

O‘ZBEKISTON RESPUBLIKASI
OLIY VA O‘RTA MAXSUS TA‘LIM VAZIRLIGI
TERMIZ DAVLAT UNIVERSITETI
O‘ZBEK FILOLOGIYASI FAKULTETI
O‘ZBEK TILSHUNOSLIGI KAFEDRASI

**5120100- Filologiya va tillarni o‘qitish: o‘zbek tili ta‘lim
yo‘nalishi**

ALIMARDONOVA SEVARANING

**“ KOMPYUTER LINGVISTIKASIDA MODELLASHTIRISH VA
ANALOGIYA METODLARI ”**

mavzusida yozilgan

BITIRUV MALAKAVIY ISHI

Ilmiy rahbari: Kat. O‘q. O‘. Xoliyorov

Termiz –2016

KIRISH	3
BIRINCHI BOB	9
KOMPYUTER LINGVISTIKASI AMALIY TILSHUNOSLIKNING ISTIQBOLI SIFATIDA	9
1.1. Amaliy tilshunoslik rivoji – ijtimoiy ehtiyoj.....	9
1.2. Kompyuter lingvistikasi yo‘nalishlari.....	14
IKKINCHI BOB	20
KOMPYUTER LINGVISTIKASINING TILNI TADQIQ ETISH METODLARI VA LINGVISTIK MODELLASHTIRISH	20
2.1. Formal grammatika.....	20
2.2. Matematik statistika va kvantitativ metod.....	31
2.3. Tilni modellashtirish – tilni kompyuterda tekshirishning optimal usuli.	36
2.4. Tilni analogiya metodi asosida formallashtirish.....	43
XULOSA	46
ADABIYOTLAR RO‘YXATI	50

KIRISH

Mavzuning dolzarbligi. Ijtimoiy-iqtisodiy sharoit fan oldiga qo‘ygan vazifalar hozirgidagidek ko‘p bo‘lgan davr insoniyat tarixida kuzatilmagan. Bugun gumanitar fanlar oldiga ham aniq vazifalar qo‘yilmoqda va hatto bu vazifalarni hal etishning aniq muddatlari belgilanmoqda. Bu talabni bajarmaslik muayyan fanning o‘z obyekti va predmetidan mahrum bo‘lib, boshqa fanlar tarkibiy qismiga aylanib qolishiga olib kelishi mumkin.

Ma‘lumki, XXI asr kompyuter texnologiyalari, ilg‘or pedagogika va axborot texnologiyalari, yuksak tafakkur, fan va texnikaning jadal taraqqiyot asri bo‘ladi. XXI asr boshiga kelib, “Elektron ta‘lim”, “Elektron boshqaruv”, “Ochiq ta‘lim”, “Masofaviy ta‘lim”, “Axborotlashgan ta‘lim” kabi tushunchalar hayotimizga kundan-kun singib bormoqda. Prezidentimiz I.A.Karimov so‘zlari bilan aytganda, “hozirgi axborot, kommunikatsiya, kompyuter texnologiyalari asrida, internet kundan-kunga hayotimizning barcha jabhalariga tobora chuqur va keng kirib borayotgan bir paytda, odamlarning ongi va tafakkuri uchun kurash hal qiluvchi ahamiyat kasb etayotgan bir vaziyatda bu masalalarning jamiyatimiz uchun naqadar dolzarb va ustuvor bo‘lib borayotgani haqida gapirib o‘tirishga hojat yo‘q”¹.

Hozirgi paytda injenerlik lingvistikasi, hisoblash lingvistikasi, matematik lingvistika va “elektron - tarjimon”, kompyuter lingvistikasi yuzasidan bir qator yangi muammolar o‘rtaga tashlanmoqda. Darhaqiqat, zamonaviy kompyuter juda ulkan imkoniyatlarga egaki, uning uchun shaxmat o‘ynashdan tortib, tibbiy tashxis qo‘yish, korxonalarni boshqarish, qadimgi xalqlarning sirli xatlarini “o‘qib”, mazmunini so‘zlab berish, she‘r yozish, kuylar ijro etish, katta hajmdagi kitoblarga annotatsiyalar yozish, loyihalar tarxini chizish kabi yuzlab

¹ Karimov I.A. Bizning bosh maqsadimiz – jamiyatni demokratlashtirish va yangilash, mamlakatni modernizatsiya va isloh etishdir. O‘zbekiston demokratik taraqqiyotning yangi bosqichida. – Toshkent: O‘zbekiston, 2005. –B. 38-39.

mehnat faoliyati turlarining barchasida, albatta, zamonaviy kompyuter texnologiyalari ham ishchi kuchi, ham vaqt jihatidan juda yaqin yordamchi sanaladi.

Bunlay ilmiy tadqiqotlar sirasiga murakkabligi, dolzarbligi bilan xarakterlanuvchi kompyuter lingvistikasi fani ham kiradi. Kompyuter lingvistikasi oldida turuvchi eng muhim masalalar quyidagilar:

- tilning aksiomatik nazariyasini ishlab chiqish;
- formal grammatika yaratish;
- tillarning matematik modellarini ishlab chiqish.¹

Kompyuter lingvistikasi kelgusida til qoidalarining matematik modellarini yaratish, o'zbek tilining formal nazariyalarini ishlab chiqish, shu asosda berilgan materiallarni kompyuter yordamida tahlil qilish, o'qitish, bilimlarni baholash, xatolarni tuzatish algoritmlarni yaratish uchun xizmat qiladi.

Kompyuter lingvistikasi kelajak fani bo'lib, uning yutuqlari mamlakatimizda amaliy va nazariy tilshunoslikning uzviyligini ta'minlash, nazariy tadqiqotlar natijalarini jamiyat manfaatlari yo'lida qo'llash uchun katta imkoniyatlarni ochib beradi. Darhaqiqat, respublikamiz prezidenti I.A.Karimov ta'kidlaganlaridek, "fanning vazifasi kelajagimizning shakl-shamoyilini yaratib berish, ertangi kunimizning yo'nalishlarini, tabiiy qonuniyatlarini, uning qanlay bo'lishini ko'rsatib berishdan iborat. Odamlarga mustaqillikning afzalligini, mustaqil bo'lmagan millatning kelajagi yo'qligini, bu tabiiy bir qonuniyat ekanligini isbotlab, tushuntirib berish kerak. Fan jamiyat taraqqiyotini olg'a siljituvchi kuch, vosita bo'lmog'i lozim"².

Shuning uchun biz ham kompyuter lingvistikasida muhim o'rin tutgan modellashtirish masalasini o'rganish va tahlil qilishni maqsad qilib oldik.

¹ Raximov A. Kompyuter lingvistikasi asoslari. –Toshkent: Akademnashr, 2011. –B. 5-6.

² Karimov I.A. Adolatli jamiyat sari. – Toshkent: O'zbekiston, 1998. – B. 58.

Adabiyotlar tahlili. Kompyuter lingvistikasi haqidagi g'oyalarning paydo bo'lganiga yetmish yildan oshdi. XX asrning 50-yillarida til materialini o'rganishda matematik usullardan foydalanishga qiziqish kuchaydi. 1951-yilda matematik lingvistikasi bo'yicha dastlabki tadqiqotlar AQShda va sobiq Ittifoqda vujudga kelgan edi.

Kompyuter lingvistikasi hozirgi vaqtda nazariy va amaliy fan sifatida to'la shakllanib bormoqda. Uning nazariy fan ekanligi nutqiy jarayonni, matnni o'rganishning turli xil model (gipoteza)larini yaratishning nazariy asoslarini ishlab chiqishida ko'rinadi. Uning amaliy xususiyati esa tarjima mashinalarini yaratish va uning ishlash jarayonini tashkil etishi bilan belgilanadi. Buning uchun matematik lingvistika fani quyidagi uch usul va metodlardan keng foydalanadi:

- a) mantiqiy-matematik usullar: masalan, inkor amali, konyunksiya, dizyunksiya amallari kabilar;
- b) nazariy-informatsion usullar;
- d) ehtimollik-statistik usullari.

Matematik lingvistikaning ana shunday nazariy va amaliy asoslarini belgilashda Daniya olimi L.Elmslev, amerikalik olimlar N.Xomskiy, K.Shenon kabilarning xizmati kattadir. Matematik lingvistikaning amaliy muammolari XX asrning o'rtalarida vujudga keldi. Bu narsa birinchi navbatda mashina tarjimasi bilan aloqadordir. 1954-yilda AQShda IBM – 701 rusumli tarjima qurilmasi yaratildi va bu qurilma 250 ta so'zdan iborat 60 ga yaqin ruscha gaplarni ingliz tiliga tarjima qilish jarayonini amalga oshirdi. Bu ishga P.Gavren, L.Dostert, P.Sheridanlar ishtirok etdilar.

1955- yilga kelib sobiq Ittifoqda BESM rusumli mashina qurilmasi yordamida inglizcha gaplarni rus tiliga tarjima qilish jarayoni amalga oshirildi (I.K.Belskaya, L.N.Korolev, S.N.Razumovskiy kabilar). Shundan keyin jahonning Angliya, Fransiya, Avstriya kabi ko'pgina mamlakatlarida olimlar mashina tarjimasi muammolari bilan maxsus shug'ullana boshladilar. Natijada tilshunos

olimlar ma'lum bir tilni tipologik jihatdan tasniflash, uning materiallarini analisis va sintez qilish, mashina tiliga o'tkazish bilan shug'ullandilar va muayan natijalarga erishdilar.

Bunday muammolarni lingvistik tadqiqotlarni avtomatlashtirish orqali hal etish mumkin. Bu masalaga doir tilshunoslarning maxsus ilmiy anjumanlari o'tkazilgan. Masalan, 1962-yilda Kembridj shahrida jahon tilshunoslarining IX-kongressi o'tkazildi. Bu anjumanda struktural tilshunoslik va matematik lingvistika muammolarini muhokama qilishga alohida o'rin ajratilgan edi.

1967-yilda Buxarest shahrida jahon tilshunoslarining X kongressi o'tkazilib, bunda ham matematik lingvistika va avtomatik tarjima muammolari maxsus muhokama qilindi. 1960 yilda Parij shahrida xalqaro kongress o'tkazilib, bunda maxsus algoritmik til – ALGOL joriy etilgan edi. Bu algoritmik til hozirgi kunga qadar EHMga matnlarni kiritish, ularni dasturlash uchun xizmat qilib kelmoqda.

1996-yil 20-21-mayda Samarqand shahrida sohibqiron Amir Temurning 660 yilligiga bag'ishlangan "Injenerlik tilshunosligi va til o'qitish jarayonini kompyuterlashtirish" muammolariga doir 2-xalqaro ilmiy anjuman o'tkazildi. Bu anjumanni akad.R.G.Puotrovskiy "Injenerlik tilshunosligi va mustaqil davlatlarda informatsion industriyaning tiklanishi" mavzuidagi ma'ruzasi bilan boshlab berdi. Shuningdek, bu anjumanda "Turkiy tillar mashina fondining dolzarb masalalari", "Injenerlik tilshunosligi va lug'atshunoslik", "Nutq sistemasi va uning unsurlarini kompyuter yordamida andozalash" singari o'ttizdan qiziqarli ma'ruzalar tinglandi hamda muhokama qilindi. Bu anjuman materiallari alohida to'plam shaklida nashr etilgan.

XX asrning 80-yillaridan boshlab til o'qitishni avtomatlashtirish va unga kompyuterni joriy etish muammolariga qiziqish kuchaydi. 1988-yilda R.G.Piotrovskiyning "Chet tilini o'qitishda kompyuterlashtirish masalalari" nomli asari nashr etilgan edi¹. Bu asar "Kirish" va "Xulosa"dan tashqari besh bobni o'z ichiga oladi. Asar boblarining nomlanishi quyidagicha: 1. Tabiiy til va

uning didaktik modeli, 1. Inson miyasi va kompyuter “miyasi”, 3. Kompyuter yordamida lingvodidaktik masalalarni yechishning tamoyillari, 4. Nutqiy va til materialini tanlash hamda tillarni o‘qitishni optimizatsiyalash, 5. Til o‘rgatuvchi lingvistik avtomat va tillarni o‘qitishni optimizatsiyalash.

