

KOMPYUTER VA KORPUS LINGVISTIKASI TALABLARI ASOSIDA SO'Z TURKUMLARI VA GRAMMATIK KATEGORIYALAR TAVSIFIGA YONDASHUV

Salomova Mahliyo Baxtiyorovna

Termiz davlat universiteti II bosqich magistranti

salomovamahliyo@gmail.com

Eshkulova Nargiza Normo'minovna

Termiz davlat universiteti II bosqich magistranti

eshkulovanargiza8@gmail.com

Annotatsiya. Ushbu maqolada, kompyuter va korpus lingvistikasi talablari asosida so'z turkumlari va grammatik kategoriyalar tavsifiga yangicha yondashuv, o'zbek tili grammatik kategoriyalari hamda so'z turkumlari tizimidagi ayrim o'ziga xos otlarga xos qarashlilik-tegishlilik - posesivlik kategoriyasi mavjudligi masalasi, o'zbek tili kelishiklar tizimi masalasi, ularning soni muammosini hal qilish, yuklama va modal so'zlar munosabatiga aniqlik kiritish kabi holatlarga aniqlik kiritilgan.

Kalit so'zlar: kompyuter lingvistikasi, korpus lingvistikasi, so'z turkumlari, grammatik kategoriyalar, posesivlik kategoriyasi, konkordans, til korpusi.

Annotation. In this article, based on the requirements of computer and corpus linguistics, a new approach to the description of word groups and grammatical categories, the issue of the existence of a category of possessiveness-belonging-possession specific to some specific nouns in the system of grammatical categories and word groups of the Uzbek language, is presented. Issues such as the issue of the Uzbek language agreement system, solving the problem of their number, clarifying the relationship between loading and modal words have been clarified.

Key words: computational linguistics, corpus linguistics, word groups, grammatical categories, possessive category, concordance, language corpus.

Аннотация. В данной статье, исходя из требований компьютерной и корпусной лингвистики, по-новому подходит к описанию словосочетаний и грамматических категорий, ставится вопрос о существовании категории притяжательности-принадлежности-владения, специфической для некоторых конкретных существительных в системе. Представлены грамматические категории и словосочетания узбекского языка, выяснены такие вопросы, как вопрос о договорной системе узбекского языка, решение проблемы их количества, уточнение соотношения между загрузочными и модальными словами.

Ключевые слова: компьютерная лингвистика, корпусная лингвистика, группа слов, грамматические категории, притяжательные категории, конкорданс, лингвистический корпус.

Bugungi kunda jamiyatning barcha sohalarini shiddat bilan qamrab olayotgan

axborot texnologiyalari jarayoni dunyo miqyosida tilshunoslikning ham zimmasiga yangi-yangi vazifalarni yuklamoqda. Shu sababdan tildan amaliy foydalanish samaradorligi fanning asosiy vazifasiga aylandi. Buning natijasida til va shaxs umumiyligi masalasi bilan shug'ullanadigan amaliy yo'nalishlar muhim ahamiyatga ega bo'lmoqda. Filologiya va tillarni o'qitish sohalarida vakillarining zimmasida tildan oqilona, samarali foydalanish natijadorligini oshirish va filologik bazani vujudga keltirish kabi masalalar bilan shug'ullanish mutlaqo kechiktirib bo'lmaydigan dolzarb vazifa sifatida ko'rilishi lozim.

Demak, amaliy filologik faoliyatni yo'lga qo'yish uchun amaliy filologiya bakalavri va amaliy filologiya magistri mutaxassislarini tayyorlash zaruriyatini keltirib chiqaradi. Bundan tashqari, amaliy filologiya bo'yicha olib boriladigan tadqiqotlar bu yo'nalishda zamonaviy ruh singdirilgan bo'lishni talab qiladi. Jonkuyar olimlarimizdan biri professor B.Mengliyev "Til egalarining lisoniy imkoniyatlaridan pragmatik foydalanish samaradorligi masalasi bilan shug'ullanadigan o'zbek amaliy tilshunosligining milliy xarakterdan kelib chiqqan holda sosiolingvistika, etnolingvistika, psixolingvistika, pragmalolingvistika, lingvokognitologiya, lingvokulturologiya, neyrolingvistika, kompyuter lingvistikasi, sud lingvistikasi kabi zamonaviy fan yo'nalishlarining rivojlanishi ijtimoiy taraqqiyot talablari bilan uyg'unlashishini, ayniqsa, axborot texnologiyalari va globallashuv asrida o'zbek tilining global til korpuslariga qo'shiluvini ta'minlash tarixiy zarurat sifatida namoyon bo'layotganini aytadi.

