

TILSHUNOSLIKDA LINGVOSTATISTIK TAHLIL MUAMMOLARINING O‘RGANILISHI

Xolbekova Dilrabo Xusniddin qizi

O‘zbekiston Milliy universiteti
Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasini o‘qituvchisi.

Ibodullayeva Kunduzoy

O‘zbekiston Milliy universiteti
Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasini o‘qituvchisi.

Annotatsiya: Maqolada jahon tilshunosligida kompyuter texnologiyalari yordamida olib borilgan tadqiqotlar hamda o‘zbek tilshunosligida lingvostatistik tahlil imkoniyatlarini kuzatib va shu kuzatish natijalarini qayta ishlab, ilmiy ishlar olib borgan olimlarning amalga oshirgan nazariy va amaliy ishlari haqida ma’lumot beriladi.

Kalit so‘zlar: Kompyuter lingvistikasi, lingvostatistika, modellashtirish, eksperimental fonetika, areal lingvistika, korpus lingvistikasi, izohli kombinator lug‘at, chastotali lug‘at.

Аннотация: В статье представлены сведения о теоретической и практической работе ученых, наблюдавших возможности лингвостатистического анализа в узбекском языкознании и возможности лингвистического анализа в мировом языкознании с помощью компьютерных технологий.

Ключевые слова: компьютерная лингвистика, лингвистика, моделирование, экспериментальная фонетика, ареальная лингвистика, корпусная лингвистика, аннотированный комбинаторный словарь, частотный словарь.

Abstract: The article provides information about the theoretical and practical work of scientists who have observed the possibilities of linguo-statistical analysis in Uzbek linguistics and the possibilities of linguistic analysis in world linguistics with the help of computer technologies.

Keywords: Computational Linguistics, Linguistics, Modeling, Experimental Phonetics, Areal Linguistics, Corpus Linguistics, Annotated Combinatorial Dictionary, Frequency Dictionary.

Bugungi kunda o‘zbek tilshunosligi yildan yilga jadal sur‘atlarda rivojlanib bormoqda. Yetuk mutaxassislarimiz tilshunoslikning sohalari ustida ilmiy ish va tadqiqotlar olib borishmoqda. Badiiy matnni turli tahlillar asosida o‘rganish natijasida, yozuvchining adabiy ijodini chuqurroq tushunish va baholashga imkon beradigan juda boy materiallar to‘planmoqda.

Ilm-fan rivojlanishining hozirgi bosqichida tilshunoslikda ishlab chiqilgan barcha ishlar asosan tabiiy tildagi matnlarning mashinada ishlash muammosini hal qiladi. Jumladan, matnni o‘rganishning yana bir usuli – lingvostatistik usul ham kompyuter amaliyoti bilan bog‘liq.

Matn korpusini har qanday kompyuter statistik qayta ishlashni amalga oshirishda ma’lum bir kompyuter texnikasining ishonchliligini baholashning uslubiy murakkabligi mavjud. Mashina tahlili ishonchli yoki ishonchli emasligini qanday bilishimiz mumkin? Bir tomondan, avtomatik tahlil natijalarini “qo‘lda” bajarilgan shunga o‘xshash ishlar bilan taqqoslab ko‘raylik. Rasmiy sinov va taqqoslash-inson mehnatining katta harajatlarini va harakatini talab qiladi. Bundan

tashqari, avtomatik tahlilning muvaffaqiyatini belgilashda bunday tekshirishning ishonchligi masalasi munozarali. Chunki u aniq matnlar korpusi taqdim etadigan va qo'lda tahlil qilish uchun mo'ljallanmagan ma'lumotlar asosida amalga oshiriladi. Aynan shu miqdordagi matnli ma'lumotlar lingvistik modellashtirish bo'yicha yangi tasavvurni taqdim etadi. [Clear., 1993: 282; Siliakus., 1983: 422].

XX asrning 60-70 yillaridagi jahon tilshunosligining taraqqiyoti eksperimental fonetika, fonologiya, poetik nutq nazariyasi, areal lingvistika, nutq statistikasi, kompyuter lingvistikasi, korpus lingvistikasi kabi bir qator yangi ilmiy yo'nalishlar paydo bo'lganligi bilan xarakterlanadi. [Urinbayeva D., 2022: 10].

Tabiiy tillardagi matnlarni tahlil qilishda kompyuterlardan foydalanish yangi fanni – kompyuter lingvistikasi fanini yuzaga keltirdi, uning asosiy maqsadi kompyuterda lingvistik belgilar, tizimlar va funksiyalarni modellashtirishdan iborat. [Piotrovskiy., 1975:4]. Kompyuterda amalga oshiriladigan statistik tahlilning yutug'i shundan iboratki, bir muallifning hayoti va ijodining turli davrlarida yaratilgan asarlari uslubini o'rganishda yordam beradi.

