matnlar korpusiga asoslanib amalga oshiriladi.

Ikkinchi yondashuvda matn ortiqchaliklarga ega koʻp bosqichli tizim sifatida baholanadi, bunga koʻra matnning yuqori sathida koʻproq informatsiya tashuvchi leksik birliklar, quyi sathida esa kam informatsiya beruvchi birliklar joylashadi. Bunda soʻzma-soʻz tarjima amalga oshirilib, ikki tildagi gaplar strukturasi yagona superstrukturaga birlashtiriladi. Bunda avtomatik tarjima modeli ikki holatda boʻladi: a) *generatsiya holati* tarjimaning ishlashini – lugʻatlarni jamlash, analiz algoritmlari, aniq bir maqsadga yoʻnaltirilgan sintezni taʼminlaydi; b) *translyatsiya holati* sistemaning mustaqil ravishda tarjima ustida ishlashini taʼminlaydi.

Ikkinchi yondashuvga koʻra, avtomatik tarjima quyidagi protseduralar asosida ishlaydi:

1-protsedura – qidiruv tizimi bilan bogʻliq boʻlib, bunda lugʻat bazasidan soʻzlar, birikmalar, iboralar izlanadi, soʻz-shakllarning morfologik analizi amalga oshiriladi.

2-protsedura – soʻzlarning sintaktik vazifasi, grammatik kategoriyalar -zamon, shaxs-son kabilar aniqlanadi.

3-protsedura – bir yoki koʻp ma/noli soʻzlar tarjimasi bayon qilinadi.

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

5-protsedura – transformatsion oʻzgarishlar (matn yoki jumla strukturasi uning mazmuniga jiddiy taʼsir etmagan holda oʻzgartirish; masalan, *Kompyuter yoqildi va musiqa yangradi* jumlasini «*Kompyuter yoqilgach / kompyuter yoqilishi bilan / kompyuter yoqilib musiqa yangradi*» tarzida oʻzgartirish) hosil qilinadi.

6-protsedura yakuniy bosqich boʻlib, bunda murakkab grammatik formalarning sintezi amalga oshiriladi.

Mashina tarjimasi strategiyasi sekin-astalik bilan oʻzgarib borgan. Ilk tajribalarda qoʻllangan tarjima strategiyasi «*togʻridan-to gʻri tarjima strategiyasi*» deb ataladi, bunda faqat tarjima jarayonida soʻzma-soʻz tarjima mufoviqligi eʼtiborga olinadi. Ushbu strategiya tarjima sifati va samaradorligiga salbiy taʼsir koʻrsatadi. 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 oʻrtasidagi avtomatik tarjima tizimlariga faol tatbiq etiladi.

Avtomatik tarjimada inson hamda kompyuter ishtiroki quyidagi tarzda boʻladi: a) *postediting* – matn (mashina-tarjimon) kompyuter yordamida

tarjima qilinadi, so‘ng inson – muharrir (inson-tarjimon) uni tahrir qiladi; b) *preediting* – bunda inson matni mashina – tarjimonga moslaydi, so‘ng uni kompyuterga havola etadi; c) *interediting* – bunda tarjimada murakkablik tug‘diruvchi holatlarga duch kelganda inson mashina – tarjimon ishiga aralashadi.¹

Matni bir tabiiy tildan boshqasiga tarjima qilish jarayonidagi murakkabliklar va muammolarga nazariy va amaliy jihatdan yechim topish yuzasidan izlanishlar davom etmoqda. Mashina tarjimasiga an‘anaviy yondashuv «**Rule Based Translation**» («Qoidalariga asoslanadigan tarjima») deb ataladi. 1984-yilda Yaponiyaning Kioto universiteti professori, mashina tarjimasi bo‘yicha yapon davlat dasturi rahbari Makoto Nagao avtomatik tarjima yo‘nalishida yangi konsepsiya ishlab chiqdi.² Ushbu konsepsiyaga ko‘ra, matnlar inson qo‘li bilan tarjima qilingan matnlarga tayanib analogiya tamoyili asosida tarjima qilinishi lozim. Buning uchun kompyuterda 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.

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») konsepsiyasiga ko‘ra, muayyan matni bir tildan ikkinchisiga tarjima qilish uchun dastlab 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 jummalarni 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

¹*Hutchins J.* Retrospect and prospect in computer-based translation // Proceedings of MT Summit VII «MT in the great translation era».–Tokyo: AAMT, 1996.-P.32. Bu haqida qarang: Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

²*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. - North Holland, 1984. - P. 173 - 180. Bu haqida qarang: *Rahimov A.* Kompyuter lingvistikasi asoslari. –T., 2011.

yol bilan «tarjimon xotirasi» asta-sekinlik bilan boyitib boriladi va tarjima samaradorligi oshib boradi.

Ushbu yoʻnalishning afzalliklari bilan birga ayrim kamchiliklari ham mavjud. Birinchidan, bu yoʻl bilan har qanday turdagi matnni tarjima qilib boʻlmaydi. U, asosan, muayyan tor yoʻnalishdagi matnlarga moʻljallangan. Ikkinchidan, maʼlum 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 toʻldirib BORISH amaliy jihatdan koʻp qiyinchiliklarni yuzaga keltiradi.

Rossiyada professor G.G.Belonogov frazeologik mashina tarjimasi kon-sepsiyasini asoslab berdi. Unga koʻra, matndagi tushunchalar koʻpincha alohida soʻzlar yordamida emas, balki soʻz birikmalari yordamida ifodalanadi. Oʻz-oʻzidan kelib chiqadiki, matnni bir tildan boshqasiga tarjima qilish jarayonida ham mazmunni yuzaga chiqaruvchi asosiy til birliklari hamda alohida soʻzlar emas, balki tipik vaziyatlarni ifodalovchi frazeologik birikmalar hisoblanadi. Mazkur yondashuvga muvofiq matnda tez-tez uchraydigan soʻzlar, iboralar, soʻz birikmalari, nutqiy etiket namunalarning tarjima ekvivalentlari bilan boyitilgan maʼlumotlar bazasi frazeologik mashina tarjimasi tizimining asosini tashkil etadi. Bunda, asosan, soʻzlarning distributsiyasiga eʼtibor beriladi. Bu metod polisemantik soʻzlarning tarjimada mos keluvchi toʻgʻri ekvivalentini tanlashga, omonim soʻzlarning kontekstdagi holatini aniqlashga, iboralarning 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 moʻljallangan edi. 1998 - 2000-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 Explorerning Web-brauzerida ishlaydi; c) MS Notepad, MS Wordpadda *Clipboard* (matnni koʻchirish)ga asoslanib ishlaydi. RETRANS tizimi tabiiy va texnika fanlari, iqtisodiyot, siyosat, qonunchilik hamda harbiy sohaga oid 4 mln lugʻat maqolasiga ega boʻlgan politematik soʻzlik bilan taʼminlangan. Kiritilgan til birliklarining 80% qismini soʻz birikmalari tashkil etadi, ular 2 ta soʻz uzunligidan tortib to 17 ta soʻz uzunligiga ega boʻlgan birikmalardir. Tizim tarkibida 400 ming lugʻat maqolasiga ega boʻlgan qoʻshimcha tematik lugʻat ham mavjud.

RETRANS tizimida ruschadan inglizchaga tarjima qilish jarayonida dastlab matn avtomatik tarzda morfologik, soʻng semantik-sintaktik tahlil qilinadi, frazeologik birliklar aniqlanadi. Matnning barcha soʻzlari hamda frazeologik birliklari qidiruv obrazlari shakllanadi. Lugʻat bazasidagi joylashuv tartibiga koʻra soʻz va birikmalar qidiruv obrazlarining yaqin kodlari topiladi, ularning dekodlanishi asosida qidiruv natijalari namoyon boʻladi. Keyingi bosqichda matnning har bir fragmentiga mos yagona yoki bir nechta tarjima ekvivalentlari tanlanadi.

Matnning rus tilidan ingliz tiliga tarjimasi ikki rejimda amalga oshiriladi:

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

Shuni alohida taʼkidlash zarurki, yuqoridagi konsepsiyalar ham avtomatik tarjima sifati va samaradorligini oshirish, til bilan bogʻliq toʻsiqlarni yoʻqotish yoʻlidagi muvaffaqiyatli izlanishlar hisoblanadi.

Savol va topshiriqlar:

1. Avtomatik tarjimadagi yondashuvlarni tavsiflang.
2. Tarjimon dasturlariga misol keltiring.
3. Avtomatik tarjimaning ishonchlilik darajasi haqida mustaqil mulohaza yuriting.
4. Avtomatik tarjimaning mukammal boʻlishiga toʻsiq boʻluvchi omillarni ayting.
5. Avtomatik tarjimaning ishlash protsedurasini izohlang.
6. Avtomatik tarjimaning maʼlumotlar bazasi haqida fikr bildiring.
7. Avtomatik tarjimaning yutuq va kamchiliklarini “T-sxema” asosida yoriting.

¹ Poʻlatov A., Muhamedova S. Kompyuter lingvistikasi. - T., 2007; Rahimov A. Kompyuter lingvistikasi asoslari.-T., 2011.

Mashina tarjimasining lingvistik muammolari

Reja:

1. Mashina tarjimasi – sun’iy intellekt yo‘nalishi sifatida.
2. Sun’iy intellekt bo‘yicha olib borilgan tadqiqotlar.
3. Neyron tarmoqlarining ishlash tizimi.

Asosiy tushunchalar: *sun’iy intellekt, neyron tarmoqlari, intellektual mashinalar, intellektual kompyuter dasturlari, intellektual sistema, bilimlar bazasi, masalani yechuvchi tizim, intellektual interfeys, Wikipedia, semiotik yondashuv, ekspert tizimlari, antiqiy hukm chiqarish tizimi, yuqori darajadagi psixik jarayonlar (tafakkur, nutq, emotsiya, ijod), simvolli yondashuv, simvolik mantiq qoidalari, ijodiy tafakkurni modellashtirish, “Turing testi”, neyron tarmoqlari.*

Kompyuter lingvistikasi fanining eng istiqbolli yo‘nalishlaridan biri sun’iy intellekt hisoblanadi. Fanda hali-hanuz sun’iy intellektning aniq va mukammal ta’rifi mavjud emas, chunki fan bu tizimga hali to‘liq erishilgani yo‘q. Mavjud ta’rif va tavsiflarning hammasi mazkur sohada erishilgan yutuqlarga asoslanib keltirib chiqarilgan. Falsafada inson intellektining statusi va tabiati haqidagi 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.

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, ma’lumotlarni umumlashtirish kabilarni yechishga yo‘naltirilgan bo‘ladi. Bu termin mazkur xususiyatlarga ega sistemalarni ishlab chiqish bilan bog‘liq kompyuter texnologiyalarining bir bo‘limiga nisbatan ham ishlatilmoqda».¹ Sun’iy intellekt bo‘yicha tadqiqotlar olib borgan olim Jon Makkarti shunday yozadi: «*Sun’iy intellekt – intellektual mashinalar, intellektual kompyuter dasturlarini yaratish texnologiyasi va u haqidagi fan. Intellektual sistema strukturasi uch asosiy blokni o‘z ichiga oladi: bilimlar bazasi, masalani yechuvchi tizim va intellektual interfeys*»².

¹ http://www.wikipedia.org/wiki/artificial_intelligence.html

² <http://www-formal.stanford.edu/jmc/whatisai.html>

Ekspert tizimlari modelida, intellektual informatsion tizimlarda, ba'zi mashinalarda, hayvonlarda va odamlarda intellektning turlicha darajalari va ko'rinishlari mavjud. Mazkur har xillik bilimlar bazasi va qanday operatsiyalarni bajara olishga 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, modellashtirilgan tizimlar texnologiyasi demakdir. Ayni shu ma'noda «artificial intelligence» terminini birinchi marta Jon Makkarti 1956-yilda Darmut universitetidagi konferensiyada qo'lladi va shundan buyon bu termin ishlatilib kelmoqda. 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 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 intellekt muammosi bilan shug'ullanadi.

Sun'iy intellekt tizimlarini yaratish muammosiga uch xil asosiy yondashuvni ajratib ko'rsatish mumkin:¹

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

-yuqori (inglizcha *Bottom-Up AI*), biologik yondashuv – neyron tarmoqlarini o'rganish, biologik unsurlarga asoslangan holda intellektual salohiyatni modellashtirish va neyrokompyuter yoki biokompyuterlarni yaratish;

-gibrid tizimli yondashuv – semiotik va biologik yondashuvning sintezi bo'lgan umumlashtirish.

¹Рассел 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. – С.8. Бу haqida qarang: Po'latov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2007. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

Mazkur yondashuvlardan tashqari, 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 programmalashtirish 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. Unga 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 sezadi¹.

Umuman olganda, robototexnika va sun‘iy intellekt bir-biri bilan aloqador tushunchalar bo‘lib, kelajakda ushbu ikki yo‘nalishning integratsiyasi natijasida intellektual robotlar yaratilishi ko‘zda 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.

Sun‘iy intellekt tizimi insonning ijodiy tafakkurini modellashtirishni ham o‘z ichiga oladi. Shu bilan bog‘liq ravishda san‘at asarlarini avtomatik tarzda yaratish jarayonini aks ettirgan “mashina ijodi” tushunchasi yuzaga keldi. Bugungi kunda ijodni to‘la modellashtirishga erishilgani yo‘q, lekin bu sohada sezilarli yutuqlar qo‘lga kiritilgan. Jumaladan, kompyuter yordamida musiqalar 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‘yinlari) 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

¹ http://www.wikipedia.org/wiki/artificial_intelligence.html

bo'yicha jahon chempioni Garri Kasparov ustidan g'alaba qozondi. «Deep Blue» nomi «Deep Thought» («Chuqur mulohaza, tafakkur» ma'nosini anglatadi) loyihasiga aloqadordir.

- MYCIN tizimi – bir qator kasalliklarga 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 toppish so'raladi, so'ng foydalanuvchiga 20 ta savol beriladi, berilgan javoblardan kelib chiqqan holda o'sha narsa topiladi. Ayniqsa, bu tizim internet paydo bolgandan so'ng juda mashhur bo'ldi.¹

- Via Voice – nutqni tovush orqali tanish tizimi.

- RoboCup – futbol bo'yicha robotlar turniri.

“Sun'iy intellekt tizimini baholash bo'yicha ikki xil nuqtayi nazar mavjud: 1) sun'iy intellektни 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 bilish qobiliyatining, tafakkurning barcha qirralarini qamrab ololmaydi. 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*². Xuddi shu pozitsiyada turgan olimlardan yana biri Rodjer Penrouz o'zining «*Qirolning yangi aqli*» kitobida tafakkurni formal tizimlar asosida mutlaq modellashtirishning imkoni yo'qligini bayon qilgan.³

Ayrim olimlar sun'iy intellekt bo'yicha erishilayotgan yutuqlarning keyingi natijasidan qo'rqmoqdalar. AQSHda Sun'iy intellekt singular instituti (SIAI) tadqiqotchisi E.Yudkovskiy insoniyatni sun'iy intellektlar

¹ (<http://www.20q.net> sayti mavjud)

² Searle J.R. Minds, brains, and programs // Behavioral and Brain Sciences. 1980. -P.417- 437. Bu haqida qarang: Po'latov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2007. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

³ Пенроуз Р. Новый ум короля. О компьютерах, мышлении и законах физики. - М.: УРСС, 2005. - P.14.

bilan bog'liq global xavf kutayotganidan ogohlantiradi. Olimning fikricha, agar kelajakdagi sun'iy intellekt tizimiga insonlarga nisbatan do'stlona munosabatda bo'lishni ta'minlovchi dastur o'rnatilmasa, butun insoniyat uchun katta xavf paydo bo'ladi. 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 singularlik deb atalmoqda. Futurologlarning fikricha, bu davr 2030-yilda boshlanadi.¹

Sun'iy intellektga doir izlanishlar XX asrning 40-yillarida boshlangan, sun'iy intellekt muammosiga bag'ishlangan «*Mashina fikrlay oladimi?*», «*Mashina o'z yaratuvchisidan aqlliroqrni?*» 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, u «Tyuring testi» deb nomlanadi. Ushbu empirik test g'oyasi 1950-yilda «Mind» falsafiy jurnalida «*Hisoblash mashinalari va tafakkur*» (*Computing Machinery and Intelligence*) nomli maqolada e'lon qilindi. Bu testdan ko'zlangan maqsad sun'iy tafakkurning inson tafakkuriga yaqin kelishi mumkinligini aniqlashdir. Unga 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

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*».²

Avtomatlashtirilgan ekspert tizimlar mantiqiy masalalarni kompyuter yordamida yechishga imkon beradi. Mazkur tizimlar ma'lum sohadagi bilimlar majmuyiga asoslangan implikativ qoidalar – «agar ...sa, u holda

¹ Kurzweil R. The Singularity Is Near. - N.Y.: Viking, 2005. - P.3. Bu haqida qarang: Po'latov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2007. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

² Jurafsky D., Martin J.H. Speech and language processing. - New Jersey, 2000.-P.7.

...» 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 etadi.

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 «Semantik» korporatsiyasi tomonidan Q&A dasturi, «Karnegi grupp» kompaniyasi tomonidan **Language Craft** dasturiy ta‘minoti ishlab chiqildi,¹ Bu turdagi dasturlar hozirda ingliz tilining standartlashgan so‘zlashuv uslubida uncha katta bo‘lmagan lu‘g‘at zaxirasi 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 tizimlari yoki neyron tarmoqlari 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 modellashtirish masalasi imkonsiz qolayotganligini ta‘kidlamodalar. Mazkur jihatlar insonning ichki dunyosi, intuitsiyasi, ong osti jarayonlari, ilhom kabi sirli fenomenlarni kiritish mumkin.

Inson miyasi murakkab biologik to‘r hisoblanadi, u yuz mlrd nerv tolalaridan - neyronlardan tarkib topgan. Neyronlarning bir-biriga uzviy aloqadorligi, zanjirli bog‘lanishi natijasida miya faoliyati me‘yoriy darajada ishlaydi. Zamonaviy sun‘iy neyron tarmoqlari inson miyasidagi

¹ *Анисимов А.В.* Компьютерная лингвистика для всех: мифы. Алгоритмы. Язык. - Киев: Наукова думка, 1991. - С. 10 - 11.

neyronlar modeli asosida yaratilmoqda. Hozirgi kunda neyron tarmoqlari bir qator amaliy masalalarni hal etishda keng qo'llanmoqda. Jumladan, ob-havo prognozini aniqlashda, obrazlarni ajratishda (xususan, og'zaki va yozma nutqni aniqlashda), robotlarni boshqarishda, tibbiyotda tashxis qo'yishda ishlatilmoqda. Amerikaning NASA kosmik agentligida neyron tarmoqlaridan fazoda robotlarni boshqarish, hududga tasodifan kelib tushgan obyektlarni tutishda foydalanilmoqda. «General Dynamics» kompaniyasi dvigatel shovqiniga qarab kemalarni tanishga imkon beruvchi tizim ishlab chiqdi. AQSH Pensilvaniya universitetida uchuvchi obyektlarni (NUJ) aniqlash uchun neyron tarmoqlar tizimi yaratildi.

Neyron tarmoqlarining ishlash tizimi matematik ifodalar yoki mantiqiy qoidalarga emas, balki namuna (**example based approach**) asosida o'qitishga asoslangan. Bunda muayyan masalani yechish uchun namunalar (misollar), ya'ni o'qituvchi moslamalar tayyorlanadi va tarmoq shu asosda tekshiriladi. Bunda turli xil metodlardan foydalaniladi. Eng keng tarqalgan metod – bu dastlabki yechim va sinov natijalarini e'tiborga olgan holda neyron tarmoqlari strukturasi korrektirovka qilishga asoslangan metod (**back propagation method**) bo'lib, unda oldindan ko'plab o'qitish materiallari 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 strukturasi 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.¹

Sun'iy intellekt tizimiga doir amaliy ishlardan biri ekspert tizimlari hisoblanadi. Ekspert tizimi – bu muayyan sohadaagi ekspertlik yoki mavjud bilimlarning imitatsiyasi tarzida yaratilgan kompyuter dasturi. Boshqacharoq aytganda, ekspert tizimi maxsus sohaga oid barcha bilimlar jamlangan qomus yoki 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

¹Anderson 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. Bu haqida qarang: Po'latov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2007. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

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 dasturlash texnikasidan ko'ra muammoli sohaga doir bilimlar bazasiga bog'liq ekanligini asoslab berdi.¹ Dastlab ekspert tizimlari diagnostika va kasallikni davolash sohasiga tatbiq etildi. Ekspert tizimlari keyinchalik kimyo, biologiya, geologiya sohalarida ham qo'llana boshladi.

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 dasturlash tillari hisoblanadi. Bunda yagona qabul qilingan bilimlar bazasi uchun deklarativ formalizm hamda qaror qabul qilish uchun protsedura tizimi qo'llaniladi. Ekspert tizimlari ichiga ekspertlar hamda bir qator yordamchi dasturlar kiradi. Jumladan:

- savol-javobli dastur foydalanuvchi bilan tabiiy tilda munosabatga kirishishni ta'minlab beradi;

- bilimlar ombori dasturi foydalanuvchiga bilimlar bazasiga o'zgartirishlar kiritish hamda to'ldirib, boyitib BORISHga imkon beradi;

- tushuntiruvchi dastur – javoblar tizimini tekshiradi;

- interpretator-dastur muayyan sohadagi terminlarni sharhlash, izohlashga yordam beradi;

- ijro etuvchi dastur barcha tizimlar ishi ta'minlab beradi.

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.

Informatsiya shiddat bilan ko'payib borayotgan bugungi davrda ekspert tizimlarining bilimlar bazasini uzluksiz to'ldirib, yangi bilimlar bazasi bilan boyitib BORISH muhim vazifa bo'lib qolmoqda. Bilimlarning mavjud holatlar bilan muvofiqligi hamda bilimlar bazasining boyitilishi ekspert tizimlarining tushuntiruvchi dasturlari asosida amalga oshiriladi.

¹ Белоногов Г.Г. Компьютерная лингвистика и перспективные информационные технологии. - М.: Русский мир, 2004. - С.226;

Keyingi paytlarda 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 informatsiya 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; 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, «Qaldirg'och qush hisoblanadi» 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» 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 sinekdoxa termini bilan ataladi. Masalan, kompyuter – butun; monitor, klaviatura, sichqoncha, protsessor kabilar qismdir. Butun uchun *xolonim*, qism uchun *meronim* atamaları ishlatiladi. Demak, motor – avtomobil uchun meronim, avtomobilning o'zi esa motor uchun xolonim sanaladi.

Bilimlar strukturasi o'ld Asosiy 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 gavdalanitirishning bir usuli sanala-

di. Ya'ni freymga modellashtirilayotgan 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'ruzalarga 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.¹ 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.Fillmor qarashlariga L.Tenyening aktantlar (fe'ning bo'sh o'rinlarini to'ldiruvchilar) haqidagi g'oyasi katta ta'sir ko'rsatgan. Fillmorning Case (kelishiklar) g'oyasiga asosan aktantlar (fe'ning hamrohlari) fe'l yordamida ifodalanadigan holat yoki harakat munosabatlaridagi bajariladigan «rollar» sifatida talqin qilinadi. Ch.Fillmor ajratgan kelishiklar quyi-dagilar:

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 yakuniy holat (inglizcha «goal» so'zidan olingan).

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 – ssenariylar, freym–vaziyatlar. Freymlar tizimi birlashib, semantik tarmoqni 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 majmuyi hisoblanadi. Slotlarda obyektarga xos belgilar ma'nosi qayd etilgan bo'ladi.

¹Minsky M. A. Framework for Representing Knowledge / The Psychology of Computer Vision. McGraw-Hill, New York (USA), 1975; Po'latov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2007. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

«Stol» freym fragmentining jadval ko‘rinishidagi reprezentatsiyasi

Slot nomi	Slot mazmuni
Oyoqchalar miqdori	To‘rtta, va undan ko‘proq mumkin
Materiali	Yog‘och, plastmassa, shisha
Yuzasi, sirti	Oval, kvadrat, aylana
Tumbaning mavjudligi	Ixtiyoriy tarzda
Vazifalari	Ovqatlanish uchun, dars tayyoriash uchun mo‘ljallangan

Freym-senariy ba’zi harakat, voqea-hodisalarning xarakterli 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 amal qiladi, ya’ni har bir harakat natijasi shartli sanaladi, uning ta’siri va oqibatida navbatdagi ish-harakat sodir boiishi mumkin. Har bir senariyda rol bajaruvchilar mavjud. Senariylar real hayotda uchraydigan vaziyatlarni aks ettiruvchi holatlar asosida yuzaga keladi.

Ko‘pchilik mutaxassislarining 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 afzallik va kamchilik tomonlari mavjud. Shuni e’tiborga olgan holda aytish mumkinki, neyron tarmoqlari va ekspert tizimlarining birgalikda ishlashini ta’minlovchi gibridd 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 «qoida bo‘yicha» (implikatsiya -if..., then... - agar... -sa, u holda...) hal etilsa, ekspert tizimlariga murojaat qilish 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.

Savol va topshriqlar

1. Ekspert tizimlarining mohiyatini yoriting.
2. Sun’iy intellekt tizimi bo‘yicha erishilgan yutuqlarni bayon eting.
3. Sun’iy intellekt tizimini yaratish borasidagi yondashuvlarni ayting.

4. Avtomatik tahrirning pragmatik jihatlarini tushuntiring.
5. Sun'iy intellektning takomillashgan texnik turlarini yaratishdagi o'rnini yoriting.
6. Kompyuter lingvistikasi imkoniyatlarini umumlashtiring.
7. "Sun'iy intellekt" tushunchasini "T-sxema" asosida tavsiflang.

13-§. KOMPYUTERDA MATNLARNI QAYTA ISHLASH TIZIMLARI

Reja:

1. Qidiruv va ekspert tizimi haqida tushuncha.
2. Informatsion qidiruv tizimlari.
3. Gipertekst texnologiyasi.

Asosiy tushunchalar: *gipertekst, kontekstli qidiruv, avtomatik referatlash, transliteratsiyalash, informatsion qidiruv tizimlari, manzilli qidiruv, semantik qidiruv, faktografik qidiruv, hujjatli qidiruv, indekslash, gipermedia.*

Kompyuter lingvistikasining muhim yo'nalishlaridan yana biri informatsion qidiruv tizimlari (IQT) hisoblanadi. Ular formal belgilarga ko'ra informatsiyani qidirishga mo'ljallangan. IQTlar 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.¹ Informatsion qidiruv tizimi – ma'lumotlarni qisqa vaqt ichida qidirib topish, kerakli hujjatlarni ma'lumotlar bazasidan qidirishni optimallashtirish tizimidir. Bu yo'nalish ilmiy-texnikaviy axborotlarning ko'payib ketishiga javoban 1960-yil oxirlarida yuzaga keldi. Axborot ko'paygan sari uni saralash, keraklisini qidirib topish muammosi yuzaga keldi. Mazkur muammoni bartaraf etish

¹ Manning C, Raghavan P., Schutze H. Introduction to Information Retrieval // Cambridge University Press. 2008. - P. 5. Po'latov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2007. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

uchun olimlar informatsion qidiruv tizimlari g'oyasini taklif etdilar. Bu yo'nalish internet tarmog'i yuzaga kelgandan so'ng yanada rivojlandi. Masalan, internetda **Google, Yandex, Rambler, Yahoo** kabi qidiruv tizimlari mavjud.

Informatsiyani qidirish quyidagi 4 bosqichni o'z ichiga oladi:

- informatsiyaga bo'lgan talabni aniqlashtirish hamda informatsion so'rovni shakllantirish;

- so'ralgan informatsiya saqlangan manbalarni (informatsion massivlarni) aniqlash;

- informatsion massivlardan zaruriy informatsiyani ajratib olish;

- informatsiya bilan tanishish va qidiruv natijalarini baholash.

Qidiruvning quyidagi turlari mavjud:

1.To'liq matnli qidiruv – hujjatning to'liq mazmun-mundarijasi asosida amalga oshiriladigan qidiruv. Masalan, internetdagi qidiruv tizimlari: www.yandex.ru. www.google.com.

2.Metama'lumotlar bo'yicha qidiruv – hujjatning muayyan atributlari (hujjat nomi, yaratilgan vaqti, hajmi, muallifi) asosida amalga oshiriladigan qidiruv. Qidiruvning bu turiga MS Windows tizimida fayllar bilan ishlashda foydalaniladigan qidiruv dialogi misol bo'la oladi.

3.Tasvir qidiruvi – tasvirning mazmuni asosida amalga oshiriladigan qidiruv. Bunda qidiruv tizimi rasmning mazmunini o'zlashtiradi va natijada shunga o'xshash rasm topiladi. *Polar Rose, Picollator* qidiruv tizimlari mana shu tamoyil asosida ishlaydi.

4.Multimediali informatsiya qidiruvi – bunda ma'lumotlar audio va videofayllar bo'yicha qidiriladi.