Tilshunoslikda EHMning qo‘llanilishi avtomatik tarjima muammolari bilan shug‘ullanish, mashina yordamida matn tuzish, uni qisqartirish, sintezlash jarayonini amalga oshirish uchun keng imkoniyatlar yaratdi. Natijada mashina til materiallarini formallashtirish, chastotali lug‘atlar tuzish, turli qisqartmalar va simvollardan foydalanish, avtomatik tarjima algoritmini tuzish bilan shug‘ullanmoqda.

Keyingi yillarda mamlakatimizda Компьютер лингвистикаси muammolariga oid ishlar yaratildi. Bu o‘rinda S.Muxamedov va R.G.Piotrovskiy¹, A.K.Po‘latov², S.Muxamedova³, S.Rizayev⁴ kabi olimlarning ishlarini alohida ta’kidlash lozim.

So‘nggi yillarda A.Raximovning⁵ va A.K.Po‘latovlarning⁶ kompyuter lingvistikasiga bag‘ishlangan fundamental asarlari nashrdan chiqdi. Bu asarlarda fanning eng asosiy tushunchalari, yo‘nalishlari atroflicha yoritilgan.

Yuqorida aytilgan asarlar va tadqiqot ishlarida modellashtirishga alohida urg‘u beriladi va kompyuter lingvistikasi fanida tilni tekshirishning eng asosiy metodi ekanligi atroflicha yoritib beriladi.

¹ Мухамедов С.А., Пиотровский Р.Г. Инженерная лингвистика и опыт системно-статистического исследования узбекских текстов. – Ташкент: Фан. 1986. – 164 с.

² Пулатов А.К., Жураева Н. Разработка формальной модели грамматики узбекского языка // Узбекский математический журнал, 2002, № 1. – С. 47-54; Пулатов А.К. Тексты лекций по математической и компьютерной лингвистике (электронный вариант). – Ташкент: УзМУ. 2003.

³ Muhamedova S. O‘zbek kompyuter lingvistikasi: rivojlanish tarixi, buguni va kelajagi // Til va adabiyot ta’limi. – Toshkent: 2004. № 2. – B. 77-81; Muhamedova S. Kompyuter lingvistikasi (metodik qo‘llanma). – Toshkent: ToshDPU nashri, 2007. – 74 b.

⁴ Rizayev S. Kibernetika va tilshunoslik. – Toshkent: O‘zbekiston. 1976; Rizayev S. O‘zbek tilshunosligida lingvostatistika muammolari. - Toshkent: Fan, 2006 va boshqalar.

⁵ Raximov A. Kompyuter lingvistikasi asoslari. –Toshkent: Akademnashr, 2011.

⁶ Po‘latov A. Kompyuter lingvistikasi. –Toshkent: Akademnashr, 2011.

Ishning maqsadi amaliy tilshunoslikning eng asosiy yoʻnalishlaridan biri boʻlgan kompyuter lingvistikasi haqida tushunchalarni chuqurlashtirish va lingvistik modellashtirishning turlari va uni amalga oshirishni tahlil etish, oʻrganish.

Ishning vazifalari. Ishning maqsadini amalga oshirish uchun quyidagi vazifalar qoʻyildi:

- amaliy tilshunoslikning kelajakdagi istiqbolini tahlil etish;
- amaliy tilshunoslik yoʻnalishlarini kompyuter lingvistikasi bilan aloqasini oʻrganib chiqish;
- kompyuter lingvistikasining yoʻnalishlarini izohlash;
- modellashtirish tushunchasi va uning turlarini yoritish;
- lingvistik modellashtirish kompyuter lingvistikasining eng asosiy tilni tadqiq etish metodi ekanligini koʻrib chiqish;
- lingvistik modellashtirish va analogiya metodlarini qiyoslash;
- formal grammatika va lingvistik modellashtirishning munosabatini izohlash.
- Modellashtirish boʻyicha slaydlar tayyorlash.

Ishning ilmiy va amaliy ahamiyati. Malakaviy ishda lingvistik modellashtirish boshqa tilni tekshirish metodlari bilan qiyoslab oʻrganib chiqildi. Amaliy tilshunoslikning kelajakdagi istiqbolli yoʻnalishlari haqida maʼlumot berildi. Ushbu ish oliy taʼlim muassasalarida “Kompyuter lingvistikasi” boʻyicha maxsus kurslarda foydalanish mumkin. Ishda keltirilgan dasturlar, slaydlar asosida oʻzbek tilidan taʼlimiy-elektron ishlanmalar tayyorlash mumkin.

Ishniig tuzilishi va hajmi. Ish kirish, ikki bob, umumiy xulosalar, adabiyotlar roʻyxatidan iborat boʻlib, jami _____ sahifani tashkil etadi.

BIRINCHI BOB.

KOMPYUTER LINGVISTIKASI AMALIY TILSHUNOSLIKNING ISTIQBOLI SIFATIDA

1.1. Amaliy tilshunoslik rivoji – ijtimoiy ehtiyoj.

Nazariy bilimlarning amalga tatbiqi doimo har qanday tadqiqot oldiga qo'yiladigan talab bo'lsa-da, bu maqsadni amalga oshirish har doim ham oson emas. Tilshunoslik fanining amaliy ahamiyati haqida gap ketganda, bir necha jabhalarni darrov sanab berish mumkin. Xususan, horijiy tillarni o'rganish va o'rgatish, og'zaki va yozma nutq madaniyatini shakllantirish kabilar.

Insonning ona tili insonni boshqa mavjudotlardangina emas, insonlarning o'zlarini ham bir-biridan farqlab turuvchi, tabiat tomonidan berilgan eng sodda, shu bilan birga eng murakkab tasnif asosidir. Aynan inson tili uning milliy, jinsiy, shaxsiy belgilari haqida ma'lumotlarni shakllantiradi, saqlaydi va boshqalarga yetkazishda asosiy vosita bo'lib xizmat qiladi.

Jahon madaniyatiga buyuk hissa qo'shgan mutafakkirlarning barchasi ham ta'lim va tarbiya borasida fikr yuritar ekanlar, bu jarayonda ona tilining o'rni beqiyos ekanligini qayd etadilar.

O'zbekiston Respublikasi Prezidenti I.A.Karimov o'zlarining "Yuksak ma'naviyat – yengilmas kuch" asarlarida: "Ma'lumki, til o'zlikni anglash, milliy ong va tafakkurning ifodasi, avlodlar o'rtasidagi ruhiy-ma'naviy bog'liqlik til orqali namoyon bo'ladi. Jamiki ezgu fazilatlar inson qalbiga, avvalo, ona allasi, ona tilining betakror jozibasi bilan singadi. Ona tili bu millatning ruhidir"¹ - deb ta'kidlaydilar.

XX-XXI asr tilshunosligida til falsafasining eng yetuk namoyandalaridan hisoblangan Vaysberger va uning izdoshlari tomonidan ona tili fenomenining inson tafakkuri va hayotiy tamoyillari shakllanuvdagi o'rni borasida g'oyatda

¹ Karimov I.A. Yuksak ma'naviyat - yengilmas kuch. - T.,2008.83-b.

qiziqarli tadqiqotlar olib borildi va borilmoqdaki, bu alohida yirik muhohasa uchun mavzu bo'la oladi.

Jahonda amaliy tilshunoslikning barq urib rivojlanayotgan bir qator jabhalari bor. Ular yutuqlarining o'zbek tiliga tatbiq qilinishi ham o'zbek tilshunosligining ijtimoiy va iqtisodiy samaradorligini yuksaltirishga xizmat qilishi shubhasiz.

Tilshunoslikning sof nazariy tadqiqotlar doirasidagi tadriji yangi amaliy shakllar kashf etishi zaruriyati bundan ancha yillar muqaddam yuzaga kela boshlagan edi. Ammo ushbu zaruriyatning konkret ko'rinishlari bahstlabgina emas, hali bu borada bahslar ham boshlanmagan edi. Tilshunos Sh.Shahobiddinova "Ma'rifat" gazetasidagi maqolasida tilning amaliy tomonlarini tadqiq etishga ehtiyoj tug'ilganiga quyidagi voqea asosida tushuntiradi: "15 yillar muqaddam maqola muallifi mudir bo'lib xizmat qiluvchi o'zbek tilshunosligi kafedrasiga bir yigit kirib kelib, o'zini tergovchi deb tanishtirdi va bir masalada yordam so'rab kelganini aytdi. Ma'lum bo'lishicha, o'z joniga qasd qilgan erkak xat qoldirgan bo'lib, uning ushbu nomasida hayotdan o'z xohishi bilan ketayotgani, bu borada hech kimni ayblamasligi aytilgan. Tergovchi shu xat borasida tilshunos sifatida fikrimizni bilmoqchi ekan. Men taajjulanib: "Xatni o'z qo'li bilan yozganiga shubhangiz bormi? Buni aniqlaydigan o'z mutaxassislaringiz yo'qmi?" - deb so'radim. "Mutaxassislarimiz bor.U xatni o'z qo'li bilan yozgani aniq. Faqat o'z hohishi bilan yozganmi, birov majburlaganmi? Xat yozayotganda, uning ruhiy holati qanday bo'lgan, shularni bilmoqchiman", - deb javob berdi u. Men bizda bunday savollarga javob berish imkoni yo'qligi, bunday masalalar bilan tilshunoslik shug'ullanmasligini aytdim. U esa Moskvada tahsil olganini va amaliyot chog'ida shunga o'xshash muammolar bilan Moskva davlat universitetining filologiya fakultetiga murojaat qilishlariga guvoh bo'lganini aytib berdi. Ushbu voqeadan bir necha yil keyin "Iste'dod" jamg'armasi orqali Moskva davlat universitetida malaka oshirayotib bir narsa meni hayron qoldirdi. Filologiya

fakultetining dars jadvali devorda juda katta maydonni egallab turardi. Ma'lum bo'lishicha, talabalar kichik - kichik (4-6 kishilik) guruhlariga bo'linib, filologiyaning muayyan yo'nalishlarida tahsil olishar ekan. Ushbu yo'nalishlar qatorida biz uchun tanish bo'lgan atamalar – tarixiy tilshunoslik, qiyosiy tilshunoslik, struktur tilshunoslik, matnshunoslik, tarjimashunoslik va hokazolar qatorida sud lingvistikasi atamasiga ko'zim tushib qoldi. Ana shunda tergovchi yigitning nima uchun filologlardan yordam so'raganini tushunib yetdim. Keyinchalik ushbu yo'nalish uchun mas'ul bo'lgan kafedra professor–o'qituvchilari bilan suhbatda sud lingvistikasi sohasida o'qish istagida bo'lganlar ko'pligi, sababki, universitetni bitirgach, ish topishda muammo yo'qligini, aksincha, bunday kadrlarga ehtiyoj katta va maoshlari yuqori ekani haqida ma'lumot oldim.”¹ Haqiqatdan ham sud lingvistikasi bo'yicha mutaxassislar Rossiyaning bir qator universitetlarida tayyorlanishga qaramay, bu sohada kadrlar yetishmovchiligi mavjud. Shu o'rinda eslatib o'tish joizki, bizning respublikasizda bu sohada umuman kadrlar tayyorlanmaydi. Barcha tatbiqiy yo'nalishlar qatori sud lingvistikasi ham pragmatik harakterda bo'lib, uning amaliy asoslari muayyan millat uchun alohida ishlab chiqilishi lozim. Tatbiqiy tilshunoslik yo'nalishlarining barchasiga xos bo'lgan yana bir xususiyat ularning chambarchas bog'liqligidir. Tilshunoslikda sathlararo bog'liqlik, til birliklarining o'zaro munosabatlari doim ham tan olinib kelingan. Tatbiqiy tilshunoslik yo'nalishlari aro bog'liqlik esa o'zaro taqozo hamda maqsad va vazifalarga ko'ra tarmoqlanishga asoslangan murakkab ko'rinish kasb etadi.

