O‘ZBEKISTON MILLIY UNIVERSITETI XABARLARI, 2026, [14]
ISSN 2181-7324(Print)

FILOLOGIYA
http://journals.nuu.uz
Social sciences
ISSN 3093-8996(Onlayn)

UDK: 81’367.6(=512.133):81’322

Ikramjon ABDULLAYEV,
Tayanch doktorant, Andijon davlat chet tillari instituti, Andijon, O‘zbekiston
E-mail: ikramjona@inbox.ru, ORCID: 0009-0007-9353-1036

Filologiya fanlari doktori, dotsent J.Egamberdiyev taqrizi asosida

QUANTITATIVE DESCRIPTION OF CROSS-CATEGORY SHIFT IN UZBEK: BOUNDARY UNITS AND DIRECTIONS OF TRANSPOSITION

Annotation

Cross-category shift has long been described in Uzbek linguistics, but it has never been counted, because counting requires a procedure that yields a single decision for every token. The article presents such a procedure: a three-valued flag system marking each unit as typical, boundary or transpositional. The system was applied to a balanced, manually annotated microcorpus of 12,048 tokens from three functional styles. Boundary units account for 2,333 tokens and transpositional units for 575, together covering nearly a quarter of the material. The distributions differ: 78.6 per cent of boundary units fall on the verb, whereas 84.7 per cent of transpositional units originate from the noun.

Key words: Parts of speech, transposition, boundary units, corpus linguistics, operational flag, tagging, adjectivalisation, non-finite forms, functional style, microcorpus.

КОЛИЧЕСТВЕННОЕ ОПИСАНИЕ МЕЖКАТЕГОРИАЛЬНОГО ПЕРЕХОДА В УЗБЕКСКОМ ЯЗЫКЕ: ПОГРАНИЧНЫЕ ЕДИНИЦЫ И НАПРАВЛЕНИЯ ТРАНСПОЗИЦИИ

Аннотация

Межкатегориальный переход давно описан в узбекском языкознании, однако он не был измерен, поскольку измерение требует процедуры, дающей однозначное решение для каждого токена. В статье предлагается такая процедура – система трёхзначных флагов, при которой каждая единица корпуса помечается как типичная, пограничная или транспозиционная. Система применена к сбалансированному микрокорпусу объёмом 12 048 токенов, размеченному вручную по трём функциональным стилям. Пограничные единицы составляют 2 333 токена, транспозиционные – 575; вместе они охватывают почти четверть материала. Распределения различны: 78,6 процента пограничных единиц приходится на глагол, тогда как 84,7 процента транспозиционных имеют именную основу.

Ключевые слова: Части речи, транспозиция, пограничные единицы, корпусная лингвистика, операциональный флаг, разметка, адъективация, нефинитные формы, функциональный стиль, микрокорпус.

O‘ZBEK TILIDA TURKUMLARARO KO‘CHISHNING MIQDORIY TAVSIFI: CHEGARAVIY BIRLIKLAR VA TRANSPOZITSIYA YO‘NALISHLARI

Annotatsiya

Turkumlararo ko‘chish o‘zbek tilshunosligida ancha oldin tavsiflangan, biroq u hech qachon sanalmagan edi, chunki sanash uchun har bir tokenga bir qiymatli javob beradigan tartib zarur. Maqolada shunday tartib taklif etiladi: uch qiymatli bayroq tizimi, unda korpusning har bir birligi tipik, chegaraviy yoki transpozitsion sifatida belgilanadi. Tizim uch funksional uslubdan olingan, qo‘lda teglangan 12 048 tokenli muvozanatlashgan mikrokorpusda qo‘llandi. Chegaraviy birliklar 2 333 tokenni, transpozitsion birliklar 575 tokenni tashkil etadi; ular birgalikda materialning deyarli chorak qismini qamraydi. Taqsimot bir xil emas: chegaraviy birliklarning 78,6 foizi fe’lga, transpozitsion birliklarning 84,7 foizi ot asosiga to‘g‘ri keladi.