Matnni tadqiq qilishda statistik metodlardan foydalanilar ekan, bunday metodlar muayyan bir matnda so'z yoki ibora, bo'g'in yoki gap singari til birliklarining qo'llanishida o'ziga xos qonuniyatlarni aniqlashga qaratiladi.[Yo'ldoshev M., Muhamedova S., Saparniyazova M., 2021: 155].

Nutq oqimida takrorlanuvchi til birliklarining bo'lishi til va nutq muammolarini o'rganishda statistik vositalardan foydalanishga imkoniyat yaratadi. Lingvostatistik tahlil imkoniyatlari til birliklarining qonuniyatlarini kuzatish va shu kuzatish natijalarini qayta ishlash jarayonida aniqlanadi. [Yo'ldoshev M., Muhamedova S., Saparniyazova M., 2021: 156].

Jahon tilshunosligida XX asrning 50-yillaridan e'tiboran til va matn muammolari bilan bog'liq masalalarni kompyuter texnologiyalari yordamida hal qilish tadqiqi bilan, nafaqat tilshunos, balki boshqa soha olimlari ham shug'ullanib kelmoqda. Mazkur masalalar har qanday matnni lisoniy o'rganishning yangi-yangi imkoniyatlari til materialini axborot texnologiya dasturlari orqali tanlash, tahlil qilish va uni qayta ishlash jarayonini muayyan darajada tezlashtirishga hamda matnning turli xususiyatlari haqida matematik jihatdan to'g'ri va aniq ma'lumotlarni qo'lga kiritishga zamin yaratmoqda. [Urinbayeva D., 2022: 5].

Jahon tilshunosligida lingvostatistikaning nazariy asoslarini ishlab chiqish, matnlarni lingvostatistik jihatdan o'rganishga yo'naltirilgan tadqiqotlar negizida quyidagi ilmiy natijalar olingan: lingvostatistikaning nazariy asoslari ishlab chiqilgan va statistik tahlil masalalari amalga oshirilgan; korpus lingvistikasi masalalari bilan shug'ullangan; badiiy matnni statistik va kvantitativ usulda o'rganish bo'yicha tadqiqotlar olib borilgan; so'zlarning semantik xususiyatlari batafsil tavsiflangan izohlikombinator lug'at yaratilgan; til tizimini lingvostatistik modellashtirish masalalari tadqiq etilgan, jumlaning tashkil etuvchi alohida so'zlarning ma'nolari butun jumlaning ma'nosini tashkil etuvchi yagona bir butunga bog'lanish xususiyatlari o'rganilgan; tilning morfologik modelini avtomatik tuzilishini statistik tahlili yuzasidan tadqiqotlar olib borilgan, tilning leksik-morfologik strukturasi lingvostatistik jihatdan tahlil qilingan; so'z turkumlari ichida fe'l va uning lug'aviy ma'no guruhlari valentliklari tasnifi va mashina tarjimasiga bag'ishlangan bir qancha tadqiqot ishlari bajarilgan. [Urinbayeva D., 2022: 7]

Bu esa kompyuter lingvistikasi bo'yicha jahon tilshunoslari R.Bailey, R.Allan, R.Oakman, V.Cherubini, R.Merritt, V.Kuraskiyevich, G.Altmannlarning, rus tilshunoslari A.Juravlev, M.Kirillov, V.Kirnosov, L.Kovriginalarning va o'zbek tilshunoslari S.Muhammedov,

S.Musaxo'jayev, S.Rizayev, M.Ayimbetov, H.Arzikulov, S.Karimov, S.Muxamedova, N.Abdurahmonova kabi tadqiqotchilarning nomi bilan bog'liq. Mazkur olimlarning ishlari asosan publisistik va badiiy matnning lingvostatistik tahlili doirasida olib borilgan, barcha uslubdagi matnlar tilining lisoniy birliklarini o'rganish, leksik qatlamlarining tarkibini aniqlash va tasniflash, matnning lingvostatistik, sinergetik tahlilini amalga oshirish, qadimiy turkiy til leksikasining ba'zi qoldiqlarini, muayyan lahja va shevaga xos bo'lgan lug'aviy birliklarini aniqlash, uslublar leksikasi o'rtasidagi mushtaraklik hamda xususiylik, o'ziga xoslik tomonlarini kuzatish, ijodkorlarning so'z qo'llash mahoratini belgilash, matnlarni kompyuterda tekshirish, tadqiq etish, istalgan matnni avtomatik ishlov berish orqali leksikografik manbalarini yaratish masalasi esa o'z tadqiqini kutmoqda. [Urinbayeva D., 2022: 8]