Qidirish jarayoni turli metodlar asosida yuz beradi: manzilli qidiruv, semantik qidiruv, hujjatli qidiruv, faktografik qidiruv. Manzilli qidiruv so'rovda ko'rsatiladigan sof formal belgilar bo'yicha hujjatlarni qidirish jarayoni bo'lib, bunda hujjatning aniq manzili ko'rsatilishi lozim. Hujjatning manzili web-server va web-sahifa manzili yoki bibliografik qayd elementlari, hujjat saqlangan joy sanaladi. Semantik qidiruv hujjatlarning mazmuni asosida qidirish usuli bo'lib, bunda hujjatning umumlashma mazmuni, sarlavhasi deskriptor ko'rinishida shakllantirilishi lozim. Hujjatli qidiruv matnli ma'lumotlar hamda ular haqidagi bibliografik qaydlarni qidirish jarayonidir. Faktografik qidiruv so'rovda nazarda tutilgan ma'lumotga mos keluvchi faktlar (matnli hujjatlar tarkibidan ajratilgan ma'lumotlar)ni qidirish jarayonidir.

Qidiruv tizimi relyatsion ma'lumotlar hamda gipertekstli ma'lumotlar bazasida mavjud bo'lgan hujjatlar, hujjatlar haqidagi metama'lumotlar, matnlar, rasmlar, video va audiofayllarni qidirishni o'z ichiga oladi.

Informatsiyaning eksponensial tarzda ko'payib ketayotganligi, shuningdek, internet tizimida ham kundan kunga ma'lumotlar hajmining ortib borayotganligi kompyuter lingvistikasi oldiga bir qator vazifalarni qo'yimoqda:

- qidiruv tizimining optimal modellarini yaratish;
- hujjatlarni tasniflash (klassifikatorlar tizimini yaratish);
- hujjatlarni klasterlash;
- foydalanuvchi interfeyslari va qidiruv tizimlari arxitekturasini loyihalash;
- zaruriy informatsiyani ajratib olish, hujjatlarni avtomatik referatlash va annotatsiyalsh;
- deskriptorli tillar va tezauruslarni takomillashtirish.

Saqlangan ma'lumotning xarakteri hamda qidirish xususiyatiga ko'ra informatsion qidiruv tizimlari ikki katta guruhga bo'linadi: hujjatli va faktografik informatsion qidiruv tizimlari. Hujjatli informatsion qidiruv tizimlarida matn, tavsif va ta'riflar saqlanadi, faktografik informatsion qidiruv tizimlari esa jadvallar, formulalar, grafik ko'rinishidagi ma'lumotlarni o'z ichiga oladi. Shuningdek, aralash tabiatli informatsion qidiruv tizimlari ham mavjud bo'lib, unda ham hujjatli, ham faktografik ma'lumotlar saqlanadi. Informatsion qidiruv tizimlarida qidiruvni ta'minlash uchun maxsus informatsion qidiruv tillari mavjud. Informatsion qidiruv tillari formal til bo'lib, informatsion qidiruv tizimlarida saqlangan hujjatlarning mazmunini tavsiflash va so'rov uchun mo'ljallangan.

Informatsion qidiruv tilida hujjatlarga ishlov berish protsedurasi *indekslash* deb ataladi. Indeksash natijasida har bir hujjatga informatsion qidiruv tizimidagi ta'rif shakli yoziladi. Mazkur ko'p bosqichli jarayonda deskriptorli informatsion qidiruv tillari ishlatiladi. Hujjat mavzusi deskriptorlarning birlashuvi asosida yoziladi. Deskriptorlar, kalit leksika sifatida so'zlar, terminlar, muayyan sohadagi boshlang'ich tushuncha va kategoriyalardan foydalaniladi. 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.

Matnning asosiy mazmun-mundarijasini aks ettiradigan kalit so'zlar, terminlar, asosiy tushunchalar bazasi asosida kompyuterda lug'at tuziladi.

Bunday lug‘at *informatsion qidiruv tezauruslari* deb ataladi. Tezauruslar ensiklopedik va izohli lug‘atlardan farqli ravishda til birliklarining matnda qo‘llanish chastotasi va matnning predmet mundarijasini aks ettirish darajasiga qarab tuziladi. Bunda bazadagi ma‘lumotlar terminning mazmuniy tavsifi orqali qidirishga mo‘ljallangan bo‘ladi. Terminlar tezaurusga qat‘iy semantik prinsiplar asosida kiritiladi, bunda *naming gipo-geperonimik (tur-jins)*, *holo-meronik (butun-bo‘lak)*, *iyerarxik (pog‘onali)* munosabatlar hamda assotsiativ mazmuniy bog‘lanishlar e‘tiborga olinishi lozim.

Virtual ensiklopedia sanalgan Wikipediada keltirilgan so‘zga bag‘shlangan lug‘at maqolasida semantik munosabatlar, assotsiativ bog‘lanishlar hamda iyerarxiya munosabati yaqqol ko‘zga tashlanadi. Shu o‘rinda ta‘kidlash zarurki, keyingi paytlarda qidiruv tizimlari bazasidagi tezauruslar giperhavolalar bilan ham ta‘minlanmoqda, bu esa user uchun qulayliklar yaratadi. Ya‘ni qidirilayotgan terminlarga aloqador bo‘lgan boshqa tushunchalar bilan bog‘liq ma‘lumotlarni ham oson topish imkonini beradi.

Ilm-fan, texnika taraqqiyoti intensiv ravishda rivojlanyotgan davrda terminlar dinamikasini e‘tiborga olgan holda tezauruslar bazasini muttasil to‘ldirib, boyitib BORISH zarur. Chunki til ochiq dinamik tizim hamda jamiyat taraqqiyotining «barometri» hisoblanadi, ya‘ni ayrim tushunchalar eskiradi, ularning o‘rniga yangi tushuncha va terminlar yuzaga keladi. Xalqaro INFOTERM (Xalqaro terminologik markaz, Avstriya) tashkiloti ma‘lumotiga ko‘ra, hozirgi kunda tillardagi terminlar miqdori 50 mln, mahsulot nomlari miqdori 100 mlanni tashkil etadi. Lekin ushbu statistik ma‘lumot turg‘un emas, u jamiyat taraqqiyoti bilan yonma-yon o‘zgarib turadi.

Kompyuter lingvistikasida inqilobiy o‘zgarish yasagan yangilik gipertekst texnologiyasi hisoblanadi. Chunki bu tizim Guttenberg asos solgan kitob bosish texnologiyasiga zid ravishda yangi turdagi matn strukturasi vujudga keltirdi. 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 xayoliy qurilma haqida «As We May Think» («Biz o‘ylaganimiz kabi») maqolasini e‘lon qildi. V. Bush o‘sha qurilmani kutubxonadagi har qanday hujjatlar, yozishmalar, kitoblarni displayda ko‘rsatib bera oladigan, cheklanmagan mikrofilmlarga

ulangan elektromexanik stol sifatida tasvirlaydi. Kompyuter texnikasi yo‘qligi bois loyihaning amalida qo‘llanishi bir oz keyinga surilgan edi.

Bushning g‘oyasi gipertekst haqidagi nazariy va amaliy izlanishlarga juda katta ta’sir ko‘rsatdi, Bushning kelajak loyihasi 1960-yilda T.Nelsonning «Ksanadu» tizimi orqali qayta ko‘tarildi. Bu texnik tizim kiritilgan matnlarni va uning qismlarini turli xil usullarda, har xil qatorda, ixtiyoriy ketma-ketlikda o‘qishga imkon beradi. Bu esa ko‘rilgan matnlarning o‘rnini, ketma-ketligini eslab qolishga, xohlagan vaqtda kerakli matnni tanlab olishga va tezkor murojaat etishga yordam beradi. Bunday xususiyatga ega bo‘lgan matnlarni T.Nelson *gipertekst* deb nomladi. 1968-yil dekabrda amerikalik olim Duglas Engelbart ham o‘zi yaratgan gipertekst interfeysini e‘lon qildi. Olim uni «The Mother of All Demos» deb atadi. 1983-yilda Ben Shneyderman ham o‘z gipertekst loyihasini yaratdi, ushbu loyiha The Interactive Encyclopedia System (TIES) deb nomlangan.¹ 1992-yilda birinchi marta gipertekst World Wide Web (xalqaro to‘r)ga tatbiq etildi.

Gipertekst yunoncha «hyper» – «ostida», «orqasida», «ortidan» ma’nolarini anglatuvchi old qo‘shimcha hamda lotincha *tekst* – «to‘qima» ma’nosini anglatuvchi so‘zlardan olingan.² 1965-yilda gipertekst va gipermedia terminlari T.Nelson tomonidan fanga kiritildi. Olim bu haqda o‘zining «Litarary Machines» kitobida shunday yozadi: «*Gipertekst termini gipermedia terminiga qaraganda keng ommalashib ketdi. Aslida har ikki termin bir tushunchani anglatmaydi. Ya’ni gipertekst faqat matnlarning tarmoqlanishiga nisbatan ishlatiladi, gipermedia termini esa grafiklar, audio va videofayllar, shuningdek, matnlarning qismlarga ajratilishini ham o‘z ichiga oladi. Bir muddat har ikki termin o‘rniga «interaktiv multimedia» atamasi ham ishlatildi. Lekin negadir gipertekst termini hammasidan ko‘ra keng e’tirof etildi va ommalashdi*».³

Gipertekst matnni 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 qilishga yordam beradi. 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

¹ <http://www.sigweb.org/conferences/ht-cover.shtml>

² Субботин М.М. Гипертекст. Новая форма письменной коммуникации. - ВИНТИ. Сер. Информатика. 1994. - С. 18.

³ <http://www.iw3c2.org>.

unda bir olamdan ikkinchi olamga hech qanday to'siqlarsiz o'tish mumkin bo'ladi. Mana shu virtuallikni ta'minlab beruvchi tizim gipertekst g'oyasidir. Internet saytlari manzilini ko'rsatishda HTTP (Hyper Text Transfer Protocol – Giper matnni o'tkazish yo'lagi)dan foydalaniladi.

Tekst va gipertekst tushunchalari o'rtasidagi farqni quyidagicha umumlashtirish mumkin:

TEKST	GIPERTEKST
Tugallanganlik	Tugallanmaganlik
Chiziqlilik	Nochiziqlilik
Aniq muallifning mavjudligi	Aniq muallifning mavjud emasligi
Bir tomonlama yo'nalganlik	Ko'p tomonlama yo'nalganlik
Bir jinslik (bir xilda tuzilganlik)	Bir jinsda tuzilmaganlik (turfa xillilik)
Yopiq tuzilishga egalik	Ochiq tuzilishga egalik

Gipertekst tuzilishiga ko'ra iyerarxik (pog'onali) yoki tarmoqli bo'lishi mumkin. Daraxt ko'rinishidagi iyerarxik gipertekstlar foydalanuvchining navigatsiya (uzellar bo'ylab bir qismdan ikkinchi qismga o'tish) imkoniyatini cheklaydi, gipertekst texnologiyasining barcha imkoniyatlarini yuzaga chiqarmaydi. Bunday gipertekstda komponentlar o'rtasidagi aloqalar jins-tur munosabatiga asoslangan tezaurus strukturasi eslatadi. Tarmoqli gipertekst esa birmuncha keng tarqalgan va imkoniyati kengroq tizim hisoblanadi, u jins-tur munosabatidan tashqari boshqa assotsiativ munosabatlarni ham o'z ichiga oladi. Shuning uchun tarmoqli gipertekstlar internet tizimida faol ishlatiladi.

Gipertekst ikki xil holatda bo'lishi mumkin: statik va dinamik gipertekstlar. Statik gipertekstlar oldindan tayyorlangan va saqlangan bo'lib, ular turg'un holatda bo'ladi. Masalan, gipertekst texnologiyasi asosida CDga yozilgan kitoblar statik gipertekstga misol bo'la oladi. Dinamik gipertekstlar murakkab va tizimli bog'lanishlarga tayangan bo'lib, u ko'proq tarmoq tizimida ishlaydi. Ayniqsa, internet tizimida dinamik gipertekstlar faol qo'llanadi. Masalan, Arizona axborot tizimi (AAIS) dinamik gipertekst sanaladi, uning bazasiga har oyda 300 - 500 referat ko'rinishidagi ma'lumotlar kiritib turiladi.

Demak, gipertekst texnologiyasi informatsion 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 yaratadi.

Ko'pgina tadqiqotchilar gipertekst tizimini yangi informatsion eraning boshlanishi, virtual olamga dastlabki qadam sifatida baholaydilar. Haqiqatan ham gipertekst tizimi kompyuter lingvistikasida, internet tizimida o'ziga xos inqilob yasadi, deyish mumkin.

Savol va topshiriqlar

1. Avtomatik tahrirning ishonchlilik darajasi nimalarga bog'liqligini ayting.
2. Avtomatik tahrirning cheklangan jihatlarini tushuntiring.
3. Gipertekst texnologiyasi va an'anaviy kitob bosish texnologiyasini solishtiring.
4. Informatsion qidiruv tizimlari haqida ma'lumot bering.
5. *Deskriptor* termini mohiyatini yoritng.
7. Informatsion qidiruv tizimlaridan foydalanib, til o'qitish dasturlarini guruhlang.

Matlarni avtomatik tahrirlash

Reja:

1. Avtomatik tahrir yo'nalishi.
2. Matn tahririda BORIS tizimi
3. Nogrammatik ifodalar.

Asosiy tushunchalar: *matn tahriri, mexanik xatolar, matn muharrirlari, avtomatik tahrir, BORIS tizimi, nogrammatik ifodalar, tezaurus, deskriptor.*

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 texnik xatolarning avtomatik tarzda to'g'rilanishi va xato ekanligi haqidagi signallarning userga taklif etilishi bilan bog'liq. Microsoft Word dasturida so'zlarning yozilishi bilan bog'liq xatolar bo'lsa, ularning ostiga qizil yoki ko'k rangda chizilishi, katta harf bilan yozilishi kerak bo'lgan so'zlarning avtomatik tarzda bosh harfga o'tkazilishi, hujjatga avtomatik numeratsiya qo'yilishi kabilar avtomatik tahrir erishgan yutuqlar hisoblanadi.

Avtomatik tahrir yo'nalishi XX asr o'rtalariga kelib rivojlandi. U matn muharriri dasturlari bilan birgalikda yangicha imkoniyatlar, qo'shimcha effektlar bilan rivojlanmoqda. Avtomatik tahrirning oddiy muharrirlardan farqi shundaki, unda tahrir avtomatik tarzda qisqa vaqt ichida hajman katta

bo'lgan matnni tez tekshirish va xatolarni to'g'rilash imkoniyati bo'ladi. Avtomatik tahrirning pragmatik ahamiyati shundaki, u kelajakda taraqqiy etsa, yetarli ma'lumotlar bazasi yaratilsa va maxsus dasturlar ishlab chiqilsa, muharrir faoliyati yengillashadi.

Ma'lumki, kompyuter yaratilmasdan avval matnlar yozuv mashinkalarida terilgan, bu jarayonda orfografik, punktuatsion, stilistik, grammatik xatolarga yo'l qo'yilgan bo'lsa, matnni boshdan oxirigacha qayta yozib chiqishga to'g'ri kelgan, bu esa foydalanuvchiga qiyinchilik tug'dirgan. Kompyuterda matnni tahrirlash birmuncha qulay bo'lib, unda foydalanuvchi dialogli rejimda ishiydi va o'z o'rnida xatolarni to'g'rilash imkoniyatiga ega bo'ladi. Shu o'rinda alohida ta'kidlash zarurki, avtomatik tahrirlash tizimi faqat matn muharrirlari dasturlari bazasiga qo'shilgan ingliz, rus, nemis, fransuz kabi tillarda amal qiladi. Agar bazada mavjud bo'lmagan tilda matn kiritilsa, uni kompyuter avtomatik tahrirlay olmaydi, ayrim internatsional so'zlar bundan mustasno. Avtomatik tahrirlash jarayonida MS Word lug'atida mavjud bo'lmagan so'z uchrasa, ostiga qizil chiziq bilan ajratib ko'rsatiladi. Mazkur holatda dastur so'z imlosi haqida variant taklif eta olmaydi, chunki Ms Word lug'atida bu so'z mavjud emas. Ehtimol u so'z tog'ri yozilgan bo'lishi mumkin, bunda insonning ishtiroki zarur bo'ladi. Agar so'zning imlosi to'g'ri bo'lsa, matnda tez-tez qayta ishlatilsa, uning avtokorreksiyada to'g'ri deb sanalishi uchun Ms Word lug'atiga «kiritib qo'yish» (kontekst menyudagi «add» yoki «добавить» komandasi yordamida) mumkin. Matn kiritish davomida so'z yoki so'zlar birikmasi ko'k chiziq bilan ajratib qo'yiladi, bunda so'zning imlosida xato mavjud bo'ladi. Bu holda foydalanuvchi xatoni o'zi tog'rilab qo'yishi zarur yoki ostiga chizilgan so'zda sichqonchani o'ng tomoni bosilsa, kontekst menyusida so'zning to'g'ri varianti taklif etiladi va uni tanlashi mumkin.

O'zbek tilshunosligida ham avtomatik tahrir yo'nalishida izlanishlar olib borilgan. Professor H.Arziqulov rahbarligida Samarqand davlat chet tillari institutida injener tilshunosligi markazi tashkil etilgan. Bu markazda davlat tilidagi matnlarga kompyuter yordamida qayta ishlov berish masalalari ustida ilmiy izlanishlar olib borilgan. Bunda matnlarga qayta ishlov beradigan kompyuter dasturlarini yaratish, o'zbek tilining barcha sathlariga oid matnlarni yig'ish, ularni muayyan bir tizimga keltirish, algoritmlarni tuzish, ularga statistik jihatdan ishlov berish, inglizcha-o'zbekcha mashina tarjimasini yaratish kabi sohalarida ilmiy-tadqiqot ishlari amalga oshirilgan. O'zbekcha matnlarga qayta ishlov beruvchi avtomatik tizim "Uzlington", lotin alifbosidagi matnlarni kirill

alifbosiga va krill alifbosidan lotin alifbosiga o'tkazuvchi "Spellchecker", kompyuter ko'magida korrektura qiladigan, imloviy xatolarni tuzatadigan kompyuter dasturlarini yaratishga birinchilardan bo'lib kirishilgan.¹

Avtomatik tahrirda uch tarkibli asosiy vazifa e'tiborga olinishi lozim:

1. Orfografiyani tekshirish – bu vazifa morfologik tahlil yordamida amalga oshiriladi, bunda etalon lug'atdagi baza asosida tekshiriladi.

2. Sintaksisni tekshirish orfografiyani tekshirishdan ko'ra murakkabroq. Chunki orfografiyani tekshiruvchi dasturlar yaratilgan, lekin norasmiy matnlarning sintaktik jihatdan avtomatik tahriri masalasi haligacha hal etilgani yo'q. Rasmiy matnlarning avtomatik tahriri bo'yicha muhim natijalarga erishilgan.

3. Semantik tahrirlash inson tafakkurini modellashtirishga asoslangan sun'iy intellekt tizimi bilan bog'liq ravishda o'rganilmoqda. Bunda mazkur ko'p aspektli muammoni semantik maydon, freym yondashuv, tezauruslar tizimini boyitish, so'z va so'z birikmalari o'rtasidagi paradigmatic munosabatlar (sinonimiya, gipo-giperonimik munosabat – tur-jins munosabati)ni modellashtirishga tatbiq etish asosida hal etish mumkin. Chunki inson bunday assotsiativ munosabatlarni o'z intellektiga tayanib o'rnatadi, biroq bu jarayon kompyuter uchun murakkabdir.

Avtomatik tahrirlash tizimi quyidagi vazifalarni ham o'z ichiga oladi:

-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 abzaslar, belgilar, so'zlarning 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).

O'zbek tilidagi matnlarni avtomatik ravishda tahrir va tahlil qilish dasturini yaratish bugungi kunning dolzarb masalasiga aylangan². Bunday dastur tizimining katta imkoniyat va yuqori sifati lingvistik modullarning

¹ *Rahimov A.* Kompyuter lingvistikasi asoslari. –T., 2011, 65-bet

² *Abjalova M.A.* Matnlarni avtomatik tahrir va tahlil qilish dasturining lingvistik bosqichlari//Yosh olim va talabalarning respublika ilmiy-amaliy anjumani materiallari. –T., 2015, 7-bet.

mukammal ishlab chiqilgani bilan belgilanadi, ular o‘z navbatida dasturning bosqichlarini tashkil etadi.

Lingvistik tahrir – bu turli ko‘rinish (ilmiy, badiiy, publisistik va rasmiy uslublar)dagi matnlarning orfografik, grammatik, stilistik hamda mantiqiy qurilishdagi xatolarni to‘g‘rilashni bildiradi.

Matnlarni tahrir qilish ikki usulda amalga oshiriladi:

1. Inson tomonidan amalga oshiriladigan tahrir (bevosita tahrir – muharrirlik ishi).

2. Zamonaviy texnika vositasida amalga oshiriladigan tahrir (bilvosita, ya’ni avtomatik tahrir).

Keyingi paytlarda axborot almashinuvining jadallashuvi va ma’lumotlar ko‘lamining tobora kengayib ketayotgani sababli bilvosita tahrirga bo‘lgan talab oshib bormoqda.

Matnlarni to‘liq avtomatik tahrir va tahlil qilish uchun dastur quyidagi modullar asosidagi lingvistik tizimdan tarkib topishi lozim:

Grafematik tahrir bosqichida matndagi xatboshi, raqam, punktuatsion va boshqa ramziy belgilar aniqlanadi, ya’ni so‘z va so‘zshakllarga qadar tekshiruv amalga oshiriladi. Grafematik tahrir bosqichining maqsadi – matndagi eng kichik birliklarni ham aniqlash va tasniflashdan iborat. Bunday birliklarga quyidagilar kiradi: so‘z, xatboshi, tinish belgilari, sanalar, pul birliklarining ramzlari, so‘z-raqamli birikmalar, sonlar, IP-manzillar va fayl nomlari, telefon raqamlari. Grafematik tahrir qilish dasturining ilk bosqichi hisoblanib, u keyingi bosqichlarning to‘laqonli va samarali bo‘lishi uchun zamin yaratadi.

Morfologik tahrir bosqichi asos va shu asosdan yuzaga kelgan so‘zshakllarni tekshiradi. Morfologik tahrirning maqsadi – faqat orfografik xatolarni tekshirishdan iborat, deb hisoblash uning vazifa ko‘lamini toraytiradi. Ushbu bosqichda kiritilayotgan so‘zshaklning qanday asosiy shakldan, qaysi lingvistik parametrlardan yuzaga kelgani aniqlanadi. Bu keyinchalik bir so‘zning barcha shakllari ustida ishlashga emas, balki asosiy shaklning qanday lingvistik parametrlarni olish imkoniyatiga egaligi ustida ish olib BORISHga turtki beradi. Masalan, ot turkumidagi *kitob* so‘zining qanday shakl hosil qiluvchi va sintaktik shakl yasovchi qo‘shimchalarni qabul qilishi hamda ularning kombinatsiyasini biriktirib, so‘zshakllarni hosil qilish imkoniyati, ya’ni parametrlari dastur bazasiga kiritiladi. Natijada *kitob* so‘zining barcha shakllari kiritilib, dastur ta’minotidan katta hajm egallanmaydi, balki asosiy shakl va unga birikishi me’yoriy hisoblangan lingvistik parametrlar kiritilib, hajm iqtisod qilinadi.

Avtomatik tahrir algoritmlari matnni qayta ishlash uchungina ahamiyatli emas. Tadqiqotlarda til materialining avtomatik qayta ishlash bosqichlari matnni tushunish bilan bir qatorda korpus yaratuvchilari uchun hamda mavjud razmetkalar asosida tadqiqot olib boruvchi lingvistlar uchun yangi imkoniyatlar yaratishi ta'kidlangan. Tabiiy tilni qayta ishlashning asosiy protsedurasi sifatida: tokenizatsiya, lemmatizatsiya, stemming, parsinglar ajratiladi.

Tokenizatsiya berilgan matnni tokenlarga bo'lib tashlashdan iborat. Bu algoritmni amalga oshiruvchi dastur *tokenayzer* deb ataladi. Leksik birliklarni ifodalash uchun "token" ishlatiladi. Tokenayzer matnni dastlab so'zlar orasidagi chegara asosida qismlarga bo'ladi. Bunda matn tarkibidagi tinish belgilari olib tashlanadi, chunki dastur tinish belgisini ham token sifatida ajratadi.

Tokenizatsiya, tabiiy tildagi belgilar oqimini alohida mustaqil birliklarga (token, so'zshakllarga) ajratish tabiiy tilni keyingi qayta ishlashda zaruriy shart sifatida qaraladi¹. Tokenlar (so'zshakllari) <> belgisi bilan ajratilgan. Masalan,

[<Xalq> <tomonidan> <yaratilgan> <tipik> <hayotiy> <barqaror> <takrorlanmas> <va> <xarakterli> <bo'lgan> <so'zlarning> <barchasi> <umumxalq> <tili> <tarkibini> <tashkil> <etganidek> <ijod> <sohiblari> <tomonidan> <tozalanib> <qayta> <ishlanib> <jilo> <berilib> <ma'lum> <qo'llanish> <qonun> <-> <qoida> <va> <o'lchovlariga> <bo'ysundirilgan> <til> <birliklari> <va> <vositalari> <adabiy> <til> <tarkibiga> <kiradi>]

Har bir token tarkibida lemmalar mavjud va bu morfologik tahrirning keyingi jarayonlarida aniqlashtiriladi:

"ishlanib" tokeni "ishla" lemmasiga, "o'lchovlariga" tokeni "o'lchov" lemmasiga mansubdir:

'<ishlanib >'

'ishla' <f> <hf> <o'f> <II> <bm> ("ishla" lemmasining fe'l turkumiga xos kategorial belgilari: fe'l>harakat fe'li>o'timli fe'li>II shaxs >buyruq mayli).

'<o'lchovlariga >'

'o'lchov <ot> <jot> <bk> ("o'lchov" lemmasining ot turkumiga xos kategorial belgilari: ot>**jarayon oti**>bosh kelishik>birlik).

Yuqoridagi matn tokenizatsiyasida tokenlarga ajratish dasturining cheklangan jihatlari ham ko'zga tashlanadi: - (chiziqcha)ning alohida

¹ Захаров В.П., Богданова С.Ю. Корпусная лингвистика.-Иркутск:ИГЛУ,2011.-С.38.

olinishi yoki juft soʻzlarni ikkita token holda ajratish semantik jihatdan muammolarni yuzaga keltirishi mumkin: <qonun> <-> <qoida>.

“Soʻz asoslarini (stemlarni) topishga qaratilgan jarayon *stemming* hisoblanadi. Stemming stem (asos)larni topishga qartiladi. Stemmer soʻzni kontekstga bogʻliq boʻlmagan holda qayta ishlaydi, turli turkumga oid boʻlgan turli maʼnodagi soʻzlarni farqlay olmaydi. Stemmerlar nisbatan oddiy, maʼlumotlarni tez surʼatda qayta ishlaydi. Lekin koʻp hollarda hal qiluvchi ahamiyatga ega boʻlmaydi¹.

Lemmatizatsiya – morfologik analizning spetsifik masalasi, soʻzning boshqa soʻzshakllaridan kelib chiqadigan dastlabki shakllanish jarayoni. *Lemmatizatsiya* – soʻzning dastlabki, boshlangʻich formasini (lugʻatdagi shaklini – lemmasini) tashkillashtirish texnikasi boʻlib, bu jarayon shu soʻzning boshqa soʻzshakllaridan kelib chiqqan holda amalga oshiriladi. Lemmatizatsiya morfologik tahlil metodi tarkibiga kiradi, u ikki bosqichni oʻz ichiga oladi: 1) deklarativ bosqich – bunda muayyan soʻzning mumkin boʻlgan barcha shakllari (soʻzshakllar) belgilanadi; 2) protsedura bosqichi – bunda soʻz asos va qoʻshimchalarga, yaʼni leksemalarga yoki morfemalarga boʻlinadi. Lemmatizatsiya soʻzlarning grammatik valentligi, qaysi affikslar bilan birika olish imkoniyatini ham belgilab beradi. Masalan, oʻzbek tilida soʻzlarning lugʻatdagi shakli – lemmasi quyidagicha:

- ot soʻz turkumi uchun – bosh kelishik, birlik shakli (*non, gul, hayot*);

- feʼl soʻz turkumi uchun – harakat nomi shakli (*yozmoq, kulmoq, ranjimoq*);

- sifat soʻz turkumi uchun – oddiy daraja shakli (*yaxshi, yashil, ogʻir*).

Gullar, gullarni, gullarga, gullarda, gullardan, gullarning; buyurmoq, buyurdi, buyurgan, buyuryapti, buyurmoqchi; qizil, qizgʻish, qizilroq².

Lemmatizatsiya tahlil davomida bir soʻz sifatida qaraladigan soʻzning turli flektiv shakllari guruhini namoyon etadi. Asos shakl *lemma* hisoblanadi. Lemmatizatsiya soʻz turkumlarini farqlash (idenfikatsiya) bilan bogʻliq, soʻzning qisqartmasidan to unga muvofiq keladigan soʻzgacha boʻlgan shakllarni oʻzida ifoda etadi. Aynan lemmatizatsiya tadqiqotchiga muayyan leksemaning barcha mumkin boʻlgan variantlarini kiritmasdan ajratishga va oʻrgatishga imkon beradi³.

¹ Захаров В.П., Богданова С.Ю. Корпусная лингвистика.-Иркутск:ИГЛУ,2011.-С.39.

² Poʻlatov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2007. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

³ Захаров В.П., Богданова С.Ю. Корпусная лингвистика.-Иркутск:ИГЛУ,2011.-С.39.