O'zbek tilining milliy korpusini yaratish bugungi kunning dolzarb vazifasi. Chunki busiz o'zbek tilshunosligining kelajagini ko'rib bo'lmaslik kun sayin ayon bo'lib bormoqda. Bunda jamiyatdagi barcha sohalarning jadal sur'atda o'sib borayotganligini inobatga oladigan bo'lsak, shular jumlasiga, o'zbek tilining zamonaviy axborot-kommunikasiya tizimida keng qo'llanishiga erishish kechiktirib bo'lmaydigan, dolzarb vazifa ekanligini alohida ta'kidlash mumkin. Bugungi kunda mislsiz rivojlanib borayotgan axborot-texnologiya jarayoni ham tilga funksional imkoniyatlaridan foydalanish borasida o'z eshiklarini ochib bermoqda. Ayniqsa, tilning imkoniyatlarini namoyon qilish va egallash borasida til korpusining roli beqiyos ekanligini ta'kidlash mumkin.

Yuqorida sanab o'tilgan holatlardan aksariyat qismlari, bugungi kun kompyuter dasturlari ichida faol ishlab kelmoqda. Lekin kompyuterga ovoz orqali kiritish amalga oshirilgani yo'q. O'zbek tiliga o'qitadigan va bilimlarni baholaydigan kompyuter dasturi deganda, faqat test asosida baholash, ya'ni ko'rsatilganlardan ma'lum miqdorini belgilash ishlatilayotgan bo'lsa, chizgan surat, matematik formuladagi xatolik yoki matnda yo'l qo'yilgan orfografik, uslubiy hamda boshqa xatoliklarni aniqdovchi dasturni mukammal qilib yaratish oldimizda turgan asosiy vazifalardan biri. To'g'ri, bu vazifalarni bajarish juda qiyinday tuyiladi, lekin bugungi kunda o'zbekcha ishlaydigan "tslo/z" operasion tizimi yaratilgan, biroq ishlash darajasi 100%ni tashkil etmasa-da, demak, bu vazifalarni uddalasa bo'ladi. Tarjimon dasturlarni ham ko'rishimiz mumkin, lekin mukammal ishlamasligiga guvoh bo'lamiz. O'zbekcha matnlarni tahrirlovchi

kompyuter dasturlarni yaratish esa, oliy maqsad bo'lishi shart.

Korpus lingvistikasi ham jahon kompyuter lingvistikasining juda tez rivojlanib borayotgan sohasi, bu borada sezilarli darajada yutuqlarga erishilmoqda. Rivojlangan davlatlardagi oliy ta'lim muassasalarida korpus lingvistikasi fan sifatida o'qitib kelinmoqda. Demak, bu sohaning predmeti korpus yaratish nazariyasi va amaliyoti bo'lsa, fan sifatida korpusning o'ziga xosligi, dasturlash asoslari kabi jihatlari o'qitilishi ko'zda tutiladi. Korpus lingvistikasi kompyuter lingvistikasining tarkibiy qismi, til korpusini yaratish, kompyuter texnologiyasi yordamida ulardan foydalanishning umumiy nazariyasi va amaliyoti bilan shug'ullanadi.¹ Korpus lingvistikasining predmeti - til korpusi. U ingliz tilida Nidshze yeogriz yoki !yex! yeogriz kabi atamalar bilan qo'llanib, ilmiy adabiyotlarda turlicha ta'riflangan.² "Korpus -ma'lum maqsadda yig'ilgan matnlar majmuini tashkil etuvchi til birliklarining yig'indisi", "korpus - alohida tilning yozib olingan og'zaki hamda yozma matnlari asosidagi lingvistik ma'lumotlar majmui", "korpus-istalgan tabiiy (real) tildagi elektron shaklda saqlanadigan yozma yoki og'zaki, kompyuterlashtirilgan qidiruv tizimiga joylashtirilgan matnlar yig'indisi kabi. Demak, til korpusi³ ma'lum tilning belgilangan davrdagi, xilma-xil janr, rang-barang uslub, hududiy holda ijtimoiy

variantdagi matnlarning elektron shakldagi maxsus dasturiy ta'limot asosidagi yig'indisidir.