Rus tilshunosligida XX asr boshlarida statistik metod yordamida tilning tovush tarkibini o'rganish ilk bor yo'lga qo'yildi. Masalan, rus tilshunoslari V.Petrov va Aleksandrovlar 1911-yilda Qozon shahrida fransuz va nemis tillarining tovush tarkibini statistik jihatdan o'rganishga bag'ishlangan maqolalarini e'lon qilgan edilar. [Yo'ldoshev M., Muhamedova S., Saparniyazova M., 2021: 157]

V.A. Bogoroditskiy, A.M.Peshkovskiy, A.Bulaxovskiy kabi olimlar 20-30-yillarda, rus tilidagi ilmiy va badiiy matnlarda unli hamda undosh tovushlarning qo'llanilish darajasini aniqlashga oid ishlarni amalga oshirdilar. Masalan, A.M. Peshkovskiy rus tili og'zaki matnida qo'llangan unli va ayrim undosh tovushlarning ohangdorligini I.I.Turgenevning «Sadaqa» hikoyasi matnidagi ohangdorlik bilan qiyoslagan edi. Uning ta'kidlashicha, I.I.Turgenev hikoyasi matnidagi ohangdorlik og'zaki matndan ancha yuqori. Chunki hikoya matnida shovqinli undoshlarga nisbatan jarangli va sonor undoshlar ancha ko'p qo'llangan. [Yo'ldoshev M., Muhamedova S., Saparniyazova M., 2021: 157]

1929-yilda Krasnodar shahrida V.Chistyakov va B.Kramorenkolarning «Tilshunoslikda statistik metodni qo'llash tajribasidan» nomli asari 350 nusxa nashr etilgan edi. Bu sobiq Ittifoqda lingvostatistika bo'yicha nashr etilgan dastlabki alohida risoladir. [Yo'ldoshev M., Muhamedova S., Saparniyazova M., 2021: 158]

Hozirgi axborotlar ko'lami jadal sur'atda kengayib borayotgan davr esa yangi ilmiy metodlar va texnologiyalarni qo'llashni taqozo etmoqda. Lingvostatistika bo'yicha tadqiqotlarning markazida chastotali lug'atlar yaratish turadi. [Urinbayeva D., 2022: 10]

Ivanov davlat universiteti professori Kirilov Mixail Anatoliyevich ilmiy dissertatsiya ishini "Badiiy matnda lingvostatistik tahlil"(F.S.Fitsjeraldning qisqa hikoyalari asosida) mavzusida himoya qildi. Ushbu tadqiqot ishida mashhur amerikalik yozuvchilardan biri F.S.Fitsjeraldning to'rtta sakson to'rt sahifali qisqa hikoyalarining chastota lug'atini yaratishga harakat qiladi. Bu lug'at tilshunos tadqiqotchilar uchun ham, ingliz tili lug'atini o'rganuvchi filolog talabalar uchun ham mo'ljallangan[Iskandarova, A.2009,2018,2019]. Har bir hikoyaning chastota ro'yxatlarini taqqoslash qiziqarli bo'lib tuyuladi, bu yozuvchi uslubining evolyutsiyasini kuzatish, matnning mavzusi va janriga qarab ma'lum so'zlardan foydalanishning bog'liqligini aniqlash imkonini beradi. Shu bilan birga, bu tadqiqot ishida nutqning o'zgaruvchan qismlarining(ism, fe'l, sifat, ergash gap) alifbo-chastota ro'yxatlari taqdim etildi, bundan maqsad matnshunoslik tadqiqotlarini osonlashtirish va qo'shimcha o'quv ma'lumotlarini taqdim etishdir. Mazkur ishning barcha amaliy natijalari talabalar, o'qituvchilar va adabiyotshunoslar tomonidan foydalanishga yo'naltirilgan magnit disketasiga yozib olingan. [Кириллов М.А.,2002]

Chastotali lug'atlar asosida esa statistik lingvistikaning alohida shaxobchasi – leksikostatistika shakllandi. [Фрумкина Р., 1973: 158].