Kalit so‘zlar: So‘z turkumlari, transpozitsiya, chegaraviy birliklar, korpus lingvistikasi, operatsional bayroq, teglash, adyektivatsiya, nofinit shakllar, funksional uslub, mikrokorpus.

Kirish. So‘z turkumlari o‘rtasidagi chegara o‘zbek tilshunosligida uzoq muhokama qilingan. Muammoning mohiyati shundaki, tasnif mezonlari – ma’no, shakl va vazifa – barcha birlikda bir xil javob bermaydi: harakat nomi fe’l asosidan yasaladi, ammo ot vazifasini bajaradi; ot esa qo‘shimchasiz aniqlovchi bo‘lib kela oladi. Bunday holatlar nazariy tavsifda «oraliq hodisa» deb qayd etiladi va shu bilan cheklanadi.

Korpus tadqiqotida bunday cheklanish mumkin emas: annotator har bir tokenga bitta teg qo‘yishi shart, tokenlar soni esa o‘n minglab. Shu bois ko‘chishni tavsiflash yetarli emas – uni o‘lchash tartibi kerak. O‘zbek tili materialida bunday o‘lchov hozirgacha o‘tkazilmagan: chegaraviylik va transpozitsiya sifat jihatidan yaxshi tavsiflangan, ammo korpusdagi hajmi va yo‘nalishlari son bilan berilmagan.

Masalaning amaliy tomoni davlat siyosati darajasida ham qo‘yilgan. Davlat tilining nufuzini oshirish to‘g‘risidagi Farmonda o‘zbek tilining axborot texnologiyalari sohasidagi imkoniyatlarini kengaytirish vazifasi belgilangan [1], til siyosatini takomillashtirish haqidagi Farmonda esa milliy korpus va elektron resurslarni yaratish alohida yo‘nalish sifatida ko‘rsatilgan [2]. Sun’iy intellekt strategiyasida til modellari uchun sifatli, teglangan ma’lumot bazasi zarurligi ta’kidlanadi [3]. Chegaraviy holatlar esa teglangan korpusda eng ko‘p xatoga sabab bo‘ladigan zona, ya’ni ularni o‘lchash amaliy ehtiyoj.

Mazkur maqolaning maqsadi – turkumlararo ko‘chishni operatsional bayroqlar tizimi orqali miqdoriy o‘lchash va uning funksional yo‘nalishlarini aniqlash. Tadqiqot uch savolga javob izlaydi: ko‘chish hodisasi korpusning qancha qismini qamraydi; u qaysi turkumlarda to‘planadi; qaysi funksional yo‘nalish ustun turadi.

Mavzuga oid adabiyotlar tahlili. O‘zbek tilshunosligida turkumlararo ko‘chish alohida tadqiqot obyektiga J.Eltazarovning ishida aylangan. Muallif ko‘chishni turkumlar tizimining ichki xususiyati sifatida qaraydi va uning bosqichlarini ajratadi [4]. Z.Isaqov esa masalaga boshqa tomondan yondashib, turkumlarni belgilarni ifodalash darajasiga ko‘ra o‘zaro ziddiyatda turuvchi guruhlar deb talqin qiladi [5]. Ikki qarash bir-birini to‘ldiradi: birinchisi ko‘chish jarayonini, ikkinchisi uning tizimdagi sababini tushuntiradi. Ammo ikkalasida ham hodisaning hajmi o‘lchanmagan.

Nazariy grammatikada bu masala tizim tushunchasi orqali qo‘yiladi. A.Nurmonov va hammualliflari turkumni yakka birlik

emas, munosabatlar to‘plami sifatida ta’riflaydi [6]; bunday talqinda oraliq holat istisno emas, tizimning tabiiy qismi bo‘lib chiqadi. Sh.Shoabdurahmonov va hammualliflarining akademik grammatikasida turkum uch mezonning uyg‘unligida belgilanadi [7], ya’ni mezonlar to‘qnashgan holat alohida qaror talab qiladi. Bizningcha, aynan shu qaror korpus tadqiqotida rasmiylashtirilishi kerak – agar u annotator sezgisiga qoldirilsa, natija takrorlanmaydi.