Aytib o'tilganlaridan tashqari, o'z nomidan ko'rinib turganidek, bu xil yo'nalishlar tilga xos bo'lgan xususiyatlar, tilshunoslik qo'lga kiritgan nazariy natijalarni boshqa fanlar yoki fan yo'nalishlariga tatbiqan amaliy qo'llaydi.

¹ Шаҳобиддинова Ш. Амалий тилшунослик истиқболлари. “Маърифат” газетаси, 2012 йил, 6- сон.

G'arb tillari, shuningdek, rus tili misolida ham tatbiqiy tilshunoslik borasida yirik tadqiqotlar olib borilmoqda. Tarixga nazar tashlasak, hatto sistemaviy tilshunoslik "otasi" F. de Sossyur ham til sistemasi manzarasini har bir individ o'zicha tiklab borishini, o'sha "o'zida va o'zi uchun" sistemada individualizm o'rnini, ya'ni tatbiqiy lingvistikaning jahonda barq urib rivojlanayotgan yo'nalishi psixolingvistika sistemaviy tilshunoslik uchun qarama-qarshi yo'nalish emasligini tan olgan. Uning izdoshlari E.Benvenist, R.Bart, J.Derrida, J.Lakan, M.Fuko va boshqalar ijtimoiy olam manzarasini yaratuvchi qurol – til haqida fikr yuritib tatbiqiy tilshunoslikning nazariy asoslari poydevorini qo'yganlar.

Rus tilshunosligida tatbiqiy tilshunoslikning turli jabhalari bo'yicha ish olib borgan o'nlab olimlarni sanab o'tish mumkin. V.N.Teliya, I.V.Zikova, A.V.Artemova, G.G.Slishkin, I.A.Sternin, G.YE.Kreydmen, F.I.Kartashkova, M.V.Tomskaya, L.V.Polubichenko, I.G.Serova kabilar ularning juda kam qismi, xolos.

O'zbek tilshunosligida keyingi davrda reklamalar, qonunchilik tiliga bag'ishlangan bir qator tadqiqotlar paydo bo'ldi. "Tilshunoslik va tabiiy fanlar", "Kompyuter lingvistikasi", "Matematik lingvistika" fanlari oliy o'quv yurtlari filologiya yo'nalishlari o'quv rejalariga (nomuntazam) kiritildi. Ushbu yo'nalishlarda muayyan tadqiqotlar paydo bo'ldi. Biroq ushbu tadqiqotlar, balki shu sohada dastlabki tajribalardan bo'lgani sababli, umumiy ma'lumotlar berish va jahon tilshunosligida mavjud tajribalarni o'zbek tiliga tatbiq etishga urinishdan iborat bo'lib qoldi.

Shulardan kompyuter lingvistikasi hozirgi kunda eng tez rivojlanayotgan va juda ahamiyatli tilshunoslik sifatida baholanishi mumkin.

Insonning kompyuterga nisbatan qaramlik holatining shakllanib borayotgani bugun jahon miqiyosidagi muammo ekanligi sir emas. Bu dastyorning quliga aylanib qolish insonning ruhi va ruh ifodasi bo'lgan tilida ham aks etayotganini ijtimoiy saytlarga qisqa ekskursiya ham ko'rsatib bera

oladi. Hozirda biz kuzata olgan kompyuter lingvistikasiga oid tadqiqotlarda bu narsa kompyuter lingvistikasi oldiga qo'yilgan asosiy vazifalar sirasida odatda sanalmasa-da, yaqin kelajak psixolingvistika va kompyuter lingvistikasining kesishmasida turuvchi asosiy masala sifatida aynan shu muammoni qo'ya boshlashi shubhasiz. Hozirda ham kompyuter lingvistikasi fani oldida jamiyatimizni yangi ijtimoiy-iqtisodiy bosqichlarga ko'tarishda zarur bo'lgan vazifalar anchagina. Xususan, kompyuter uchun o'zbekcha to'liq va to'laqonli lingvistik ta'minot yaratish, avtomatik tahrirlash, o'zbek tilidan boshqa tilga yoki aksincha tarjima qiluvchi mukammal dasturlarni yaratish, sodda qilib aytganda, kompyuterni o'zbekcha "gapirtirish" amaliy tilshunoslikning bu yo'nalishi oldida turgan eng dastlabki vazifalardir. Keyingi yillarda tatbiqiy yo'nalishning bu sohasiga tilshunoslar orasida qiziqishning ortib borayotgani quvonarli hol bo'lsa-da, ko'p sonli bo'lmagan bu tadqiqotchilar faoliyati bu o'ta dolzarb yo'nalishda maxsus mutaxassislar tayyorlash zaruriyatini hal etib bera olmasligi ham aniq.

Hozirgi kunda amaliy tilshunoslikning o'zbek tilini tatbiqiy tadqiq etishning metodologik asoslari ishlab chiqish, metod va prinsiplari belgilash muhim masala. O'zbek tilining tatbiqiy yo'nalishlarini amalga oshirishning dasturi ishlab chiqilishi lozim. Unda amaliy tilshunoslik taraqqiyoti uchun dolzarb bo'lgan quyidagi muammolarning hal etilishi ko'zda tutiladi:

- tilshunoslikning sof nazariy fan doirasidan chiqib, iqtisodiy samarador yo'nalishlar taraqqiy topgan fan mavqeini ham olishida zaruriy tadqiqotlar metodologiyasi va metodlari belgilab berish;

- amaliy tilshunoslikning bugungi ijtimoiy-iqtisodiy taraqqiyot davri uchun ayniqsa dolzarb bo'lgan ustuvor yo'nalishlari belgilash va asoslash;

- ishlab chiqilgan metodlar va metodologiya asosida muayyan amaliy yo'nalish tavsifini yaratish.

Kelgusida amaliy tilshunoslikka oid tadqiqotlar natijalari asosida darsliklar, qo'llanmalar yaratilishi ma'naviy yetuk, tadbirkor va zamonaviy

barkamol yosh avlodni tarbiyalash ishida muhim hissa bo'lib xizmat qilishi shubhasiz.

1.2. Kompyuter lingvistikasi yo'nalishlari.

Ma'lumki, tilshunoslik fani XIX (1816-yilda) asrda mustaqil fan sifatida shakllandi. Shundan boshlab u turli aspektlarda, yo'nalishlarda rivojlanib kelmoqda. Keyingi yillarda barcha fanlardagi kabi tilshunoslikda ham ikki fanning «chorrahasida» (kesishuvida) yuzaga kelgan fanlar jadal rivojlanmoqda. Jumladan, ana shunday fanlar sirasiga sotsiolingvistika (sotsiologiya va tilshunoslik), psixolingvistika (psixologiya va tilshunoslik), etnolingvistika (etnografiya va lingvistika), neyrolingvistika (nevrologiya va tilshunoslik), matematik lingvistika va kompyuter lingvistikasi fanlarini kiritish mumkin. Bunday holni boshqa fanlar doirasida ham kuzatish mumkin: biokimyo, astrofizika, matematik fizika, matematik logika kabi. Buni fanlar tizimida bir necha fanlarning o'zaro hamkorligi, integratsiyasi deb baholash lozim bo'ladi. XX asrning 50-yillaridan boshlab tilshunoslikda «mashina tarjimasi», «mashina tilshunosligi» atamaları qo'llanila boshlandi. Mazkur asrning buyuk kashfiyoti bo'lgan kompyuter texnologiyalari tilshunoslikka ham kirib kelganining isboti edi. Mashina tarjimasi yoki avtomatik tarjima deyilganda bir tildagi matnni ikkinchi bir tilga EHM (kompyuter) vositasida, tez vaqt ichida tarjima qilish nazarda tutiladi. Mashina tarjimasining asoschilari kibernetika va matematika sohasi vakillari bo'lib, keyinchalik bu ishda tilshunoslar ham faol qatnasha boshlagan. Shu tariqa mashina tarjimasi g'oyalari butun dunyoda nazariy va amaliy tilshunoslikning rivojlanishida katta ahamiyat kasb etdi. Bu yo'nalish bilan parallel ravishda formal grammatika nazariyasi yuzaga kelib, til va uning alohida aspektlari modelini yaratishga e'tibor qaratildi. Tilning bu jihatlari matematik lingvistika fanida ishlab chiqildi, bu, o'z navbatida, kompyuter lingvistikasi fanining yuzaga kelishi uchun poydevor bo'ldi. Demak, shu asosda tilshunoslikning yangi yo'nalishi -

kompyuter lingvistikasi va tilshunoslikning bir qator nazariy va amaliy yo'nalishlari vujudga keldi.

Matematik lingvistika fani XX asrning 50-yillarida (1952-yilda) tilshunoslikning alohida yo'nalishi sifatida yuzaga keldi. Bu fanning shakllanishida Kopengagen struktural tilshunoslik maktabi (glossematika)ning asoschisi Lui Yelmslevning g'oyalari o'ziga xos «turtki» vazifasini o'tagan. U hatto til 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 Chomskiyning formal grammatika, transformatsion grammatika haqidagi qarashlari bevosita matematik lingvistikaning alohida yo'nalish sifatida yuzaga kelishiga sabab bo'lgan.² Mana shunday qarashlar ta'sirida matematik lingvistika fani shakllandi. Matematik lingvistika - bu tabiiy tillarning matematik modellarini (bunday formallashtirilgan til metatil deb ataladi) ishlab chiqish, xususan, sun'iy tillarni yaratish algoritmini tuzish bilan shug'ullanuvchi fandır. Matematik lingvistika oldida turuvchi eng muhim masalalar quyidagilardir:

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

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

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

¹ Нурмонов А. Структур тилшунослик: илдизлари ва йуналишлари. - Тошкент, 2008. - Б.145.

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

Kompyuter lingvistikasi matematik lingvistikaning mantiqiy davomi bo'lib, u amaliy tilshunoslikning eng muhim qismini tashkil etadi. Kompyuter lingvistikasi 1954-yil AQShda Jorjtaun universitetida mashina tarjimasida bo'yicha dunyoda o'tkazilgan birinchi tajriba asnosida yo'nalish sifatida shakllana boshladi, 1960-yilga kelib mustaqil fan sifatida shakllandi. Kompyuter lingvistikasi inglizcha «computational linguistics» so'zining kalkasidir. XX asrning 80-yillariga qadar bu fan turlicha nomlar bilan atalgan: hisoblash lingvistikasi, matematik lingvistika, kvantitativ lingvistika, injener lingvistikasi kabi. Bu fanning asosiy maqsadi lingvistik masalalarni yechishning kompyuter dasturlarini ishlab chiqish, inson va mashina (kompyuter) muloqotini optimallashtirish, tabiiy tilni qayta ishlashdir. 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. Tabiiy tilni qayta ishlash bo'yicha yaratilgan dasturiy ta'minotlar quyidagilar: AlchemyAPI, Expert System S.p.A., General Architecture for Text Engineering (GA TE), Modular Audio Recognition Framework, Monty Lingua, Natural Language Toolkit (NLTK).