Lemmatizatsiya ham soʻzning asosini aniqlashdan iborat jarayon boʻlib, faqat unda berilgan soʻz shakli qaysi soʻz turkumiga tegishli ekanligi ilgaridan berilgan boʻladi. Masalan, stemmer *yozuv*, *yozuvning*, *yozish*, *yozmoq* soʻzlari uchun asos sifatida *yozni* oladi. Lemmatayzer esa feʼl shakllari *yozish*, *yozmoq*dan uchun asos deb *yozni*, ot turkumidagi *yozuv*, *yozuvning* soʻzlari uchun asos deb *yozuv* leksemasini aniqlaydi. Lemma tushunchasi leksemani ifodalaydi. Lemmatizatsiya masalasi esa bir leksemaga mos keluvchi soʻz shakllarini aniqlashdan iborat boʻladi. Lemmatayzer asosida *daraxtlarimning* soʻzi analiz qilinganda soʻz asosi *daraxt*, son qoʻshimchasi *-lar*, egalik qoʻshimchasi *-im*, kelishik qoʻshimchasi: *-ning* morfologik maʼlumotni ifodalaydi.

Elektron lugʻatlar tuzish jarayonida maʼlumotlar ombori, matnlar korpusini yaratish, qidiruv tizimi, kodlash, lingvistik va dasturiy taʼminot unsurlari qatoridagi bosqich hisoblanadi.

Parsing leksemalarning (soʻz, tokenlarning) chiziqli ketma-ketligini uning formal grammatikasi bilan chogʻishtirishdir. Natijada tobelik daraxti (sintaktik daraxt) yuzaga keladi. Katta hajmli korpuslar uchun avtomatik sintaktik analizatorlarning qurilishi kompyuter lingvistikasining muhim sohalaridan biri hisoblanadi¹.

Masina tahririda, asosan, imloviy xatolarni tekshirish koʻzda tutiladi va asosiy eʼtibor soʻzshakllarga qaratiladi. Ushbu maqsadga erishish uchun quyidagi ishlar amalga oshirilishi talab qilinadi: 1) tilning lugʻat tarkibini kompyuterga kiritish, yaʼni elektron lugʻatni tuzish; 2) lugʻatdagi adabiy tilga mansub soʻzlarni ajratib olish; 3) ajratib olingan soʻzlarni uslubiy (ilmiy, badiiy, publitsistik, rasmiy-idoraviy va soʻzlashuv) jihatdan guruhlash; 4) adabiy soʻzlarni turkumlarga ajratish; 5) soʻzlarni turkum doirasida guruhlash (masalan, ot turkumidagi soʻzlarni shaxs otlari, oʻsimlik nom (fitonim)lari, joy nom (toponim)lari kabi guruhlarga ajratish); 6) guruhlangan soʻzlarning affikslar kombinatsiyasini tuzish; bunda affikslar kombinatsiyasining amalda adabiy til doirasida foydalaniladigan hajmi qamrab olinadi. Kombinatsiyalar qatorini tuzishda affikslar ketma-ketligiga eʼtibor qaratiladi. M-n, ot turkumida *koʻplik+egalik +kelishik* kategoriyasi va h.k.²

Sintaktik tahlil bosqichi – matn tahririning eng qiyin bosqichi. Uning asosiy vazifasi soʻzlarning oʻrnini va ularning oʻzaro birikishini aniqlash

¹ Захаров В.П., Богданова С.Ю. Корпусная лингвистика.-Иркутск:ИГЛУ,2011.-С.41.

² Абжалова М. Ўзбек тилидаги матнларни тахрир ва тахлил қилувчи дастурнинг лингвистик модуллари (расмий ва илмий услубдаги матнлар тахрири дастури учун). Филол. фан...(PhD) дисс.-Тошкент, 2019.-Б.15.

hisoblanadi. Shu bois ushbu bosqichni *tahrir bosqichi* deb emas, balki *lingvistik tahlil bosqichi*, deyish mumkin. Bir soʻzshakl bir qancha soʻzning grammatik shakliga toʻgʻri kelishi mumkin. Bunday hollarda soʻzlarni farqlash uchun matn kirituvchiga shunday soʻzlar varianti koʻrsatiladi va matn teruvchi soʻzlardan muvofiqini tanlab qoʻyadi. Ushbu jarayon morfologik tahrirda ham kuzatiladi. Yaʼni soʻzning xato yozilgani aniqlangach, matn teruvchiga tanlab qoʻyish imkonini beruvchi soʻzlar variantini taklif etadi. Sintaktik tahlilda asosiy maqsad soʻzlarning oʻzaro bogʻlanish meʼyorlarini eʼtiborga olishdir.

Matnlarni grammatik tahrir qilishda (Grammar checking) kompyuterga kiritilayotgan matndagi gap strukturasi grammatik jihatdan toʻgʻri tuzilganligi tekshiriladi. Buning uchun dasturga gaplarning kombinatsiyalarini kiritish talab etiladi. Shunda dastur gapni kiritilgan koʻplab meʼyoriy kombinatsiyalar doirasida tekshiradi. Agar gap tuzilish kombinatsiyalariga muvofiq kelmasa, dastur uni xato deb hisoblaydi va foydalanuvchiga bu haqda xabar beradi. Grammar checking dasturida faqat grammatik meʼyorlar emas, balki semantika, soʻzlarning maʼnosi hamda mantiqiylikiga ham eʼtibor beriladi. Masalan, *Qush uchmoqda. Odam uchmoqda. Devor uchmoqda* gaplari Grammatik jihatdan toʻgʻri yozilgan, yaʼni ega+kesim munosabati mavjud (III shaxs, birlik). Ammo semantikasiga eʼtibor bersak, gaplar xato tuzilgan. Chunki odamning qanoti yoʻq, faqat badiiy matndagina insonning uchishi haqida yozish mumkin. Matnlarni kompyuter vositasida qayta ishlashda ilmiy, rasmiy-idoraviy, axborot doirasidagi matnlar eʼtiborga olinadi¹. Keltirilgan *devor uchmoqda* gapining semantikasi esa notoʻgʻri va gap hech bir uslubga mos kelmaydi. Chunki devor muqim bir joyda turuvchi, binoning, uyning tomini koʻtarib turuvchi yoki xonalarga ajratuvchi tik koʻtarilgan qismi².

Matnlarni semantik tahlil qilish bosqichida tushunchaga ega soʻzning oʻrinli qoʻllanilayotgani tekshirib boriladi. Buning uchun dastur taʼminotiga tezaurus lugʻatlarni kiritish talab qilinadi. Tezaurus lugʻatning oʻziga xos maxsus koʻrinishi boʻlib, unda lugʻaviy birliklar oʻrtasida semantik munosabat (sinonimlar, antonimlar, paronimlar, giponimlar, giperonimlar va boshqalar) koʻrsatilgan boʻladi.

Zamonaviy texnologiyalarning kundalik turmushda muhim oʻrin egallashi, ilm-fanning rivojlanishi avtomatik tahrir va tahlilning

¹ Абжалова М.А. Ўзбек тилидаги матнларни автоматик таҳрирлашда услуб масаласи// Таълим тизимидаги ижтимоий-гуманитар фанлар, №2, Т.: 2013, 141-145-бетлар; L. Abduhamidova . Kompyuter lingvistikasining avtomatik tahrir yoʻnalishi. BML.-T.,2016. 33-36

² Қаранг: Ўзбек тилининг изоҳли луғати. 1-жилд. Т., 2006., 588-б.

ahamiyatini oshirmoqda. Avtomatik tahrir faqat matnlardagi mexanik xatolarni bartaraf etibgina qolmay, matnning savodli yozilishini ham ta'minlaydi, matn kirituvchiga jumlada xato so'z borligi haqida xabar berib, xatoni bartaraf etishga undaydi, natijada to'liq lingvistik me'yorlar darajasidagi matnni yuzaga keltirishda vaqt tejiladi.

So'zni turkumlarga bo'lishda ikki umumiy super kategoriya mavjud: yopiq sinf va ochiq sinf. Yopiq sinflar boshqa a'zolar bilan aloqa o'rnatgan. Masalan, predloglar yopiq sinf hisoblanadi, chunki ana shu aloqalar natijasida o'rnatilgan, yangi predloglar kamdan-kam hollarda hosil bo'ladi. Bundan farqli ravishda, otlar va fe'llar ochiq sinf hisonlanadi, chunki yangi otlar va fe'llar davomiy tarzda paydo bo'lib boradi yoki boshqa tillardan o'zlashtiriladi.

Dunyo tillarida to'rta katta ochiq guruh mavjud: otlar, fe'llar, sifatlar va ravishlar.

So'zlarni turkumlarga ajratish algoritmi so'zlarni bir tizimga tizilishi va ma'lum bir tartibda belgilanishidir. Ingliz tilida so'zlarni turkumlarga ajratishda Brown korpusi (1979) dastlabki asosiy ajratish namunasi bo'lib, turli janrlarda yozilgan 500 dan ortiq matndan 1 million ta so'zlar to'plamini o'z ichiga oladi. Bu korpus birinchi turkumlarga ajratishda dastlabki tadqiqotlardan biri hisoblanadi. Unda 87 ta ajratilgan to'plam berilgan. Keyinchalik 45 ta ajratilgan to'plamga ega Penn Treebank korpusi (Marcus, 1993), Britaniya Milliy Korpusi (British National Corpus, 1997) kabi so'z turkumlariga ajratishdagi to'plamlar berilgan.

Ba'zi turkumlarga ajratishdagi farqlar inson uchun ham, mashina uchun ham murakkabdir. Masalan, predloglar, yordamchi so'zlar, ravishlar katta o'xshashlikka ega. Masalan, *around* so'zi yuqorida sanalgan uchta vazifada ham kelishi mumkin:

1. Mrs. Shafaer never got **around** (particle) to joining.
(Mister Shafaer birlashtirishda ortga qaytmagan).
2. All we gotta do is go **around in** (determiner) the corner.
3. Chateau Petrus costs **around** (adverb) 250.
(Chateau Petrusi taxminan 250 narxlanadi).

Ajratish (Tag)	Tavsif (Description)	Namuna
CC	Coordin. Conjunction	<i>and, but, or</i>
CD	Cardinal number	<i>one, two, three</i>
DT	Determiner	<i>a, the</i>
EX	Existential 'there'	<i>There</i>
FW	Foreign word	<i>Mea culpa</i>
IN	Preposition/sub-conj	<i>of, in, by</i>

JJ	Adjective	<i>Yellow</i>
JJR	Adj., comparative	<i>Bigger</i>
JJS	Adj., superlative	<i>Wildest</i>
LS	List item marker	<i>1, 2, One</i>
MD	Modal	<i>Can, should</i>
NN	Noun, sing. or mass	<i>Ilama</i>
NNS	Noun, plural	<i>Ilamas</i>
NNP	Proper noun, singular	<i>IBM</i>
NNPS	Proper noun, plural	<i>Carolinas</i>
PDT	Predeterminer	<i>all, both</i>
POS	Possessive ending	<i>'s</i>
PRP	Personal pronoun	<i>I, you, he</i>
PRP\$	Possessive pronoun	<i>your, one's</i>
RB	Adverb	<i>quickly, never</i>
RBR	Adverb, comparative	<i>Faster</i>
RBS	Adverb, superlative	<i>Fastest</i>
RP	Particle	<i>up, off</i>
SYM	Symbol	<i>+, %, &</i>
TO	“to”	<i>To</i>
UN	Interjection	<i>ah, oops</i>
VB	Verb, base form	<i>Eat</i>
VBD	Verb, past tense	<i>Ate</i>
VBG	Verb, gerund	<i>Eating</i>
VBN	Verb, past participle	<i>Eaten</i>
VBP	Verb, non-3sg pres	<i>Eat</i>
VBZ	Verb, 3sg pres	<i>Eats</i>
WDT	Wh-determiner	<i>which, that</i>
WP	Wh-pronoun	<i>what, who</i>
WP\$	Possessive wh-	<i>Whose</i>
WRB	Wh-adverb	<i>how, where</i>
\$	Dollar sign	<i>\$</i>
#	Pond sign	<i>#</i>
“	Left quote	<i>(‘ or “)</i>
”	Right quote	<i>(’ or ”)</i>
(Left parenthesis	<i>([, (, { , <</i>
)	Right parenthesis	<i>(],) , } , ></i>
,	Comma	<i>,</i>
.	Sentence-final punc	<i>(. ! ?)</i>

:	Mid-sentence punc	(: ; ... --)
Penn Treebank so‘z turkumlariga ajratish korpusi (puntuatsiyani o‘z ichiga olgan holda). ¹		

Bunday imkoniyatga erishish uchun, albatta, o‘zbek tili asosida ko‘p bosqichli avtomatik tahrir dasturining lingvistik ta‘minotini mukammal darajada ishlab chiqish talab etiladi.

Matn mazmunini o‘zlashtirishda *BORIS* tizimining o‘rni

Reja:

1. Avtomatik tahrir jarayonida avtomatik tizimlardan foydalanish.
2. Kompyuter lingvistikasida mavjud avtomatik tizimlar.
3. *BORIS* tizimining ishlash prinsipi.

Tayanch tushunchalar: *kompyuter lingvistikasi, BORIS tizimi, kompyuter dasturlari ishlash prinsipi, savol-javob tizimi, start, METAKOP (sabab zanjiri), PTRANS (ma‘lum bir obyektning bir joyda ikkinchi joyga ko‘chishi), ATRANS (ma‘lum bir shaxsning muayyan narsaga ega bo‘lishi), ATTEND (voqeaga turtki bo‘lgan holat), PROPEL (ma‘lum narsaga erishish uchun harakat qilish), ATB (affektiv mazmunli bloklar, insonning psixologik holati).*

Insoniyatning rivojlanish tarixida qo‘lga kiritgan barcha yutuqlari zamirida hayotiy faoliyatini yaxshilash, qisqa fursatlarda ko‘p ishlarni amalga oshirish, aql-zakovatining tuganmas imkoniyatlarini namoyon qilish, jamiyat manfaatlarini uchun xizmat qilish maqsadi yotadi. “Kompyuter texnologiyalari” deb nom olgan asrimizda fan-texnika yutuqlari, deyarli, barcha sohaning ajralmas qismiga aylanib ulgurdi. Jumladan, tilshunoslik sohasi ham ayni shu jarayondan mustasno emas.

Matematika, informatika, lingvistika fanlarining kesishgan nuqtasida shakllangan yangi soha – kompyuter lingvistikasini axborot texnologiyalarisiz, kompyuter dasturlarisiz tasavvur qilib bo‘lmaydi. Endilikda bu soha lingvistika oldiga yangi talablarni qo‘yib, fan doirasida turli yo‘nalishlarning paydo bo‘lishiga zamin tayyorlamoqda.

Dunyo afkor ommasi tez fursatlarda samaradorligi yuqori bo‘lgan holatda ma‘lumot olishni xohlaydi. Qaysi manbaga murojaat qilmaylik, unda ifodalangan asosiy fikrni tezlikda anglash, matn mazmunini o‘zlashirish diqqat markazida turadi.

¹*Daniel Jurafskiy & James H. Martin. Speech and Language Processing: An introduction to natural language processing, computational linguistics and speech recognition. Qarang: Abduhamidova L. Kompyuter lingvistikasining avtomatik tahrir yo‘nalishi. BMI.-T., 2016.-B.36-38.*

START dunyodagi birinchi savol-javob tarzida ishlaydigan on-line boshqaruv tizimi hisoblanadi. Bu tizim 1993- yil dekabr oyidan boshlab doimiy ravishda qo'llana boshlagan. BORIS Katzvauning hamkorlari Info Lab Group tomonidan rivojlantirilgan. Bu tizimning maqsadi foydalanuvchilarga "soniyada to'g'ri ma'lumot"ni yetkazish bo'lgan. Bu tizim, dastlab, milliondan ortiq joylar (shaharlar, poytaxtlar, ko'llar, koordinatalar, ob-havo, xaritalar, demografik holat, siyosiy va iqtisodiy tizim), filmlar (nomlari, aktyorlari, rejissyorlari), odamlar (tug'ilgan yili, joyi), lug'atlar haqidagi savollarga javoblarni o'z ichiga qamrab olgan. Ular shartli ravishda geografiya, madaniyat va san'at, ilm-fan kabi tarmoqlarga bo'lingan.

Namuna tariqasida quyidagilarni keltiramiz.

1. Q: How far is Mount Kilimanjaro from Mount Everest?

A: The distance between Kathmandu, the capital of Nepal (where Mount Everest is located) and Dar Es Salaam, the capital of Tanzania (where Mount Kilimanjaro is located) is 3,892 miles (6,277 kilometers).

Source: START KB

S: Klimanjaro tog'i Everest tog'idan qancha masofa uzoqlikda?
J'

: Nepal (Everest tog'i joylashgan) poytaxti Katmandu va Tanzaniya (Klimanjaro tog'i joylashgan) poytaxti Dodoma orasidagi masofa 3,892 mil (6,277 kilometr)

Manba: START KB

2. Q: Show the capital of the second largest country in Asia.

A: I know that the the second largest country in Asia is India (source: The World Factbook).

Using this information, I determined the capital of India:

New Delhi is the capital of India.

Source: START KB

S: Osiyoda kattaligi jihatdan ikkinchi o'rinda turuvchi mamlakat poytaxtini ko'rsating.

J: Osiyodagi kattaligi jihatdan ikkinchi o'rinda turuvchi mamlakat Hindiston ekanligini bilaman (manba: The World Factbook). Shu ma'lumotdan foydalanib, aytishim mumkin: Yangi Dehli Hindistonning poytaxti.

Manba: START KB.

¹ Bu o'rinda inglizcha Q, o'zbekcha C belgilari "savol", inglizcha A , o'zbekcha J delgilari "javob" tushunchasini ifodalalaydi.

Axborot uzatishda faqat xabar olish emas, balki xabar atrofidagi tegishli ma'lumotlarni ham o'zlashtirish, bilish lozim bo'ladi. Shu ma'noda, jahon tilshunosligida kompyuter tizimi asosida bir qator dasturlar yaratilganini ta'kidlash joiz. Ana shunday dasturlardan biri ingliz tili asosidagi BORIS tizimidir.

BORIS – darak gaplardan iborat matnlarni avtomatik tahlil qilish tizimi. BORIS matni tushunadigan va savollarga javob beradigan eksperimental tizim, u bir qator bilim manbasiga asoslanadi. Matni uning asosiy mazmuninigina o'zlashtirib, qolgan jihatlarini e'tibordan chetda qoldiradigan yuzaki tizimlarga zid o'laroq, BORIS tizimi matni chuqur tushunishga qaratilgan. Matn qismlarini BORIS dasturi asosida qayta ishlash natijasida mazmuni chuqur o'zlashtirish imkoni yaratiladi.

BORIS savol-javob asosida amalga oshadigan dastur bo'lib, matn haqida axborot olishga, mazmunini chuqur o'zlashtirishga xizmat qiladi. BORIS tizimida mashinaning o'zi dastur asosida savol qo'yadi, javob beradi va matni tahlil qiladi. Tizim javob berishi mumkin bo'lgan savollar botiniy matn talqiniga aloqador bo'lgan, fikr asosida turuvchi tasavvurdagi sabab-oqibat munosabatlarini aks ettiradi.

BORIS tizimi asosida tahlil qilingan va “Новое в зарубежной лингвистике” seriyasi XXIV sonida e'lon qilingan rus tilidagi matn aslida ingliz tilidan tarjima qilingan. Matni shu tilda berishni ma'qul topdik.

(BORIS РАЗВОД-1)

ОБРАБАТЫВАЕТСЯ АБЗАЦ 1

Ричард несколько лет не имел никаких известий о Поле, своем товарище по комнате во время учебы в университете. В то время Ричард одолжил у Поля некоторую сумму, которая так и не была возвращена. Теперь же он совершенно не представлял себе, где он может найти своего старого друга. Когда вдруг из Сан-Франциско пришло письмо, Ричард загорелся желанием узнать, как живет Поль.

СИСТЕМА ГОТОВА К РАБОТЕ В РЕЖИМЕ В/О

Какое событие произошло у Ричарда дома?¹

РИЧАРД ПОЛУЧИЛ ПИСЬМО ОТ ПОЛЯ.

Кто такой Поль?

ДРУГ РИЧАРДА.

¹ Dastur tizimi asosidagi savollar kichik harflarda, javoblar katta harflarda ifodalandi.

Хотел ли Ричард видеть Поля?
ДА, РИЧАРД ХОТЕЛ ЗНАТЬ, КАК ЖИВЕТ ПОЛЬ.
Помогал ли когда –либо Поль Ричарду?
ДА, ПОЛЬ ДАЛ РИЧАРДУ ДЕНЬГИ В ДОЛГ.
Почему Ричард не вернул долг Полю?
РИЧАРД НЕ ЗНАЛ, ГДЕ НАХОДИТСЯ ПОЛЬ.
Что почувствовал Ричард, когда пришло письмо?
РИЧАРД БЫЛ ОБРАДОВАН, ПОТОМУ ЧТО ПОЛЬ И ОН
БЫЛИ ДРУЗЬЯМИ.

ОБРАБАТЫВАЕТСЯ АБЗАЦ 2

К сожалению, новости не были хорошими. Жена Поля Сара настаивала на разводе. Она хотела также получить машину, дом, взять детей и добиться алиментов. Поль хотел развода, но он не хотел, чтобы к Саре перешло все их имущество. Его жалование в государственной школе было очень мало. Не зная, к кому обратиться, он надеялся на содействие единственного знакомого ему адвоката. Поль указал номер своего домашнего телефона на тот случай, если Ричард сочтет, что может быть ему полезным.

СИСТЕМА ГОТОВО К РАБОТЕ В РЕЖИМЕ В/О

Что сообщалось в письме?
ПОЛЬ И САРА ЗАТЕЯЛИ ДЕЛО О РАЗВОДЕ.
Написал ли Поль письмо Саре?
НЕТ, НЕ САРЕ. ПИСЬМО ПОЛУЧИЛ РИЧАРД.
Почему Поль был расстроен в связи с разводом?
ПОЛЬ И САРА СПОРИЛИ ПО ПОВОДУ СЕМЕЙНОГО
ИМУЩЕСТВА.
Чем Поль зарабатывал себе на жизнь?
ПОЛЬ БЫЛ УЧИТЕЛЕМ.
Чего хотела Сара?
САРА ХОТЕЛА ПОЛУЧИТЬ МАШИНУ И ДОМ, ВЗЯТЬ
МАЛЫШЕЙ И ДОБИТЬСЯ АЛИМЕНТОВ.
Почему Поль написал Ричарду?
ПОЛЬ ХОТЕЛ, ЧТОБЫ РИЧАРД БЫЛ ЕГО АДВОКАТОМ.

ОБРАБАТЫВАЕТСЯ АБЗАЦ 3

Ричард нетерпеливо схватил трубку телефона и набрал номер. После короткого разговора Поль договорился с ним позавтракать вместе на следующий день. В его голосе слышались явное облегчение и благодарность.

СИСТЕМА ГОТОВО К РАБОТЕ В РЕЖИМЕ В/О

Почему Ричард позвонил Полю?

РИЧАРД ХОТЕЛ ДОГОВОРИТЬСЯ С ПОЛЕМ О ВСТРЕЧЕ.

Установил ли Ричард связь с Полем?

ДА, РИЧАРД ПОЗВОНИЛ ПОЛЮ.

Как Ричард установил связь с Полем?

ПО ТЕЛЕФОНУ.

Что почувствовал Поль, когда позвонил Ричард?

ПОЛЬ БЫЛ ДОВОЛЕН, ТАК КАК РИЧАРД СОГЛАСИЛСЯ БЫТЬ АДВОКАТОМ ПОЛЯ.

ОБРАБАТЫВАЕТСЯ АБЗАЦ 4

На следующий день, когда Ричард ехал на машине в ресторан, он едва не сбил на улице старика. Он чувствовал себя совершенно потрясенным после этого происшествия и в ресторане выпил три рюмки спиртного. Когда пришел Поль, Ричард был изрядно пьян. Когда принесли еду, Ричард пролил чашку кофе на Поля. Это, видимо, сильно раздосадовало Поля, и поэтому Ричард предложил отвезти его домой, чтобы он мог переодеться.

СИСТЕМА ГОТОВА К РАБОТЕ В РЕЖИМЕ В/О

Почему Ричард завтракал вместе с Полем?

РИЧАРД ХОТЕЛ ОБСУДИТЬ С ПОЛЕМ ДЕЛО О РАЗВОДЕ.

Что случилось с Ричардом на пути в ресторан?

РИЧАРД ЕДВА НЕ ЗАДАВИЛ СТАРИКА.

Почему Ричард напился?

РИЧАРД БЫЛ ПОТрясен тем, что он едва не задавил СТАРИКА.

Что случилось в ресторане?

РИЧАРД ПРОЛИЛ КОФЕ НА ПОЛЯ.

Что почувствовал Ричард, когда пролил кофе?

РИЧАРД ПОЧУВСТВОВАЛ СЕБЯ ВИНОВАТЫМ.

Почему Ричард пролил кофе?

РИЧАРД БЫЛ ПЬЯН.

Когда Ричард едва не сбил старика?

КОГДА РИЧАРД ЕХАЛ НА МАШИНЕ В РЕСТОРАН.

Где Ричард завтракал с Полем?

В РЕСТОРАНЕ.

ОБРАБАТЫВАЕТСЯ АБЗАЦ 5

Когда Поль вошел в спальню и обнаружил там Сару с другим мужчиной, его едва не хватил удар. Затем он сообразил, что это невероятная удача. Поскольку рядом был свидетель в лице Ричарда, то дело о разводе для Сары было очевидным образом проиграно. Ричард поздравил Поля и предложил отпраздновать удачу и пообедать вместе. Поль охотно согласился.

СИСТЕМА ГОТОВА К РАБОТЕ В РЕЖИМЕ В/О

Почему Ричард отвез Поля домой?

ЧТОБЫ ПОЛЬ МОГ ПЕРЕМЕНИТЬ ОДЕЖДУ.

Что произошло у Поля дома?

ПОЛЬ ЗАСТАЛ САРУ В МОМЕНТ НАРУШЕНИЯ
СУПРУЖЕСКОЙ ВЕРНОСТИ.

Что почувствовал Поль?

ПОЛЬ БЫЛ ПОРАЖЕН.

Почему Сара проиграла дело о разводе?

САРА НАРУШИЛА СУПРУЖЕСКУЮ ВЕРНОСТЬ.

Почему Ричард поздравил Поля?

ПОЛ ВЫИГРАЛ ДЕЛО О РАЗВОДЕ¹.

Tahlil qilingan matndan quyidagilar asosida qo‘shimcha bilim hosil qilinadi:

КОР (aloqadorlik): qarzdorlik;

МЕТАКОР (sabab zanjiri): ajrim;

¹ Венди Ленерт, М.Г.Дайер, П.Н.Джонсон, К.Дж.Яна, С.Харли. BORIS-экспериментальная система глубинного понимания повествовательных текстов. «Новое в зарубежной лингвистике». – М., 1989. – С. 107-159. Настоящая работа отчасти финансировалась Агентством перспективных исследовательских проектов (ARPA), а отчасти- Национальным научным фондом США.

PTRANS (ma'lum bir obyektning bir joydan ikkinchi joyga ko'chishi) : restoranga, restorandan uyga borish;

ATRANS (ma'lum bir shaxsning muayyan narsaga ega bo'lishi, nazorat qilishi): Polning mol-mulkka ega bo'lishi;

ATTEND (voqealarga turtki bo'lgan holat): qahfaning to'kilib ketishi;

PROPEL (ma'lum narsaga erishish uchun harakat qilish): Polning uy-joy, mashina, mol-mulkning ma'lum qismiga da'vo qilishi);

ATB (affektiv mazmunli bloklar, insonning psixologik holati):

1. Mamnunlik.
2. Asabiylashish.
3. G'azablanish.
4. To'satdan anglash.

O'zbek tilshunosligida kompyuter lingvistikasiga oid tadqiqotlarda matni BORIS tizimi asosida tahlil qilishga qaratilgan lingvistik ta'minot shakllantirilmoqda. Shulardan biri borasida to'xtalamiz.

PECHENYE O'G'RISI

Birinchi abzasni qayta ishlash:

Bir ayol aeroportda samolyot uchishini kutardi. Hali vaqt bor edi. Bekor o'tirishni istamagan ayol do'kondan bitta kitob va pechenye sotib oldi. Bo'sh kursilardan biriga o'tirib, kitobini o'qiy boshladi. Kitobni berilib o'qir ekan, nogoh yonida o'tirgan odam pechenyesidan olishga jur'at qilayotganini sezib qoldi. Ammo birgina pechenye uchun betakalluflik qilgisi kelmadi.

Tizim savol-javob rejimida ishlashga o'tadi:

- *Kim aeroportda samolyot uchishini kutardi?*
- Bir ayol.
- *Ayol do'kondan nimalar sotib oldi?*
- Bitta kitob va pechenye.
- *Ayol kitobni o'qiyotganida yonida o'tirgan odam nima qildi?*
- Ayolning pechenyesidan olishga jur'at etdi.