Til korpusi bugun jahonda tez taraqqiy etayotgan soha - korpus lingvistikasi mutaxassisleri tomonidan ilmiy tadqiqotlar olib borish, til o'rgatish maqsadida yaratilmoqda. Jahon tillarining juda ko'pchiligi mukammal darajali, matnning ilmiy, qayta ishlash imkoniyati bilan farq qiluvchi milliy korpusiga ega. Zero, "bugungi kunda lingvistik tadqiqot va amaliy topshiriqlar yechimi uchun til korpuslari zamonaviy tilshunoslikning inkor etib bo'lmas ish quroliga aylandi, chunki korpus turli lingvistik topshiriqlarni yechishga xizmat qiladi.

Qidiruv natijasi, odatda, konkordans - qidirilayotgan birlik leksik bog'lanish holatida manbaga ilova qilingan kontekstlar ro'yxati shaklida namoyon bo'ladi, korpus menejerining qidiruv imkoniyati aniq so'zshakli, so'zning lemma (korpusda leksema shunday aytiladi) shakli, birikmali bo'lingan va bo'linmagan sintagma, morfologik belgilarga asoslangan jamlanma shaklini qamrab oladi. Talab qilingan hajmdagi kontekst ko'rsatkichlari qidiruv natijasi, saralangan materialni alohida faylda saqlash imkonini beradi⁴.

Korpus yordamida bajariladigan amalning samaradorligini oshirish maqsadida korpus menenjeriga quyidagi umumiy talablar qo'yiladi:⁵

- 1) to'liq konkordans ro'yxatini tuzish;
- 2) nafaqat alohida so'zni, balki murakkab birlik qidiruvini amalga oshirish;
- 3) chiqarilgan ro'yxat (qidiruv natijasini) bir necha imkon asosida saralash;

¹ Захаров В.П. Корпусная лингвистика // Учебно-методическое пособие. - Санкт-Петербург: СПбГУ, 2005. -

² Кутузов А.Б. Корпусная лингвистика. - Москва, 2008. - С. 26.

³ Национальный корпус русского языка // Вопросы языкознания. - Москва: Наука, 2006. - С. 149-155.

⁴scoyroua.ru, Плунгян В. Зачем мы делаем корпус?

⁵ Кутузов А.Б. Курс. Корпусная лингвистика. - Москва, 2008. - С. 26.

- 4) foydalanuvchiga qidirilgan so'z shaklini kengaytirilgan ko'rinishda taqdim etish;
- 5) korpusning har bir elementi haqida sintaktik axborot berish;
- 6) lemmalarni aks ettirish;
- 7) so'zshakllarning morfologik xususiyati va qo'shimcha axborotini (bibliografik, tipologik) berish (bu parametr korpusga berilgan qo'shimcha izohga bog'liq);
- 8) natijalarni saqlash va nashr etish uchun alohida faylda ishlay olish;
- 9) nafaqat korpus, balki alohida fayllar bilan ham yuqori tezlikda ishlay olish;
- 10) matnning barcha turdagi formatini (1x1, glG. s1s, 11sh1, xt1) o'zida saqlay olish;
- 11) tajribali va yana foydalanuvchilar uchun birdek yengil va intuitiv tushunarli bo'lish.