Kompyuter lingvistikasining asosiy yo'nalishlariga quyidagilar kiradi:

- avtomatik o'qitish tizimini ishlab chiqish;
- bilimlarni tekshirish;
- matnlarni turli jihatdan avtomatik tahrirlash;
- matnlarning avtomatik tarzda morfologik, sintaktik va semantik tahlilini (inglizcha parsing) ta'minlovchi tizimlar yaratish (parser dasturlar);
- mashina tarjimasida uchun mo'ljallangan dasturlarni ishlab chiqish;
- lug'atlarni va kompyuterdagi matnni statistik tahlil qilish;
- lingvistik muammolarni hal qilishga yo'naltirilgan optimal dasturlar yaratish;

- muloqotning kompyuter modelini ishlab chiqish;
- matn strukturasi gipertekst texnologiyasini yaratish;
- elektron lug'atlar va tezauruslar yaratish;
- sun'iy intellekt tizimini ishlab chiqish;
- matnlar korpusini yaratish;
- matnni avtomatik referatlash;
- sujet strukturasi modellashtirish;
- nutqni avtomatik tushunish tizimini ishlab chiqish (inglizcha ASR - Automatic Speech Recognition);
- matndan faktlarni avtomatik ajratib olish (inglizcha fact extraction, text mining).

Kompyuter lingvistikasida qo'llaniladigan va yaratiladigan lingvistik vositalarni shartli ravishda 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 vositalarini (algoritm tuzish, dasturlar yaratish, kompyuter analizi va sintezi kabilar) o'z ichiga oladi.

O'zbek tili materiallari bo'yicha kompyuter lingvistikasiga oid tadqiqot olib borgan olimlar sifatida H.Arziqulov, S.Rizayev, S.Muhamedov, A.Po'latov, S.Muhamedova, N.Jo'rayevalarni ko'rsatish mumkin. Olimlar asosan statistik tahlil, algoritm tuzish, o'zbek tilining aksiomatik nazariyasi, fe'llarning kompyuter analizi va sintezi kabi yo'nalishlar bo'yicha izlanishlarni amalga oshirganlar. O'zbek tilshunosligidagi kompyuter lingvistikasining boshqa yo'nalishlari esa o'z tadqiqotchilarini kutmoqda. Shunga ko'ra, bugungi kunda o'zbek tilshunosligining kompyuter bilan bog'liq holda hal etilishi lozim bolgan quyidagi vazifalarni ta'kidlab ko'rsatish mumkin:¹

¹ Po'latov A., Muhametlova S. Kompyuter lingvistikasi. - Toshkent, 2007. - B.5-6.

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

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

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

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

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

- mumtoz tilshunoslikning tahlil obekti tabiiy til hisoblanadi va talilida uning mavjud barcha imkoniyatlari e'tiborga olinadi. Kompyuter lingvistikasida esa ko'proq sun'iy tillar (programmashtirish 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 yashay olmaydi. Jumladan, kompyuter lingvistikasi fani ham bundan mustasno emas. Kompyuter lingvistikasi mantiq, informatika, statistika, semiotika, kibernetika, ehtimollar nazariyasi kabi turli fan sohalari bilan o'zaro aloqadorlikda ish ko'radi.

IKKINCHI BOB.

KOMPYUTER LINGVISTIKASINING TILNI TADQIQ ETISH METODLARI VA LINGVISTIK MODELLASHTIRISH

2.1. Formal grammatika.

Grammatika tabiiy tillar strukturasi haqidagi fan bo'lib, u morfologik kategoriya va shakllar, sintaktik ketegoriya va konstruksiyalar hamda so'z yasash usullari tizimini tashkil etadi. Til qurilish birliklarining paradigma- tik va sintagmatik munosabatga kirishuvi (vertikal va gorizonta l yo'nalishda) muayyan qoidalar asosida yuz beradi. Ushbu qoidalar yigindisi tilning grammatik qurilishi va grammatik sistemasi deb yuritiladi. Grammatika ikki yirik sistemani o'z ichiga oladi:

1. Morfologik sistema - so'z formalari, grammatik shakllar, grammatik kategoriyalar tizimi.
2. Sintaktik sistema - sintaktik kategoriyalar va konstruksiyalar, so'zlarning o'zaro birikish qonuniyatlari.

Grammatika yunoncha so'z bo'lib, «o'qish va yozish san'ati» ma'nosini bildiradi. Bu tushuncha dastlab xat-savod, imloni o'rgatuvchi fan ma'nosida ishlatilgan. Keyinroq grammatika tilning fonologik, morfologik, sintaktik va semantik belgilarini tavsiflovchi til qoidalari majmuasi ma'nosiga ko'chgan. Grammatika haqidagi ilk qarashlar falsafa, mantiq, germeneytika (diniy matnlarni sharhlash, talqin etish) qobig'ida boigan. Bunda tilning paydo boiishi, til va tafakkur munosabati, lingvistik belgi motivatsiyasi, diniy matnlar mazmunini to'g'ri tushunish bilan bog'liq masalalar o'rganilgan, til strukturasi esa yetarli darajada tadqiq etilmagan. Faqatgi- na miloddan avvalgi II - I asrlarda shakllangan Aleksandriya grammatika maktabining eng yirik vakili Dionisiy Traks (miloddan avvalgi 170 - 90- yillarda yashagan) o'zigacha boigan tadqiqotchilarning ishlari va tajriba- larini o'rganib, ulardan foydalanib rimliklar uchun «Grammatika san'ati» («Grammatike techne») nomli

sistemalashtirilgan dastlabki yunon grammatikasini yaratdi. Shundan buyon grammatika falsafadan mustaqil soha sifatida ajralib chiqdi.¹

Antik davrda grammatika preskriptiv (to'g'ri gapirish, to'g'ri jumla tuzishni ko'rsatuvchi yo'riqnoma) xarakterga ega boigan. 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 aspekt-da tadqiqotlar olib borildi.

XX asr boshlarida Ferdinand de Sossyur grammatikani tavsiflashda sinxroniya va diaxroniyani farqladi. Bunda grammatikani tarixiy (diaxron) o'rganishni emas, balki sinxron (muayyan vaqtdagi holati) tadqiq qilish lozim, deb hisobladi. XX asrning 50-yillarida Noam Chomskiy grammatikani tavsiflashning yangi metodini tavsiya etdi. Bu tilshunoslikda «formal grammatika nazariyasi» deb ataladi. Olimning tadqiqot natijalari kompyuter lingvistikasining shakllanishiga asos bo'ldi. N.Chomskiyning formal grammatika haqidagi qarashlari «Syntactic Structures» (1957) hamda «Aspects of the Theory of Syntax» nomli kitoblarida yoritilgan.

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

Grammatik tavsifning eng sodda (primitiv) formal modeli gap bo'laklariga ajratish modeli sanaladi. Tabiiy tillar grammatikasida sodda gap

¹ Сусов И.П. История языкознания. - Тверь: Тверский гос. ун-т. 1999. -С.24 - 25.

ikki qismdan - ega va kesimdan tashkil topgan sintaktik qurilma sifatida tavsiflanadi. Ega nutqning predmetini, kesim esa uning belgisini 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 ifodalan- sa, 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 aniqlashtiruvchilari vazifasida keladi. Eganing aniqlashtiruvchilari ot, sifat, son, olmosh, kvantifikatorlar (miqdor bildiruvchi so'zlar), determinatorlar (artikllar) orqali ifodalangan aniqlovchi 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 hoi kesimning aniqlashtiruvchilari bo'lib keladi. Mazkur modelda ega va kesim gapning bosh boiaklari, gapning semantik-sintaktik yadrosi, aniqlovchi, to'ldiruvchi, hoi esa ikkinchi darajali bo'laklar deb nomlanadi. Ushbu talqin mumtoz sintaksis tamoyillariga muvofiq keladi.

XX asrning ikkinchi yarmida struktur sintaksis, semantik sintaksis yo'nalishlari paydo bo'lishi bilan gapning sintaktik tuzilishiga, gap bo'laklariga bo'lgan yondashuv ham o'zgardi. Yuqoridagi ikki cho'qqili nazariyaning o'rniga bir cho'qqili nazariya maydonga keldi. Unga ko'ra, faqat kesimgina gapning grammatik yadrosi sifatida e'tirof etildi. Ushbu yondashuv fransuz tilshunosi Lyusen Tenerning verbosentrizm (lotincha «verbum» - «fe'l», «kesim» ma'nolarini anglatadi, «centrum» - «markaz», unga ko'ra, gapning markazi gap bo'lagi sifatida kesim, so'z turkumi sifati- da fe'l-dir) konsepsiyasi bilan bog'liqdir.¹ Ingliz tilshunosi O.Yespersen mazkur konsepsiyaga qarshi o'laroq gapdagi mutlaq hokim bo'lak sifatida egani e'tirof etgan. Gap bo'laklari

¹ Теньер Л. Основы структурного синтаксиса. - М.: Прогресс, 1988. -С.12.

iyerarxiyasida egani markazga qo'yish tendensiyasi mantiq fanidan kelib chiqqan. Chunki mantiqda ism (subekt) 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 bo'lak sifatida talqin etiladi.

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

Tilshunoslikda an'anaviy gap bo'laklariga ajratishdan tashqari, yana gapni kommunikativ nuqtayi nazardan qismlarga ajratish ham mavjud, u gapning aktual bo'linishi deb ataladi. Bunda gap tema va rema qismlariga ajratiladi. Bunda tema nutqiy muloqotning bizga ma'lum qismiga nisbatan ishlatiladi, 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 bo'lishi ham mumkin. Masalan, «*Botirjon universitetda ishlayapti*» gapida ega vazifasida kelgan *Botirjon* tema hisoblanadi, chunki nutq vaziyati uchun Botirjon aniq va avvaldan tanish bo'lgan shaxsdir. 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.¹ Konstruksiya atamasi ostida ma'noli qismlarning ketma-ket munosabatidan tashkil topgan butunlik tushuniladi. Ishtirokchilar deb kattaroq konstruksiya tarkibiga kirgan so'z yoki konstruksiyaga aytiladi. Muayyan konstruksiyaning bevosita shakllanishida ishtirok etgan bir yoki bir nechta ishtirokchilarga bevosita ishtirokchilar deyiladi. Masalan, *A'lochi talabalar o'z vazifalarini aniq va puxta bajaradilar*. Bu jumla konstruksiya hisoblanadi, uning tarkibidagi barcha ma'noli so'zlar («va» bog'lovchisidan tashqari) ishtirokchilar, bevosita aloqaga kirishgan so'zlar bevosita ishtirokchilar deyiladi. Bunda *a'lochi* so'zi *talabalar* so'zi bilan, *aniq*, *puxta* so'zlari *bajaradilar* so'zi bilan, *o'z* so'zi *vazifalarini* so'zi bilan o'zaro zich bog'langan. Demak, bevosita ishtirokchilar sintaktik modelida dastlab so'z birikmalari ajratiladi, so'ngra ular gap ichida birlashtiriladi.

O'zbek tilshunosligi darg'alaridan biri, o'tkir nigohli olim, professor Nurmonovning 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 boiadi. Ko'rinadiki, har bir konstruksiya binar tamoyilga ko'ra qismlarga bo'linadi va bu bo'linish konstruksiya doirasida graduallik (darajalanish) xususiyatiga ega boiadi.²

Bevosita ishtirokchilarga ajratish modeli tanqidiy idrok etilib, undagi ayrim g'oyalar 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 transformatsiya xususidagi qarashlari ta'sirida yuzaga kela boshladi, bu yo'nalish Z.Harrisning shogirdi N.Chomskiy tomonidan rivojlantirildi. Unga ko'ra, dastlabki yadro

¹ Nurmonov A. Lingvistik tadqiqot metodologiyasi va metodlari. - Toshkent: Akademnashr, 2010. - B.88.