Ikkinchi abzasni qayta ishlash:

Dam-badam u ham soatga qarab qo'yardi. Ayol kitob o'qib, bittadan pechenyeni o'g'rincha olayotgan odamga ko'z qirini tashlar ekan, asabiylashmaslikka urinardi. Pechenye o'g'risiga e'tibor bermaslikka va soat millari yurishini kuzatib, uning chiqqillashini eshitishga harakat

qilardi. O‘ziga o‘zi: “Agar madaniyatli bo‘lmaganimda, shu odamning ko‘zlarini ko‘kartirardim”, – deb qo‘ydi. Har gal pechenye olaman, deganida yonidagi odam ham qo‘lini uzatardi. Oxiri paketda bitta pechenye qolgach, “Qani endi nima qilasan?” – dedi ichida. Odam istehzoli kulimsirab, oxirgi pechenyeni oldi va ikkiga bo‘lib, yarmini og‘ziga soldi; yarmini esa ayolga uzatdi. Ayol jahl bilan pechenyeni olar ekan, “Bu qanday surbetlik! Bunchalar qo‘pol bo‘lmasa, bu odam. Bir og‘iz rahmat ham demadi!” – deb gapirib ketdi o‘z-o‘ziga.

Tizim savol-javob rejimida ishlashga o‘tadi:

- *Ayol o‘zini qanday tutar edi?*
- Soatga qarab qo‘yardi.
- *Kim “Agar madaniyatli bo‘lmaganimda, shu odamning ko‘zlarini ko‘kartirardim”, deb aytadi?*
- Ayol o‘ziga-o‘zi.
- *Paketda bita pechenye qolgach, ayol ichida nima dedi?*
- “Qani endi nima qilasan?”
- *Pechenye olayotgan odam qanday harakat qildi?*
- Oxirgi pechenyeni ikkiga bo‘lib, yarmini og‘ziga solib, yarmini ayolga berdi.

Uchinchi abzasni qayta ishlash:

Ayol hech qachon bunday asabiylashmagandi. Samolyotning uchish vaqti yetgach, chuqur tin oldi. Yuklarini olib, eshik tomon yurdi. “Pechenye o‘g‘risi”ga qayrilib ham qaramadi. Samolyotga chiqdi va o‘rniga joylashdi. “Pechenye o‘g‘risi”dan qutulganiga shukr qildi. Biroq, uning adabini bermaganiga afsuslandi. “Janob, – deyishi kerak edi u. – Shukr qiling, sizdek betakalluf emasman. Aks holda yuzingizga shapaloq tushirgan bo‘lardim. Ammo, siz qilgan ish rasman o‘g‘rilik sanaladi. Pechenye o‘g‘riligi... Bilingizmi?!”

Tizim savol-javob rejimida ishlashga o‘tadi:

- *Ayol samolyotning uchish vaqti kelgach nima qildi?*
- Chuqur tin oldi.
- *Ayol nimaga shukr qildi?*
- Pechenye o‘g‘risidan qutulganiga.
- *Nimaga afsuslandi?*
- Uning adabini bermaganiga.

To‘rtinchi abzasni qayta ishlash:

Shularni xayol qilib, qo‘lini sumkaga uzatdi. Chala qolgan kitobini olib o‘qimoqchi bo‘ldi. Ammo, ne ko‘z bilan ko‘rsinki, sumkasida do‘kondan sotib olgan bir paket pechenye turardi. Bir og‘iz ham gapirmay, eng oxirgi pechenyesini ham u bilan baham ko‘rgan odamdan borib uzr so‘rashga endi kech edi...

“Qo‘pol va jur‘atli pechenye o‘g‘risi”ning samolyoti osmonga ko‘tarilgan edi...

Tizim savol-javob rejimida ishlashga o‘tadi:

- *Ayolning sumkasida nima bor edi?*
- *Do‘kondan sotib olgan bir paket pechenyesi.*
- *Kimning samolyoti osmonga ko‘tarildi?*
- *“Qo‘pol va jur‘atli pechenye o‘g‘risi”ning .*

Yuqoridagit matnni quyidagicha sistemalashtirish mumkin:

METAKOP (sabab zanjiri): o‘g‘rilik

●PTRANS (ma‘lum bir obyektning bir joydan ikkinchi joyga ko‘chishi): samolyotning uchishi;

●ATRANS (ma‘lum bir shaxsning muayyan narsaga ega bo‘lishi, nazorat qilishi): ayolning yuklariga egalik qilishi;

●ATTEND (voqealarga turtki bo‘lgan holat): pechenye sotib olish;

●PROPEL (ma‘lum narsaga erishish uchun harakat qilish): Sumkadan kitobning olinishi.

ATB (affektiv mazmunli bloklar, insonning psixologik holati):

2. Asabiylashish.

3. Shukr qilish.

4. Afsuslanish.

5. Kulimsirash¹.

Xulosa qilib aytganda, BORIS tizimini takomillashtirish komyuter lingvistikasining istiqbollaridan biri bo‘lib, matn mazmunini o‘zlashtirishda yuqori samaradorligi bilan ahamiyatlidir.

Savol va topshiriqlar

1.Avtomatik tahrir yo‘nalishi haqida ma‘lumot bering.

2.Avtomatik tahrir yo‘nalishi rivojlangan tillar haqida bilganlaringizni ayting.

¹ *Abduhamidova L.* Komyuter lingvistikasining avtomatik tahrir yo‘nalishi. BMI.-T., 2016.-B.36-38.

3. “Start” tizimi haqida so‘zlang.
4. “BORIS tizimi” haqida tushunchangizni bayon eting.
5. Matn tanlang va “BORIS tizimi” asosida tahlil qiling.
6. Imlo qoidalarini o‘zlashtirish va avtomatik tahrir yo‘nalishi munosabatini yoriting.
7. Avtomatik tahrir yo‘nalishining ahamiyati haqida fikrlaringizni bayon eting.

14-§. KOMPYUTER LINVODIDAKTIKASI

Reja:

1. Kompyuter texnologiyalari yordamida dars berishning didaktik konsepsiyasi.
2. Elektron darslik yaratish tamoyillari.
3. Lug‘atlardan o‘quv materialini sifatida foydalanish.

Asosiy tushunchalar: *lingvodidaktika, elektron darslik, elektron-multimedia darsliklari, taqdimot texnologiyasi, elektron nusxa, o‘rgatuvchi audiodasturlar, elektron lingafon kurslari, elektron virtual kutubxona, elektron lug‘atlar, kompyuter lug‘atlari, tezaurus lug‘atlardan o‘quv materialini sifatida foydalanish, tezaurus –elektron manba sifatida, olamining lisoniy manzarasini o‘zlashtirish darajasi, tilning aktual holati, lingvomadaniy tajriba.*

“Kompyuter” lingvodidaktikasi tushunchasini yoritish uchun, avvalo, “lingvodidaktika” terminiga to‘xtalish lozim bo‘ladi.

Lingvodidaktika (lot. *lingua* – “til”+ grek. *didaktikos* – “ta’lim”) til o‘rgatishning umumiy nazariyasi. Termin akad. N. M. Shanskiy tomonidan tilni o‘quv maqsadida tavsiflash jarayonida kiritilgan. Tillarning muvofiq keluvchi va farqli jihatlarini tadqiq qilish, o‘rganilayotgan til strukturasini tahlil qilish, lingvistika va pedagogika chegarasida turgan muammolarni hal qilish lingvodidaktikaning asosiy vazifasidir.

Uzluksiz ta’lim tizimida til o‘rgatishga alohida e’tibor beriladi. Zamonaviy bozor iqtisodiyoti va axborot texnologiyalari taraqqiyoti sharoitida chet tilini o‘rganishga ehtiyoj tug‘iladi. Xorijiy tillarni bilish, madaniyatlararo muloqot vositalarini o‘rganish davr talabiga aylandi. 2012- yil 10-dekabrda O‘zbekiston Respublikasi birinchi Prezidenti qabul qilgan “Chet tillarni o‘rganish tizimini yanada takomillashtirish chora-

tadbirlari to‘g‘risida”¹gi Qarori bu sohaning istiqbolini belgilovchi davlat hujjatidir.

Lingvodidaktika til o‘rgatish, aniqrog‘i xorijiy tillarni o‘rganish jarayonini ifoda etish uchun qo‘llanmoqda. Lingvodidaktika va metodika terminlarini sinonim tarzda baholash hollari ham mavjud.

Xorijiy tillarni o‘rganish metodikasi XX asr o‘rtalarida shakllangan mustaqil fandır. Metodikani didaktikaning pedagogika, pedagogik psixologiya oraliq‘ida turuvchi bir qismi sifatida baholagan I.L.Bim metodikaning obyekt sifatida dastur, darslik, o‘quv qo‘llanma, o‘qituvchi va o‘quvchi faoliyatini hamda ular o‘rtasidagi o‘zaro hamkorlikning asosini tashkil etadigan shakllarni o‘z ichiga oluvchi ta‘lim-tarbiya jarayonini keltiradi.²

A.A.Mirolyubov xorijiy tillarni o‘qitish metodikasining nafaqat pedagogika, didaktika, pedagogik psixologiya, balki lingvistika va yosh pedagogikasi kesimida rivojlanishini ta‘kidlagan edi. Bu sohani o‘rganish bilan bog‘liq masalalar va taraqqiyot muammolari metodikaning predmetini tashkil etadi. Xorijiy tillarni o‘rganish metodikasi muammolarini A.A. Mirolyubov to‘rtga bo‘ladi:

1) xorijiy tilni o‘quv predmeti sifatida belgilash, ya‘ni xorijiy tilni o‘rganish maqsadi va mazmunini aniqlash bilan bog‘liq muammolar;

2) o‘qituvchi faoliyati bilan bog‘liq muammolar: ta‘lim va tarbiyaning metod va usullarini ishlab chiqish, ta‘lim vositalarini shakllantirish va ulardan foydalanish, ta‘lim jarayonini tashkil etish shakllari, ta‘limda erishilgan yutuqlarni nazorat qilish;

3) o‘quvchi faoliyati bilan bog‘liq muammolar: ta‘lim vositalari va o‘quvchilarni tarbiyalash va bilim olishning samaradorligi;

4) metodika fanlarini rivojlantirishga yo‘naltirilgan muammolar: fan tarixi, uning hamkor fanlar bilan aloqadorligi, tadqiqot metodlari.³

Yuqoridagilardan ma‘lum bo‘ladiki, xorijiy tillarni o‘qitish metodikasini xorijiy tilga o‘rgatish usullari qiziqtiradi, ya‘ni N.D. Galskova ta‘kidlaganidek, “metodika fan sifatida tilni o‘qitish va o‘rgatish jarayonini tadqiq qiladi”.⁴

¹ PQ-1875-son. 2012- yil 10- dekabr, Toshkent sh.

² Бим И.Л. Методика обучения иностранным языкам как наука и проблемы школьного учебника: опыт системно-структурного описания. - М.: Русский язык, 1977. –С.21.

³ Миролюбов А.А. История отечественной методики обучения иностранным языкам. - М: СТУПЕНИ, ИНФРА-М, 2002. –С.12-13.

⁴ Гальскова Н.Д. Еще раз о лингводидактике / Н.Д. Гальскова // Иностранные языки в школе. 2008. - № 8. - С. 2.

Bunda o'quvchilarning tilga munosabati, yangi lingvomadaniyat doirasidagi muloqotni e'tiborga olish talab qilinadi. Shuning uchun xorijiy tillarni o'rgatish metodikasi maqsadini o'zga til qurilishi tizimini nafaqat shakliy tuzilish – kodlar asosida, balki shu tildagi olam manzarasini anglash va mental xususiyatlarni o'zlashtirish negizida o'rganish usullari tashkil etadi.

Lingvodidaktika G.I.Bogin, I.I.Xaleeva, N.D.Galskova, G.I.Boginlar qarashlari asosida shakllangan yangi soha bo'lib, o'zida lingvistika fani mohiyatini namoyon etadi. I.I. Xaleeva «lingvodidaktika» terminini «ta'lim, o'quv mazmuni komponentlarini asoslovchi, til tabiati va tilning sotsial fenomeni sifatidagi individning o'zaro harakatlari mexanizmi» deb ta'riflaydi¹.

Lingvodidaktika til falsafasi, psixologiya, lingvistika, psixolingvistika, til o'zlashtirish nazariyasi, madaniyatlararo muloqot va boshqa fanlardan farqli o'laroq metodikaning o'rganish masalalariga kirmagan quyidagi muammolarni hal qiladi:

1) turli o'quv sharoitlarida o'rganilayotgan tilni va til sohiblari olamining lisoniy manzarasini o'zlashtirish darajasini izohlash;

2) tilning aktual holati va taraqqiyoti bilan belgilanuvchi lingvomadaniy tajribani o'rganish va o'zlashtirishni analiz qilish;

3) yo'l qo'yilgan (madaniy, lingvoo'lkashunoslikka oid va tilning o'ziga xos) kamchiliklar tabiatini asoslash;

4) til xususiyatlarini o'rganuvchilarning ko'p tillilik, ko'p madaniyatlilik, yosh va individual xususiyatlari kontekstida tadqiq qilish;

5) tilni mukammal va nomukammal egallashni belgilovchi omillarni analiz qilish va asoslash.²

Keyingi yillarda lingvodidaktika kompyuter tizimida faol qo'llanmoqda. Lingvodidaktikaga doir anjumanlar tashkil qilinmoqda.³

4.Kompyuter lingvodidaktikasi va ta'lim jarayonida kompyuter texnologiyalaridan foydalanishning ahamiyati.

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

¹Xaleeva I.I. Основы теории обучения пониманию иноязычной речи (подготовка переводчиков): Моногр.– М.: Высш. школа, 1989.–С.199.

² Гальскова Н.Д. Еще раз о лингводидактике / Н.Д. Гальскова // Иностранные языки в школе. 2008. - № 8. - С. 8-9.

³Тил системаси ва hozirgi zamon lingvodidaktikasi. Республика илмий-амалий анжумани материаллари (Самарқанд шаҳри, 2011 йил, 25-26 март). – Самарқанд: СамДЧТИ нашри, 2011. – 270 б.

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 asosiy yo'nalishga aylanib qolmoqda.

Til o'rganishning yana bir muhim usullaridan biri kompyuter texnologiyalari yordamida amalga oshiriladi. Fan-texnika rivoji natijasida til o'rgatish uchun maxsus kompyuter dasturlari yaratildi. Ularda fe'llar, zamonlar, modal fe'llar tavsifi va zaruriy gap qoliplari keltirilgan".¹ Zamonaviy ta'lim texnologiyasida til o'rgatuvchi kompyuter dasturlari muhim o'rin tutadi. Bu dasturlarning ayrimlari elektron o'qituvchi bilan ta'minlangan.

O'qituvchi lingvistik avtomatlardan foydalanishda ikki xil yondashuv mavjud:

1) bixevioristik yondashuv – mavjud an'anaviy o'qitish tizimining barcha jabhalarini e'tiborga oluvchi o'qitish mexanizmlari, grammatikaning ma'lum bolimi 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» (turki) 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, **avtomatik korrektorlar, speller** bilan ta'minlash, **audio** va **vizual effektlar** bilan boyitish kabi amaliy ishlar majmuyidir.

Til o'rgatishda quyidagi kompyuter dasturlarini qo'lladi:

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 slaydlar yaratiladi va ko'rgazmalar paketi tayyorlanadi.

– o'zbek tili fani bo'yicha mavjud bo'lgan muayyan darslikning kompyuterdagi varianti (nusxasi) bo'lib, unda multimedia vositalaridan faqat ovoz va matn qatnashadi.

¹ *Borbala Richter*. First steps in theoretical and applied linguistics.-Budapest, 2006.-B.19.

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

Multimedia («ko‘p qavatli muhit» degan ma‘noni bildiradi) zamonaviy kompyuterlarning, deyarli, barchasida mavjud bo‘lgan imkoniyatlar majmuyini, kompleks tushunchani anglatadi. Multimedia axborotning turli ko‘rinishlari – matn, jadval, grafika, ovoz, animatsiya (multiplikatsiya), videotasvir, musiqa yordamida ma‘lumotni yig‘ish, saqlash va qayta uzatish vazifalarini bajaradi. Multimedia «inson-kompyuter» interaktiv (dialogik) muloqotining yangi, takomillashgan pog‘onasi bo‘lib, bunda foydalanuvchi juda keng va batafsil axborot oladi. Bir qator tadqiqotlar shuni ko‘rsatadiki, ta‘lim oluvchi birinchi marta eshitgan mavzusining faqat to‘rt dan 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.¹ Multimediali texnologiya dastur va texnik moddiy ta‘minot asosida kompyuterda bir vaqtning o‘zida matnli tasviriy axborotlarni tovushli-ovozli, rangli, harakatga keltirilgan holda ifodalash imkoniyati demakdir. Hozirgi kunda tilshunoslik fanlarini o‘rganishda birgina darslik, o‘quv qo‘llanma va ma‘ruza matnlari bilan cheklanmasdan, shu fan sohalariga oid multimediali yoki animasiyali o‘quv ko‘rsatuvlari va elektron darsliklar yaratish tilshunoslik sohalarini chuqurroq o‘rganish uchun aniq imkoniyatlarni vujudga keltiradi. Bu esa talabalarning tilshunoslik fanlariga bo‘lgan qiziqishini oshirishga, ilmiy ish olib BORISH ko‘nikmalarini shakllantirishga xizmat qiladi.

Elektron ensiklopediya – kuchli qidiruv tizimiga ega bo‘lgan ma‘lum yo‘nalishdagi illyustratsiyali maqolalar, elektron xaritalar, fotosuratlar, giperhavola (giper murojaat)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‘at – odatdagi lug‘atlar asosida tuzilgan, multimedia vositalari bilan boyitilgan dastur. Elektron lug‘at turli yo‘nalishda tuzilishi mumkin, undan talabaning so‘z boyligini oshirish, tarjima qilishda foydalaniladi.

O‘rgatuvchi audiodasturlar, elektron lingafon kurslari umuman ikkinchi tilni o‘rganuvchilar uchun samarali ta‘lim texnologiyalaridir.

¹ Po‘latov A., Muhamedova S. Kompyuter lingvistikasi. - T., 2007. -B.66-67.

To'g'ri talaffuzni o'rgatuvchi kompyuter dasturlarida animatsiya yordamida tovushning artikulyatsiya o'rni, usuli ko'rsatiladi, etalon nutq eshittiriladi, talabani 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 qulaydir.

Universal test dasturi o'zbek tili darslarining 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.

Elektron virtual kutubxona – multimedia xonasining keng imkoniyatlaridan biri, elektron nusxa yoki elektron multimedia darsliklari mujassamlangan, tarmoqqa ulangan kutubxona.

Internet tarmog'idan o'zbek tilini o'rganuvchilar unumli foydalanishlari mumkin: mustaqil ta'lim jarayonida ma'lumotlar to'plash, mavzularni o'zlashtirish, berilgan mavzuda referat, insho yozish va boshqalar. Shuningdek, talaba elektron pochta – axborotni bir kompyuterdan ikkinchisiga yuborishni ta'minlovchi tizim orqali o'qituvchi yoki tengdoshlari bilan muloqotda bo'lishi mumkin.

Videomateriallar bilan ishlash til o'qitishda alohida ahamiyat kasb etadi. Axborot texnologiyalarining bu vositasi turli ko'rinishda qo'llanilishi mumkin.

Masofaviy ta'lim – masofadan internet tarmog'i orqali o'qitish usuli bo'lib, bu borada respublikamizda ma'lum sohalar bo'yicha test-tajribalar o'tkazilmoqda¹.

Tillarni avtomatik o'qitish tizimi hisoblangan kompyuter lingvodidaktikasida 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 tarjimasini va tahriri bo'yicha mashqlar;
 - e) grammatikaning turli bo'limlari bo'yicha diagnostik testlar o'tkazish;
- a) CD-ROMda standart o'qitish kurslaridan foydalanish.

¹ Po'latov A., Muhamedova S. Kompyuter lingvistikasi. - T., 2007. -B.66-67.

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 beradi.

15-§. TIL O'QITISH TIZIMINI AVTOMATLASHTIRISH

Reja:

1. Ta'lim jarayonida kompyuter texnologiyalaridan foydalanishning ahamiyati.
2. Til o'rgatishda qo'llanadigan kompyuter dasturlari.
3. Ta'limiy ishlanmalar konsepsiyasi.

Asosiy tushunchalar: *til o'qitishni avtomatlashtirish, o'quv dasturlari, test dasturi, ta'limiy dastur, elektron o'qituvchi, elektron ishlanmalar, multimedia dasturi, elektron multimedia darsligi, elektron nusxa, CALL, avtomatizatsiya, bixevioristik yondashuv, kognitiv-intellektual yondashuv, avtomatik korrektorlar, speller, audio va vizual effektlar, interaktiv multimedia texnologiyasi, virtual ensiklopediya, virtual kutubxona.*

Avtomatlashtirish, kompyuterlashtirish jarayoni nafaqat ishlab chiqarishning turli sohalari, balki madaniyat va ta'lim sohasini ham qamrab oldi. Kompyuter texnologiyalarining tez sur'atlarda 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 yuzaga keltirdi.

Xususan, ta'lim jarayonida bilimlarni diagnostika qilishning, xorijiy tillarni o'rgatishning samarali avtomatik tizimi ishlab chiqilgan.

XX asr tarixga ilm-fan inqilobi davri, kompyuter, avtomatizatsiya asri sifatida kirdi. Bu davrda insoniyat og'ir mehnatini o'z aqli bilan osonlashtirishga erishdi. Ya'ni ishlab chiqarish jarayonida foydali ish koeffitsiyenti bir necha barobarga oshdi, mehnat unumdorligi o'sdi. Insoniyat yirik zavod, fabrikalarni, korxonalarini avtomatlashtirishga erishdi.

«Avtomatizatsiya» so'zi lotincha «o'zi harakatlanadi» degan ma'noni bildiradi. Tizim mustaqil harakatlanishga asoslangan, turli operatsiyalar bajarishga yo'naltirilgan va shu tamoyilda ishlashga moslab dasturlashtirilgan mexanizmlar yig'indisidir. Hammaga ma'lumki, yirik korporatsiyalarda, masalan, avtomobil ishlab chiqaruvchi zavodlarda

avtomatizatsiya tizimi keng yo'lga qo'yilgan. Ushbu tizim optimalligi sababli jamiyatning boshqa sohalariga ham asta-sekinlik bilan tatbiq etilmoqda. Hozirda til o'rgatishning avtomatik tizimlari ishlab chiqilgan. Til o'rgatish 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 alohida “Computer Assisted Language Learning (CALL)” yo'nalishining yuzaga kelishiga asos bo'ldi.¹ 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'rtaqa tashlandi, o'qitishda audio va video vositalardan foydalanish samarali ekanligi ta'kidlandi.

Dunyo miqyosida kompyuter tizimini o'quv jarayoniga tatbiq qilish bo'yicha yetarli tajriba to'plangan, bu borada olib borilgan muhim ilmiy-metodik tadqiqotlar, ko'pgina ta'limiy-nazorat dasturlari, multimedia darsliklari yaratilgan.

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

TOEFL tizimi – ingliz tilini avtomatik o'qitishga mo'ljallangan, eng rivojlangan xalqaro dastur. Inglizcha Testing Of English as a Foreign Language so'zlari bosh harfidan olingan qisqartma bo'lib, ingliz tilidan xorijiy til sifatida sinov standarti hisoblanadi. Dunyoning, deyarli, barcha mamlakatlarida amal qiladi. Bu tizim ingliz tilini o'rgatish bilan birgalikda o'zlashtirilgan bilimlarni diagnostika qila olish imkoniyatiga ham ega. Amerika universitetlari yoki kollejlari kiruvchilar uchun sinov vazifasini o'taydi.

IELTS – International English Language Testing System, ya'ni Xalqaro ingliz tili sinovi tizimi. TOEFL tizimiga o'xshaydigan avtomatik til o'rgatish va bilimlarni diagnostika qilish tizimi. U grammatika, lug'at, o'qib va tinglab tushunish, to'g'ri jumla qurish kabilar yuzasidan bilimlarni baholaydi.

IELTS – International English Language Testing System, ya'ni Xalqaro ingliz tili sinovi tizimi. Imtihonning maqsadi ingliz tilida o'qish,

¹ [http:// www.yazykoznanie.ru](http://www.yazykoznanie.ru)

yozish, tinglab tushunish va gapirish ko'nikmalarining qay darajada shakllanganligini baholash hisoblanadi.

TOEFL – Testing of English as a Foreign Language so'zlarining bosh harflaridan yasalgan qisqartma so'z bo'lib, Amerika universitetlari yoki kollejlari kiruvchilar uchun sinov vazifasini o'taydi.

Mazkur dasturlarning bosh maqsadi ingliz tilini mukammal o'rganishni ta'minlashdan iboratdir. IELTS tizimi Buyuk Britaniya ingliz tiliga asoslangan bo'lib, aynan ingliz tili grammatikasiga tayangan holda o'qitishni nazarda tutadi. TOEFL dasturi Amerikada foydalaniladigan ingliz tili qonun-qoidalariga asoslanadi. Ushbu ikki dastur test olinish usuli, vazifalari, turlari, baholash, test o'tkazilish jarayoni bilan bir-biridan farqlanadi.

IELTS test dasturining 2 turi mavjud bo'lib, Umumiy (IELTS General Training Module) va Akademik (IELTS Academic Module) test sinovi qismlaridan iborat. Umumiy yo'nalishda test topshirganlar o'rta yoki o'rta maxsus ta'lim asosida o'quv faoliyatlarini davom ettirishi mumkin. Oliy ma'lumotga ega bo'lish uchun esa Akademik test sinovini topshirish talab etiladi.

TOEFL o'qishga kirish uchun yagona sinov-test tizimi emas. Amerika o'quv yurtlarida TOEFLga qo'shimcha ravishda o'quvchining avval o'qigan ta'lim muassasidan olgan baholari va ingliz tiliga bog'liq biror dastur bo'yicha qo'lga kiritgan yutuqlari ham hisobga olinadi. Ingliz tilida amalga oshirilgan jamoaviy tadbirlardagi faoliyat ham nazarda tutiladi.

Har ikkala dasturda ham sinov tizimi to'rt xil shaklda bosqichma-bosqich amalga oshiriladi. Mazkur dasturlarning farqi bosqichlarni amalga oshirish usullari hamda baholash tizimida namoyon bo'ladi.

IELTS (2 soat 45 min):

➤ **listening (tinglash)** – 30-40 daqiqa davomida, ma'lum axborotni tinglab, unga tegishli bo'lgan 40ta savolga javob berish;

➤ **reading (o'qish)** – 60 daqiqa davomida muayyan sohaga oid bo'lgan matnni o'qish hamda shunga oid 40ta savolga javob berish;

➤ **writing (yozish)** – 60 daqidada umumiy mavzuda esse yoki insho yozish;

➤ **speaking (suhbat)** – 10-15 daqiqa ichida shaxsiy hayoti, faoliyati, kelajakdagi rejalari haqida so'zlab berish.

TOEFL(4 soat 30 min):

➤ **reading (o'qish)** – 60-80 daqiqada 700 so'zdan iborat, 3-4 fragmentli matnni o'qib tahlil qilish;

➤ **listening (tinglash)** – ikki xil ko‘rinishda: 60-90 daqiqa davomida biror dialogdagi suhbat yoki axborotni tinglab tushunish;

➤ **speaking (suhbat)** – 6 xil vazifadan iborat: 20 daqiqa ichida 2ta o‘zi istagan tanish mavzu, 4ta eshitgan yoki o‘qigan mavzularidan biri haqida so‘zlab berish;

➤ **writing (yozish)** – 50 daqiqa mobaynida o‘qiganlari hamda eshitganlari bo‘yicha savollarga javob tarzida yoki o‘zi bilgan mavzuni dalillar, ma‘lumotlarga tayangan holda yozma tarzda asoslab berish.

Baholash tizimi:

IELTS → 9ball, TOEFL → 450 (ba‘zan 500-550) ball.

Ballar orasidagi tafovut katta. Buyuk Britaniyaning iqtisodiyot yo‘nalishidagi universitetlariga topshirish uchun kamida 6 ball, yuridik yoki gumanitar sohaga yo‘naltirilgan o‘quv yurtlariga topshirish uchun kamida 7.5ball to‘plash kerak. IELTS dasturida tanlangan institut yoki yo‘nalishdan qat’i nazar, hamma bir xil tartib va talab asosidagi sinovlardan o‘tadi. Farq faqat to‘plangan ballar va ularning qayerda, qanday tartibda qabul qilinishi bilan izohlanadi. Amerika magistratura mutaxassisligida o‘qish uchun TOEFL tizimidan to‘plangan ball 61-109 oralig‘ida bo‘lishi kerak. Masalan, Harvard hamda Kolumbiya universitetlariga shu talab asosida talabalar qabul qilinadi.

Bulardan tashqari, **FCE** (First Certificate in English), **CAE** (Certificate in Advanced English), **GRE** (Graduate Record Examinations), **GMAT** (Graduate Management Admissin Test) kabi til o‘rgatuvchi va sinov asosida chet elda o‘qishni ta‘minlovchi maxsus dasturlar ham mavjud.

O‘zbekistonda o‘rta, o‘rta maxsus va oliy o‘quv yurtlarida o‘qitilayotgan ingliz tili ta‘limi Buyuk Britaniya ingliz tili grammatikasiga asoslangan. Shuning uchun IELTS dasturi mavqeyi va bu tizimda sinov topshiruvchilar soni TOEFLga nisbatan ko‘pchilikni tashkil etadi. Bu dastur talablari asosida tayyorlanayotganlar soni tobora ko‘paymoqda. IELTS test tuzuvchilarining ta‘kidlashicha, ushbu testlar sinovdan o‘tuvchining umumiy bilimini ham, uning maxsus tayyorgarligini ham tekshirmaydi. Test dasturining asosiy maqsadi test topshiruvchining ingliz tilini kelajakda yetarli darajada o‘zlashtirishini ta‘minlashdir. Lekin, jarayon shuni ko‘rsatmoqdaki, sinov davomida talabgor savollarga to‘g‘ri javob topish uchun mantiqiy fikrlashi, ma‘lumotlarni analiz, sintez qilishi, o‘z ona tili ko‘nikmalari asosida anglab, qayta ishlashi va ularni xorijiy tilga o‘girib, to‘g‘ri ifodalay olishi kerak. Bu jarayon talabgordan muayyan bilimga ega bo‘lishni, maxsus tayyorgarlik ko‘rishni talab etadi. Demak,

IELTS ingliz tilini mukammal egallash ko'nikmalarini shakllantirishga xizmat qilibgina qolmasdan, talabgorlarning ingliz tili bo'yicha bilim darajalarini aniqlashga ham xizmat qiladi.