Xorijiy tilshunoslikda oraliq holatlar gradiyentlik tushunchasi doirasida o‘rganiladi. R.Kvirk va hammualliflari turkumlarni ochiq va yopiq sinflarga ajratib, ular orasidagi chegara keskin emasligini ko‘rsatgan [8]. V.Kroft bundan uzoqroqqa borib, kategoriyalarni til-xususiy qurilmalar sifatida talqin qiladi [9]. Ushbu fikrga qisman qo‘shilish mumkin: kategoriyaning til-xususiyligi o‘zbek materialida ham ko‘rinadi, ammo undan tasnifdan voz kechish emas, tasnifni har til uchun operatsional shaklga keltirish zarurati kelib chiqadi.

So‘nggi yillarda o‘zbek tili uchun korpus resurslari faollashdi. N.Abduraxmonovaning qo‘llanmasida korpus tuzilishi va annotatsiya qatlamlari tavsiflangan [10]. M.Sharipov va hammualliflari ko‘p sohali teglangan tanlanmani e’lon qildi [11], L.Bobojonova va hammualliflari neyron model asosidagi teglagichni taqdim etdi [12]. Bu ishlarda teglash sifati va qamrovi oshdi, biroq chegaraviylik alohida annotatsiya maydoni sifatida ajratilmagan. Sh.Zarmasovning ot turkumini modellashtirishga bag‘ishlangan tadqiqotida turkumni rasmiy model shaklida berish tajribasi keltiriladi [13], ammo unda ham oraliq holatlar o‘lchov predmeti emas.

Adabiyotlar tahlili bo‘shliqni ko‘rsatadi: turkumlararo ko‘chish nazariy jihatdan puxta tavsiflangan, korpus resurslari ham yaratilgan, biroq ikkalasi birlashtirilmagan.

Tadqiqot metodologiyasi. Tadqiqot korpus tahlili va distributiv tahlil usullariga tayanadi. Material – muvozanatlashgan mikrokorpus: uch funksional uslubdan o‘n manbadan taxminan teng hajmda parcha olingan, jami 12 048 token (badiiy 4 006, ilmiy 4 004, publitsistik 4 038). Har manbadan matn boshidan uzluksiz parcha olindi, chunki tasodifiy gaplar to‘plami qo‘shni birliklar munosabatini yo‘qotadi va chegaraviylikni aniqlash imkonini bermaydi.

Teglash ikki bosqichda bajarildi: avtomatik dastlabki teglash va undan keyin har bir tokenni qo‘lda tekshirish. Qo‘lda tekshirish davomida 1 065 nuqson aniqlanib tuzatildi. Annotatsiya ishonchliligi 400 tokenlik nazorat tanlanmasida o‘lchandi; turkum tegi bo‘yicha Koen kappa koeffitsiyenti 0,850 ni, bayroq bo‘yicha 0,615 ni tashkil etdi. Ikkinchi ko‘rsatkichning pastroq bo‘lishi tabiiy: chegaraviylikni belgilash turkumni belgilashdan qiyinroq va bu farqning o‘zi natija hisoblanadi.

O‘lchov vositasi – uch qiymatli operatsional bayroq tizimi. Har bir tokenga turkum tegidan tashqari uning turkum ichidagi maqomini ko‘rsatuvchi belgi qo‘yiladi. OK belgisi birlik o‘z turkumining tipik vazifasida kelganini bildiradi. CHEG belgisi birlik turkumlararo oraliqda turganini, ya’ni shakl paradigmasi bo‘yicha bir turkumga, vazifasi bo‘yicha boshqasiga tortilayotganini qayd etadi. TRANS belgisi esa birlik boshqa turkumning funksional o‘rnini egallaganini, ammo lug‘aviy asosi o‘zgarmaganini bildiradi.

Tizimning operatsionalligi shundan iboratki, uch qiymatning har biri matndan bevosita ko‘riladigan belgiga tayanadi va boshqa tadqiqotchi tomonidan qayta tekshirilishi mumkin. Qaror lug‘aviy asos tamoyiliga qurildi: turkum birlikning barqaror asosiga qarab beriladi, matndagi siljish esa alohida maydonda qayd etiladi. Vazifaga qarab turkum belgilash yo‘lidan voz kechildi – u chastota hisobini beqaror qilardi.