² Nurmonov A. Lingvistik tadqiqot metodologiyasi va metodlari. - Toshkent: Akademnashr, 2010. - B.124.

strukturalar birikma strukturalar qoidalar tizimi orqali tug'iladi. Ya'ni har qanday til grammatikasi yadro strukturalar (kernel structures) yig'indisidir. Yadro gaplar deb 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.	NVpN	Noun Verb Preposition Noun	Ot fe'l predlogli ot	John looked at Mary
3.	NVN	Noun Verb Noun	Ot fe'l ot	John saw Mary
4.	N is N	Noun is Noun	Ot is ot	John is a teacher
5.	N is p N	Noun is Preposition Noun	Ot is predlogli ot	John is in bed
6.	N is D	Noun is Adverb	Ot is ravish	John is out
7.	NisA	Noun is Adjective	Ot is sifat	John is angry

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.

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

² Херрис З.С. Совместная встречаемость и трансформация в языковой структуре // Новое в лингвистике. Вып.11. - М., 1962. - С.627 - 628.

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

Transformatsion grammatika matematik lingvistika va kompyuter lingvistikasi fanining rivojlanishida muhim ahamiyat kasb etadi. 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 ta'minlaydi. Uchinchidan, transformatsion grammatika til o'qitishni modellashtirishda ham muhim ahamiyatga molik. To'rtinchidan, transformatsion grammatika tabiiy tillardagi paradigmatic va sintagmatik munosabatlarning me'yoriyligi va grammatik to'g'ri va noto'g'ri jummalarni ajratishning aniq mezonlarini belgilaydi.

N.Chomskiy tavsiya etgan grammatikada gaplarni «tug'dirish» (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:

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

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

Gap hosil qilish jarayoni noterminal simvollarni o'ng qismga davomiy o'rin almashtirish asosida yuz beradi. Bunda boshlang'ich simvol S o'ziga mos keluvchi simvollar birikuvi NP hamda VP ga almashtiriladi. So'ng NP va VP simvollari ham, o'z navbatida, Det + N hamda Verb + NP ga almashtiriladi. Bu jarayon barcha simvollar noterminal bo'lguncha davom eta- veradi. Masalan, ingliz tilidagi quyidagi jumlaning bevosita ishtirokchilar modeli asosida tahlil

etaylik. Birinchi bosqichda jumla NP va VPga ajratiladi: *The old man | saw a black dog there* - *Keksa odam u yerda qora itni ko'rdi*. Keyingi bosqichda yuqoridagi ajralish, o'z navbatida, yana ikki qism- ga ajratiladi: *The || old man \ saw a black dog || there*. So'nggi bosqichda jumla bevosita ishtirokchilar modeli asosida «daraxt» («sentence tree») ko'rinishida tasvirlanishi mumkin:¹

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

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

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

N.Chomskiy qoidalarning ifodalanish turiga qarab formal grammatikani

¹ Бархударов И.С. Структура простого предложения английского языка. М.: Высшая школа, 1966. - С.21.

4 tipga ajratgan va bu Chomskiy iyerarxiyasi deb nomlanadi:

1.0 grammatika (unrestricted phrase-structure grammar) - juda murakkab generativ grammatika, bunda $G=(N, \Sigma, P, S)$ belgilari yordamida ish ko'riladi. N - 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 ($V=T \cup N$). Bu grammatika o'ta abstraktlashgan matematik modellarga, informatika nazariyasiga tatbiq etilgan.

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

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

4.3 grammatika (regular grammar) - regular grammatika deb ataladi, u juda sodda va cheklangan grammatika bo'lib, tilning sodda unsurlari uchun ishlatiladi (miqdor, konstantalar, o'zgaruvchilar uchun).

Formal grammatikada 2 ta tushuncha asosiy: a) generation - tug'dirish, hosil qilish, porojdeniye; b) recognition - tanish, bilish, raspoznavaniye. Shunga ko'ra, grammatika 2 ga bolinadi: 1) tug'dirish grammatikasi (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. Sir emaski, har bir inson o'z ona tili grammatikasini bolaligidanoq o'zlashtiradi. Bu jihatlar esa grammatikani formallashtirish masalasi juda murakkab ekanligini tasdiqlaydi. Grammatikani ideal darajada

tavsiflovchi qoidalar sistemasining mavjud emasligi, har qanday tavsiya etilayotgan model nomukammal ekanligi xususida mashhur tilshunos olim K.Fossler shunday yozadi: «Grammatikani fonetika, morfologiya va sintaksisga bo'lish hech kimga sir emas. Tilni uning shakllanish jarayoni ichida emas, balki uning o'z holati ichida o'rganish zarur. Uning ustida anatomik operatsiya o'tkaziladi. Jonli nutq gaplarga, gap bo'laklariga, so'zlarga, bo'g'inlarga va tovushlarga ajratib chiqiladi. Ushbu metod to'liq o'zini oqlaydi va ahamiyatga molik natijalarga olib borishi mumkin, biroq shu bilan birga xatolarning manbai hisoblanishi ham mumkin. Xatolar ajratilgan bo'linishning nutq organizmi uchun asos bo'lishini da'vo qilingan paytda boshlanadi, aslida bu bo'linish insonning ixtiyoriga bog'liq ravishda mexanik va majburiy tarzda parchalash edi. Mazkur holat anatomiyadagi holatni eslatadi: anatom o'zining yorish, kesish ishlarini albatta o'zboshimchalik bilan asossiz amalga oshirmaydi, i o'ziga birmuncha qulaylik tug'diruvchi joylarni tanlab oladi. Xuddi sliuningdek grammatist ham nutqni tovush, bo'g'in, asos, qo'shimcha kabilarga ajratilida tabiiy bo'Imagan, lekin o'ziga qiday yo'ldan boradi».¹

Yuqoridagilardan kelib chiqib aytish mumkinki, shu kungacha olimlar tomonidan ishlab chiqilgan barcha grammatik modellar o'ta shartli va nisbiydir. Keltirilgan barcha modellar - gap boiaklari modeli ham, bevosita ishtirokchilar modeli ham, tobelik daraxti modeli ham, transformatsion va generativ grammatika modeli ham mutlaq va benuqson emas. Ularning hammasi til 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). Bu xususda nemis olimi 1931-yilda K.Gedel shunday yozadi: «Yetarli darajada mazmunga boy nazariya formallashtirilsa, bu nazariyani formal tizimda to'la aks ettirish mumkin emas. Bunda aniqlanmagan,

¹ Фослер К. Позитивизм и идеализм языкознания // Звегинцев В.А. История языкознания XIX – XX очерках и извлечениях. Част 1. –М.: Просвещение, 1964. –С. 328 – 329.

formallashtirilmagan qoldiq qoladi».¹ Demak, har qanday formallashtirilmagan model til grammatikasining butun qirralarini qamrab ololmaydi.

2.2. Matematik statistika va kvantitativ metod.

Statistik tahlil metodi boshqa fanlar qatori tilshunoslik fani uchun ham muhim ahamiyatga ega. Bunda tilda mavjud hodisalar aniq matematik parametrlarda baholanadi, tekshiriladi. Statistika lotincha «status» soʻzidan olingan boʻlib, hodisalarning holati, ahvolini bildiradi. «Status» soʻzidan «stato» - davlat, «statusta» - davlatni biluvchi, «statustica» - davlat toʻgʻrisida muayyan bilim, maʼlumotlar yigʻindisi maʼnosini bildiruvchi soʻzlar kelib chiqqan.² Bundan 3500 yil oldin Qadimgi Misrda oʻtkazilgan aholi hisobi, Rossiyada Pyotr I davridagi «aholi taftishi», Angliyadagi «Dahshatli sud kitobi» kabilar dastlabki statistik maʼlumotlardir.

XVII asrda Angliyada yuzaga kelgan «siyosiy arifmetika» fani statistikaning shakllanishiga asos boʻlgan. Statistika ijtimoiy hodisalarning miqdoriy tomonlarini ularning sifat tomonlari bilan uzviy aloqadorlikda oʻrganuvchi fandır. Hodisalarning sifat tomonlarini maxsus fanlar oʻrganadi. Xususan, tovushlar, soʻzlar, gaplarni lingvistika, qofiya, ritm, vazn, sheʼr tuzilishini adabiyotshunoslik, organik hayotni biologiya, yer qatlamlari va boyliklarini geologiya kabi fanlar oʻrganadi. Statistika esa ushbu hodisalarni miqdoriy parametrlar asosida oʻrganadi. Bunda statistika hodisalarni turli omillar taʼsiridagi oʻzgaruvchanlik xususiyatini - variatsiyasini eʼtiborga oladi. Shu jihatdan statistik maʼlumotlar hodisalar rivojining dinamikasini aks ettira oladi. Lekin shu oʻrinda aytib oʻtish lozimki, statistikaga yagona mutlaq haqiqat deb yondashish, statistik maʼlumotlarga haddan tashqari ishonish ham toʻgʻri emas. Ayrim yolgʻon faktik materiallarga asoslanib tayyorlangan statistik maʼlumotlar hodisalar haqida xato tushunchalar berishi mumkin. Shu maʼnoda

¹ Анисимов А.В. Компьютерная лингвистика для всех: Мифы. Алгоритмы. Язык. - Киев: Наук, думка, 1991. - С.77.

² Abdullayev Y. Statistika nazariyasi. - Toshkent, 2002. - B.4 - 5.

statistika yuzasidan shunday hazilomuz gap aytiladi: «Yolgʻonning uch turi bor: a) yolgʻon; b) gʻirt yoʻlgʻon; c) statistika».

Statistikaning til hodisalariga tatbiq etilishi natijasida lingvostatistika sohasi shakllandi, unda til birliklari, lisoniy hodisalar, matnlar statistik tahlil etiladi. Matnning statistik tahlili - tilni ehtimollik belgilariga koʻra tahlil qilish, til hodisalarining statistik xarakteristikalarini asosida baholash demakdir. Til birligining gapda yoki matnda qoʻllanish chastotasi (surʼati) ehtimollik deyiladi. Til hodisalarini haqidagi haqiqatlar lingvostatistik tadqiqotlar yordamida aniqlanadi. Masalan, oʻzbek tili uchun qaysi boʻgʻin strukturasi tipik ekanligini, gap modullaridan qaysi biri eng asosiy konstruktsiya ekanligini, bosh kelishikdagi otlarning asosan qaysi gap boʻlagi vazifasida kelishini, tanlangan matnda soʻz turkumlarining qay darajada ishlatilishini aniqlash uchun lingvostatistik tahlillarga ehtiyoj seziladi.