IELTS dasturi o'qish, yozish, eshitish va gapirish ko'nikmalarining qay darajada shakllanganligini aniqlashda ingliz tilida so'zlovchi davlatlarda ta'lim olish jarayonining zarur qismi hisoblanadi. Shu bois IELTS sinovlariga tayyorgarlik ko'rishning eng asosiy yo'nalishlari ingliz tilini kundalik turmushda qo'llash ko'nikmalarini shakllantirishga qaratilgan bo'ladi. Ingliz tili muloqot tili bo'lgan davlatlarda ta'lim olishni maqsad qilgan talabalar bilan ishlash maqsadida borayotgan fuqarolarning tilni bilish ko'nikmalari farq qilishi mumkin, ammo IELTS sinovlaridagi topshiriqlarda bu farq belgilovchi rol o'ynamaydi.

Ingliz tilida gaplashadigan davlatlarda ta'lim olish jarayonida quyidagi vaziyatlarga duch kelish mumkin:

- ✓ Kitob va jurnallar o'qish;
- ✓ Yozma topshiriqlarni bajarish;
- ✓ Ma'ruzalar tinglash;
- ✓ Seminarlarda qatnashish.

Ta'lim olish jarayonida sohaga oid turli adabiyotlarni o'qish, aniq bir mavzuga doir ma'lumotlar berilgan gazeta va jurnallarni mutolaa qilish; matnlar, jadvallar va chizmalardan kerakli ma'lumotlarni olish; olingan ma'lumotlarni tahlil qilish talab etiladi.

Yozma topshiriqlarni bajarish o'rtaga tashlangan muammoli masalani analiz qilish; yozma ishning rejasini tuzish; ijodiy yondashib yozish orqali mavzuni tushunarli tarzda ifodalash, o'z fikrini yozma ravishda tahlil qila olish uchun kerak.

Ma'ruzalarni tinglash davomida nutqni tinglab tushuna olish; turli sheva elementlari aralashgan nutqni idrok etish; ma'ruzaning asosiy qismlarini, urg'u berilayotgan ma'nolarini anglash; zarur ma'lumotlarni yozib olish kerak. Bulardan tashqari radio va video orqali ifodalangan turli darajadagi tezlik va shevada gapirilgan nutqni tinglab tushuna olish, nutqning asosiy mazmunini anglay olish hamda zaruriy holatlarda ular orqali yetkazib berilayotgan ma'lumotlarni yozib olish ko'nikmalarini shakllantirish lozim bo'ladi.

Seminarlarda qatnashish tinglovchilardan matnlar muhokamasida ishtirok etishga tayyor bo'lishni, ma'ruzachi va talabalarning berayotgan savollarini tushuna olishni, savollarni ingliz tilida tushunarli tarzda bera olishni talab etadi.

Yuqoridagilardan kelib chiqqan holda, IELTS dasturi asosida ta'lim olishni ikki bosqichga bo'lib amalga oshirish kerak:

1. Nutq, yozish, o'qish va tinglab tushunishni uzviy o'rgatish. Talabgorning so'z boyligini imkon qadar oshirgan holda ingliz tili grammatikasining eng zaruriy va murakkab jihatlarini o'rgatish. IELTS sinovini topshirishga oid yo'l-yo'riqlar ko'rsatish.

2. Nutq, yozish, o'qish va tinglab tushunish ko'nikmalarini keng miqyosda rivojlantirish; gazeta va jurnallar o'qish ko'nikmasini shakllantirish, radio va audioeshittirishlar asosida tinglab tushunish ko'nikmasini mustahkamlash, imkon qadar jonli muloqotni amalga oshirish asosida og'zaki nutq malakasini rivojlantirish.

IELTSni o'rgatish kursi oxirida ta'lim oluvchilar quyidagi ko'nikmalarni to'liq egallagan bo'lishlari zarur:

1) ingliz tilidan ona tili sifatida foydalanuvchilarning o'rtacha tezlikdagi nutqida ifodalanayotgan ma'lumotni tinglab tushuna olish va qabul qilingan ma'lumotni yuqori darajadagi aniqlik bilan (diktant ko'rinishida) eshitilganlarni aynan takrorlamasdan, mazmunini gapirib bera olish;

2) hujjatlarni tashkil etish (masalan, hisobotlar, rasmiy va norasmiy xatlar, chizma va diagrammalardagi ma'lumotlarni matn ko'rinishida ifodalash, referatlar yozish), so'rovnomaga ko'rinishidagi hujjatlarni to'g'ri to'ldira olish. Hujjatlarni ma'lum cheklangan vaqt ichida aniq to'ldirish hamda ma'lumotlarni yozish jarayonida kerakli bo'lgan (rasmiy yoki norasmiy) uslubdagi ingliz tilidan foydalana olish;

3) ma'lum cheklangan vaqt ichida ijtimoiy darajasiga ko'ra muhim bo'lgan turli mavzularga asoslangan insholarni yoza olish;

4) ko'rib chiqish, ma'lumot qidirish va o'rganish usullari orqali anglash, talab darajasida o'zlashtirish. Cheklangan vaqt ichida aytilayotgan (og'zaki nutq) yoki yozilgan (matn) axborot ichidan kerakli bo'lgan ma'lumotlarni topa olish;

5) ma'lum bir mavzuning asosiy mazmunini og'zaki tarzda bayon qila olish; mavzu asosida berilishi mumkin bo'lgan savollarga javob berish va buni mos uslubdagi ingliz tili asosida amalga oshira olish;

6) ingliz tilida asosiy bilimlarning puxta o'zlashtirilganligini namoyon eta olish (masalan, til grammatikasi, leksikasi, orfografiyasi va boshqalar).

IELTS test sinovi barcha ingliz tilini o'rganuvchilar uchun butun bir imkoniyatlar majmuyidir. Bu sinov tizimi ingliz tilining mavqeyini yanada oshiradi. IELTS va shu tipdagi test sinovlarini kuzatib, xulosa chiqarish mumkin-ki, o'zbek tilini o'rganishni xohlagan, shu tilda o'qishni davom

ettirmoqch bo‘lgan chet ellik talabalar uchun sinov turlari ishlab chiqilsa, maqsadga muvofiq bo‘lar edi.

Xorijiy tillarni o‘rganish kun tartibidagi dolzarb masalalardan biridir. Yoshlar shuni anglashlari lozimki, xorijiy tilni o‘rganish milliy til imkoniyatlarini to‘liq o‘zlashtirish negizida amalga oshirish lozim.¹

ENGLISH FOR ALL (“Hamma uchun mo‘ljallangan ingliz tili”) – O‘zbekistonda yaratilgan avtomatik o‘qitishga asoslangan dastur. Unda o‘quvchiga ingliz tilidagi tovushlar va ularning talaffuzi audioeffektlar yordamida tushuntiriladi, qisqacha ingliz tili grammatikasi bayon etiladi, mashqlar va testlar yordamida bilim tekshiriladi. Ushbu o‘quv dasturining afzal tomoni shundaki, unda o‘qitish tizimi videoeffektlar bilan ham boyitilgan, ingliz tilidagi 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, multimedia asosida ish ko‘radi. Dasturning yuqoridagilardan farqli tomoni shundaki, unda dastur «elektron» («virtual») o‘qituvchi bilan ta‘minlanganidir.

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

PROFESSOR HIQGTNS – АНГЛИЙСКИЙ БЕЗ АКЦЕНТА – fonetik tizim asoslarini o‘rganishga qaratilgan dastur, 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 – ingliz tilini uch yil davomida bosqichma-bosqich (Elementary, Intermediate, Upper intermediate, Advanced English), tizimli tarzda o‘rgatishga asoslangan avtomatik o‘qitish tizimi.

Bulardan tashqari, til o‘qitish dasturlari, videodarslari jamlangan internet saytlari ham mavjud:

BOND.UZ – ingliz tilini o‘rgatish dasturlari, videodarslari jamlanmasi. Masalan, mazkur saytdagi “On kun ichida ingliz tilini o‘rganamiz” videodars jamlanmasi “Asosiy tushunchalar”, “Sodda gaplar”, “Noto‘g‘ri fe‘llar”, “Kishilik olmoshlarining to‘ldiruvchi shakli”, “Savollar va so‘roq so‘zlar”, “*To be* fe‘li”, “Qo‘shimchalar va so‘z

¹Sharopova S. Til o‘qitish dasturlarining amaliy ahamiyati. Metodist. Ilmiy maqolalar to‘plami. – Toshkent, 2017. -B.8-12.

yasalishi”, “Sifat va sifat darajalari”, “Predloglar”, “Yakunlovchi qism” kabi o‘nta mavzudagi ixcmalashtirillgan, jadallashtirilgan kursni qamrab olgan.

Bugungi kunda o‘qitishni kompyuterlashtirish nazariyasini ishlab chiqish o‘quv faoliyatining samaradorligini oshiradigan, o‘rganilayotgan mavzuga nisbatan ijobiy munosabat, 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, uini yaratish, tekshirish va qo‘llash tartibi belgilanadi. Mutaxassislarining fikricha, ta’limiy-elektron ishlanmalar, asosan, texnik, pedagogik, metodik talablarga javob berishi zarur.

Til o‘qitishning avtomatik tizimi yildan yilga rivojlanib, yangi effektlar bilan boyib bormoqda. Bu tizimni yanada rivojlantirish, kompyuter lingvodidaktikasini yangi dasturiy ta’minotlar va pedagogik texnologiya, bilimlar diagnostikasi bilan boyitish kompyuter lingvistikasi fanining dolzarb vazifalaridan sanaladi.

Yaratilgan dastur va boshqa elektron ishlanmalarning sifati, albatta, mualliflarning shaxsiy qarashlariga, ilmiy-metodik saviyasiga va umumiy madaniyatiga bog‘liqdir. Bu o‘ziga xos murakkabliklarga ega. Masalan, ba’zi ta’limiy – elektron ishlanmalar ko‘zlangan maqsadga to‘la-to‘kis erishishga xizmat qilsa, ba’zilari o‘ta qiziqarli, estetik jihatdan yuqori sifatli bo‘lishiga qaramay, muayyan mavzu bo‘yicha kerakli bilim va ko‘nikmalarni egallashga xizmat qilmaydi, informatika va informatsion madaniyat kursini o‘rganishgagina yordam beradi.

Ta’lim jarayoniga, ayniqsa, til o‘qitishga kompyuter texnologiyalarining joriy etilishi XX asr ta’lim tizimining muhim talablaridan biri sifatida maydonga chiqdi. XXI asr – axborot asrida kompyuter texnologiyasidan quyidagi maqsadlarda til o‘qitishda keng foydalanish yo‘lga qo‘yildi:

a) o‘zbek tilini ona tili sifatida, o‘zbek tilini chet tili sifatida va chet tillarini o‘qitish jarayonida kompyuter dasturlaridan foydalanish. Bu muammoni hal etish bo‘yicha mamlakatimizda muayyan ishlar amalga oshirilmoqda;

b) kompyuter ma’lumotlaridan tilning grammatika tizimi, fonetikasi va leksikasini o‘qitishda foydalanish;

v) kompyuter dasturidan o‘zbek tili va chet tillari bo‘yicha olingan bilimlarni baholash maqsadida foydalanish kabilar.

Kompyuter lingvistikasi yordamida til o‘qitishni maqsadga muvofiq yo‘lga qo‘yish uchun, dastavval, o‘zbek tilidagi barcha so‘z shakllarini o‘z ichiga olgan kompyuter dasturlari yaratilishi lozim. Shu dasturni amalga oshirish maqsadida 2003-yilda Toshkentda “Dunyoviy o‘zbek tili” asari birinchi jildi yaratildi. O‘zbek tili negizidagi mashina tilini yaratish asoslari ishlab chiqilgan. Bu asarda o‘zbek, rus va ingliz tillarini qiyoslash orqali o‘zbek tili grammatikasidagi afzalliklar va ayrim ziddiyatlarni inobatga olgan holda ma’lumotlar bazasi keltirilgan.

O‘zbek tilidagi 500 dan ortiq fe’lni LSG (lug‘aviy semantik guruh)ga ajratib, har bir fe’lning grammatik xususiyatlari, valentlik imkoniyatlari ko‘rsatilgan holda dasturlar yaratish bo‘yicha muayyan ishlar amalga oshirilgan va ular hozirgi kunda til o‘qitishda sinovdan o‘tkazilmoqda. Masalan, M.Y.To‘xtamirzayev tadqiqoti milliy maktablarda rus tilini teleekran yordamida o‘qitish, A.N.Maxammatovning nomzodlik dissertatsiyasi esa musiqa darslarini modellashtirish muammolariga bag‘ishlangan. Sh.Yusupova nomzodlik dissertatsiyasida ona tili ta’limi samaradorligini oshirishda noan’anaviy usullar va EHMdan foydalanish haqida so‘z yuritiladi. Mazkur ishda umumta’lim maktablarida ona tili bo‘yicha noan’anaviy dars o‘tish metodlari keng ko‘lamda aks etgan bo‘lib, kompyuter texnologiyalarining birgina passiv, nafaol turi qisman yoritilgan. S.Adilova tadqiqotida rusiyzabon guruhlardagi o‘zbek tili ta’limi mazmuni uzviylik va uzluksizlik tamoyillariga amal qilingan holda belgilangan, bu fanni o‘qitishda mashg‘ulotlarni kompyuter vositasida tashkil etishning nazariy asoslari yaratilgan, o‘zbek tili dasturiga kirgan til sathlari bo‘yicha ta’limiy elektron ishlanmalarning loyihasi (maketi) ishlab chiqilgan.

O‘zbek tilini o‘qitishda quyidagi turdagi topshiriqlardan ham foydalansa bo‘ladi:

1. “Bo‘sh kataklarni to‘ldiring” topshirig‘i :

Masalan, **Men til o‘rganish... yaxshi ko‘raman.**

2. Elektron pochtdan foydalanib xabarlar jo‘natish;

Masalan, Dear students! You must send little story in uzbek language about you to this mail address: lobar.abduhamidova@mail.ru;

3. Matnlar tarjimasiga doir mashqlar:

Masalan, **I promise you to come at this time tomorrow.**

O‘zbek tilini o‘rganishga doir elektron lug‘atlar va qo‘llanmalar mavjud. Hozirgi kunga qadar o‘zbek tilini o‘rganuvchilar uchun ingliz

tilidagi IELTS, TOEFL, ENGLISH FOR ALL kabi universal ekeptron dasturlar yaratilmagan.

Quyida internet ma'lumotlaridan ayrimlarini keltiramiz:

1. Word formation suffixes.

Nominals and verbals with new lexical meanings can be created in Uzbek language through suffixation. So by adding suffixes we create new lexical meaning:

Bosh – head (anatomy)

Boshliq – chief, foreman

Boshlamoq – to begin (infinitive)

Adding a suffix to a verbal root or word also produces a nominal or a verbal with a new lexical meaning:

Boshlanmoq – to be begun (infinitive)

boshlang'ich – beginning elementary

2. Grammatical suffixes

Grammatical functions like number, case, mood or tense is accomplished by adding suffixes to words. These grammatical suffixes change the relationship of one word to other words in a sentence, but they do not change the basic lexical meaning of the word:

Lexical Grammatical

uy home ; **uyda** at home

uylamoq to marry; **uyga keldi** he came home

uylanmoq to get married; **uylandi** he got married

3. Plural suffix

The Uzbek language plural suffix *-lar* is added to nominals to indicate that there is more than one subject or object, but also to verbals to indicate that there is more than one subject. But NOTE that after a numeral, the plural suffix is not added to nominals:

qiz daughter – **keldi** she came

qizlar daughters – **keldilar** they came

uch qiz three daughters

4. Case suffixes

Case suffixes express relationships between nominals and verbals and are equivalent to English 'to', 'in', 'from' and other ideas:

o'g'il – son; **o'g'ilga** to the son

xalq –people ; **xalqdan** from the people

kitob –book; **kitobda** in the book

5. Possession suffixes

Uzbek has several means of expressing possession. One means require adding possessive suffixes to a nominal:

kitobim my book; **kitobimiz** our book

kitobing your book ; **kitobingiz** your book
(singular) – (plural)

A second means requires the possessive relationship construction.

In Uzbek, the possessor of an object is placed first, often with the suffix -ning, while the person or object possessed is placed second, always with the suffix -i or -si (plural -lari). The Uzbek equivalent of an English phrase like 'my daughter's book' is the following: **qizimning kitobi** daughter+my+of book+her= my daughter's book

6. Infitive suffix

The Uzbek equivalent of the English infinitive of verbs ('to enter', 'to see', etc.) is created by adding suffix *-moq* to the verbal root. Without this or other suffixes, the verbal root express the familiar or at times impolite imperative mood:

kirmoq to enter (infinitive)

kir! enter (singular, familiar, impolite)

7. Mood suffixes

To express moods like the imperative or the conditional, Uzbek adds suffixes to verbal roots:

kiring! come in! (singular, familiar, polite)

kirsa if he enters

8. Tense suffixes

In Uzbek, tenses like past, present, and future are formed by adding suffixes to a verbal root, including the past tense suffix -di:

kelmoq to come (infinitive)

keldi he came

Uzbek combines the meanings of English present and future tenses into one idea that is expressed by the suffix -a or -y. This present-future tense indicates that an action occurs as a habit in the present and may occur at some point in the future:

o‘qimoq to read (infinitive);

o‘qiydi he reads (as a habit), he will read (tomorrow);

Enjoy learning and practicing, if there are questions, please don't hesitate to post in the comments!

Coming up: Quiz and examples (Suffixations), and some entertainment videos about Uzbekistan.¹

¹ L.Abdusamidova. Kompyuter lingvistikasining avtomatik tahrir yo‘nalishi. BMI. –T., 2016.

Savol va topshiriqlar

1. “English for all” dasturining ahamiyatini yoriting.
2. “Way to perfectness” dasturida multimedyaning o‘rni haqida mulohaza yuriting.
3. “Reward” dasturining ishlash prinsipi haqida ma’lumot bering.
4. “Professor Hiogtns” dasturida aksentga e’tibor berilish sabablarini izohlang.
5. “Repetitor english” dasturida zamonaviy metodlarning o‘rni haqida fikr bildiring.
6. “Euro plus” dasturining o‘qitish bosqichlari haqida bilganlaringizni so‘zlang.
7. Til o‘qitish dasturlari asosida “keys” topshirig‘ini tuzing.

16-§. O‘ZBEK TILINING AXBOROT USLUBI

Reja:

1. Tilning axborot uslubi nazariyasi.
2. O‘zbek tilining axborot – kompyuter uslubini yaratish mezonlari.
3. Axborot-kompyuter uslubiga qo‘yiladigan talablar.

Asosiy tushunchalar: *tilning axborot uslubi nazariyasi, o‘zbek tilining axborot-kompyuter uslubi, aniqlik, qisqalik, lo‘ndalik; bir qiymatlilik; jahon andazalariga moslik; axborot - kompyuter uslubini yaratish mezonlari, sinonimlar, frazeologizmlar va polisemantik so‘zlarning ifodalanish darajasi, modellar, kodlardan foydalanish.*

Til insongagina in’om etilgan ulug‘ ne’matdir. Shu sababli lisoniy hodisalar tadqiqi boshqa ijtimoiy hodisalar tadqiqidan tubdan farq qiladi. Barcha tabiiy va ijtimoiy hodisalar tizimli tuzilishga va o‘ziga xos tarkibiy xususiyatga ega. Tilning ham tizimli hodisa ekanligi e’tirof etilgan. Tilshunoslikka doir adabiyotlarda til tizimi strukturasi, til birliklarning o‘zaro munosabati masalalari talqin etilgan. Til tizimiga muvofiq ravishda kompyuter uslubini yaratish muloqotning tezkorligi va to‘liqligini ta’minlaydi. Kompyuter uslubi inson manfaatlariga tayangan holda yaratilgan atomatik muloqot tilidir.

Jahon miqyosida kechayotgan integratsiya va globallashuv jarayonida o‘zbek tilini dunyoviy tillar darajasiga olib chiqish hayotiy zaruratdir. Ushbu 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 dunyoviy tillardan biriga aylanishiga, uni o'rganish va o'rgatish ishlarining optimallashtirishiga zamin yaratadi.

Jahon andazalari 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 ma'lumotlarning tezkor yetkazib berilishi va qisqa muddatlarda qabul qilinishini talab qiladi. Axborot tezligini oshirish qisqalik, aniqlik, muayyan qolip asosidagi jummalarni taqozo qiladi. Axborot -kompyuter uslubi sun'iy tilga asoslanadi. Tabiiy til muloqot vositasi bo'lishi bilan bir qatorda, emotsional-ekspressiv, akkumulyativ funksiyalarni ham bajaradi. Tilning ijtimoiy funksiyalari tabiiy tilda to'laligicha namoyon bo'ladi. Axborot uslubida kommunikativ funktsiya birinchi darajaga chiqadi, emotsional-ekspressivlik, deyarli, reallashmaydi. Tabiiy til va sun'iy til (kompyuter-axborot uslubi) o'z o'rnida ahamiyatlidir. Muayyan tabiiy tilni kompyuter axborot uslubiga mos emas, deyish yoki qayta ishlangan, modellashtirilgan tabiiy tilni imkoniyatlari cheklangan deb baholash nisbiydir. Kompyuter- axborot uslubi nutq uslublari bilan parallel ravishda shakllanib, rivojlanib boraveradi, ammo ular qo'llanish sohasi bilan farqlanadi. Bu uslublardan turli soha vakillari foydalanishadi: og'zaki muloqotda so'zlashuv uslubidan, fan-texnika taraqqiyotiga oid ma'lumotlar berishda ilmiy uslubdan, ish yuritishda rasmiy uslubdan, ijtimoiy-siyosiy, ommabop ma'lumotlarni yetkazishda publitsistik uslubdan, adabiyot, san'at, madaniyat sohalarida badiiy uslubdan foydalaniladi. Kompyuter uslubidan kompyuter mutaxassislari, reklama, biznes, tabiiy soha vakillari va eng muhimi o'zbek tilida ish yurituvchilar (yurisprudensiya vakillari) foydalanishlari mumkin.

Kompyuter uslubi tabiiy tilning jahon andazalari talablariga monand holda fikrni aniq, qisqa, ixcham, qoliplashtirilgan ko'rinishda qayta ishlash asosida shakllantiriladi.

O'zbek tilining kompyuter uslubini shakllantirish milliy til mavqeyini yuksaltirishga, ma'lumotlarni ona tilida qabul qilish imkoniyatiga zamin hozirlaydi. O'zbek tili kompyuter uslubining mezonlari sifatida quyidagilarni keltirish mumkin:

- a) aniqlik, qisqalik, lo‘ndalik; bir qiymatlilik;
- b) jahon andazalariga moslik;
- d) barcha uchun birdek tushunarli bo‘lish (muayyan bir fikrning umum tomonidan bir xil tarzda qabul qilinishi);
- e) turli badiiy tasvir vositalari, modallik, sinonimlikdan xolilik (sinonimlardan dominanta so‘zlar tanlab olinadi);
- f) so‘zlar, asosan, bir ma’nodlilik kasb etadi, ko‘chma ma’nolar e’tiborga olinmaydi (ya’ni bu uslubda so‘zlarning faqat denotativ semalari qamrab olinadi);
- g) kompyuter talablari qat’iy hisobga olinadi, simvollar, modellar, kodlardan foydalaniladi.

O‘zbek tilining kompyuter uslubini yaratish ijtimoiy-iqtisodiy hayotda ham muhim ahamiyatga ega. Bu quyidagilarda o‘z ifodasini topadi:

-o‘zbek tili axborot uslubining yuzaga kelishi shu tilda ma’lumotlar bazasini shakllantirishga xizmat qiladi;

-kompyuter tizimining o‘zbek tili asosida ishlashini ta’minlaydi; kompyuter tizimidagi buyruqlarning shu tilda amalga oshirilishiga zamin yaratadi;

-o‘zbek tilidan rus tiliga (rus tilidan o‘zbek tiliga), o‘zbek tilidan ingliz tiliga (ingliz tilidan o‘zbek tiliga) tarjima dasturlari takomillashadi;

-o‘zbek tilidagi matnlarni tahrir qilish imkoniyati yaratiladi; ijtimoiy sohalarga doir matnlarda yo‘l qo‘yilayotgan imloviy nuqsonlar bartaraf etiladi;

-barcha soha mutaxassislarining savodxonligi ortadi.

Insoniyat tarixida ma’lum davrlarda turli predmetlar - toshko‘mir, tilla, tuz kabilar eng katta boylik hisoblangan. Insoniyat uchinchi ming yillik boshlarida axborot asriga qadam qo‘ydi. Bugungi kunda axborot olish hayotiy zaruriyatga aylanadi. Shuning uchun jahon bozorida ma’lumot oltinidan ham qimmatliroq baholanmoqda. Axborotni topish, saqlash, qayta ishlash va boshqalarga yetkazishning qulay usullariga bo‘lgan ehtiyoj kun sayin ortib bormoqda. Dunyodagi eng badavlat kishi kompyuter, ya’ni axborotni qabul qilish, saqlash va qayta ishlash sohasi vakili Bill Geyts hisoblanadi. Bu holatni Bill Geytsning o‘zi shunday izohlaydi: *«Kim axborotga ega bo‘lsa, u hamma narsaga ega bo‘ladi»*. 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»*. Kompyuter lingvistikasi tilshunos, adabiyotshunos, qolaversa, barcha soha vakillari uchun ham ma’naviy, ham moddiy boylik manbayidir. Kompyuter lingvistikasi

tillarni, xususan, rus va ingliz tillarini o'rganishda ham asosiy vosita hisoblanadi. Shuningdek, Oliy va o'rta maxsus ta'lim muassasalarida ilmiy, badiiy, biznes va reklama bo'yicha ixtisoslashgan tarjimonlar tayyorlashni yo'lga qo'yishda, o'zbek tilida ingliz tilini va ingliz tilida o'zbek tilini to'liq o'rgatadigan o'quv qo'llanmalari va kompyuter dasturlarini yaratishda ham kompyuter lingvistikasi muhim o'rin tutadi.

«Dunyoviy o'zbek tili» kitobi o'zbek tilidagi ma'lumotlar bazasini yaratish yo'lidagi birinchi qadamdir. Kitob muallifi fizika-matematika fanlari doktori, professor Abdumajid Po'latov o'zbek kompyuter lingvistikasining asoschisi, o'zbek tilidagi axborot uslubi modelining yaratuvchisi hisoblanadi. «Dunyoviy o'zbek tili» kitobi ko'p jildli bo'lib, uning 1-jildida o'zbek tilidagi fe'l shakllari (fe'lning leksik-grammatik paradigmalari) va ularning rus, ingliz tillaridagi muqobillari keltirilgan. Mazkur kitob fe'l -kesim va fe'lning xoslangan shakllari – sifatdosh, harakat nomi, ravishdosh kabilarni uch tilda to'liq o'zlashtirishga yordam beradi: kitob bilan tanishgan har bir o'quvchi fe'l shakllarini rus va ingliz tillarida yozishni o'rganib oladi. Eng muhimi, o'zbek tilidagi fe'l asosida yaratilgan mazkur tadqiqot asosida butun jahon kompyuter lingvistikasi assotsiatsiyasida ma'lumotlar bazasi yaratiladi.

Savol va topshiriqlar

1.O'zbek tilining rasmiy-idoraviy va kompyuter uslublarini ishlab chiqish haqida fikr bildiring.

2.O'zbek tili grammatikasining kompyuter modelini yaratish haqida mulohazalaringizni bayon eting.

3.O'zbek-rus-ingliz tillari tarjima dasturi haqidagi fikrlaringizni ayting.

4.O'zbek tilidagi Internet sahifalarini (site) yaratish asoslarini ishlab chiqish haqida tavsiyalaringsizni bering.

5.O'zbek tilidagi matnlarni tahrir qilish dasturlarini yaratish haqida fikr bildiring.

6.O'zbekcha-ruscha-inglizcha grammatik atamalar lug'atini yaratishning ma'lumotlar bazasi sifatidagi ahamiyatini tushuntiring.

7.Ijtimoiy va psixologik omillarning tilga ta'sirini klaster usulida yoriting.

17-§. KOMPYUTER LINGVISTIKASI METODLARI

Reja:

1. Kompyuter lingvistikasida matematik statistika.
2. Kvantitativ metod va uning tilga tatbiqi.
3. Flektiv tillarda o‘zakni, polisemantik so‘zlarni va sinonimik qatordagi dominant so‘zni aniqlash.