Matnni matematik usullarda tekshirish Qozog'istonda ham izchil rivojlangan. Bunda prof.Q.B.Bektayevning xizmati katta deb aytilish mumkin. Q.B.Bektayev rahbarligidagi «Lingvostatistika va avtomatlashtirish» guruhining a'zolari M.Avezovning «Abay yo'li» romani (4 kitob) tilining chastotali lug'ati so'zligini kompyuter yordamida 50 soat davomida tuzib chiqqanlar. Mazkur chastotali lug'atda 20000 dan ziyodroq leksema va 60000 so'z shakli («glossema») mavjud bo'lib, u 466 000 marta qo'llanilgan. Guruh a'zolari bu lug'atni tuzish uchun atigi sakkiz oy vaqt sarflaganlar. Vaholanki, shu ish oddiy qo'l kuchi bilan bajarilganda, mazkur guruh a'zolari tinimsiz 10 yil, bir tilchi esa kamida 100 yil ishlagan bo'lar edi. Shu tariqa ushbu ulkan ish qozoq tilshunosligida kompyuter lingvistikasining rivojlanishiga asos bo'ldi. [Yo'ldoshev M., Muhamedova S., Saparniyazova M., 2021: 160]

Rus va o'zbek tillarining qiyosiy fonetikasi bo'yicha dastlabki statistik tadqiqot V.M.Popovning "O'zbek va rus tillarini qiyosiy o'rganishda statistik metodni qo'llash" nomli maqolasi (1949) sanaladi. Bu maqolada har ikki tildagi tovushlarning nutqdagi o'rni, ishlatilishi, miqdori, yumshoqlik va qattqlik xususiyatlarini aniqlash bo'yicha olib borilgan tekshirishlarning natijalari keltirilgan. [Yo'ldoshev M., Muhamedova S., Saparniyazova M., 2021: 160]

Tilshunos olim B.Bafoyevning Alisher Navoiy asarlari leksikasiga oid monografiyasining dastlabki bobi "Alisher Navoiy asarlari leksikasini statistik o'rganish natijalari" deb nomlanadi. Ishning "So'z boshi" qismida Alisher Navoiy asarlari tilini statistik jihatdan o'rganish bo'yicha amalga oshirilgan tadqiqotlar, e'lon qilingan lug'atlar, "Abushqa", "Sangloh", "Badoe ul-lug'at" singari bir qator qo'lyozma lug'atlardagi izohlangan so'zlar miqdori haqida qisqacha ma'lumot berilgan. [Yo'ldoshev M., Muhamedova S., Saparniyazova M., 2021: 161]

Sh.Egamova Alisher Navoiy asarlari tilidagi qadimgi turkiy leksikaning lingvostatistik tahlilini amalga oshirdi. Olib borilgan bu lingvostatistik tahlil natijasida Navoiy asarlari tilida tahminan 1400 ga yaqin qadimgi turkiy va eski turkiy tilga taalluqli so'zlarning qo'llanishda davom etganligi aniqlandi. Ot turkumiga oid leksemalar yuqorida keltirilgan 1400 ta leksik birlikning deyarli 25,1 % ini, ya'ni 583 tasini tashkil qiladi. Sifat kategoriyasiga oid 116 ta qadimgi turkiy leksema Navoiy asarlarida qo'llanishda davom etgan. Qadimgi turkiy tildagi son turkumiga oid so'zlarning 44 tasi Navoiy asarlarida ham qo'llangan. Statistik tahlil Navoiy asarlari so'z boyligida 33 ta olmoshlar keng qo'llanganligidan dalolat beradi. Navoiy asarlarida qo'llanishda davom etgan leksemalarning 557 tasi (24,0 %)ni fe'l kategoriyasiga oid so'zlar tashkil etadi, bu fe'llarning 183 tasi tub, qolgan 374 tasi esa yasama ekanligi aniqlandi. [Йўлдошев Б., 2008: 29].

O'zbek tili materiallari bo'yicha statistik tadqiqotlar olib borgan yana bir olimlar S.Rizayev va S.Muhamedovlardir. Mazkur olimlar ham asosan kompyuter lingvistikasining statistik tahlil yo'nalishi bo'yicha ishlarni amalga oshirishgan[Iskandarova, A.2009,2018,2019]. S.Muhamedov o'zining R.R.Piotrovskiy bilan hammualliflikda yozgan «Инженерная лингвистика и опыт системно-статистического исследования узбекских текстов» nomli kitobida lingvistik modellar, modellashtirish va uning umumiy tamoyillari haqida fikr yuritib, o'zbekcha matnlarning kvantitativ modellarini keltiradi. Mualliflar, shuningdek, mazkur asarda turk gazeta matnlarning leksikmorfologik mashina tarjimasini ham berganlar. [Yo'ldoshev M., Muhamedova S., Saparniyazova M., 2021: 163]