Transpozitsion tokenlarning funksional yo‘nalishini aniqlash uchun ikki bosqichli tartib qo‘llandi. Avval Izoh, XPOS, UPOS va FEATS maydonlari asosida dastlabki shakliy tasnif tuzildi. Keyin har bir token alohida ko‘rib chiqildi: uning tayanch birligi, tobelik yo‘nalishi va almashtirilish imkoniyati tekshirildi. Birlik otni aniqlab, «qanday?» so‘rog‘iga javob bersa – adektivatsiya; fe’lni yoki butun predikativ mazmunni hol jihatidan belgilasa – adverbializatsiya; kelishik va son ko‘rsatkichlari bilan mustaqil argument o‘rnida tursa – substantivatsiya; ikki qism o‘rtasidagi aloqani ta’minlasa – bog‘lamalashuv.

Izoh. Maqolani tayyorlash jarayonida kvantitativ hisob-kitoblarni qayta tekshirishda sun’iy intellekt vositalaridan foydalanildi. Korpus materiali, teglash qarorlari, tahlil va xulosalar muallifga tegishli.

Tahlil va natijalar. Korpusning barcha birliklari bayroq bo‘yicha taqsimlanganda quyidagi manzara hosil bo‘ldi: 9 140 token (75,9 foiz) tipik holatda, 2 333 token (19,4 foiz) chegaraviy, 575 token (4,8 foiz) transpozitsion. YA’ni turkumlararo ko‘chish hodisasi materialning deyarli chorak qismini – 2 908 tokenni – qamrab oladi. Bu miqdor tasnifning ishonchsizligini emas, oraliq zonaning haqiqiy hajmini ko‘rsatadi.

Ikki bayroqning turkumlar bo‘yicha taqsimoti bir-biriga o‘xshamaydi va bu – tadqiqotning asosiy kuzatuvi. Chegaraviylik fe’lda to‘plangan, transpozitsiya esa otda. 1-jadval buni son bilan ko‘rsatadi.

1-jadval.

Chegaraviylik va transpozitsiyaning turkumlar bo‘yicha taqsimoti

Turkum	CHEG	TRANS	Turkumdagi ulushi, %
Fe’l	1 833	47	70,1
Ot	195	487	11,8
Ko‘makchi	105	0	25,1
Modal so‘z	69	0	56,6
Olmosh	39	4	6,8
Sifat	30	8	3,5
Ravish	—	21	6,0
Bog‘lovchi	—	8	1,8
Jami	2 333	575	24,1

Manba: muallif hisob-kitoblari, korpus v14.

Fe’l turkumining 2 681 tokenidan 1 833 tasi chegaraviy bayroq oldi. Bu – turkumning 68,4 foizi va ayni paytda korpusdagi barcha chegaraviy tokenlarning 78,6 foizi. Sabab o‘zbek fe’lining shakl tizimida: harakat nomi, sifatdosh va ravishdosh fe’l asosidan yasaladi, ammo matnda ot, sifat yoki hol vazifasini bajaradi. Biz kuzatgan misollarda bu ikki tabiat bir tokenda birlashadi:

(1) …Saraton oftobi ayovsiz qizdirayotgan, oyoq ostida qum-beton qorishmasining kukunlari sochilib yotgan maydonchadan o‘tib… (B1_Hoshimov, ID 71, 78).

Bu parchada qizdirayotgan sifatdoshi otni aniqlaydi, sochilib ravishdoshi esa yotgan fe’liga bog‘lanadi. Ikkalasi ham fe’l paradigmasida qoladi, chunki zamon va nisbat ko‘rsatkichlarini saqlaydi. Agar ularni turkum darajasida ko‘chgan deb hisoblasak, fe’l qatlamining hajmi sun’iy qisqaradi.