Asosiy tushunchalar: *kvantitativ metod, matematik parametr, statistika, indeks, absolut chastota, nisbiy chastota, lisoniy tip, flektiv tillar, agglyutinativ tillar, matnning statistik tahlili, Zif qonuni, matn atributsiyasi, anonim va psevdonim matn-lar, idiosstil, polisemantik so‘zlar, sinonimik qator, dominant, to‘plam, ling-vostatistika.*

Statistik tahlil metodi boshqa fanlar qatori tilshunoslik fani uchun ham muhim ahamiyatga ega. Bunda mavjud til hodisalari aniq matematik parametrlarda baholanadi, tekshiriladi. Statistika lotincha «status» so‘zidan olingan bo‘lib, hodisalarning holati, vaziyatini bildiradi. «Status» so‘zidan «stato» – davlat, «statusta» – davlatni biluvchi, «statustica» – davlat to‘g‘risida muayyan bilim, ma’lumotlar yig‘indisi ma’nosini bildiruvchi so‘zlar kelib chiqqan.¹

Bundan 3500 yil oldin Qadimgi Misrda o‘tkazilgan aholi hisobi, Rossiyada Pyotr I davridagi «aholi taftishi», Angliyadagi «Dahshatli sud kitobi» kabilar dastlabki statistik ma’lumotlardir.

XVII asrda Angliyada yuzaga kelgan «*siyosiy arifmetika*» (U.Petti va Jon Grauntlar nomi bilan bog‘liq) fani statistikaning shakllanishiga asos bo‘lgan. Statistika ijtimoiy hodisalarning miqdoriy tomonlarini ularning sifat tomonlari bilan uzviy aloqadorlikda o‘rganuvchi fandır. Hodisalarning sifat tomonlarini maxsus fanlar o‘rganadi. Xususan, tovushlar, so‘zlar, gaplarni lingvistika, qofiya, ritm, vazn, she’r tuzilishini adabiyotshunoslik, organik hayotni biologiya, yer qatlamlari va boyliklarini geologiya kabi fanlar o‘rganadi. Statistika esa ushbu hodisalarni miqdoriy parametrlar asosida tahlil qiladi. Bunda statistika hodisalarning turli omillar ta’siridagi o‘zgaruvchanlik xususiyatini – variatsiyasini e’tiborga oladi. Shu jihatdan statistik ma’lumotlar hodisaier rivojining dinamikasini aks ettira oladi.

Statistikaning til hodisalariga tatbiq etilishi natijasida lingvostatistika sohasi shakllandi, unda til birliklari, lisoniy hodisalar, matnlar statistik

¹ Abdullayev Y. Statistika nazariyasi. - T., 2002. - B.4 - 5.

tahlil etiladi. Matnning statistik tahlili – tilni ehtimollik belgilariga ko‘ra tahlil qilish, til hodisalarining 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.

Til hodisalarini statistik metod asosida tahlil qilish uchun doim ma‘lum bir matn yoki matnlar to‘plami olinadi. Ular badiiy adabiyotdan, gazeta va jurnallardan, publitsistik maqolalardan, ilmiy asarlardan, dialektologlarning jonli nutq orqali yozib olgan materiallaridan tanlanadi. Matnning hajmi tekshirilayotgan hodisalarining umumiy sonidir, uni «H» bilan belgilash mumkin. Masalan, 100 ta so‘z; 37ta bosh kelishik shaklidagi ot kabi. Aniqlanayotgan hodisani ko‘rsatadigan miqdor (masalan, ot oldida aniqlovchi vazifasida keladigan takrorlangan otlar miqdori) absolyut chastota hisoblanadi, uni «M» harfi bilan belgilash mumkin. 37ta otdan 30tasining aniqlovchi vazifasida kelishi absolyut chastotadir. Absolyut chastota (M)ning "matn hajmi (H)ga nisbati (M/H)dan nisbiy chastota kelib chiqadi.¹ Nisbiy chastota «P» harfi bilan belgilanadi. Nisbiy chastota uch xil usul bilan aniqlanadi:

- a) oddiy kasr hisobida: $P = M/H$
- b) protsent asosida: $P = M \cdot 100/N$
- c) promil hisobida: $P = M \cdot 1000/H$

Matnning statistik tahlilida Zif qonunidan o‘rinli foydalaniladi. Katta hajmdagi matnlarda har bir so‘zning qanchalik tez takrorlanishini (chastotasini) va ularning takrorlanish tezligini hisoblashda Zif qonuni amal qiladi.² 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 so‘z $r = 2$, undan keyingisi $r = 3$ tarzida belgilanadi. Bundan kelib chiqadiki, matndagi ixtiyoriy so‘z uchun (f) chastotaga (r)ning ko‘paytmasi natijasi o‘zgarmas miqdor - konstanta (C)ga teng bo‘ladi ($r^4 f = C$), bu yerda C matn hajmiga bog‘liq. Masalan, amerikalik yozuvchi Mark Tvenning «Tom Soyerning sarguzashtlari»

¹ Усмонов С. Умумий тилшунослик. – Т.: Ўқитувчи, 1972. - Б.199.

² Плат У. Математическая лингвистика // Новое в лингвистике. - М.: Прогресс, 1965. Вып.IV. - С.204.

asari Zif qonuni asosida tahlil etilgan. Asarda jami 71 370 ta belgi mavjud bo‘lib, takrorlanmagan so‘z tiplari 8018 ta. Matnda ishlatilgan so‘zlarning o‘rtacha takrorlanish chastotasi 8,9 dir, ya’ni matndagi so‘zlar taqriban 9 marta takrorlanadi. Lekin bunda eng katta muammo shuki, so‘z tiplarining matn ichida qo‘llanish darajasi bir xil emas. Ayrim so‘zlar 700 marta takrorlanadi, bunday so‘zlar romanning 1% qismini tashkil etadi. Yana shunday so‘zlar ham borki, ular bir marta uchraydi. Bunday so‘zlar *hapas legomena* (yun. «faqat bir marta o‘qi») deb ataladi, ular asarning qariyb yarmini tashkil etadi. Qariyb 90% so‘zlar 10 marta yoki undan kam, 10% ga yaqin so‘z tiplari 3 marta yoki undan kam miqdorda takrorlanadi.

Matnning statistik tahlili matn atributsiyasi muammosiga ham oydinlik kiritishda muhim tadqiqot usuli hisoblanadi. 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 kvantitativ 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.¹ Bu aspekt kriminalistika sohasida muhim ahamiyat kasb etmoqda.

Amerikalik olim Jozef Grinberg kvantitativ metodni tilshunoslikka - morfologik tiplar nazariyasiga tatbiq etgan. Tadqiqot natijalarini «*Quantitative approach to the morphological typology of languages*» («*Tillar morfologik tipologiyasiga kvantitativ yondashuv*») nomli maqolasida umumlashtirgan (1960). Olim o‘zigacha amalga oshirilgan ishlar haqida shunday yozadi: «*Tillarning tipologik tasnifi mantiqiy-ratsional mezonlar asosida yuzaga kelgan. Tasnif namunalari nazariy jihatdan asoslangan, ammo lingvistik fakt bilan uzviyligi yaqqol ko‘zga tashlanmaydi. Bu esa tasnifni amaliyotga tatbiq etib tekshirishni taqozo etadi*»²

¹ Баранов А.Н. Введение в прикладную лингвистику. - М.: Эдиториал УРСС, 2001.-С.24.

² Гринберг Дж. Квантитативный подход к морфологической типологии языков // Новое в лингвистике. - М., 1963. Вып. III, - С.60 - 94. Bu haqida qarang: Rahimov A. Kompyuter lingvistikasi asoslari.-T.,2011.

Shu nuqtayi nazardan olim matematik aniqlikka erishish maqsadida til hodisalarini, mavjud lisoniy tiplarning uchrash darajasini miqdoriy parametrlar asosida qayta baholadi. J.Grinberg kvantitativ metodni 8 tilga (sanskrit, anglo-saks, yoqut, vyetnam, fors, ingliz, suaxili, eskimos tillariga) tatbiq qilgan. Bunda olim o‘zi belgilagan formulalar asosida lisoniy tiplarning va tillarda kuzatiladigan tipologik belgilarning miqdoriy tavsifini, statistik ma’lumotlarni ilmiy jamoatchilikka havola etdi. Masalan, sintetiklikning dunyo tillaridagi darajasini aniqlash uchun M/W , agglyutinatsiyani aniqlash uchun A/J , kompozitsiyani aniqlash uchun R/W , derivatsiyani aniqlash uchun D/W , prefiksatsiyani aniqlash uchun P/W , suffiksatsiyani aniqlash uchun S/W formulalaridan foydalandi. Bu yerda M – ma’lum tipdagi chekli (100 so‘zdan iborat) matn ichida uchragan morflar soni, A – agglutinatsiya, R (root) - o‘zak, D (derivational) – yasovchi qism, P – prefiks, S – suffiks, W (Word) - matndagi so‘zlar sonidir. Hisoblash natijalariga ko‘ra, vyetnam tili uchun – 1,06 (ya’ni 100 so‘zda 106 morf uchraydi), ingliz tili uchun – 1,68, sanskrit tili uchun – 2,59, eskimos tili uchun – 3,72 ko‘rsatkichlari belgilandi. Olim Natijalardan kelib chiqib tillarni quyidagicha baholaydi: «2 dan past ko‘rsatkichga ega bo‘lgan tillar (vyetnam, ingliz, fors, xitoy, italyan, nemis tillari) analitik tillar, 2 dan 3 gacha miqdoriy ko‘rsatkichga ega bo‘lgan tillar (rus, sanskrit, qadimgi yunon, lotin, eski slavyan, chex, polyak, yoqut, suaxili tillari) sintetik tillar, 3 dan yuqori ko‘rsatkichga ega bo‘lgan tillar (eskimos, ayrim kavkaz tillari, Amerika hindulari tillari) polisintetik tillar hisoblanadi». Rus olimi V.Z.Demyankovning fikricha, Grinberg tavsiya etgan meiod tilshunoslikda mavjud bo‘lgan kartotekalash usulidan ishonchliroq va aniqroqdir¹.

Statistikadan, tovushlar, so‘zlar, birikmalar chastotasini belgilashdan so‘z tarkibidagi o‘zakni aniqlashda ham foydalanish mumkin.

Agglyutinativ tillarda o‘zak va qo‘shimcha orasidagi chegar aniq sezilib turadi. Bu tillarda o‘zakni aniqlash oson.

Flektiv tillarda o‘zak va qo‘shimcha orasidagi chegara sezilmaydi, bu esa ularda o‘zakni aniqlash muammosini keltirib chiqaradi. Ingliz, arab tillari flektiv tillar guruhiga mansubdir. Bu tillarda o‘zakni aniqlashda yuqoridagi usuldan foydalanish mumkin. Bunda eng ko‘p takrorlanuvchi shakllar «o‘zak» («root») sifatida olinadi.

¹ Po‘latov A., Muhamedova S. Kompyuter lingvistikasi. –T., 2007. Rahimov A. Kompyuter lingvistikasi asoslari. –T., 2011.

Savol va topshiriqlar:

1. Statistika tahlilining ahamiyatini tushuntiring.
2. Mustaqil ravishda amalga oshirgan statistik tadqiqotingiz haqida ma'lumot bering.
3. Matematik statistikaning ahamiyati, mazmun-mohiyatini yoritib.
4. Muayyan matnning so'z turkumlari bo'yicha statistik tahlil asosida tekshiring.
5. Matn atributsiyasi tushunchasini izohlang.
6. Badiiy asardan bir sahifa ko'chiring, o'z va o'zlashgan qatlam leksikasini statistik tahlil qiling.
7. Kartochka usulidagi an'anaviy hisoblash bilan kompyuter statistikasini "Venn diagrammasi" asosida tahlil qiling.

18-§. KOMPYUTER LINGVISTIKASINING AHAMIYATI

1. Informatsion jamiyat va kompyuter lingvistikasi.
2. "Informatsion jamiyat" tushunchasi.
3. Kompyuter lingvistikasi istiqboli.

Asosiy tushunchalar: *informatsion jamiyat, istiqbolli yo'nalishlar, neyron tarmoqlar, ekspert tizimlari, lingvistik protsessorlar, tabiiy tilli interfeys, biokompyuter, semiotik va biologik yondashuv, mashina ijodi, intellektual (ratsional) agent, bilimlar injeneriyasi.*

Insoniyat uchinchi ming yillikda axborot asriga qadam qo'yishi bilan bog'liq holda informatsion jamiyat konsepsiyasi ishlab chiqildi. Bu konsepsiyaning asoschisi amerikalik futurolog olim E. Toffler hisoblanadi. U jamiyatni turli bosqichlarga bo'lib o'rganishni taklif etib, **sivilizatsiyaning uch to'liqini** ajratdi: a) birinchi to'liqin – industrial jamiyatgacha bo'lgan davr (qishloq xo'jaligi sivilizatsiyasi). Bunda asosiy boylik yer va tabiatdan olingan zaxiralar hisoblanadi; b) ikkinchi to'liqin – industrial sivilizatsiya. Bunda asosiy boylik ishlab chiqarish vositalari sanaladi; d) uchinchi to'liqin -postindustrial sivilizatsiya (informatsion jamiyat). Bunda kapital va mehnat o'z o'rnini informatsiya va bilimga bo'shatib beradi.

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 Deklaratsiya¹ hamda information jamiyat qurish bo'yicha ish-reja.² Mazkur hujjatlarda belgilanishicha, endilikda jamiyatning taraqqiyoti kompyuter texnologiyalari bilan chambarchas bog'liq bo'lib qoladi. Shu o'rinda alohida ta'kidlash lozimki, kompyuter lingvistikasi informatsion texnologiyalarning 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. Kompyuterlardan teoremlarni isbotlashda, rasm chizishda, shaxmat o'yinida foydalanilishi sun'iy intellekt bo'yicha izlanishlarning ilk samarasidir. Insonning aqliy faoliyatini modellashtiruvchi dasturlarning yaratilishi muayyan sohalardagi murakkab masalalarni yechuvchi ekspert tizimlarining yuzaga kelishiga olib keldi.

Kompyuter lingvistikasi ijtimoiy-iqtisodiy, fan taraqqiyoti asosidagi soha hisoblanadi.

Kompyuter lingvistikasi fanidan kutilayotgan natijalar quyidagicha:

1) kompyuterda sun'iy intellektni yaratish ehtiyoji yuzaga keldi (sun'iy intellekt vaziyatdan kelib chiqib, manfaatli va ijobiy samaraga ega bo'lgan funksiyalarni bajaradi);

2) hozirga qadar kompyuter lingvistikasi shakl va belgilar orqali muloqotni ta'minlagan бўлса, endilikda ovozli buyruqlarni qabul qilishga ixtisoslashgan dasturlar ustida ish olib borish lozim bo'ladi;

3) mashina tarjimasini rivojlantirish, avtomatik tarjima mukammalligiga erishish asosiy maqsadlardan biridir;

4) kompyuter lingvistikasida buyruqlarni belgi va simvollarsiz mashinaga yetkazib berish usullarini ishlab chiqish va tajribalar o'tkazish talab etiladi.

5) kompyuter lingvistikasining ijtimoiy sohalar taraqqiyotini ta'minlovchi vosita sifatidagi ahamiyatini yoritish lozim bo'ladi.

Kompyuter lingvistikasi fanining bosh maqsadi – odam va mashina o'rtasidagi muloqotni ta'minlashdan iboratdir.

¹World 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.

²World summit on the information society. Plan of Action. Document wsis-03/ geneva/doc/5-e. 12 December, 2003.

Sohada erishilgan yutuqlar sirasiga sun'iy til va mashina tarjimai borasidagi yutuqlarni keltirish mumkin. Endilikda "sun'iy ong" shakllantirish asosiy vazifa bo'lib, u mashina – kompyuterning vaziyatdan kelib chiqib harakatlanishini ta'minlaydi.

O'zbek kompyuter lingvistikasini to'laqonli, chinakam fan sifatida shakllantirish uchun bosqichli va uzviy reja asosida ilmiy tadqiqotlar olib borish lozim.

Keyingi yillarda amaliy lingvistika sohasiga ijtimoiy taraqqiyotni belgilovchi omillardan biri sifatida e'tibor qaratilmoqda. Amaliy lingvistika birinchi navbatda o'qitish metodikasini takomillashtirishga asoslangani, grammatik xususiyatlarni umumlashtirgan holda chet tillarini, milliy tilni ona tili va chet tili sifatida o'qitish bilan shug'ullangani uchun ham unga bo'lgan ijtimoiy-iqtisodiy ehtiyoj ortib bormoqda.

Amaliy lingvistika til birliklaridan kundalik faoliyatda foydalanish darajasini aniqlovchi, til sohibining amaliy faoliyati bilan bog'liq bo'lgan soha hisoblanadi. O'zbek tilshunosligida shu kunga qadar amaliy lingvistika borasida tadqiqot olib borilmagan yoki ilmiy izlanishlar yaratilmagan, degan fikrdan yiroqmiz. Negaki, mavjud adabiyotlardagi fikr-mulohazalarda amaliy lingvistika masalalari qisman bo'lsa-da, aks etgan. Faqat bu masalalar shunchaki axborot tarzida berib o'tilgan. Hozirgi davr amaliy lingvistikaning alohida soha sifatida rivojlanishini taqozo qilmoqda. Axborotni tez sur'atda yetkazish, ketma-ketlikda o'zlashtirish, ma'lumotlar oqimini boshqarish amaliy lingvistikaning dolzarbligini belgilaydi. Amaliy lingvistika til imkoniyatlaridan foydalanish darajasini, til birliklarining nutqiy jarayondagi semantik xususiyatlarini o'zlashtirish masalalarini ham ifoda etmoqda. Shu jihatdan, amaliy lingvistika tamoyillari bugungi kunda ijtimoiy sohalarda, ishlab chiqarish birlashmalarida til birliklarining pragmatik, psixolingvistik xususiyatlari bilan bog'liq muammoli vaziyatlarda, lingvistik ekspertiza jarayonlarida hal qiluvchi ahamiyatga egadir.

Savol va topshiriqlar

1. Kompyuter lingvistikasi fanining zamonaviy yo'nalishlari haqida internetdan ma'lumot oling.

2. Lingvistik kompyuter dasturlarining ishlash samaradorligini oshirish tadbirlarini ayting.

3. Kompyuter lingvistikasi yo'nalishlarining ijtimoiy-iqtisodiy hayotdagi o'rnini yoriting.

4. Kompyuter lingvistikasi va globallashtirish jarayoni borasidagi

mulohazalaringizni bayon eting.

5. Kompyuter lingvistikasining o‘zbek tili taraqqiyotidagi ahamiyatini ko‘rsating.

6. Til o‘qitish dasturlarini yod ayting.

7. Kompyuter lingvistikasining til taraqqiyotidagi ahamiyatini “swod” tahlili orqali yoriting.

“Kompyuter lingvistikasi” fani yuzasidan blis-so‘rov

Mashina tarjimasi bo‘yicha birinchi tajribalar qaerda va qachon amalga oshirildi? –1954- yilda AQShning JORJtaun universitetida.

Kompyuter lingvistikasi XX asrning 80-yillariga qadar qanday nomlar bilan yuritildi?

– *“hisoblash lingvistikasi”, “matematik lingvistika”, “kvantitativ lingvistika”.*

NLP (Naturel lenguige prosessing)ning ma‘nosi?

– *tabiiy tilni qayta ishlash.*

Inglizcha *parsing* (*pirsing*) so‘zining ma‘nosi?

– *“tahlil”.*

Til va nutq birliklari lug‘ati, grammatik ma‘lumotnomalar, matnlar korpusi qanday qismni tashkil etadi?

– *deklarativ qismni.*

“Kompyuter lingvistikasi” tilni insondan tashqaridagi mavhum sistema, sof munosabatlar tizimi sifatida tavsiflagan olim?

– *L. Yelmslev*

“Uchinchi mustasno” qonuni qaysi yo‘nalishda ustuvor?

– *formal mantiqda: “ikki zid fikrdan biri rost, ikkinchisi xato, uchinchi holat bo‘lishi mumkin emas”.*

Yunon madaniyatida qaysi sohalar ta‘lim poydevori hisoblangan?

– 7 erkin san‘at (*artesliberales*) ta‘lim poydevori sifatida qaralgan. Uchtasi *grammatika*, *ritorika* (notiqlik san‘ati) va *mantiq* “formal san‘atlar (*artesformales*) sifatida *trivia* (uchlik)ga birlashgan; *arifmetika*, *geometriya*, *musiqa*, *astronomiya* “real san‘atlar” (*artesreales*) sifatida *kvadrivia* (to‘rtlik)ka birlashgan.

Semiotikada belgilar to‘rt aspekt asosida tekshiriladi. Bu aspektlar qaysilar?

– 1. *Sintaktika* – belgilar strukturasi. 2. *Semantika* – belgilarning mazmun tomoni. 3. *Sintagmatika* – belgi va uni aks ettiruvchi obyekt orasidagi munosabat. 4. *Pragmatika* – belgilarning insonga bo‘lgan munosabati, foydasi, amaliy jihati.

“Kompyuter lingvistikasi”da “matla” yoki “mabda”ning ma’nodoshi bo‘lgan termin?

–algoritm (lot. “ibtido”).

Lingvistik ta’minot bazasini boshqarish vositalari (algoritm tuzish, dastur yaratish, kompyuter analizi va sintezi) lingvistik vositalarning qanday qismini tashkil etadi?

–*protsedura qismini*

Kvantitativ tavsif nima?

–*miqdoriy tavsif, algoritmlash, modellashtirish, statistik tahlil.*

To‘rt jildli “Mantiq fani” nomli monografiya muallifi?

–*Fridrix Gegel.*

“Oraliq uchinchi” qoidasini mantiqqa kim olib kirgan?

–*Fridrix Gegel.*

Semiotika qachon vujudga keldi, asoschisi kim?

–*XIX asr oxirida, asoschisi Charlz Pirs*

“Signifikat” deganda nima nazarda tutiladi?

– “leksik ma’no”.

“Kompyuter lingvistikai”dagi qaysi tushuncha hasharot nomi bilan omonimlik hosil qiladi?

– *bit (binit inglizcha buhapy unit” – “ikkilik birligi” deb ataladi).*

Algoritm ko‘rinishlari qanday holatda bo‘lishi mumkin?

– *a) so‘z ko‘rinishida; b) grafik ko‘rinishda; d) psevdokod ko‘rinishida– algoritmlarning yarim formallashtirilgan tavsifi, programmalash, dastlabki bosqichda qo‘llanadi; e)programma ko‘rinishida.*

4 raqamiga Charlz Friz konsepsiyasi asosida yondashilsa...

–ravishni anglatadi.

Bola uchun teshik kulcha, undan kattaroqlar uchun aylana, ximik uchun modda, tilshunos uchun tovush, kompyuter lingvistikasi modellashtirish yo‘nalishida to‘ldiruvchi, matematik uchun “qiymatsiz”...Gap nima haqida?

–*0 (nol) haqida.*

Topshiriq: O‘zingiz ham xuddi shu yo‘nalishda blis-so‘rovlar tuzing va bilimlaringizni mustahkamlashda foydalaning.

“Kompter lingvistikasi” fanidan test savollari

1. *“Kompter lingvistikasi”ning hozirga qadar qo‘llangan nomlarini belgilang*

A) hisoblash lingvistikasi, matematik lingvistika, kvantitativ lingvistika, injener lingvistikasi

B) hisoblash texnikasi, matematik lingvistika, kvantitativ lingvistika, injener lingvistikasi

C) hisoblash lingvistikasi, matematik lingvistika, kvantitativ matematika, injener lingvistikasi

D) hisoblash texnikasi, matematik lingvistika, kvantitativ lingvistika, til injeneriyasi

2. *“Kompyuter lingvistikasi”ning vazifalari nimadan iborat?*

A) lingvistik masalalarni kompyuterda hal qilishga o‘rgatish

B) dunyoviy yagona til yaratish

C) kompyuterdan foydalanishga o‘rgatish

D) kompyuter dasturlari tuzishni o‘rgatish

3. *Mashina tarjimasi qaysi sohaga tegishli?*

A) amaliy lingvistika

B) informatika

C) sotsiolingvistika

D) Matematik lingvistika

4. *O‘zbek tili axborot uslubi qaysi sohalar doirasida faol qo‘llanadi?*

A) biznes, huquq, tabiiy fanlar, meditsina, sport, ta'lim, diplomatiyada

B) adabiyot va san'atda

C) hamma sohalarda

D) faqat tilshunoslikda

5. *Mashina tarjimasi haqidagi birinchi fikr kimga tegishli edi?*

A) Ch. Bebbidj

B) DJ. Mey

C) Ch. Fergyusson

D) DJ. Taun

6. *Birinchi mashina tarjimasi qachon amalga oshirildi?*

A) 1954

B) 1952

C) 1960

D) 1953

7. *Ilk mashina tarjimasi qaysi sohada amalga oshirilgan edi?*

A) ximiya

- B) tilshunoslik
- C) bilologiya
- D) badiiy adabiyot

8. *Ilk mashina tarjimasi qaysi kompaniya tomonidan amalga oshirildi?*

- A) "IBM"
- B) "IBM" va "Apple"
- C) "Apple"
- D) "Microsoft"

9. *Birinchi mashina tarjimasi qayerda amalga oshirildi?*

- A) AQSh
- B) Fransiya
- C) Germaniya
- D) Angliya

10. *XX asrning 50- 60 yillaridagi mashina tarjimasi, asosan, qaysi tillar o'rtasida amalga oshirilgan?*

- A) rus- ingliz
- B) rus-fransuz
- C) ingliz-nemis
- D) ingliz-xitoy

11. *O'zbek tili axborot uslubi asosiy prinsiplari qaysilar?*

- A) formallik, aniqlik, qisqalik, bir ma'nolilik
- B) badiiylik, tasviriy ifoda
- C) ko'pqirralilik, ko'p ma'nolilik
- D) qisqalik, badiiylik

12. *R.Piotrovskiy yetakchiligidagi mashina tarjimasini amalga oshiruvchi ilmiy guruh qanday nomlangan?*

- A) «nutq statistikasi»
- B) «tarjima markazi»
- C) «til statistikasi»
- D) «mashina tarjimasi»

13. *Birinchi mashina tarjimasi qaysi shahar nomi bilan bog'liq?*

- A) Jorjtaun
- B) Koliforniya
- C) Texas
- D) Kentukki

14. *AQShda qaysi tarjima tizimi amaliyotga tatbiq qilish uchun tavsiya qilingan?*

- A) rus-ingliz

- B) ingliz- yapon
- C) fransuz- ingliz
- D) nemis-ingliz

15. «*Birinchi mashina tarjimasining tarjimon dasturlari qanday nomlangan?*

- A) MARK, GAT
- B) MARK, PROMT
- C) MARK, LINGVO
- D) MARK, SOKRAT

16. *O‘zbek tili axborot uslubining formal modelida tub so‘z deganda nima tushuniladi?*

- A) kompyuter lug‘atida tub so‘z deb ko‘rsatilgan so‘z
- B) qo‘shimcha va affikslari bo‘lmagan so‘z
- C) ikkitta so‘zga ajralmaydigan so‘z
- D) imlo lug‘atida berilgan so‘zlar

17. *Mashina tarjimasi» yana qanday nomlanadi?*

- A) avtomatik tarjima
- B) avtomatlashtirish
- C) elektron tarjima
- D) kibernetik tarjima

18. *Kompyuter lingvistikasi nuqtayi nazaridan yagona dunyoviy til yaratish muammosini qanday hal etish mumkin?*

- A) hamma tabiiy tillar asosida to‘liq kompyuter ma’lumotlar bazasini yaratib, elektron tarjimonlardan foydalanishga o‘tish asosida
- B) hamma tabiiy tillar asosida to‘liq kompyuter ma’lumotlar bazasini yaratib, ingliz tiliga o‘tkazish asosida
- C) butun dunyoda esperanto tiliga o‘tish asosida
- D) muammoning yechimi yo‘q.

19. «*Mashina tarjimasi* » nima?

- A) matnlarni maxsus kompyuter dasturlari yordamida tarjima qilish
- B) tabiiy tildan mashina tiliga o‘girish
- C) mashina tilidan tabiiy tilga o‘girish
- D) sun’iy tildagi matnni tabiiy tilga o‘girish

20. *Matematik mantiqda “mulohaza” deganda nima tushuniladi?*

- A) faqat rost yoki yolg‘on qiymat qabul qiluvchi fikr
- B) ixtiyoriy fikr
- C) faqat rost qiymat qabul qiluvchi fikr
- D) faqat yolg‘on qiymat qabul qiluvchi fikr

21. *O'zbek tili axborot uslubida xalqaro so'zlarni qo'llash muammosining yechimi...*

- A) imkon darajasida xalqaro so'zlarni qo'llash
- B) faqat xalqaro so'zlardan foydalanish
- C) xalqaro so'zlarni qo'llamaslik
- D) xalqaro so'zlarning tarjimasini qo'llash

22. *EHMdan tarjima maqsadida foydalanish fikri, dastavval, qachon aytilgan?*

- A) 1949- y.
- B) 1950-y.
- C) 1973-y.
- D) 1954-y.