Samarqand viloyati pedagoglarni yangi metodikalarga o'rgatish milliy markazi Tillarni o'qitish metodikasi kafedrasini mudiri, professor Dilbar Bazarovna Urinbayeva "Xalq og'zaki ijodi

asarlari matning janriy-lisoniy va lingvostatistik tadqiqi” deb nomlangan monografiyasining obykti sifatida o‘zbek xalq og‘zaki ijodi asarlari – doston, maqol, topishmoq, ertak, qo‘shiq matnlari, jumladan, «Alpomish» va «Ravshan», «Topishmoqlar», «O‘zbek xalq maqollari», «O‘zbek xalq ertaklari», «Boychechak» asarlari tanlandi. Ushbu matnlar asosida tuzilgan «O‘zbek xalq dostonlari tilining alfavitli-chastotali lug‘ati» (2006), «O‘zbek xalq dostonlari tilining chastotali lug‘ati» (2006), «O‘zbek xalq dostonlarining chastotali izohli lug‘ati» (2008) singari leksikografik tadqiqotlar monografiyaning yozilishi jarayonida tayanch manba vazifasini o‘tadi. [Urinbayeva D., 2022: 8]

Ushbu monografida morfologik nuqtai nazardan asosiy o‘rinni quyidagi so‘z turkumlari egallagan: Ot – D: 7; M:21; T:10 ; E: 7 ; Q:9. Fe’l – D:19; M:10; T:7; E:12; Q:8. Sifat – M: 2; T:9; Q:6. Olmosh – D: 11; M:2; T:7; E:8; Q:7. Son – D:1; M:2; T:4; E: 2; Q:2. Ravish: D:1; E:2. Yordamchi so‘zlar turkumi - D:2; M:4; T:2 ; Q:8. Raqamlarga qaraganda ot so‘z turkumida maqol janri, fe’lda doston, sifatda topishmoq, olmoshda doston, sonda topishmoq, ravishda ertak, yordamchi so‘zlarda maqol dominant bo‘lib kelyapti. Hech bir janrda so‘z qo‘llash mutanosibligi mavjud emas. Ko‘pgina tilshunos olimlarni ot+sifat, ot+fe’l, ravish+fe’l, ot+ko‘makchi kabi birikmalar o‘zbek tili morfologiyasiga xosligi qiziqtiradi. Ammo raqamlar bunda hech bir yaqinlikni ko‘rsatmayapti.[Urinbayeva D., 2022: 58].

Professor D.B.Urinbayeva tahlili natijalariga ko‘ra, ertak, topishmoq janrlari barcha zonalarda nisbiy to‘plash chastotasining yuqori darajasini ko‘rsatgan bo‘lsa, maqol va qo‘shiq keyingi pog‘onani egallagan. Eng past ko‘rsatkich doston matniga to‘g‘ri kelgan. [Urinbayeva D., 2022: 69]

Bundan tashqari monografiyada ushbu janrlarda eng faol bo‘lgan tovushlar haqida statistik ma’lumot berilgan. Umumiy va 1000 so‘zli matnlardagi tovushlar qo‘llanishidan darak beruvchi harflarning qo‘llanish chastotasidagi farqlar juda katta emas. Masalan, a va I tovushlarining doston, maqol, topishmoq, ertak matnlarida uchrashi 13-14 % atrofida, ammo qo‘shiqda a 20 %ni, i esa 9%ni tashkil etadi. Yoki b va l tovushlarining qo‘llanishi 3,5 %dan 7%gacha bo‘lgan oraliqda joylashgan, ammo l undoshining qo‘shiqda qo‘llanilishi 14 %ni ko‘rsatmoqda. [Urinbayeva D., 2022: 75]

Qo‘shiq matnidagi alla (193), bolam (47), alla-yay (24), alla-yo (19) birliklari eng yuqori chastotali so‘zlar sifatida namoyon bo‘lib, lug‘at eng yuqori beshligida turibdi.[Urinbayeva D., 2022: 77]

Janrlardagi rang bildiruvchi sifatlar statistik qiyoslanishi natijasida chastotalar yig‘indisining eng yuqori ko‘rsatkichi topishmoqlarga to‘g‘ri kelib, boshqa janrlarga nisbatan 4 barobar ko‘p. Qo‘shiq topishmoqning mos keluvchi nisbati 2:1, qo‘shiq – maqolda 2:1, doston janrlarining mos keluvchi nisbati 1:3, qo‘shiq – ertak janrining mos keluvchi nisbati 4:1. [Urinbayeva D., 2022: 91]

Yuqoridagi statistik ma’lumotlar monografiyada formulalar, jadvallar, grafiklar ko‘rinishida yanada aniq ochib berilgan.