Ot turkumida manzara teskari: 5 757 tokendan atigi 195 tasi chegaraviy, 487 tasi esa transpozitsion. Ot oraliq holatda kam turadi – u boshqa turkumning vazifasini to‘g‘ridan-to‘g‘ri egallaydi:

(2) …Quruvchi xalqi nozik. Maoshning barakasi uchdimi, darrov o‘z ixtiyori bilan ariza beradi-da… (B1_Hoshimov, ID 94).

Quruvchi birligi bu yerda xalqi otini aniqlaydi va «qanday?» so‘rog‘iga javob beradi. Almashtirish sinovi shuni ko‘rsatadiki, uning o‘rniga sifat qo‘yish mumkin – mehnatkash xalqi – va gap tuzilishi o‘zgarmaydi. Biroq birlik sifat yasovchi qo‘shimcha olmagan, ya’ni lug‘aviy asosi ot bo‘lib qolgan; aynan shu holat transpozitsiya deb belgilanadi.

Ko‘makchi va modal so‘z turkumlaridagi ulush ham diqqatga sazovor. Ko‘makchining 418 tokenidan 105 tasi (25,1 foiz), modal so‘zning 122 tokenidan 69 tasi (56,6 foiz) chegaraviy. Bu ikki turkum kichik, lekin ularda oraliq holat nisbati eng yuqori. Ko‘makchida sabab – ot-ko‘makchi qurilmalarning ikki tabiati:

(3) …oyoq ostida qum-beton qorishmasining kukunlari sochilib yotgan maydonchadan… (B1_Hoshimov, ID 73).

Ostida birligi bu yerda sof grammatik munosabatni bildiradi, ammo predmetlik ma’nosini butunlay yo‘qotmagan – boshqa kontekstda uyning osti tarzida mustaqil ot bo‘lib kela oladi. Shu bois u ko‘makchi sifatida teglanib, chegaraviylik bayrog‘i bilan qayd etildi.Transpozitsion tokenlarning funksional yo‘nalishlari bo‘yicha to‘liq kontekstual audit o‘tkazildi. Uning natijasi 2-jadvalda keltiriladi.

2-jadval

Transpozitsiya yo‘nalishlarining taqsimoti (575 token)

Yo‘nalish	Token	Ulushi, %	Asosiy manba turkumi
Adyektivatsiya	458	79,7	ot (445 token)
Bog‘lamalashuv	37	6,4	ravishdosh
Adverbializatsiya	15	2,6	ot va sifat asosi
Substantivatsiya	2	0,3	sifat asosi
Yuklamalashuv	0	0,0	–
Tasnif chegarasida qolgan	63	11,0	–
Jami	575	100,0	–

Manba: muallif hisob-kitoblari, korpus v14.

Jadvalning asosiy natijasi shundaki, transpozitsiyaning markazi ot–sifat yo‘nalishida: 575 tokenning 458 tasi, ya’ni 79,7 foizi adyektivatsiyaga to‘g‘ri keladi. Qolgan yo‘nalishlar birgalikda ham shu ko‘rsatkichga yetmaydi. Adyektivatsion tokenlarning 445 tasi ot manbali, ya’ni yo‘nalish deyarli bir manbadan oziqlanadi.

Bizningcha, bu natijaning sababi tipologik. O‘zbek tilida ot qo‘shimchasiz aniqlovchi bo‘la oladi, chunki bitishuv yo‘li ochiq: temir eshik, bozor munosabatlari tipidagi birikmalarda birinchi komponent hech qanday shakl ko‘rsatkichini olmaydi. Shu bois yangi sifat leksemasi yasashga ehtiyoj kamayadi va til mavjud ot asosini atributiv pozitsiyaga qo‘yish bilan cheklanadi. Boshqacha aytganda, eng faol model yasalish emas, pozitsiya almashinuvi.