23. *Aksioma nima?*

- A) nazariyaning isbot talab qilmaydigan boshlang'ich tasdiqi
- B) ixtiyoriy teorema
- C) nazariya formulasi
- D) teoremaning aksi

24. *Aksiomatik nazariya simvollari...*

- A) nazariyada ishlatiladigan barcha belgilar
- B) nazariyada ishlatiladigan harflar
- C) nazariyada ishlatiladigan tinish belgilari
- D) nazariyada ishlatiladigan ramzlar

25. *Mashina tarjimasiga oid ishlar hali amalga oshirilmagan mamlakatni ko'rsating.*

- A) Avstriya
- B) Italiya
- C) Xitoy
- D) AQSh

26. *Aksiomatik nazariya aksiomalari...*

A) nazariyaning boshlang'ich isbotsiz qabul qilingan mutlaqo to'g'ri formulalari

- B) nazariyaning ixtiyoriy teoremalari
- C) nazariyaning ixtiyoriy tenglamalari
- D) nazariyaning tizimli holda qabul qilingan ixtiyoriy formulalari

27. *Matnlarni tarjima qilish jarayoni– bu ...*

- A) mashina tarjimasi
- B) elektron tarjima
- C) avtomatlashtirish
- D) yozma tarjima

28. *Tabiiy tillardan avtomatik tizimda informatsiya uchun foydalanish masalalari bilan shug'ullanadigan fan...*

- A) kompyuter lingvistiksi
- B) informatika
- C) hisoblash lingvistikasi
- D) internet

29. *Aksiomatik nazariya keltirib chiqarish qoidalarini...*

A) nazariyada aksiomalardan teoremlar yaratishda foydalaniladigan qoidalar

- B) nazariyaning ixtiyoriy qoidalarini
- C) nazariyaning tizimli holda qabul qilingan ixtiyoriy formulalari
- D) nazariyaning ixtiyoriy tenglamalari

30. *Aksiomatik nazariya teoremlari...*

A) nazariya aksiomalardan keltirib chiqarish qoidalarini asosida olinadigan formulalar

B) isbotga ega bo'lgan tasdiqlar

C) nazariyada aksiomalardan teoremlar yaratishda foydalaniladigan qoidalar

D) nazariyaning formulalari

31. *O'zbek tili axborot uslubi grammatikasining aksiomatik nazariyasi modeli...*

A) ega + to'ldiruvchi + hol + kesim

B) ega + hol + to'ldiruvchi + kesim

C) ega + to'ldiruvchi + kesim + hol

D) ega + to'ldiruvchi + aniqlovchi + kesim

32. *Kompyuter lingvistikasi qachon mustaqil yo'nalish sifatida shakllandi?*

A) XX asrning 60- yillarida

B) XIX asrning 90 -yillarida

C) 2000 – 2005- yillarda.

D) XIX asrning 70 -yillarida

33. *Kompyuter lingvistikasi nima bilan shug'ullanadi?*

A) tabiiy tillardan kompyuter tizimida foydalanish masallari bilan

B) elektron lug'atlar yaratish bilan

C) ma'lumot uzatish va axborot berish bilan

D) mashina tarjimasini tuzilishi bilan

34. *Kompyuter lingvistikasiga oid tadqiqotlar dastlab qaysi davlatlarda olib borildi?*

A) Rossiya, Xitoy, Ispaniya

- B) AQSh, Germaniya, Yaponiya
- C) Angliya , Fransiya, AQSh
- D) Ispaniya, Germaniya, Rossiya

35. *Algoritm nima?*

- A) sonli, mantiqiy masalani bosqichma-bosqich yechish yo‘li
- B) Masalani yechish g‘oyasi
- C) masalaning yechimi
- D) sonli, mantiqiy masalaning muhim qismini yechish yo‘li

36. *Kompyuter lingvistikasini shartli ravishda qanday qismlarga ajratish mumkin?*

- A) Deklarativ va protsedurali
- B) protsedurali va funksional
- C) stereotipli va funksional
- D) deklarativ va me‘yorlashtirilgan

37. *Algoritm so‘zi qaysi olim nomidan olingan*

- A) Al-Xorazmiy
- B) Yevklid
- C) Gilbert
- D) Al-Farg‘oniy

38. *Kompyuter lingvistikasi – bu...*

- A) masalalarning yechimi bilan bog‘liq bo‘lgan bilimlar sohasi
- B) ilmiy bilimlar sohasi
- C) nazariy bilimlar bilan bog‘liq bilimlar sohasi
- D) amaliy vazifalar bilan bog‘liq bilimlar sohasi

39. *Til va nutq birliklari lug‘ati, matnlar va grammatik tablitsalarning turli ko‘rinishi lingvistik vositalarning qaysi ko‘rinishiga mansub?*

- A) deklarativ
- B) avtomatlashtirilgan
- C) protsedurali
- D) amaliy

40. *Kompyuter lingvistikasi bo‘yicha tadqiqot olib borgan olimlarni belgilang.*

- A) J.Martin, G.Belonogov, Ch.Bebbij, R.Piotrovskiy
- B) J.Martin, G.Belonogov, Ch.Bebbij, A.Reformatskiy
- C) J.Martin, G.Ramstedt, Ch.Bebbij, R.Piotrovskiy
- D) J.Martin, G.Belonogov, Ch.Bebbij, G.Ramstedt

41. *Programma (dastur) deganda nima tushuniladi?*

- A) algoritmtning kompyuter uslubida yozilgan ko‘rinishi.
- B) biror ishni bajarish rejasi

- C) ko'rsatuv va eshittirishlar tizimi
- D) biror ishni tizimli, izchil bajarish rejasi

42. *Kompyuter lingvistikasi haqidagi fan qachon paydo bo'ldi?*

- A) XX asrning 50-yillarida
- B) XX asrning 60-yillarida
- C) XX asrning 70-yillarida
- D) XX asrning 80-yillarida

43. *Kompyuter lingvistikasi bo'yicha tadqiqot olib borgan o'zbek olimlarini belgilang.*

- A) A.Qayumov, S.Muhamedov, S.Rizayev, H.Arziqulov
- B) A.Po'latov, S.Muhamedov, S.Ahmedov, H.Arziqulov
- C) A.Po'latov, S.Mo'minov, S.Rizayev, H.Arziqulov
- D) A.Po'latov, S.Muhamedov, S.Rizayev, H.Ne'matov

44. *"Kompyuter lingvistikasi"ning asosiy ilmiy muammolarini ko'rsating.*

- A) matn mazmunini tushunish jarayonini modellashtirish
- B) xatolarni avtomatik avtomatik tarzda aniqlash
- C) nutqni analiz qilish masalalari
- D) frazeologizmlarni avtomatik aniqlash masalalari

45. *1960 - yilda qayerda avtomatik tarjimaga bag'ishlangan maxsus yig'ilish bo'lgan?*

- A) sobiq Ittifoq FAsida
- B) INFORM ELECTRO institutida
- C) amaliy matematika institutida
- D) MDUda

46. *O'zbek tilshunosligida «kompyuter lingvistikasi» tushunchasini birinchi bo'lib qo'llagan olim?*

- A) A.Po'latov
- B) A.Jumaboyev
- C) I.Mo'minov
- D) S.Muxamedov

47. *"Kompyuter lingvistikasi" tilshunoslikning qanday bo'limi hisoblanadi?*

- A) amaliy
- B) nazariy
- C) umumiy
- D) xususiy

48. *Kompyuter lingvistikasining tushunchaviy apparati qanday elementlarni o'z ichiga oladi?*

- A) ontologik va eksperimental
- B) ontologik va tatbiqiy
- C) eksperimental va tadqiqiy
- D) ontologik va nazariy

49. *Qozog‘istondagi qaysi olim rahbarligida lingvostatistik tadqiqotlar olib borilgan?*

- A) A.Bektayev
- B) N.Bo‘ronov
- C) A.Po‘latov
- D) S.Rizayev

50. *“Kompyuter lingvistikasi” qaysi xususiyatlari bilan ajralib turadi?*

- A) til hodisalarini kompyuter tizimida aks ettirish bilan
- B) sun'iy tillarni o‘rganish metodi bilan
- C) tabiiy tillarni o‘rganish metodi bilan
- D) Kompyuter lingvistikasi masalalarini hal etish bilan

51. *Sobiq Ittifoq FA sida avtomatik tarjima bo‘yicha tashkil qilingan laboratoriyaga kim rahbarlik qilgan?*

- A) A.Belonogov
- B) N.Bo‘ronov
- C) A.Bektayev
- D) A. Baranov

52. *“Kibernetika” fanining shakllanishi qaysi olim nomi bilan bog‘liq?*

- A) Vinner
- B) Amper
- C) Platon
- D) Sokrat

53. *1974- yilda INFORM- ELECTRO institutida nechta tizimlari bo‘yicha qancha ilmiy tadqiqot olib borilgan?*

- A) 3ta
- B) 2ta
- C) 4ta
- D) 5ta

54. *1974 -yilda INFORM- ELECTRO institutida qaysi tarjima tizimlari bo‘yicha ilmiy tadqiqotlar olib borilgan?*

- A) fransuzcha – ruscha, inglizcha – ruscha, nemischa – ruscha
- B) fransuzcha – nemischa, inglizcha – ruscha, nemischa – ruscha
- C) nemischa – ruscha, inglizcha – ruscha, nemischa – ruscha
- D) fransuzcha – ruscha, inglizcha – nemischa, nemischa – ruscha

55. *INFORM- ELECTRO* da *YETAR-1*, *YETAR-2*, *NERPA* tarjima tizimlari qachon yaratilgan?

- A) 1974 - yilda
- B) 1947 - yilda
- C) 1957 - yilda
- D) 1975 - yilda

56. 1959 - yilda qayerda mashina tarjimasi bo'yicha laboratoriya ochilgan?

- A) Moskvada
- B) Olma-otada
- C) Toshkentda
- D) Leningradda

57. *O.Kulagina* va *I. Melchuklar* rahbarligida qanday tarjima tizimi yaratilgan?

- A) FR-1
- B) ETAR
- C) ETAR-2
- D) Cokrat

58. Tarjima bo'yicha yaratilgan ilk algoritm bazasi necha so'zdan iborat bo'lgan?

- A) 50 000
- B) 10 000
- C) 30 000
- D) 70 000

59. 1959- yilda qaysi muassasada mashina tarjimasi bo'yicha laboratoriya ochilgan?

- A) Moskva davlat lingvistik universitetida
- B) Moskva davlat universitetida
- C) Moskva davlat tarjima universitetida
- D) Moskva davlat texnika universitetida

60. *INFORM-ELECTRO*da yaratilgan mashina yordamidagi avtomatik tizimlarda qaysi tildan qaysi tilga tarjima qilish nazarda tutilgan?

A) fransuzchadan ruschaga, inglizchadan-ruschaga, nemischadan ruschaga

B) ransuzchadan nemischaga, inglizchadan ruschaga, nemischadan ruschaga

C) nemischadan ruschaga, inglizchadan ruschaga, nemischadan ruschaga

D) fransuzchadan ruschaga, inglizchadan nemischaga, nemischadan ruschaga

61. *Kibernetika qanday fanlar sirasiga kiradi?*

A) sohalararo

B) gumanitar

C) aniq

D) ijtimoiy

62. *Moskva davlat lingvistik universitetida mashina tarjimasi bo'yicha laboratoriya qachon ochilgan?*

A) 1959- yilda

B) 1957 - yilda

C) 1954- yilda

D) 1958 - yilda

63. *Nelson tarjima tizimidagi matnlar umumlashmasini qanday nomlagan edi?*

A) gipertekst

B) ssilka (havola)

C) tezaurus

D) giperssilka (giperha-vola)

64. *O.Kulagina va I. Melchuklar rahbarligida qanday tarjima tizimlari yaratilgan?*

A) fransuzchadan ruschaga, inglizchadan-ruschaga, fransuzdan ruschaga (yangi variantda)

B) fransuzchadan nemischaga, inglizchadan ruschaga, nemischadan ruschaga

C) nemischadan ruschaga, inglizchadan ruschaga, nemischadan ruschaga

D) fransuzchadan ruschaga, inglizchadan nemischaga, nemischadan ruschaga

65. *Informatsion qidiruv tizimida ma'lumotni tavsiflash qanday nomlanadi?*

A) indekslash

B) modellashtirish

C) referatlash

D) dalillash

66. *Modellashtirish asosidagi murakkab informatsion tizim ...*

A) algoritm

B) intellekt

C) model

D) gipertekst

67. *Qaysi tushuncha kompyuter lingvistikasiga tegishli emas?*

A) matnni belgilar tizimi sifatida tushunish

B) tabiiy tildagi matnlarni sintez qilish

C) tabiiy tilni qayta ishlash

D) tabiiy tildagi matnlarni analiz qilish

68. *Gipertekst nima?*

A) o'tish tizimiga asoslangan matn

B) matn abzasi

C) matnni nomlash

D) katta hajmli matn

69. *Tabiiy tillarni qayta ishlashning nechta ko'rinishi bor?*

A) 5

B) 7

C) 4

D) 10

70. *Belgilar tizimi, uni qabul qiluvchilar va undan foydalanuvchilar o'rtasidagi munosabatga asoslangan semiotika bo'limi qanday nomlanadi?*

A) pragmatika

B) kibernetika

C) avtomatika

D) informatika

“KOMPYUTER LINGVISTIKASI”GA OID ASOSIY TUSHUNCHALAR IZOHI (GLOSSARIY)

Abzas (*nem.* abzatz < absetzen – surmoq, nari surmoq). 1. *ayn.* Xat boshi. 2. Matnning bir xat boshidan keyingi xat boshigacha bo‘lgan qismi: *ikki abzasni tashlab o‘qimoq.*

Abstrakt ot (*lot.* abstractus – uzoqlashgan, mavhum) – mavhum ot.

Avtomatik tarjima– avvaldan tuzilgan algoritm bo‘yicha elektron-hisoblash mashinasi (EHM) yordamida biror tilga oid matni boshqa tilga tarjima qilish.

Adstrat (*nem.* adstrat) – bir hududda turli tillarning uzoq davr birga amal qilishi natijasida yuzaga kelgan ikki tillilik turi.

Aksiomatik nazariya – ilmiy nazariya qurilishi, 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.

Algoritm (bu so‘zning etimologiyasi buyuk alloma Al-Xorazmiy nomi bilan bog‘liq) – ma‘lum bir tipga oid masalalarni yechishda ishlatiladigan jarayonlar tizimining muayyan tartibda bajarilishi haqidagi aniq qoida yoki qoidalar tizimi.

Algoritmik tillar – programmalashtirishga asoslangan, tabiiy tildagidek aniq interpretatorga – insonga yoki avtomatik qurilmaga yo‘nalgan sun‘iy tillar. Algoritmik tillar imperativ xarakterga ega bo‘ladi, ya‘ni, asosan, buyruqlar tizimiga tayanadi (*begin* – boshla, *end* – tugalla kabi).

Analiz – tabiiy til xususiyatlarini tushunish uchun kompyuterda amalga oshiriladigan morfologik, sintaktik va semantik tahlil. *Morfologik analiz* morfemalarga ajratish hamda grammatik ma‘lumotlar tahlilini o‘z ichiga oladi.

Akkumulyatsiya – biror tilda faol ishlatiladigan so‘z, ibora va sh.k. ta‘sirida shunday til birliklarining boshqa tilda ham faol ishlatila boshlashi.

Amaliy tilshunoslik – tilshunoslikning lingvistik masalalarni (mas., mashina tarjimasini) hal etishni amaliy yo‘l bilan o‘rganuvchi sohalari (eksperimental fonetika, leksikografiya, lingvostatistika kabi).

Asos ma‘no. 1. Biror ma‘noning yuzaga kelishida asos bo‘lib xizmat qiladigan ma‘no. Mac., *o‘yin* so‘zining «ko‘ngil ochish, dam olish maqsadidagi erkin (ermak) harakat» ma‘nosi *o‘yinga tushmoq, futbol o‘yini* birikmalaridagi ma‘nosining kelib chiqishiga asos bo‘lgan ma‘nodir.

2. Yasama so‘z asosiga (so‘z yasali sh asosiga) oid ma‘no. Mac., *terim* so‘zining *ter, terimchi* so‘zining *terim* qismiga oid ma‘no.

Avtomatik tarjima (mashina tarjimasi) – bir tildagi matnni ikkinchi bir tilga EHM (kompyuter) vositasida qisqa vaqt ichida tarjima qilish tizimi.

Avtomatik tahrir – kompyuter lingvistikasining asosiy yoʻnalishlaridan biri boʻlib, kompyuterdagi matn muharrirlari (Microsoft Word, Excel, Wordpad, Lexicon) rivojlanishi bilan bogliq boʻlgan tahrir jarayoni. Avtomatik tahrir matndagi mexanik xatolarning avtomatik tarzda toʻgʻrilanishi va xatolar haqidagi signallarning userga taklif etilishi kabilar bilan bogʻliq.

Bevosita ishtirokchilar (immediate constituents) – sintaktik modellarga ajratish. Bevosita ishtirokchilar metodiga muvofiq tayanch nuqta konstruktsiya hisoblanadi. Konstruktsiya tarkibidan ishtirokchilar va bevosita ishtirokchilar ajratiladi.

Bilimlar injenerligi (Knowledge Engineering) – bilimlar bazasi bilan shugʻullanuvchi maxsus yoʻnalish boʻlib, u sunʼiy intellekt tizimining tarkibiy qismi sanaladi. Bilimlar injenerligi kognitiv lingvistika, kompyuter semantikasi va komp-yuter lingvistikasining yutuqlariga tayanadi.

Bilingvizm (*lot. bi < bis – ikki marta + lingua – til*) – *ayn.* ikki tillilik.

Bixevioristik yondashuv – mavjud anʼanaviy oʻqitish tizimining barcha jab-halarini eʼtiborga oluvchi oʻqitish mexanizmlari, bilimlarni baholash, shuningdek, leksik minimumlar boʻyicha bilimni nazorat qilishga yoʻnaltirilgan dasturiy taʼminotga asoslangan oʻqitish tizimi. Bunda bixeviorizmning «stimul-reaksiya-mustahkamlash» formulasiga asoslaniladi. Oʻqitish usullari «stimul» (turtki) vazifasini bajaradi, oʻrganuvchi bundan taʼsirlanadi va bilimni mustahkamlaydi.

CALL (Computer Assisted Language Learning) – til oʻqitishni avtomatlashtirish – xorijiy tillarni kompyuter yordamida oʻqitish tizimini taʼminlovchi amaliy dasturlar hamda bilimlarni avtomatik baholash mexanizmlari.

Ekspert tizimi – muayyan sohadagi ekspertlik yoki mavjud bilimlarning immitatsiyasi tarzida yaratilgan kompyuter dasturi. Ekspert tizimi maxsus sohaga oid barcha bilimlar jamlangan qomus yoki oʻta professional ekspertlarning kompyuter modelidir.

Elektron ensiklopediya – kuchli qidiruv tizimiga ega boʻlgan maʼlum yoʻnalishdagi illyustratsiyali maqolalar, elektron xaritalar, fotosuratlar, giperhavola (gipermurojaat)lar, audio va videoizohlar, xronologik roʻyxat, foydalanilgan va tavsiya etiladigan adabiyotlar roʻyxati va hokazolardan iborat boʻlgan katta hajmdagi elektron dastur.

Gallitsizm (*lot. gallicus*) – Fransuz tilidan o‘zlashgan lekin yot element ekani sezilib turadigan so‘z yoki ibora – *merci pardon* kabi.

Defis (*nem. devis < lot. divisio – bo‘linish*) – juft va takroriy so‘zlarda, so‘z va unga qo‘shiladigan yuklamalar orasida qo‘yiladigan chiziqcha (-): *kecha-kunduz, oq-u qora, aytdim-ku* kabi.

Denotat (*lot. denotare «belgilamoq»*) – til birligi (mas., so‘z) yordamida nomlanuvchi (ataluvchi) voqelik, narsa-buyum. Mac., *yuz, aft, bashara, chehra, turq* so‘zlari bir predmet (denotat)ning turli nomlaridir.

Denotativ ma’no – denotat bilan bevosita bog‘liq, uni bevosita *англатувчи* ma’no.

Deskriptor – ma’lumotni qidirish uchun asos bo‘ladigan kalit so‘z yoki tushuncha.

Elektron lug‘atlar – odatdagi lug‘atlar asosida leksikografiya sohasiga kompyuter texnologiyalarining tatbiq etilishi natijasida yaratilgan, multimedia vositalari bilan boyitilgan dastur.

Elektron multimedia darsligi – o‘quv kursini kompyuter vositasida, o‘qituvchi yordamida yoki mustaqil o‘zlashtirishni ta’minlovchi matnli fayllar, turli xil animatsiyalar, audio va videoeffektlar bilan boyitilgan tizimga asoslanuvchi dasturiy-metodik kompleks.

Formallashtirish – bilimlar majmuyini, mazmunini muayyan shakllar yordamida sun’iy til belgilari orqali ifodalash.

Formal grammatika – muayyan shakllar yordamida sun’iy til belgilari orqali ifodalangan grammatik qoidalar majmuyi.

Frazeologik mashina tarjimas – avtomatik tarjimaning sifatini oshirish uchun Rossiyada professor G.G.Belonogov tomonidan ishlab chiqilgan konsepsiya; unga ko‘ra, matndagi tushunchalar ko‘pincha alohida so‘zlar yordamida emas, balki so‘z birikmalari yordamida ifodalanadi.

Freym (inglizcha “frame” – “karkas”, “ramka”, “skelet”) – bilimlar strukturasi o‘id asosiy tushunchalardan biri, sun’iy intellekt tizimida bilimlarni, stereotip (juda ko‘p kuzatiladigan, tipik) vaziyatlarni gavdalantirishning bir usuli. Freymga modellashtirilayotgan hodisa, vaziyat va jarayonlarni umumlashtirilgan holda joylashtirish mumkin. Freym turlari: freym-namunalar, freym-strukturalar, freym-rollar, freym-senariylar, freym-vaziyatlar.

Freym-senariy – ba’zi harakat, voqea-hodisalarning xarakterli jihatlarini o‘z ichiga olgan tipik strukturalar hamda vaziyatlarning rivojlanishi, ketma-ketligi (algoritmik bosqichlari).

Generativ va transformation grammatika – formallashtirishning negizi hisoblangan grammatika.

Gipertekst – kompyuter lingvistikasida inqilobiy o‘zgarish yasagan texnologiya, matnlarni giperhavolalar yordamida bog‘lashga imkon beradigan vosita.

Grammatika – (yun. “o‘qish va yozish san’ati” ; bu tushuncha dastlab xat-savod, imloni o‘rgatuvchi fan ma’nosida ishlatilgan) –tilning fonologik, morfologik, sintaktik va semantik belgilarini tavsiflovchi qoidalar majmuyi.

Indekslash –informatsion qidiruv tilida hujjatlarga ishlov berish protsedurasi.

Implikativ universalliya – Amerikalik olim Jozef Grinberg tomonidan implikatsiyalar (mantiqiy shartli hukmlar) asosida aniqlangan universalialar.

Ikki tillilik – ikki tilda bimalol fikrlay olish; ikki tilni bir xilda egallaganlik; zullisonaynlik.

Ikki tilli lug‘at – bir tilning lug‘aviy birliklari boshqa tilga tarjima qilib beriladigan lug‘at. Mac., ruscha-o‘zbekcha lug‘at, o‘zbekcha–ruscha lug‘at.

Informatsion qidiruv (ingl. “information retrieval”) – ma’lumotlarni qidirish, kerakli hujjatlarni ma’lumotlar bazasidan topishni optimallashtirish tizimi. Masalan, internetdagi *Google, Yandex. Rambler, Yahoo* qidiruv tizimlari.

Informatsion qidiruv tillari – IQTda qidiruvni ta’minlash uchun yaratilgan maxsus tillar. Informatsion qidiruv tillari formal til bo‘lib, u IQTda saqlangan hujjatlarning mazmunini tavsiflash va so‘rov uchun mo‘ljallangan.

Informatsion qidiruv tezauruslari (lot. *thesaurus* “xazina”) – matnning asosiy mazmun-mundarijasini aks ettiradigan kalit so‘zlar, terminlar, Asosiy tushunchalar bazasi negizida kompyuterda tuzilgan lug‘at.

Intellektual sistema bilimlar bazasi, masalani yechuvchi va intellektual interfeyslarni o‘z ichiga oluvchi tizim.

Interlingva (lot. *inter* “ opa”, *lingua* “til”) –1903 yilda Juzeppe Peano tomonidan taqdim etilgan xalqaro yordamchi til.

Interlingvistika – tilshunoslikning xalqaro aloqa-aralashuv vositalari sifatida turli yordamchi tillarni yaratish va ulardan foydalanish jarayonini o‘rganuvchi bo‘limi.

Internatsional leksika (lot. *inter* “opa”, *nationis* “xalq”) – baynalmilal leksika.

Izohli lug‘at – lug‘aviy birlik ma’nolari keng izohlanadigan, grammatik, uslubiy belgilari ham qayd etiladigan lug‘at. Mac., 2006-2008 yillarda nashr etilgan besh jildli «O‘zbek tilining izohli lug‘ati».

Ishtirokchilar– kattaroq konstruksiya tarkibiga kirgan so‘z yoki konstruksiya.

Bevosita ishtirokchilar – muayyan konstruksiyaning bevosita shakllanishida ishtirok etgan bir yoki bir nechta ishtirokchilar.

Kodlash va dekodlash - informatsiyani mashina (kompyuter) tushunadigan tilga o‘tkazish va, aksincha, kompyuter tushunadigan tildan inson tushunadigan tilga aylantirish jarayoni.

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 majmuyi.

Kompyuter leksikografiyasi (elektron lug‘atlar tuzish) – muayyan til leksikonidagi so‘zlarning ma’lurnotlar bazasi va maxsus programmalashtirish tillari yordamida ishlash tizimiga ega bo‘lgan dasturlar (kompyuter lug‘atlari) yaratish bilan shug‘ullanuvchi soha.

Kompyuter lingvistikasi – matematik lingvistikaning mantiqiy davomi bo‘lib, u amaliy tilshunoslikning eng muhim qismini tashkil etadi.

Kompyuterli modellashtirish – barcha fanlardagi, xususan, kompyuter lingvistikasidagi samarali metod.

Konstruksiya – mazmuniy qismlarning ketma-ket munosabatidan tashkil topgan butunlik .

Kontekstga bog‘liq bo‘lmagan grammatika (context-free grammar) – avvaldan empirik tarzda belgilab olingan, matndan xoli bo‘lgan grammatika.

Lemma – so‘zlarning lug‘atdagi shakli.

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.

Lingvistik protsessorlar – tabiiy tilda kompyuter bilan muloqot qilishni ta’minlashga yordam beradigan dasturlar.

Lingvistik universalিয়া – barcha tillarga yoki dunyo tillarining mutlaq ko‘pchilik qismiga xos umumiy qonuniyatlar.

Metatil – formallashtirilgan til.

Matematik lingvistika – XX asrning 50-yillarida yuzaga kelgan amaliy tilshunoslikning alohida yo‘nalishi bo‘lib, unda tabiiy tillarning matematik modellarini ishlab chiqish, xususan, sun‘iy tillarni yaratish algoritmini tuzish bilan shug‘ullanuvchi fan.

Matnlar korpusi (lot. *corpus* “tana”) – elektron holda saqlanadigan ma‘lum til birliklari.

Matnning statistik tahlili – tilni ehtimollik belgilariga ko‘ra tahlil qilish, til hodisalarini statistik xarakteristika asosida baholash.

Mashina ijodi – san‘at asarlarini avtomatik tarzda yaratish jarayoni.

Mashina tarjimas – Matn (tekst)ni berilgan dastur asosida elektron mashinalar yordamida bir tildan ikkinchi tilga avtomatik tarjima qilish.

Metagrammatika (yun. *metha* “keyin”, *grathmatike* “o‘qish va yozish san‘ati”) – metaqoidalarga asoslangan grammatika; tavsifiy grammatika, formal grammatika, transformatsion grammatika bosqichlaridan so‘ng yuzaga kelgan, qayta yozilgan grammatika; dasturlash tillarini tavsiflash uchun tabiiy til grammatikasi modeli asosida ishlab chiqilgan metaqoidalar to‘plami.

Model (lot. *modelus* “nusxa”, “andaza”, “o‘lchov”, “me‘yor”) – muayyan original-obyektning hosilasi, moddiy qurilma, grafik, sxema, umuman, bilish vositasi. Model tabiiy obyektlarning imitatsiyasidir (o‘xshashi, taqlidiy ko‘rinishi), u o‘zbek tilidagi qolip, andaza so‘zlariga mos keladi. Masalan, *globus* yerning modeli, *o‘yinchoq mashina* haqiqiy mashinaning modeli.

Modellashtirish – barcha fanlar uchun xos bo‘lgan umumilmiy metod.

Neyron tarmoqlar – inson miyasidagi neyronlar modeli asosida yaratilgan, sun‘iy intellektning muhim tarkibiy qismi hisoblanadigan tizim.

Morfologik analiz va sintez – kompyuterda kodlash-dekodlash tamoyili asosida ishlaydigan jarayon. Morfologik analizning mantiqiy davomi sintaktik analiz hisoblanadi. Sintaktik analiz morfologik analiz asosiga quriladi, bunda gap strukturasi turli formal modellar asosida tahlil etiladi.

On-line lug‘atlar – ochiq turdagi internet tizimida ishlaydigan lug‘atlar. Bu lug‘atlarning aktivlashishi uchun albatta internetga ulanish lozim. Ularning afzalligi shundaki, foydalanuvchilar ushbu lug‘atlar bazasini to‘ldirib BORISH, boyitish imkonini beradi.

Parallel matnlar korpusi – badiiy asar, qo‘llanma, ommaviy axborot vositasi nashrlari, turli xil hujjatlarning ikki yoki undan ko‘p tildagi elektron ko‘rinishlari.

Qomusiy lug‘at – turli ijtimoiy sohalarga oid muhim ma’lumot beradigan, voqelikni batafsil izohlaydigan lug‘at.