Statistik usul xususida A.Ibrohimov quyidagicha yozadi: «Statistik usulni tilshunoslikda qo‘llash shunchaki maqsad emas, balki tilning nihoyatda murakkab tuzilishi, uning amal qilish sirlarini bilishda an’anaviy usul bilan birgalikda bir vositadir [Iskandarova, A.2009,2018,2019]. Aniqlik ilmiy ishning talablaridan biridir. Lekin an’anaviy tilshunoslikning eng katta kamchiligi unda aniqlikning yetishmasligi, ko‘pchilik natijalarning taxminiy bo‘lishidir. Mana shuning uchun

ham tilshunoslar statistikaga asoslangan usulga murojaat qila boshladilar». [Иброҳимов А., 200: 4].

Bugungi kunda lingvostatistika fanning nazariy va metodologik muammolarini hal etuvchi metod sifatida shunday darajaga ko'tarildiki, uning zimmasiga nazariy xulosalarni dalillash, tushuncha va tadqiqot metodlarini aniqlash, o'rganilayotgan obyektning ko'zga ilg'anmas yangi qirralarini ochish, ularni tushuntirish yo'llarini topish kabi masalalar yuklatildi [Abdurakhmonova, N.2021]. Shundan so'ng kompyuter lingvistikasi yo'nalishlarining rivoji uchun lingvistik modellashtirish nazariyasi hamda tabiiy tilning ma'lumotlar bazasini yaratish masalasi qo'yildi va bir qator asarlarda aksini topdi. [Пиотровский Р.Г., Билан В.Н., Боркун М.Н., Бобков А.К.,1985.]

«Statistik usulning bir qator afzalliklari mavjud: matndagi leksikani miqdor jihatdan o'rganish natijasida matnlarni tavsif va tasnif etish, ayniqsa, muallifi noma'lum matnlarni aniqlash, tilning turli birliklari (fonema, morfema, so'zlar va h.k.)ning matnda uchrash chastotasi, ularning turli janrlar bo'yicha taqsimlanishi, boshqa til birliklari bilan birikishi kabi ma'lumotlar olish imkonini beradi». [Ибрагимов А.П., 2008: 4].

Qoraqalpoq, qozoq va o'zbekcha matnlarni lingvistik tadqiq etishga statistik metodni olib kirgan va bunda injener lingvistikasi yutuqlarini tatbiq etgan olimlardan yana biri M.Ayimbetov hisoblanadi. U dastlab nomzodlik dissertasiyasini himoya qildi va oradan ma'lum vaqt o'tgach alohida lug'atini, monografiyasini e'lon qildi. [Айимбетов М., 1991: 128].

Muhamedov gazeta-publitsistik matnlari ustidagi kuzatishlarini miqdoriy (statistik) metod asosida amalga oshirdi. U o'zbek tilshunosligida birinchi marta til va nutq munosabatini axborot-statistik usullar yordamida o'rganib, shu asosda hozirgi zamon o'zbek publisistik nutqi leksikasi va grammatikasining bir qator tipologik xususiyatlarini aniqladi. [Мухамедов С., 1980: 25.20]

Uning agglyutinatив so'zshakllari qurilishi va axborot uzatishi jihatidan hind-yevropa tillariga qaraganda o'zbek tilining farqli xususiyatlari, ma'lum o'zakdan yangi so'zshakllari yasash imkoniyatining cheksizligi baravarida bu imkoniyatning publisistik nutq uzusida nihoyatda cheklanib qolishi kabi tilni me'yorlashtirishga oid masalalarni ochib berdi [Abdurakhmonova, N.2021] Tadqiqot manbai vazifasini o'tagan til materialini keyinchalik lug'at shaklida nashr etildi. Lug'at lingvostatistik metod asosida tuzilgan va unda lug'at birligi sifatida leksemalar qabul qilinib, 200000 dan ortiq so'z qo'llanishi olingan. Mazkur lug'at turkcha matnlarning kvantativ modeli tamoyillari asosida yaratilgan. [Мухамедов С., 1982: 110; Мухамедов С., 1986: 141-142].

S.Muhamedova o'zbek tili leksikasidagi semantik makromaydon birliklari – harakat fe'llari asosida kompyuter dasturlari uchun lingvistik ta'minot yaratish namunalarini ishlab chiqdi. [Мухамедова С., 2006: 80].

Sh.S.Imyaminova nomzodlik dissertasiyasida ko'p bo'g'inli so'zlarda bo'g'in tiplarini aniqlash, muayyan me'yorlarini belgilash va struktural tadqiq qilish maqsadida statistik usuldan foydalangan.[Urinbayeva D., 2022: 20]

S.Karimov rahbarligida A.Qarshiyev va G.Isroilovalar hammuallifligida xalq yozuvchisi A.Qahhor asarlari matnini ilk kompyuter dasturlari yordamida 3 tomli (chastotali, alfavitli, ters) lug'ati tartibga keltirilgan va unda adib badiiy-publitsistik merosining leksik-semantik arsenali to'liq aks ettirilgan. [Каримов С., Қаршиев А., Исроилова Г., 2007: 434; Шу муаллифлар., 2007: 420; Шу муаллифлар., 2007: 420].