Til korpusining yana bir muhim jihati - korpus annotasiyasi. Bunga Sh.Hamroyevaning monografiyasida shunday ta'rif berilgan: "turli lingvistik topshiriqdar amalini bajarish uchun matnga lingvistik va qo'shimcha ma'lumot bilan ishlov berilgan bo'lishi lozim. Buning uchun matn ekstralingvistik annotasiyasi mavjud matn, uning komponentlariga maxsus izoh (masalan, so'z turkumi haqida ma'lumot) berilishi kerak. Bu izoh matn tegi (belgi ishora) yoki annotasiya deb ataladi.⁶

Bundan shunday xulosa chiqadiki, annotasiya mukammal, to'liq bo'lishi korpus ahamiyatini belgilovchi omil. Sh.Hamroyevaning yuqorida keltirilgan monografiyasida yana bir xususiyat - reprezentativlik to'g'risida so'z boradi. Bu atama ostida zaruriy hajm, miqdor, janr, uslub, mualliflarga tegishli matnlar proporsiyasining to'g'ri tanlanishi, tushunilishini ta'kidlagan. Bu xususiyat korpus tuzish jarayonida ham, undan foydalanish jarayonida ham saqlab turish korpusining ahamiyatini ta'minlovchi asosiy omil. Demak, korpus turli mavzudagi gazeta, jurnal materiallari, fanning barcha sohasiga tegishli ilmiy, ilmiy-ommaviy asarlari, reklama, shaxsiy yozishma matnlari hamda badiiy asarlaridan tarkib topgan bo'lishi lozim.

Shuni alohida ta'kidlash lozimki, korpus kompyuter asrigacha ham shakllanib ulgurgan edi. Lekin keyingi davrda tadqiqning korpus metodidan foydalanishga qiziqishning kuchayishini matnlarning katta massasini qayta ishlashga to'g'ri kelganligi bois, kompyuter texnologiyalarining kirib kelishiga ehtiyoj tug'ildi. Aksariyat olimlar tilni tadqiq etishda introspeksiya (o'z-o'zini kuzatish), tabiiy faktga tayanish usuli doim ham aynan bir xil natija beravermaydi, degan fikrni yoqlashadi. Mashhur korpus lingvistikasi mutaxassisi Toni Mak Enneri, Endryu Uilson ham empirik fakt, introspeksidan ham tabiiy, sun'iy faktdan foydalanish lozimligini uqtiradi.⁷ Demak, korpus lingvistikasi korpus shaklida bo'lmagan

⁶ Хамраева Ш. Ўзбек тили морфологик анализатерининг лингвистик таълимоти - Тошкент: С1о6ЕЕбй, 2020.

⁷ Кутузов А.Б. Курс "Корпусная лингвистика. Лицензия Сгеайуе еоштопз АИпьиЮп кбаге-Аьке 3.0 Ипрройеб (электрон ресурс) - //1аb314.Бг8и.Бу/ктр-1Бе/ктр-у1Бео/Сь/ СогрогеБтдуа.рШ.

nutqiy ma'lumotning muhimligi, ahamiyatini ham rad etmaydi. Korpusni tuzish, o'rganish korpus lingvistikasi sohasi rivojlanishidan oldin XVII asrlarda boshlangan, bunga Bibliyaga oid tadqiqotlar, lug'atlar, tillarni o'qitish va Kvirik korpusini misol sifatida keltirish mumkin.⁸ Faqatgina kompyuter asrigacha bo'lgan davrida noelektron shaklda yuzaga kelganligidan dalolat beradi. Bunga misol sifatida Panini tomonidan tuzilgan grammatika, Injil matnlarining tadqiq etilganligi to'g'risidagi ma'lumotlarni misol keltirish mumkin.

Keyingi asrda ham noelektron shaklda yaratilgan korpuslarga ehtiyoj borligi sezilib turdi. Chunki yangi ko'rinishdagi korpuslarning yuzaga kelishi uchun noelektron korpuslar asos bo'ldi, bundan tashqari, korpusni shakllantirish bazasi sifatida muhim manba vazifasini bajaradi.

Foydalanilgan adabiyotlar ro'yhati

1. Хамраева Ш. Ўзбек тили морфологик анализатерининг лингвистик таълимоти -Тошкент:2020
2. Захаров В.П. Корпусная лингвистика // Учебно-методическое пособие. - Санкт-Петербург: СПбГУ, 2005. -
3. Кутузов А.Б. Корпусная лингвистика. - Москва, 2008. - С. 26.
4. Национальный корпус русского языка // Вопросы языкознания. - Москва: Наука, 2006. - С. 149-

⁸ Хамраева Ш. Ўзбек тили морфологик анализатерининг лингвистик таълимоти -Тошкент: С1оьеЕйй, 2020. - Б.32