Sema (yun. *sema* “belgi”) – leksik ma’no (semema)ni tashkil toptiruvchi komponent (qism)lar. Mac, *atrof* leksemasining ma’nosi (sememasi) “hamma” va “tomon” semalaridan iborat. Ma’lum leksik-semantik guruhga oid sememalardagi semalar shu sememalar uchun umumiy yoki xususiy ekaniga ko‘ra ikki asosiy turga bo‘linadi: 1) umumiy sema – integral sema; 2) farqlovchi sema – differensial sema.

Semiotika yoki semiologiya (yun. *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. Semiotikada tabiiy tillar, sun’iy tillar, yo‘l belgilari, audio va videosignallar tadqiq etiladi.

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 munosabatlarining biri ikkinchisiga ta’sir ko‘rsatadigan nuqta) munosabati asosida modellashtiriladi.

Ssenariy – voqea-hodisalar, harakatlar ketma-ketligining kauzal bog‘liqlik tamoyili asosidagi zanjiri. Har bir ssenariyda rol bajaruvchilar mavjud. Ssenariylar real hayotda uchraydigan vaziyatlarni aks ettiruvchi holatlar asosida yuzaga keladi.

Sintez – kompyuterda matnning grammatik shakllantirilishi va generatsiyasi (hosil qilinishi). *Morfologik sintez* matnga avtomatik ishlov berish jarayonida yoki uning so‘nggi bosqichida (natijalarni inson hukmiga havola qilishda) so‘zlarning turli shakllarini chiqarish, analiz natijalarini umumlashtirishdan iborat.

Slot – freymni tashkil etuvchi qismlar. Slotlarda obyektarga xos belgilar ma’nosi qayd etilgan bo‘ladi.

Sun’iy intellekt – intellektual mashinalar, intellektual kompyuter dasturlari yaratish texnologiyasi haqidagi soha. Sun’iy intellekt insondagi intellekt tushunchasining aynan imitatsiyasi emas, balki inson intellektual faoliyatining ko‘plab qirralarini o‘z ichiga olgan, modellashtirilgan tizimlar texnologiyasi demakdir.

Tezaurus (yun. *thesauros* “xazina”) – tilga oid leksik tarkibni to‘la hajmda aks ettiruvchi lug‘at.

Termin (lot. *terminus* – “chek”, “chegara”, “chegara belgisi”) – fan, texnika va boshqa sohaga oid tushunchani aniq ifodalaydigan, ishlatilish doirasi shu sohalar bilan chegaralangan soʻz yoki soʻz birikmasi. Terminlar bir maʼnoli boʻlishi, ekspressivlik va emotsionallikka ega boʻlmasligi bilan ham umumisteʼmoldagi soʻzlardan farqlanadi. Mac., lingvistik terminlar: *boʻgʻin, urgʻu, ega, kesim...*; geometriyaga oid terminlar: *yuza, toʻrtburchak...*; fizikaga oid terminlar: *harakat, tezlanish, maydon*; kimyoga oid terminlar: *kumush, ishqor, tuz, rux* va b.

Tobelik daraxti (shajara) – L.Tenyer gʻoyalari taʼsirida shakllangan tobelik grammatikasi modeli boʻlib, u kompyuter lingvistikasining avtomatik analiz das-turlarida ishlatiladi. Bunda gap asosan daraxt koʻrinishida tasvirlanadi, jumla tarkibidagi soʻzlar bir-biri bilan tobelik munosabatida boʻladi (*bevosita dominatsiya munosabati*).

Translyator dasturlar (yoki kompilatorlar) –algoritmik til bilan inson yoki avtomatik qurilma (kompyuter) oʻrtasida muloqotni taʼminlashda vositachi vazifasini bajaradigan dasturlar.

Translation Memory (“Tarjimon xotirasi”) yoki **Sentence Memory** (“Gap toʻplagich”) – aytomatik tarjima samaradorligini oshirish maqsadida amerikalik olim Veb Layn tomonidan ishlab chiqilgan konsepsiya.

Tyuring testi – ingliz olimi Alan Tyuring tomonidan ishlab chiqilgan, insonni xatoga yoʻl qoʻyishga, notoʻgʻri tanlov qabul qilishga daʼvat qiladigan, chalgʻituvchi empirik test.

Zif qonuni — matnning statistik tahlilida keng foydalaniladigan, katta hajmdagi matnlarda har bir soʻzning takrorlanish miqdorini (chastotasini) hisoblashda qoʻllanadigan metod.

Chastotali lugʻatlar – matnda soʻzlarning qoʻllanish miqdori (chastotasi)dan kelib chiqqan holda tayyorlanadigan soʻzliklar. Chastotali lugʻatlar kompyuterning lingvistik informatsion bazasi hisoblanadi.

Foydalanilgan adabiyotlar ro‘yxati

Ilmiy-metodologik manbalar

1. Mirziyoyev Sh.M. Erkin va farovon demokratik O‘zbekiston davlatini birgalikda barpo etamiz. –T.: O‘zbekiston, 2016. –B.14.
2. Mirziyoyev Sh.M. Tanqidiy tahlil, qat’iy tartib-intizom va shaxsiy javobgarlik –har bir rahbar faoliyatining kundalik qoidasi bo‘lishi kerak. - Toshkent: “O‘zbekiston” NMIU, 2017.
3. Mirziyoyev Sh.M. “O‘zbekiston Respublikasini yanada rivojlantirish bo‘yicha harakatlar strategiyasi to‘g‘risida”gi PF-4947-sonli farmon. 07.02.2017. “Xalq so‘zi” gazetasi, 2017-yil, 28-(6722) soni.
4. O‘zbekiston Respublikasi Prezidentining “Oliy ta’lim tizimini yanada rivojlantirish chora-tadbirlari” to‘g‘risidagi qarori // “Xalq so‘zi” gazetasi, 2017-yil, 21-aprel.
5. Abdullayev Y. Statistika nazariyasi. – Toshkent, 2002.
6. Абдурахмонова Н. Машина таржимаси. –Тошкент, 2008.
7. Абдурахмонова Н. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти. Фило. фан... (PhD) дисс. – Тошкент, 2018.
8. Abduhamidova L. Tilshunoslikning yangi yo‘nalishi: kompyuter lingvistikasi. –Toshkent, 2015.
9. Abduhamidova L. Komyuter lingvistikasining avtomatik tahrir yo‘nalishi. ВМІ. –Toshkent, 2016. 67 b.
10. Абжалова М.А. Ўзбек тилидаги матнларни компьютер воситасида таҳрир қилиш дастурининг лингвистик таъмини. Магистр илмий даражасини олиш учун ёзилган диссертация. –Тошкент, 2011.
11. Абжалова М. Ўзбек тилидаги матнларни таҳрир ва таҳлил қилувчи дастурнинг лингвистик модуллари (расмий ва илмий услубдаги матнлар таҳрири дастури учун). Филол. фан...(PhD) дисс. – Тошкент, 2019.
12. Аврорин В.А. Проблемы изучения функциональной стороны языка (к вопросу о предмете социолингвистики). –Л., 1975. –С. 177.
13. Адилова С. Ўзбек тилини ўқитишда компьютер технологияларидан фойдаланиш. –Тошкент, 2006.
14. Алишер Навоийнинг “Ҳайрат ул-аброр” достони конкорданси. –Тошкент, ТДШИ, 2012.
15. Амосов Н.М. Моделирование мышления и психики. –Киев, 1965.
16. Андреев А.А. Дистанционное обучение в системе непрерывного профессионального образования. Автореферат. диссер.

на соис. уч. ст. доктора педагогических наук.
<http://www.iet.mesi.ru/dis/oglo.htm>

17. Анисимов А.В. Компьютерная лингвистика: Мифы. Алгоритмы. Язык. – Киев: Наукова-думка, 1991.

18. Ахманова О.С. Словарь лингвистических терминов. – М.: СЭ, 1969.

19. Баранов А.Н. Категории искусственного интеллекта в лингвистической семантике. Фреймы и сценарии. – М., 1987.

20. Баранов А.Н. Введение в прикладную лингвистику. – М.: Эдиториал УРСС, 2001

21. Бегимқулов У.Ш. Малака ошириш тизимида замонавий ахборот технологияларидан фойдаланиш. //“Халқ таълими” жур. – № 6, 2004. 132 –137 бетлар.

22. Бегимқулов У.Ш. Олий таълим муассасаларининг ягона ахборот маконини ташкил этиш ва уни ривожлантириш истиқболлари //“Халқ таълими” жур. № 4, 2006. 4–7 бетлар.

23. Базарова Б.Б. Введение в корпусную лингвистику. -Улан-Удэ, 2016.-Б.19.

24. Belonogov G., Khoroshilov A., Kuznetsov B., Novoselov A., Zelenkov Yu. Systems of Phraseological Machine Translation of Polythematic Texts from Russian into English and from English into Russian (RETRANS and ERTRANS Systems). International Forum on Information and Documentation. Vol. 20. №2. 1995.

25. Белоногов Г.Г., Зеленков Ю.Г., Новоселов А.П., Хорошилов А.А., Метод аналогии в компьютерной лингвистике // Научно-техническая информация. Сер.2. № 1. ВИНТИ. 2000.

26. Белоногов Г.Г. Компьютерная лингвистика и перспективные информационные технологии. – М.: Русский мир, 2004.

27. Бим И.Л. Методика обучения иностранным языкам как наука и проблемы школьного учебника: опыт системно-структурного описания. –М.: Русский язык, 1977. –С.21.

28. Боярский К.К. Введение в компьютерную лингвистику. Учебное пособие.-Спб.:НИУ ИТМО, 2013.

29. Валиева Г. Расмий-идоравий услубнинг лисоний бирликларини моделлаштириш. Магистрлик дисс.–Тошкент, 2003, 60 б.

30. Very A Brief. Introduction to Computational Linguistics. Weiwei Su. Institute of Computer Science and Technology Peking University. 2017.

31. Гальскова Н.Д. Еще раз о лингводидактике / Н.Д. Гальскова // Иностранные языки в школе, 2008.
32. Гладкий А.В., Мельчук И.А. Элементы математической лингвистики. –М.: Наука, 1969.
33. Городецкий Б.Ю. Компьютерная лингвистика: моделирование языкового общения. – Новое в зарубежной лингвистике. Вып. XXIV, Компьютерная лингвистика. –М., 1989
34. Григорьев С.Г., Гриншкун В.В., Демкин В.П., Краснова Г.А., Макаров С.И., Роберт И.В. Концепция образовательных изданий и ресурсов – шаг на пути к качественному образованию // ТДПУ Ахборот технологиялари ва масофавий ўқитиш маркази ахборотномаси. – 2002. –№ 9. –Б. 77-82.
35. Jurafsky D., Martin J.H. Speech and Language Processing: An introduction to natural language processing, computational linguistics and speech recognition (2007).
36. Daniel Jurafskiy & James H. Martin. "Speech and Language Processing (2008)
37. D. Yu and L. Deng "Automatic Speech Recognition: A Deep Learning Approach" (Publisher: Springer), 2014.
38. Дўсимова У. Матндаги феълларни автоматик таҳрир қилувчи дастурнинг лингвистик таъмини (расмий-идоравий услубдаги матнлар асосида). Магистрлик дисс. –Тошкент, 2002, 56 б.
39. Йўлдошев Б. Компьютер лингвистикаси. – Самарканд, 2009.
40. Кислов А.В., Колпачкова У.Н. Влияние интернета на современный китайский язык / Компьютерная лингвистика и вычислительные онтологии. Выпуск 1. 2017. – С.72 –86.
41. Кобозева И.М., Лауфер Н.И., Сабурова И.Г. Моделирование общения в человеко-машинных системах. Лингвистическое обеспечение информационных систем. –М., 1987.
42. Копотев М.В., Мустайоки А. Современная корпусная русистика// Slavica Helsingiensia 34. Инструментарий русистики: корпусные подходы. Хельсинки, 2008. –С.10-13.
43. Кревский И.Г., Селиверстов М.Н., Григорьева К.В. Формальные языки, грамматики и основы построения трансляторов. – Пенза, 2003.
44. Курбонова Ф. Компьютер луғатлари: тезаурус. –Тошкент, 2014.
45. Лагута О.Н. Логика и лингвистика. – Новосибирск, 2000.
46. Лингвистический энциклопедический словарь. –М.: СЭ, 1990.

47. Лукашевич В.Н. Тезаурусы в задачах информационного поиска. – М.: Изд-во Моск. ун.-та, 2011. 61-97.
48. Li Deng and Doug O'Shaughnessey "Speech Processing: A Dynamic and Optimization-Oriented Approach" published in 2003.
49. L. Deng and D. Yu. "Deep Learning: Methods and Applications, 2014.
50. Manfred R. Computer Speech Schroeder, second edition in 2004.
51. Mairesse F. (2011). Controlling user perceptions of linguistic style: Trainable generation of personality traits. Computational linguistics.
52. Mengliyev B., Bobojonov S., Hamroyeva Sh. O'zbek tilining milliy korpusi. – Toshkent: Ma'rifat / 2018.
53. Mikheev A. (2003). Text segmentation. In Mitkov, R. (Ed). Oxford Handbook of Computational Linguistics, Oxford University Press, Oxford.
54. Мирюлюбов А.А. История отечественной методики обучения иностранным языкам. - М: СТУПЕНИ, ИНФРА. – М, 2002.
55. Mohri M. (2000). Minimalization algorithms for sequential transducers/ Theoretical Computer Science, 234. 177–201.
56. Макьюин К. Дискурсивные стратегии для синтеза текста на естественном языке. – Новое в зарубежной лингвистике. Вып. XXIV. Компьютерная лингвистика. – М., 1989
57. Марчук Ю.Н. Проблемы машинного перевода. – М.: Наука, 1983.
58. Марчук Ю.Н. Методы моделирования перевода. – М.: Наука, 1985.
59. Марчук Ю.Н. Компьютерная лингвистика. – М.: АТС: Восток и Запад, 2007.
60. Мухамедов С. Статистический анализ лексико-морфологической структуры узбекских газетных текстов. Автореф. дисс... канд. филол. наук. – Ташкент, 1980.
61. Мухамедов С.А., Пиотровский Г.Г. Инженерная лингвистика и опыт системно-статистического исследования узбекских текстов. – Ташкент: Фан, 1986.
62. Мухамедова С. Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. – Тошкент, 2006.
63. Muhamedova S. Kompyuter lingvistikasi (metodik qo'llanma). – Toshkent: ToshDPU nashri, 2007. – 74 b.
64. Нелюбин Л.Л. Компьютерная лингвистика и машинный перевод. – М.: ВЦП, 1991.
65. Новое в зарубежной лингвистике. Вып. XXIV. Компьютерная

лингвистика. –М.: Прогресс, 1989.

66.Нурмонов А. Лисоний белги хусусиятлари ҳақида. – Андижон, 1992.

67.Нурмонов А., Расулов Р. Ўзбек тили жадвалларда. –Тошкент: Ўқитувчи, 1992.

68.Нурмонов А., Йўлдошев Б. Тилшунослик ва табиий фанлар. – Тошкент: Ўқитувчи, 2001.

69.Nurmonov A. Lingvistik tadqiqot metodologiyasi va metodlari. – Toshkent: Akademnashr, 2010.

70.Nurmonov A. Struktur tilshunoslik: ildizlari va yo‘nalishlari. – Toshkent, 2008.

71.Nurmonov A., Iskandarova Sh. Tilshunoslik nazariyasi. –Toshkent: Fan, 2008. –B. 140 -141.

72.Пенроуз Р. Новый ум короля. О компьютерах, мышлении и законах физики. I- М.: УРСС, 2005.

73.Пиотровский Р.Г. Инженерная лингвистика и теория языка. – Л.,1979.

74.Пулатов А.К., Джураева Н. Разработка формальной модели грамматики узбекского языка // Узбекский математический журнал. 2002, №1. –С.47-54.

75.Пулатов А.К., Алиходжаев Б., Джураева Н. Разработка программы компьютерного анализа и синтеза глаголов узбекского языка // ЎЗМУ хабарлари. 2002, 2-сон. –Б.17-19. 90 935 69 41

76.Пўлатов А.К., Мухамедова С.Х. Компьютер тилшунослигида автоматик таҳрир қилувчи дастурнинг лингвистик таъминотини яратиш асослари // Ўзбек тилшунослигининг долзарб масалалари. – Тошкент, ТДПУ илмий тўплами, 2003. –Б.32-36.

77.Пўлатов А., Мўминова Т., Пўлатова И. Дунёвий ўзбек тили. – Тошкент: Университет, 2004.

78.По‘латов А., Muhamedova S. Kompyuter lingvistikasi (o‘quv qo‘llanma).-Toshkent, 2007, 2008, 2009.

79.Пўлатов А. Компьютер лингвистикаси. –Тошкент: Akademnashr, 2011.

80.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.

81.Rahimov A. Kompyuter lingvistikasi asoslari. –Toshkent: Akademnashr, 2011.

82.Rahmatullayev Sh. O‘zbek tilining etimologik lug‘ati. –Toshkent, 2000.

83.Ризаев С. Ҳозирги ўзбек адабий тилининг грамматик ва фонологик системасини статистик текшириш. Филол.фан.номз...дисс. –Тошкент, 1970, 236 б.

84.Ризаев С. Кибернетика ва тилшунослик. – Тошкент:Ўзбекистон, 1976, 16 б.

85.Ризаев С., Бўронов Н.Б. А.Қаҳҳор «Синчалак» повести тилининг частотали луғати. –Тошкент: Ўқитувчи, 1986.

86.Rizayev S. O‘zbek tilshunosligida lingvostatistika muammolari. – Toshkent: Fan, 2006.

87.Richter B. First steps in theoretical and applied linguistics. – Budapest, 2006.

88.Сиокум Дж. Обзор разработок по машинному переводу: история вопроса, современное состояние и перспективы развития // Новое в зарубежной лингвистике. Вып. XXIV. Компьютерная лингвистика. – М.: Прогресс, 1989.

89.Субботин М.М. Гипертекст. Новая форма письменной коммуникации. – ВИНТИ, Сер. Информатика, 1994, т. 18

90.The Voice in the Machine. Building Computers That Understand Speech" by Roberto Pieraccini (2012).

91.VI International Conference on Computer Processing of Turkic Languages. Tashkent, 2018.www.turklang.uz, www.navoiy-uni.uz, www.complinglab.uz.

92.Убин И.И. Современные средства автоматизации перевода: надежды, разочарования и реальность // Перевод в современном мире. –М.: ВЦП, 2001.

93.Умаров Э. Алишер Навоийнинг “Хамса” дostonлари конкорданслари. –Тошкент, 2011.

94.Халеева И.И. Основы теории обучения пониманию иноязычной речи (подготовка переводчиков): Моногр.– М.: Высш. школа, 1989.

95.Хаматова А.А. Тенденции развития лексики китайского языка в начале XXIII века. Иркутск: Вестник Иркутского государственного лингвистического университета, 2012. – С.9-12

96.Хамидов В.С. «Акт ва таълим сифатини таъминлаш» модули бўйича ўқув –услубий мажмуа. –Тошкент,2015.

97.Хоназаров Қ. Глобаллашув ва тил фалсафаси. – Т., 2009. – Б.29–30.

98. Sharopova S. Til o'qitish dasturlarining amaliy ahamiyati. Metodist. Ilmiy maqolalar to'plami. –Toshkent, 2017. –B. 8–12.

99. Шаумян С.К. Теоретические основы трансформационной грамматики // Новое в лингвистике. Вып. I. – С.408.

100. Шафрин Ю. Основы компьютерной технологии. Справочник школьника. –Бишкек. 2000.

101. Швейцер А.Д. Современная социолингвистика: Теория. Проблемы. Методы. –М.: Наука, 1976. –С. 148.

102. Швейцер А.Д., Никольский Л.Б. Введение в социолингвистику. –М.: Высшая школа, 1978. –С. 117.

103. Шемакин Ю.И. Начало компьютерной лингвистики. –М.: Высшая школа, 1992.

104. Электрон университет. Масофавий таълим технологиялари. Олий таълим муассасалари учун/ А.Парпиев, А.Марахимов, Р.Ҳамдамов, У.Бегимкулов, М.Бекмурадов, Н.Тайлоқов. ЎзМЭ давлат илмий нашриёти.-Тошкент: 2008, 196 б.

105. Эшмуминов А. Ўзбек тили миллий корпусининг синоним сўзлар базаси. –Қарши, 2019.

106. Ўринбоев Д.Б. Компьютер лингвистикаси (Маърузалар матни). Самарқанд, 2009.

107. Ҳамроева Ш. Ўзбек тили муаллифлик корпуси тузишнинг лингвистик асослари. Филол. фан...(PhD) дисс. –Қарши, 2018.

Internet manbalari

The Global WordNet Association <http://global.wordnet.org/wordnets-in-the-world>.

http://encyclopedia.bigaru.ru/enc/liberal_arts/kompyuternaya_lingvistika.html

http://pmlectures.ru/article/Novyyi_tsikl_lektsiyi_Kompyuternaya_lingvistika-5

<http://il.rsuh.ru/section.html?id=9800#selegey>

<http://cl.nuu.uz>

<http://www.philol.msu.fu/~lex/corpus>

<http://www.ipcs.math.msu.ru/rus/mlogic.htm>.

www. Wikipedia.

2012 total-web. NET – Ахборот электрон журнал.

<http://arbuckle.sfs.uni-tuebingen.de/iscl/Russian/CL-ru.shtml>

<http://m.ielts-test.ru/ielts-ili-toefl.html>

<http://www.wikipedia.org/wiki/model.html>

<http://uz.m.wikipedia.org/wiki/IELTS>

<http://forum.ziyouz.com/index.php.topic=3899.0>

<http://en.wikipedia.org>. Computitional linguistics.
Bond.uz.

MUNDARIJA

Kirish.....	3
I. KOMPYUTER LINGVISTIKASINING RIVOJLANISH TENDENSIYASI	
1-§. “Kompyuter lingvistikasi” fanining maqsadi, vazifalari.....	9
2-§. Kompyuter lingvistikasining fan sifatida rivojlanishida nazariy va amaliy tadqiqotlarning o‘rni.....	18
Kompyuter lingvistikasi bo‘yicha xorijda olib borilgan tadqiqotlar Turkologiyada va o‘zbek tilshunosligidagi kompyuter lingvistikasiga doir tadqiqotlar.....	18
3-§. “Kompyuter lingvistikasi”ning matematik mantiq va boshqa sohalar bilan aloqadorligi.....	31
4-§.Mulohazalarni hisoblashning formal nazariyasi.....	40
47	
II.KOMPYUTER LINGVISTIKASIDA TILSHUNOSLIK MASALALARI	
5-§. Lingvistik bilimlar bazasi.....	53
6-§. Lingvistik protsessorning grammatik xususiyatatlari.....	68
7-§.Morfologik tahlilning avtomatik tizimi.....	72
8-§. Sintaktik tahlilning avtomatik tizimi.....	79
III. KOMPYUTER LINGVISTIKASINING MA’LUMOTLAR BAZASI	
9-§. Korpus lingvistikasi.....	87
10-§. Kompyuter leksikografiyasi.....	93
Chastotali lug‘atlar.....	97
Konkordans lug‘atlar.....	100
Tezaurus lug‘atlar.....	102
Tezaurus lug‘atlar tarkibi.....	113
Tezaurus lugatlarga oid asosiy tushunchalar tavsifi.....	119
Tezaurus atamalari manbasini aks ettirish.....	126
IV . KOMPYUTER LINGVISTIKASI YO‘NALISHLARI VA METODLARI	
11-§. Kompyuter lingvistikasi yo‘nalishlari tavsifi.....	132
12-§. Mashina tarjimasini.....	136
Mashina tarjimasining lingvistik muammolari.....	148
13-§. Kompyuterda matnlarni qayta ishlash tizimlari.....	159
Matlarni avtomatik tahrirlash.....	165
Matn mazmunini o‘zlashtirishda <i>BORIS</i> tizimining o‘rni.....	175
14-§. Kompyuter lingvodidaktikasi.....	184
15-§. Til o‘qitish tizimini avtomatlashtirish.....	190

16-§. O‘zbek tilining axborot uslubi.....	201
17-§. Kompyuter lingvistikasi metodlari.....	205
18-§. Kompyuter lingvistikasining ahamiyati.....	209
“Kompyuter lingvistikasi” fanidan test savollari.....	214
“Kompyuter lingvistikasi”ga oid asosiy tushunchalar izohi (glossariy).....	225
Foydalanilgan adabiyotlar.....	283

СОДЕРЖАНИЕ

Введение.....	3
I. ТЕНДЕНЦИЯ РАЗВИТИЯ КОМПЬЮТЕРНОЙ ЛИНГВИСТИКИ	
1-§. Цель и задачи дисциплины “Компьютерная лингвистика”..	9
2-§. Роль теоритических и прикладных исследованиях в развитии «Компьютерной лингвистики» как наука.....	18
Исследования в зарубеьем по компьютерной лингвистике....	18
Исследования в тюркологии и узбекском языкознании по компьютерной лингвистике.....	31
3-§. Связь “Компьютерной лингвистики” с наукой математики и другими дисциплинами.....	40
4-§. Формальная теория вычисления обсуждений.....	47
II. ЛИНГВИСТИЧЕСКИЕ ВОПРОСЫ В КОМПЬЮТЕРНОЙ ЛИНГВИСТИКИ	
5-§. База лингвистических знаний.....	53
6-§. Грамматические особенности лингвистического процессора.....	68
7-§. Автоматическая система морфологического анализа.....	72
8-§. Автоматическая система синтактического анализа.....	79
III. БАЗА СВЕДЕНИЙ КОМПЬЮТЕРНОЙ ЛИНГВИСТИКЕ	
9-§. Корпусная лингвистика.....	87
10-§. Компьютерная лексикография.....	93
Частотные словари.....	97
Словари – конкордансы.....	100
Словари – тезаурусы.....	102
Построение тезаурусных словарей.....	113
Описание основных понятий по словарям-тезаурусам.....	119
Изображение источников тезаурусных понятий.....	126
IV. ОСНОВНЫЕ НАПРАВЛЕНИЯ “КОМПЬЮТЕРНОЙ ЛИНГВИСТИКИ И МЕТОДЫ	
11-§. Описание направлений компьютерной лингвистики.....	132
12-§. Машинный перевод.....	136
Лингвистические проблемы машинного перевода.....	148
13-§. Система переработки текстов в компьютере.....	159
Автоматическое редактирование текстов.....	165
Роль система BORIS в освоении содержаний текста.....	175
14-§. Компьютерная лингводидактика.....	184

15-§. Автоматизация систему языкового обучения.....	190
16-§. Информационный стил узбекского языка.....	201
17-§. Методы компьютерной лингвистики.....	205
18-§. Значение компьютерной лингвистики.....	209
Тестовые вопросы по “Компьютерной лингвистике”.....	214
Значение основных понятий по “Компьютерной лингвистике” (глоссарий).....	225
Использованная литература.....	283

CONTENT

The foreword	3
I. THE DEVELOPING TENDENCIES OF THE COMPUTER LINGUISTICS. INTRODUCE TO “COMPUTER LINGUISTICS”	
1-§. The role of “Computer linguistics” as a forming subject at thereotical and practical researches. Studies on computer linguistics in foreign countries.....	9
2-§. The role of theoritical and applied research in the development of computer linguistics as a science.....	18
Overseas research in computer linguistics.....	18
Research on computer linguistics in turkology and Uzbek linguistics.....	31
3-§. The relationship among “Computer linguistics” with mathematic logics and other subjects.....	40
4-§. Formal theory of feedback calculation.....	47
II. THE LINGUISTIC TASKS AT COMPUTER LINGUISTICS	
5-§. The base of the linguistic knowledge.....	53
6-§. The grammar of the linguistic processor.....	68
7-§. The computer approach of thegrammatical analyzes. The automatic system of the morphological analyze.....	72
8-§. The automatic system of the syntax analyzes.....	79
III. БАЗА СВЕДЕНИЙ КОМПЬЮТЕРНОЙ ЛИНГВИСТИКЕ	
9-§. Corpus linguistics.....	87
10-§. The computer lexicography.....	93
Frequency dictionaries.....	97
Concordance dictionaries.....	100
Thesaurus dictionaries.....	102
Content of thesaurus dictionaries.....	113
Description of the basic concepts of thesaurus dictionaries.....	119
Reflecting the source of thesaurus terms.....	126
IV. THE COMPUTER LINGUISTIC TENDENCIES AND METHODS	
11-§. Description of computer linguistics trends.....	132
12-§. The machine translate.....	136
The problem of machine translate.....	148
13-§. Revise texts at computer.....	159

The automatic edit of texts.....	165
The role of the BORIS system in mastering the content of the text	175
14-§. The computer lingvodidactics.....	184
15-§. Automations of the process teaching languages.....	190
16-§.The method of information uzbek language.....	201
17-§. The methods of computer linguistic.....	205
18-§. Importance of computer linguistics.....	209
The test answers from subject computer linguistics.....	214
Glossary.....	225
Literatures.....	283

Xolmanova Zulxumor Turdiyevna

KOMPYUTER LINGVISTIKASI

o‘quv qo‘llanma

Adadi 100 nusxa. Hajmi 15,4 b/t. Bichimi 60x84 ¹/₁₆
«Times New Roman» garniturasida. Ofset usulida bosildi.
OOO “AKTIV PRINT” bosmaxonasida chop etildi.
Toshkent, Chilonzor 25, Lutfiy 1A.