A.Ibragimovning «Bobur «Devon»i tilidagi o‘zlashma leksika» deb nomlangan asari «Devon»dagi faqat o‘zlashma so‘zlarni qamrab olgan bo‘lib, unda, birinchidan, lug‘aviy birliklar misolida ishonarli tahlilarni bajargan, ikkinchidan, bu tipdagi lug‘atlar yordamida amalga oshirilishi mumkin bo‘lgan tadqiqotlar metodikasini ham bergan. [Ибрагимов А., 2006].

Ushbu ishning davomi sifatida A.Ibragimov doktorlik dissertasiyasini yoqladi. Tadqiqotda Bobur asarlari leksikasidagi turkiycha, arabcha, forsha, hindcha so‘zlarning genetik asosini tekshirib, ularni statistik, semantik, genetik jihatdan tahlil qilgan va jami 14784 ta so‘zshakli (takrorlari bilan 70463) ishlatilganligini aniqlagan. [Ибрагимов А.П., 2008].

N.Abdurahmonova nomzodlik dissertasiyasida «hozirga qadar o‘zbek kompyuter lingvistikasida mashina tarjimasiga oid tadqiqotlarning yetarlicha emasligi sababli o‘zbek tilidan xorijiy tilga va aksincha tarjima qiluvchi dastur uchun o‘zbek tilining lingvistik ma’lumotlar bazasini yaratish zarur» ekanligini ta’kidlagan. [Абдурахмонова Н., 2018: 4].

U tadqiqotida o‘zbek tilshunosligida birinchi marta mashina tarjimasining lingvistik ta’minotini yaratish bo‘yicha ilmiy xulosalarini keltirgan. Lingvistik ta’minotda ingliz va o‘zbek tillaridagi leksemalarning derivativ, fe’lli frazemalar, so‘z birikmalarning semantik - mantiqiy imkoniyatlari va ularning algoritmik tavsifi, lingvistik modelini yaratgan. Matnni lingvistik tahlil qilishda leksiksemantik hodisalar, grammatik kategoriyalar va ularning paradigmatic munosabatlari modellashtirilgan. Muallif yozadi: «o‘zbek va ingliz tillarining formal grammatikasi asosida tarjimon dastur uchun lingvistik ta’minot, tarjima algoritmi, gap strukturalarning matematik modellarini yaratish muhim ahamiyat kasb etadi».

Badiiy matnning lingvostatistik tahlili bo‘yicha bir qancha maqolalar va tezislar chop etilmoqda. Jumladan, No‘monova Dilafruz Mohirjon qizi va Qo‘ziyev Umidjon Yandashaliyevichlarning hammualifligidagi “Navodir un-nihoya” asarining lingvostatistik tahlili” maqolasida tarixiy, adabiy, mifologik shaxs nomlari va geografik, etnik nomlar hususida so‘z borgan. Bunday nomlar asarda 112 tani tashkil etadi. Bu nomlarning 54,4 %i tarixiy, adabiy, mifologik shaxs nomlariga tegishli. Geografik va etnik nomlar esa 45,5%ni tashkil etadi, ularning absolut chastotasi esa 51tadir. Tarixiy, adabiy, mifologik shaxs nomlarining nisbiy va absolut chastotalarning ko‘rsatkichiga ko‘ra Majnun nomi, geografik va etnik nomlarning nisbiy va absolut chastotalarining ko‘rsatkichiga ko‘ra esa Ka’ba nomi eng faol chastotaga ega bo‘lgan birlik ekanligi ma’lum bo‘ldi. [No‘monova D.M., Qo‘ziyev U.Y., 2020: 161-165]

Yuqoridagi mualliflar nomi ostida “Badiiy matnni lingvostatistik tomondan tahlil qilish” mavzusida navbatdagi maqola chop etilgan. Bu maqolada Erkin Vohidov va Abdulla Oripov asarlari lingvostatistik metod yordamida tahlil qilinadi. Asarda qo‘llanilgan yordamchi so‘z birliklari, egalik va kelishik kategoriyalari foiz miqdorida hisoblab, raqamlar orqali fakt keltirib o‘tiladi. [No‘monova D.M., Qo‘ziyev U.Y., 2020: 119-121]

Demak, o‘zbek tilshunosligida ham keyingi yillarda statistik metodlardan foydalanish sohasida muayyan natijalarga erishilgan. Bugungi kunda esa lingvostatistik jihatdan o‘rganiladigan, tadqiq etiladigan noyob asarlarimiz bisyor.