Til hodisalarini statistik metod asosida tahlil qilish uchun doim maʼlum bir matn yoki matnlar toʻplami olinadi. Ular badiiy adabiyotdan, gazeta va jurnallardan, publitsistik maqolalardan, ilmiy asarlardan, dialektologlarning jonli nutqidan yozib olgan materiallaridan tanlanadi. Statistik xususiyati oʻrganilayotgan muayyan matnlar «tanlama» deyiladi. Tanlamaning hajmi tekshirilayotgan hodisalarning umumiy sonidir, uni «N» bilan belgilash mumkin. Masalan, 100 ta soʻz; 37 ta bosh kelishik shaklidagi ot kabi. Aniqlanayotgan hodisani koʻrsatadigan miqdor (masalan, ot oldida aniqlovchi vazifasida keladigan takrorlangan otlar miqdori) absolut chastota hisoblanadi, uni «M» harfi bilan belgilash mumkin. 37 ta otdan 30 tasining aniqlovchi vazifasida kelishi absolut chastotadir. Absolut chastota (M)ning tanlama 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: $P = M/H$

¹ Усмонов С. Умумий тилшунослик. - Тошкент: У^итувчи, 1972. - Б.199.

b) protsent asosida: $P = MЧ100/H$

c) promil hisobida: $P = MЧ 1000/H$

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.¹ 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 ($r^4f = C$), bu yerda S matn hajmiga bog'liq. Masalan, amerikalik yozuvchi Mark Tvenning «Tom Soyerning sarguzashtlari» 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 barchasi 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 takrorlanadi. Bunday so'zlar hapas legomena (yunoncha so'z «faqat bir marta o'qi», degan ma'noni anglatadi) 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. Ayniqsa, badiiy matnlarning statistik tahlili natijasida o'sha matn muallifi ham aniqlanmoqda. Bunda matnda tez-tez ishlatiluvchi til birliklari (otlar, sifatlar, kalit so'zlar, fe'llar,

¹ Плат У. Математическая лингвистика // Новое в лингвистике. - М.: Прогресс, 1965. Вып.IV. - С.204.

grammatik shakllar, jumla qurilishi, bir soʻz bilan aytganda, yozuvchining idiostilini - oʻziga xos uslubini koʻrsatuvchi vositalari) lingvostatistik tahlil yordamida aniqlanadi. Turli matnlardan olingan dalillarning qiyosiy tahlili bizga oʻsha matnning mazmun-mundarijasini, matn tuzilgan davrni, dalillarni argumentlash xarakterini va hatto mualliflikni aniqlashga imkon beradi. Badiiy asarlarning sujet qurilishiga, matn strukturasiqa kvantitativ metodlarning tatbiq etilishi natijasida F.Dostoyevskiy, L.Tolstoy, M.Sholoxov kabi yozuvchilarning matnlari atributsiyasi amalga oshirilgan. Keyingi paytlarda anonim (muallifi nomaʼlum) va psevdonim (soxta muallifli) 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» (1960) («Tillar morfologik tipologiyasiga kvantitativ yondashuv») nomli maqolasida umumlashtirgan. Olim oʻzigacha 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 koʻzga tashlanmaydi. Bu esa tasnifni amaliyotga tatbiq etib tekshirishni taqozo etadi». 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, vetnam, fors, ingliz, suaxili, eskimos tillariga) tatbiq qilgan. Bunda olim oʻzi belgilagan formulalar asosida lisoniy tiplarning va tillarda kuzatiladigan tipologik belgilarning miqdoriy tavsifini, statistik maʼlumotlarni ilmiy jamoatchilikka havola etdi.

¹ Баранов А.Н. Введение в прикладную лингвистику. - М.: Эдиториал УРСС, 2001.-С.24.

Chastotali lug'atlar matnda so'zlarning ishlatilish chastotasi (qay darajada ishlatilishi, tez-tez 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 agglutinativ tillarda 64 - 89 foiz atrofidadir. Flektiv tillarda dastlabki ko'p uchraydigan 100 - 200 so'z formasi (ko'pchiligi yordamchi so'zlar) jami har xil so'z formalarining 50 foizini tashkil etsa, bu ko'rsatkich uchun agglutinativ konstruktiv tillarda 600 - 1000 ko'p uchraydigan so'z formalari zarur bo'lar ekan, buning asosiy sababi agglutinativ va flektiv tillar morfologik strukturasi va so'z formalarining tuzilishidadir.

Chappa (ters) chastotali lug'atlar tilda qo'shimchalarning qanday qo'llanilishini aniqlab beradi. Chastotali va ters chastotali lug'atlarning ma'lumotlariga ko'ra, ot, olmosh, sifat, ravish va son tillardagi jami so'zlarning 70% ni 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%).

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

- birinchidan, chastotali lug'atlar avtomatik qidiruv tizimi asosida ishlaydigan elektron lug'atlar yaratish uchun, ularning sohaviy turlari uchun tayyor lingvistik ta'minot vazifasini o'taydi;

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

- uchinchidan, 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.

2.3. Tilni modellashtirish - tilni kompyuterda tekshirishning optimal usuli.

- Model (lotincha «modelus» so'zidan olingan bo'lib, «nusxa», «andaza», «o'lchov», «meyor» ma'nolarini anglatadi) tabiiy fanlar yoki umuman fanda shunday moddiy qurilma, grafik, sxema, umuraan, bilish vositasi sifatida tushuniladiki, u muayyan original-obekt haqidagi ma'lumotlar majmui tarzida yuzaga kelgan hosila-obekt demakdir. Boshqacharoq tushuntirganda, model tabiiy obektlarning imitatsiyasidir (o'xshashi, taqlidiy ko'rinishi), u o'zbek tilidagi qolip, andaza so'zlariga mos keladi. U hodisalarning yuzaga kelishi uchun asos vazifasini o'taydi, bunda aniq yoki mavhum obektlar kichraytirilgan obektlar, sxemalar, chizmalar, fizikaviy konstruksiyalarda tadqiq etiladi. Buni oddiy hayotiy misol bilan tushuntira- digan bo'lsak, olmani xarakterlovchi belgilar, atributlar - uning dumaloqligi, mevaligi, shirinligi o'sha tushunchaning fikriy modeli hisoblanadi. Agar biz olmani loydan yoki sun'iy bir materialdan yasasak, bu uning moddiy modeli hisoblanadi. Yoki globus yerning modeli, o'yinchoq mashina haqiqiy mashinaning modeli sifatida baholanishi mumkin. Model quyidagi asoslarga ko'ra bilishda muhim hisoblanadi:

- birinchidan, u o'rganish obektini soddalashtiradi;
- ikkinchidan, uni boshqa obektlar ta'siridan ajratadi;
- uchinchidan, model obektni ta'riflashni osonlashtiradi.

Model dastlab amaliy sohalarda ishlatilgan. Keyinchalik ilm-fanning ijtimoiy sohalariga ham kirib keldi. Bu matematika, kimyo fanlariga keng tatbiq qilingani yaqqol ko'rinadi. «Model» tushunchasi fan va texnikada turli ma'nolarda ishlatilgani bois modellashtirish turlarining yagona tasnifi mavjud emas. Tasnif modelning xarakteriga ko'ra, modellashtirilayotgan obektning tabiatiga ko'ra, modellashtirish tatbiq qilinayotgan soha yoki yo'nalishga qarab amalga oshirilishi mumkin. Masalan, modellashtirishning quyidagi turlarini ajratib ko'rsatish mumkin: informatsion modellashtirish, lingvistik modellashtirish, kompyuterli modellashtirish, matematik modellashtirish, matematik-kartografik modellashtirish, molekulyar modellashtirish, mantiqiy modellashtirish, pedagogik modellashtirish, psixologik modellashtirish, statistik modellashtirish, struktur modellashtirish, fizik modellashtirish, iqtisodiy-matematik modellashtirish, imitatsion modellashtirish, evolutsion modellashtirish, kognitiv modellashtirish kabi.

Modellarni shartli ravishda quyidagicha tasniflash mumkin:

1. Tabiiy modellar – o'rganilayotgan obekt bilan bir turda bo'ladi va undan faqat o'lchamlari, jarayonlarining tezligi va ba'zi hollarda yasalgan materiali bilan farq qiladi.
2. Matematik modellar – prototipdan (asl nusxadan) jismoniy tuzilishi bilan farq qiladi, lekin prototip bilan bir xil matematik tasvirga ega bo'ladi.
3. Mantiqiy-matematik modellar – belgilardan iborat bo'lib, abstrakt model hisoblanadi va tafakkur jarayonini o'rganishda qo'llaniladi.
4. Kompyuterli modellar – matematik va mantiqiy modellashtirish metodlariga asoslanib kompyuterda algoritm va dasturlardan foydalanib yaratiladigan modellar.

Kompyuterli modellashtirish bugungi kunda barcha fanlarda, xususan, kompyuter lingvistikasida ham samarali metod hisoblanadi.¹ Kompyuter

¹ Grishman R. Computational linguistics // Cambridge University Press. 1994. - P.7 - 8.

modellashtirish quyidagi asosiy bosqichlardan iborat:

- masalaning qo'yilishi, modellashtirish obektining aniqlashtirilishi;
- konseptual (tushunchaviy, fikriy) modelning ishlab chiqilishi, tizim asosiy unsurlarining ajratib olinishi;
- formalizatsiya, ya'ni matematik model bosqichi; algoritm yaratilishi hamda dastur ishlanishi;
- 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 modellashtirish natijasi hisoblanadi.

Modellashtirish jarayoni uchta unurni o'z ichiga oladi:

- subekt (tadqiqotchi);
- tadqiqot obekti;
- o'rganuvchi subekt va o'rganiluvchi obekt munosabatini aks ettiruvchi model.

Modellashtirish har bir fan obektini soddalashtiruvchi metoddir. Lingvistik birliklarni modellashtirish bu belgilar tarkibidagi elementlarning bar-qaror munosabatlariga asoslanadi. Shuning uchun ham butunlik tarkibidagi elementlar o'rtasida munosabatlarning barqaror va beqaror turlarga ajratilishi lingvistik modellashtirish uchun katta ahamiyatga ega. Modellashtirish barcha fanlar uchun xos boigan umumilmiy metod hisoblanadi va u quyidagi tamoyillarga amal qiladi:

- deduktivlik - mantiqiy xulosa chiqarishga asoslangan boiadi, xususiylikdan umumiylikka tamoyilida boiadi;
- tafakkur eksperimentidan foydalanish;
- evristik funksiyaga ega boiishi. Ya'ni u yangi g'oyalar bera olishi va uni amaliyotda sinab ko'rish imkoniyati boiishi zarur;
- eksplanatorlik xususiyatiga, ya'ni tushuntirish kuchiga ega boiishi

kerak. Shundagina model eski nazariya tushuntirib bera olmagan muam- moni hal qiladi, obektning ilgari kuzatilmagan, ammo kelajakda amalga oshishi mumkin boigan tomonini kashf etadi;

- modelni ideallashtirilgan obekt sifatida talqin qilish.

Shu o'rinda aytib o'tish zarurki, modellashtirish obektni umumlashtirish darajasiga ko'ra quyidagicha bo'ladi:

1. Lingvistik faktni 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.Chomskiy, Z.Harris, Ch.Hokket kabi namoyandalari tomonidan kirib kelgan. Uning taraqqiyoti esa 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 jarayonini aks ettiradi. 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, o'ta abstraktlashgan va ratsionallashgan bo'ladi.

Modellashtirish metodi ayrim tillarga, jumladan, ingliz tiliga faol tatbiq qilingan. O'zbek tilida sodda gap qurilishi:

S + O + V : Men kitob o'qidim. Men xat yozdim.

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

Bundan kelib chiqib aytish mumkinki, o'zbek tilida qo'shma gapning eng kichik modeli quyidagicha boiadi:

S + V, S₂ + V₂: Bahor keldi, ishlar qizib ketdi.

Ingliz, rus tillarida sodda gap qurilishi quyidagicha belgilangan:

S + V + O:

Я пишу диссертацию. Он читает книгу.

I have read the book. I am writing a research work.

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 olg'a 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 so'z turkumlariga oid bo'lgan so'zlar zan-jiri hisoblanadi va tahlilda morfologiya bazasiga tayaniladi. Masalan, «*The young man painted the door yesterday*» jumlasini distributiv model asosida quyidagicha tahlil etiladi: D 3 I^a 2-d D I^b 4. Bu yerda D - otning aniqlovchisi (inglizcha determiner), 3 - sifat, Γ - birlikdagi ot, 2-d - o'tgan zamon shaklidagi fe'l, 4 - ravishni anglatadi. Demak, mazkur modelda turli so'z turkumlariga mansub so'z shakllarining nutq zanjiridagi distributsiyasi (tarqalishi, qurshovi) gap strukturasi modellashtirishning asosiy mezonini sanaladi.

Shunga ko'ra, so'z turkumlarini modellashtirishda quyidagi shartli belgilar keng e'tirof etilgan.

So'z turkumlari

I. Ot predmetning (u keng ma'noda tushuniladi, ya'ni jonli va jonsiz narsalarni anglatadi) nomini bildiradi. U kim, nima, qayer so'roqlaridan biriga javob bo'ladi. Ot uchun N, ko'plik shaklidagi ot uchun Ns, turdosh otlar uchun N , atoqli otlar uchun N , o'tli birikma uchun NP simvollaridan

com⁷ 1 prop

dan foydalaniladi.