Foydalanilgan adabiyotlar:

1. Urinbayeva D.B “Xalq og‘zaki ijodi asarlari matnining janriy-lisoniy va lingvostatistik tadqiqi” “Qamar media” nashriyoti Toshkent – 2022.
2. M.Yo‘ldoshev, S.Muhammedova, M.Saparniyozova “Matn lingvistikasi” Toshkent «Ishonchli hamkor» 2021

3. Абдурахмонова Н. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти: Филол. фан. фалсафа фан. д-ри (phd). дисс. – Тошкент, 2018. –Б. 4.
4. Айимбетов М. Опыт лингво-статистического анализа лексики и морфологии каракалпакского публицистического текста. – Нукус: Билим, 1991. – С.128.
5. Фрумкина Р. Роль статистических методов в современных лингвистических исследованиях // Математическая лингвистика. М.: Наука, 1973. - С.158.
6. Иброҳимов А. Бобур «Девон»и тилидаги ўзлашма лексика. –Тошкент: ТДШИ, 2006, 4-бет
7. Mengliev, D., Barakhnin, V., & Abdurakhmonova, N. (2021). Development of intellectual web system for morph analyzing of uzbek words. *Applied Sciences*, 11(19), 9117.
8. Ибрагимов А.П. Бобур асарлари лексикасининг лингвостатистик, семантик ва генетик тадқиқи ("Девон", "Мубаййин", "Аруз"): Филол. фан. д-ри ... дисс. автореф. – Тошкент, 2008. Б. 4.
9. Iskandarova, A. (2018). Vatanni olib ketib bo'lmaydi.
10. Iskandarova, A. (2009). A.Navoiyning "Lisonut-tayr" dostonidagi vodiylar ta'rifi va mohiyati haqida. Мумтоз суз таълими.
11. Iskandarova, A. (2019). INSONIYATNI SAODATGA YETAKLAGAN ADIB (Adabiyot darslarida Chingiz Aytmatov ijodini o'rganishga doir). Tafakkur Chashmasi MCHJ.
12. Кириллов М.А. Лингвостатистический анализ художественного текста (На материале коротких Ф.С. Фицджеральда) 2002
13. Каримов С., Қаршиев А., Исроилова Г., 2007: 434; Шу муаллифлар., 2007: 420; Шу муаллифлар., 2007: 420].
14. Мухамедов С. Статистический анализ лексико-морфологической структуры узбекских газетных текстов: Автореф. дис.... канд. филол. наук. -Ташкент, 1980. – С.25.20
15. Мухамедов С. Ўзбек тилининг алфавитли-частотали луғати. – Тошкент: Фан, 1982. – Б.110.2 Мухамедов С., Пиотровский Р. Инженерная лингвистика и опыт системно-статистического исследования узбекских текстов. – Ташкент: Фан, 1986. – С. 141-142.
16. Мухамедова С. Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. Методик қўлланма. - Тошкент, 2006. –Б.80.
17. Каримов С., Қаршиев А., Исроилова Г. Абдулла Қаҳҳор асарлари тилининг луғати. Алфавитли луғат. - Тошкент: «Ўзбекистон миллий энциклопедияси» Давлат илмий нашриёти, 2007. – Б.434; Шу муаллифлар. Абдулла Қаҳҳор асарлари тилининг луғати. Частотали луғат. - Тошкент: «Ўзбекистон миллий энциклопедияси» Давлат илмий нашриёти, 2007. – Б.420; Шу муаллифлар. Абдулла Қаҳҳор асарлари тилининг луғати. Терс луғат. - Тошкент: «Ўзбекистон миллий энциклопедияси» Давлат илмий нашриёти, 2007. –Б.420
18. Пиотровский Р.Г., Билан В.Н., Боркун М.Н., Бобков А.К. Методы автоматического анализа и синтеза текста. – Минск,1985
19. Йўлдошев Б. Матнни ўрганишнинг лингвостатистик методлари. –Самарқанд, 2008. –Б. 29

20. Kuvondikova G. I., & To‘rayeva, S. (2024). O‘zbek tilidagi ko‘chma ma‘noli so‘zlarni fransuz tiliga tarjima qilish xususiyatlari. *Educational Research in Universal Sciences*, 3(2), 254-257.
21. Abdunazarova M. “Genderolingvistik tahlilga dior” *Образование наука и инновационные идеи в мире*” elektron jurnalining 2022-yil 11-soni 108-114-betlar
22. Sharipbay, A., Shirinova, R., Abdurakhmonova, N., & Qodirova, Z. (2023). Ontological knowledge-based models representing medicine terminology. *Advances in Social Sciences and Economics*, 1(1), 4.