II. Sifat qanday, qanaqa so'roqlariga javob bo'lib, predmetning belgisini bildiradi. Sifat uchun Adj yoki A, sifatli birikma uchun AP simvollaridan

foydalaniladi.

III. Son qancha, nechanchi so'roqlariga javob bo'lib, predmetning miqdori va tartibini bildiradi. Son so'z turkumi uchun Q simvolidan foydalaniladi.

IV. Fe'l predmetning harakat yoki holatini bildirib, nima qilgan, nima qilyapti, nima qilmoqchi so'roqlariga javob beradi. Fe'l uchun V, o'timli fe'l uchun V_t o'timsiz fe'l uchun V, fe'lning predikativ (shaxsli shakli-finite form) shakli uchun V_p fe'lning nopredikativ shakli (shaxssiz shakli - non-finite form) uchun V_{np} hozirgi zamon sifatdoshi yoki gerundiy shaklidagi fe'l uchun V o'tgan zamon sifatdoshi yoki majhul nisbatdagi fe'l uchun V_{cn} , fe'lli birikma uchun VP, yordamchi fe'llar uchun aux, modal fe'llar uchun mod simvollaridan foydalaniladi.

V. Ravish ifodalayotgan harakat yoki holatning belgisi yoki turlicha vaziyatlarini bildiradi. Ravish asosan fe'lga, yana sifat yoki ravishga ham bog'lanib, qay tarzda, qayerda, qachon va h.k. so'roqlarga javob boiadi. Ravish uchun Adv yoki D simvolidan foydalaniladi.

VI. Olmosh gapda ot, sifat yoki ravish o'rnida qo'llanuvchi so'z turkumidir. Olmosh uchun Pron simvolidan foydalaniladi.

VII. Artikl yordamchi so'z bo'lib, otning maxsus belgilovchisi hisoblanadi. Artikl barcha tillarda uchrayvermaydi, Art simvolidan foydalaniladi.

VIII. Predlog (old ko'makchi) yordamchi so'z bo'lib, otning gapdagi boshqa so'zlarga nisbatan bo'lgan munosabatini (ya'ni makon, vaqt va sabab kabi munosabatlarni) ifodalaydi. Predlog uchun Prep yoki p simvolidan foydalaniladi.

IX. Bog'lovchi yordamchi so'z bo'lib, gap bo'laklari yoki gaplarni bog'lash uchun xizmat qiladi. Bog'lovchi uchun Conj simvolidan foydalaniladi.

X. Yuklama uchun Part simvolidan foydalaniladi.

XI. Undov so'zlar uchun Interj simvolidan foydalaniladi.

XII. Taqlid so'zlar uchun Mim simvolidan foydalaniladi.

Gap bo'laklari

Ixtiyoriy tildagi yozma matn yoki shaxsning nutqi jumla deb ataluvchi alohida segmentlardan iborat bo'ladi. Ular sodda va qo'shma (ya'ni sodda gaplardan tashkil topgan) gaplarga bo'linadi. Har qanday tilning grammatikasini o'rganish sodda gaplar bilan tadqiq etishdan boshlanadi. Har qanday tildagi sodda gaplar tarkibida ega, kesim, to'ldiruvchi, aniqlovchi, hoi kabi gapning bosh va ikkinchi darajali boiaklari mavjud bo'lishi mumkin.

I. Ega - gapning zaruriy bo'lagi bo'lib, boshqa boiaklar bilan grammatik boglanadi, u predmet yoki shaxsni bildiradi. Ega kim, nima, qayer so* roqlariga javob boiadi. Ega uchun S simvolidan foydalaniladi.

II. Kesim - gapning ham semantik, ham grammatik markaziy boiagi bo'lib, gapning yadrosi hisoblanadi, kesimsiz gap shakllanmaydi, u harakat, holat, xususiyat kabilarni anglatadi. Kesim nima qilyapti, nima qilmoqchi, nimadan tashkil topgan kabi so'roqlarga javob beradi. Kesim uchun P (ba'zan V) simvolidan foydalaniladi.

III. To'ldiruvchi deb predmetni bildiruvchi va ma'no jihatidan kesimni toidiruvchi gap boiagiga aytiladi. Vositasiz (obyektsiz) toidiruvchi kimni, nimani soioqlariga, vositali toidiruvchi esa kimga, nimaga, kimda, nima- da, kimdan, nimadan, kim haqida, nima haqida va h.k. so'roqlarga javob boiadi. Toidiruvchi uchun O simvolidan foydalaniladi.

IV. Aniqlovchi otning belgilarini tushuntiruvchi, izohlovchi gap boiagidir. Aniqlovchi predmetning belgisini bildirgan holda qanday, kim- ning, nechanchi, qancha kabi so'roqlarga javob boiadi. Aniqlovchi kompyuter lingvistikasida gap boiagi sifatida ajratilmaydi, chunki u gapda boshqa boiaklarga tobelanib keladi, u gap uchun nozaruriy (fakultativ, obligator emas) boiak hisoblanadi.

V. Hol kesimda ifodalanayotgan harakat yoki holatning o'rni, payti, maqsadi, sababi va bajarilish sharti kabilarni bildiruvchi gap boiagidir. U qanday, qay tarzda, qayerda, qachon, nima sababdan, nima uchun kabi so'roqlarning biriga javob beradi. Hoi uchun M simvolidan foydalaniladi.

Modellashtirish metodi keyingi paytlarda tilshunoslikka faol tatbiq qilina boshlandi. Bunda olimlar mazkur metodning bir qator afzalliklari va pragmatik jihatlarini nazarda tutadilar. Buni quyidagicha izohlash mumkin:

- birinchidan, modellashtirish amaliy metod hisoblanadi, tavsifiy emas;
- ikkinchidan, modellashtirish metodi har qanday sharoitda optimal («eng qulay», «eng maqbul» degan ma'nolarni ifodalaydi) hisoblanadi;
- uchinchidan, modellashtirish metodi ekonomiya prinsipiga tayanadi. Bunda so'z, uzun ta'rif va tavsiflarga ehtiyoj boimaydi;

- to'rtinchidan, obyektни tushuntirish va izohlashni osonlashtiradi va soddalashtiradi.

Borliqdagi hamma narsalarning ijobiy va salbiy tomonlari boiganidek, modellashtirish metodining ham pozitiv va pragmatik tomonlari bilan birga salbiy tomonlari ham bor. Uning kamchiliklari, salbiy tomonlarini quyidagicha umumlashtirish mumkin:

- birinchidan, 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 topgan bo'ladi;

- ikkinchidan, bizning obyekt haqidagi bilimlarimiz rivojlanib borgan sari modellar eskiradi. Avvalgi bilimlarni inkor qilish hisobiga fan rivojlanadi. Mavjud modellar bilimlarimiz ufqini cheklaydi va xato tasavvurlarga olib kelishi mumkin;

- uchinchidan, tabiat va jamiyatdagi hech bir narsa boshqa obyektlar ta'siridan xoli bo'lmaydi. Ya'ni modellar nisbiylik xarakteriga ega, biz ularni mutlaq haqiqat sifatida qabul qila olmaymiz.

2.4. Tilni analogiya metodi asosida formallashtirish.

Kompyuter lingvistikasida eng ko'p qo'llanilayotgan metodlardan biri analogiya metodi hisoblanadi. Analogiya (yunoncha so'zdan olingan, «moslik», «o'xshashlik» ma'nosini anglatadi) - munosabatlarning tengligi, muvofiqligi;

hodisa va jarayonlarning muayyan xossalari 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 mush-tarak 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. Shunda analogiyaga tayanib berilgan so'zlarning mushtarak jihatlari ajratish mumkin. Yoki ko'p odamlar «*singan oyna - baxtsizlik, noxitsh voqealarni keltirib chiqaradi*», degan irimga ishonadi. Bu irimning kelib chiqishi ham qandaydir mistik asosga ega emas, balki analogiyaga asoslangan. Ya'ni singan oyna insonni xunuk aks ettiradi, mazkur tashqi o'xshashlik asosida tasvirning buzilishi undan foydalanuvchi uchun ham shunga mos yomon oqibatlarni keltirib chiqaradi, degan xulosaga olib kelgan. Shu o'rinda ta'kidlash joizki, analogiya *induksiya* (umumiylikdan xususiylikka borish; qoidadan misolga borish) va *deduksiya* (xususiylikdan umumiylikka borish; misoldan qoidaga borish) metodlari bilan uzviy aloqadorlikda boladi. 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 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, bandan ham ko'ra kuclili matematik - nazariyalar analogiyasini sezgan odam; biroq odam o'zini eng kuchli hisoblashi uchun analogiyalar o'rtasidagi analogiyani ko'ra olishi kerak*». Ushbu mulohazalardan kelib chiqqan holda aytish mumkinki, analogiya lingvistikadagi, xususan, kompyuter lingvistikasidagi tadqiqotlar uchun ham o'ziga xos «kompass»

vazifasini bajaradi.

Akademik L.V.Sherba o'zining bir ma'ruzasida yozuv taxtasiga shunday jumlani yozadi: «Глокая куздра штеко будланула бокра и кудрячит бокренка». Talabalardan uning grammatik tahlilini talab qiladi. Mazkur jumla aslida hech qanday mazmunga ega emas, lekin u rus tilidagi mazmunga ega boigan qaysidir jumlaning eslatadi. Bu jumlaning grammatik tahlilini amalga oshirish mumkin. Masalan, «кудра» so'zi - birlik, bosh kelishik, jenskiy rod shaklidagi ot bilan ifodalangan ega, «будланула» so'zi - o'tgan zamon shaklidagi fe'l bilan ifodalangan kesimlardan biri, «кудрячит» so'zi noaniq shakldagi fe'l bilan ifodalangan boshqa bir kesim. Masalaga bunday yondashish analogiya metodi asosida bo'ladi. 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 korrelatsion aloqaga tayaniladi.¹ Ya'ni oxirgi harf tarkibi bir xil bo'lgan so'zlarning derivatsion va relatsion modellari (so'z yasalishi va shakl yasalishi) hamda grammatik ma'lumotlari (qaysi so'z turkumiga mansubligi, qanday grammatik kategoriyalarga egaligi) bir xil boiishi tabiiy. Masalan, ingliz tilida *-er, -ness, -dom* kabi, rus tilida *-ность, -щик, -ник* kabi, o'zbek tilida *-shunos, -chi, -lik* qo'shimchalarga ega bo'lgan so'zlar yuqori ehtimollik bilan 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.

¹ Белоногов Г.Г. Об использовании принципа аналогии при автоматической обработке текстовой информации // Проблемы кибернетики. 1974. №28.-С.33.

XULOSA

Ma'lumki, XXI asr kompyuter texnologiyalari, ilg'or pedagogika va axborot texnologiyalari, yuksak tafakkur, fan va texnikaning jadal taraqqiyot asri bo'ladi. XXI asr boshiga kelib, "Elektron ta'lim", "Elektron boshqaruv", "Ochiq ta'lim", "Masofaviy ta'lim", "Axborotlashgan ta'lim" kabi tushunchalar hayotimizga kundan-kun singib bormoqda. Prezidentimiz I.A.Karimov so'zlari bilan aytganda, "hozirgi axborot, kommunikatsiya, kompyuter texnologiyalari asrida, internet kundan-kunga hayotimizning barcha jabhalariga tobora chuqur va keng kirib borayotgan bir paytda, odamlarning ongi va tafakkuri uchun kurash hal qiluvchi ahamiyat kasb etayotgan bir vaziyatda bu masalalarning jamiyatimiz uchun naqadar dolzarb va ustuvor bo'lib borayotgani haqida gapirib o'tirishga hojat yo'q"¹.

Hozirgi paytda injenerlik lingvistikasi, hisoblash lingvistikasi, matematik lingvistika va "elektron - tarjimon", kompyuter lingvistikasi yuzasidan bir qator yangi muammolar o'rtaga tashlanmoqda. Darhaqiqat, zamonaviy kompyuter juda ulkan imkoniyatlarga egaki, uning uchun shaxmat o'ynashdan tortib, tibbiy tashxis qo'yish, korxonalarni boshqarish, qadimgi xalqlarning sirli xatlarini "o'qib", mazmunini so'zlab berish, she'r yozish, kuylar ijro etish, katta hajmdagi kitoblarga annotatsiyalar yozish, loyihalar tarxini chizish kabi yuzlab mehnat faoliyati turlarining barchasida, albatta, zamonaviy kompyuter texnologiyalari ham ishchi kuchi, ham vaqt jihatidan juda yaqin yordamchi sanaladi.

Bunlay ilmiy tadqiqotlar sirasiga murakkabligi, dolzarbligi bilan xarakterlanuvchi kompyuter lingvistikasi fani ham kiradi. Kompyuter lingvistikasi oldida turuvchi eng muhim masalalar quyidagilar:

- tilning aksiomatik nazariyasini ishlab chiqish;

¹ Karimov I.A. Bizning bosh maqsadimiz - jamiyatni demokratlashtirish va yangilash, mamlakatni modernizatsiya va isloh etishdir. O'zbekiston demokratik taraqqiyotning yangi bosqichida. – Toshkent: O'zbekiston, 2005. – B. 38-39.

- formal grammatika yaratish;
- tillarning matematik modellarini ishlab chiqish.¹

Keyingi yillarda mamlakatimizda Kompyuter lingvistikasi muammolariga oid ishlar yaratildi. Bu o'rinda S.Muxamedov va R.G.Piotrovskiy², A.K.Po'latov³, S.Muxamedova⁴, S.Rizayev⁵ kabi olimlarning ishlarini alohida ta'kidlash lozim.

So'nggi yillarda A.Raximovning⁶ va A.K.Po'latovlarning⁷ kompyuter lingvistikasiga bag'ishlangan fundamental asarlari nashrdan chiqdi. Bu asarlarda fanning eng asosiy tushunchalari, yo'nalishlari atroflicha yoritilgan.

Yuqorida aytilgan asarlar va tadqiqot ishlarida modellashtirishga alohida urg'u beriladi va kompyuter lingvistikasi fanida tilni tekshirishning eng asosiy metodi ekanligi atroflicha yoritib beriladi.

Hozirgi kunda amaliy tilshunoslikning o'zbek tilini tatbiqiy tadqiq etishning metodologik asoslari ishlab chiqish, metod va prinsiplari belgilash muhim masala. O'zbek tilining tatbiqiy yo'nalishlarini amalga oshirishning dasturi ishlab chiqilishi lozim. Unda amaliy tilshunoslik taraqqiyoti uchun dolzarb bo'lgan quyidagi muammolarning hal etilishi ko'zda tutiladi:

¹ Raximov A. Kompyuter lingvistikasi asoslari. –Toshkent: Akademnashr, 2011. –B. 5-6.

² Мухамедов С.А., Пиотровский Р.Г. Инженерная лингвистика и опыт системно-статистического исследования узбекских текстов. – Ташкент: Фан. 1986. – 164 с

³ Пулатов А.К., Жураева Н. Разработка формальной модели грамматики узбекского языка // Узбекский математический журнал, 2002, № 1. – С. 47-54; Пулатов А.К. Тексты лекций по математической и компьютерной лингвистике (электронный вариант). – Ташкент: УзМУ. 2003.

⁴ Muhamedova S. O'zbek kompyuter lingvistikasi: rivojlanish tarixi, buguni va kelajagi // Til va adabiyot ta'limi. – Toshkent: 2004. № 2. – B. 77-81; Muhamedova S. Kompyuter lingvistikasi (metodik qo'llanma). – Toshkent: ToshDPU nashri, 2007. – 74 b.

⁵ Rizayev S. Kibernetika va tilshunoslik. – Toshkent: O'zbekiston. 1976; Rizayev S. O'zbek tilshunosligida lingvostatistika muammolari. - Toshkent: Fan, 2006 va boshqalar.

⁶ Raximov A. Kompyuter lingvistikasi asoslari. –Toshkent: Akademnashr, 2011.

⁷ Po'latov A. Kompyuter lingvistikasi. –Toshkent: Akademnashr, 2011.

- tilshunoslikning sof nazariy fan doirasidan chiqib, iqtisodiy samarador yo'nalishlar taraqqiy topgan fan mavqeini ham olishida zaruriy tadqiqotlar metodologiyasi va metodlari belgilab berish;

- amaliy tilshunoslikning bugungi ijtimoiy-iqtisodiy taraqqiyot davri uchun ayniqsa dolzarb bo'lgan ustuvor yo'nalishlari belgilash va asoslash;

- ishlab chiqilgan metodlar va metodologiya asosida muayyan amaliy yo'nalish tavsifini yaratish.

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

Model (lotincha «modelus» so'zidan olingan bo'lib, «nusxa», «andaza», «o'lchov», «meyor» ma'nolarini anglatadi) tabiiy fanlar yoki umuman fanda shunday moddiy qurilma, grafik, sxema, umuraan, bilish vositasi sifatida tushuniladiki, u muayyan original-obekt haqidagi ma'lumotlar majmui tarzida yuzaga kelgan hosila-obekt demakdir. Boshqacharoq tushuntirganda, model tabiiy obektlarning imitatsiyasidir (o'xshashi, taqlidiy ko'rinishi), u o'zbek tilidagi qolip, andaza so'zlariga mos keladi. U hodisalarning yuzaga kelishi uchun asos vazifasini o'taydi, bunda aniq yoki mavhum obektlar kichraytirilgan obektlar, sxemalar, chizmalar, fizikaviy konstruksiyalarda tadqiq etiladi.

Modellarni shartli ravishda quyidagicha tasniflash mumkin:

1. Tabiiy modellar - o'rganilayotgan obekt bilan bir turda bo'ladi va undan faqat o'z xossalari, jarayonlarining tezligi va ba'zi hollarda yasalgan material bilan farq qiladi.

2. Matematik modellar - prototipdan (asl nusxadan) jismoniy tuzilishi bilan farq qiladi, lekin prototip bilan bir xil matematik tasvirga ega bo'ladi.

3. Mantiqiy-matematik modellar - belgilardan iborat bo'lib, abstrakt model hisoblanadi va tafakkur jarayonini o'rganishda qo'llaniladi.

Kompyuterli modellar - matematik va mantiqiy modellashtirish metodlariga asoslanib kompyuterda algoritm va dasturlardan foydalanib yaratiladigan modellar.

Lingvistik model tushunchasi struktur tilshunoslikning E.Sepir, L.Blumfeld, R.Yakobson, N.Chomskiy, Z.Harris, Ch.Hokket kabi namoyandalari tomonidan kirib kelgan. Uning taraqqiyoti esa 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 jarayonini aks ettiradi. Masalan, o'zbek tilida morfologik usul asosida so'z yasashining umumiy modeli: asos + so'z yasovchi qo'shimcha; xususiy modellari: asos + -chi; asos + -dosh; asos + -do'z kabi.

Metamodellar - bunda lingvistik modellar saralanadi, u gipotetik-deduktiv xarakterga ega, o'ta abstraktlashgan va ratsionallashgan bo'ladi.

ADABIYOTLAR RO'YXATI

1. Karimov I.A. Bizning bosh maqsadimiz - jamiyatni demokratlashtirish va yangilash, mamlakatni modernizatsiya va isloh etishdir. O'zbekiston demokratik taraqqiyotning yangi bosqichida. - Toshkent: O'zbekiston, 2005. - B. 38-39.
2. Karimov I.A. Adolatli jamiyat sari. - Toshkent: O'zbekiston, 1998. - B. 58.
3. Karimov I.A. Yuksak ma'naviyat - yengilmas kuch. - T., 2008. 83-b.
4. Raximov A. Kompyuter lingvistikasi asoslari. - Toshkent: Akademnashr, 2011. - B. 5-6.
5. Abdullayev Y. Statistika nazariyasi. - Toshkent, 2002. - B.4 - 5.
6. Grishman R. Computational linguistics // Cambridge University Press. 1994.
7. Muhamedova S. O'zbek kompyuter lingvistikasi: rivojlanish tarixi, buguni va kelajagi // Til va adabiyot ta'limi. - Toshkent: 2004. № 2. - B. 77-81; Muhamedova S. Kompyuter lingvistikasi (metodik qo'llanma). - Toshkent: ToshDPU nashri, 2007. - 74 b.
8. Nurmonov A. Lingvistik tadqiqot metodologiyasi va metodlari. - Toshkent: Akademnashr, 2010. - B.88.
9. Nurmonov A., Iskandarova Sh. Tilshunoslik nazariyasi. - Toshkent: Fan, 2008. - B.140- 141.
10. Po'latov A. Kompyuter lingvistikasi. - Toshkent: Akademnashr, 2011.
11. Po'latov A., Muhametlova S. Kompyuter lingvistikasi. - Toshkent, 2007. - B.5-6.
12. Raximov A. Kompyuter lingvistikasi asoslari. - Toshkent: Akademnashr, 2011.
13. Rizayev S. Kibernetika va tilshunoslik. - Toshkent: O'zbekiston. 1976;

14. Rizayev S. O'zbek tilshunosligida lingvostatistika muammolari. - Toshkent: Fan, 2006 va boshqalar.
15. Анисимов А.В. Компьютерная лингвистика для всех: Мифы. Алгоритмы. Язык. - Киев: Наук, думка, 1991. - С.77.
16. Баранов А.Н. Введение в прикладную лингвистику. - М.: Эдиториал УРСС, 2001.-С.24.
17. Бархударов И.С. Структура простого предложения английского языка. М.: Высшая школа, 1966. - С.21.
18. Белоногов Г.Г. Об использовании принципа аналогии при автоматической обработке текстовой информации // Проблемы кибернетики. 1974. №28.-С.33.
19. Гладкий А.В., Мельчук И.А. Элементы математической лингвистики. - М.: Наука, 1969. -С.7.
20. Мухамедов С.А., Пиотровский Р.Г. Инженерная лингвистика и опыт системно-статистического исследования узбекских текстов. - Ташкент: Фан. 1986. - 164 с.
21. Нурмонов А. Структур тилшунослик: илдизлари ва йуналишлари. - Тошкент, 2008. - Б.145.
22. Пиотровский Р.Г. Компьютеризация преподавания языков (учебное пособие по спецкурсу). - Л.: 1988. - С. 3-75.
23. Плат У. Математическая лингвистика // Новое в лингвистике. - М.: Прогресс, 1965. Вып.IV. - С.204.
24. Пулатов А.К., Жураева Н. Разработка формальной модели грамматики узбекского языка // Узбекский математический журнал, 2002, № 1. - С. 47-54;
25. Пулатов А.К. Тексты лекций по математической и компьютерной лингвистике (электронный вариант). - Ташкент: УзМУ. 2003.

26. Теньер Л. Основы структурного синтаксиса. - М.: Прогресс, 1988. -С.12.
27. Усмонов С. Умумий тилшунослик. - Тошкент: У^итувчи, 1972. - Б.199.
28. Фослер К. Позитивизм и идеализм вязыкознании \ \ Звегинцев В.А. История языкознания XIX – XX очерках и извлечениях. Част 1. –М.: Просвещение, 1964. –С. 328 – 329.
29. Шаҳобиддинова Ш. Амалий тилшунослик истиқболлари. “Маърифат” газетаси, 2012 йил, 6- сон.