Substantivatsiyaning atigi ikki token bilan cheklanishi esa boshqa izohni talab qiladi. Bu ko‘rsatkich o‘zbek tilida otlashuv kam degan xulosaga asos bo‘lmaydi. Mazkur tadqiqotda lisoniylashgan otlashuvlar allaqachon ot sifatida teglangan va transpozitsiya bayrog‘ini olmagan, ya’ni ular hisobga kirmagan. Bu teglash konvensiyasining natijasi, tilning xususiyati emas. Mazkur yondashuvning cheklangan tomoni shundaki, otlashuvni o‘lchash uchun alohida operatsional mezon ishlab chiqilishi kerak.

Bog‘lamalashuv yo‘nalishida ravishdosh ustun turadi:

(4) …Bu qanaqa chempionat, dedi yuqoriga imo qilib. Indamay turaverdim… (B1_Hoshimov, ID 334).

Qilib ravishdoshi bu yerda tarz holini emas, ikki predikativ qism o‘rtasidagi aloqani ta’minlaydi; uning o‘rniga bog‘lovchi qo‘yish mumkin, ammo ravishdosh shakli saqlanadi. Bu holat o‘zbek tilida ergashtiruvchi bog‘lovchining kam qo‘llanishini izohlaydi: ergashish munosabati alohida bog‘lovchi so‘z bilan emas, fe’lning tuslanmaydigan shakli bilan ifodalanadi.

Ikki bayroqning uslublar bo‘yicha taqsimoti qarama-qarshi yo‘nalishda o‘zgaradi. Chegaraviy tokenlar badiiy uslubda eng ko‘p (871), ilmiy uslubda 842, publitsistik uslubda esa 620 tani tashkil etadi. Transpozitsion tokenlar esa teskari tartibda: badiiy 177, ilmiy 193, publitsistik 205. Kuzatishimizcha, bu qarama-qarshilik tasodifiy emas. Badiiy matn voqealar ketma-ketligini nofinit fe’l shakllari bilan quradi va shu bois chegaraviylik ortadi. Publitsistik matn esa nominativ zanjirlarni atributiv otlar bilan kengaytiradi, natijada transpozitsiya ko‘payadi. Ya’ni ikki bayroq grammatik erkinlikning bir darajasini emas, ikki xil diskursiv strategiyani aks ettiradi.

Xulosa va takliflar. Turkumlararo ko‘chishni operatsional bayroqlar tizimi orqali o‘lchash mumkinligi va bunday o‘lchov mazmunli natija berishi tasdiqlandi. Mikrokorpusda chegaraviy birliklar 2 333 tokenni (19,4 foiz), transpozitsion birliklar 575 tokenni (4,8 foiz) tashkil etdi; ikkalasi birgalikda materialning 24,1 foizini qamradi. Bu miqdor oraliq zonaning nazariy tavsiflarda taxmin qilinganidan kengroq ekanini ko‘rsatadi. Ikki hodisaning taqsimoti bir-biridan farq qiladi va shu farq tadqiqotning asosiy xulosasini beradi. Chegaraviylikning 78,6 foizi fe’lga to‘g‘ri keladi va uning manbai fe’lning tuslanmaydigan shakllari. Transpozitsiyaning 84,7 foizi esa ot asosidan hosil bo‘ladi va uning yetakchi yo‘nalishi ot–sifat: 458 token, 79,7 foiz. Demak, o‘zbek tilida fe’l oraliq holatni shakl tizimi orqali, ot esa pozitsiya orqali yuzaga keltiradi.

ADABIYOTLAR

1. O‘zbekiston Respublikasi Prezidentining “O‘zbek tilining davlat tili sifatidagi nufuzi va mavqeyini tubdan oshirish chora-tadbirlari to‘g‘risida”gi 2019-yil 21-oktabrdagi PF-5850-son Farmoni. https://lex.uz/docs/-4561730
2. O‘zbekiston Respublikasi Prezidentining “Mamlakatimizda o‘zbek tilini yanada rivojlantirish va til siyosatini takomillashtirish chora-tadbirlari to‘g‘risida”gi 2020-yil 20-oktabrdagi PF-6084-son Farmoni. https://lex.uz/docs/-5058351
3. O‘zbekiston Respublikasi Prezidentining “Sun’iy intellekt texnologiyalarini 2030-yilga qadar rivojlantirish strategiyasini tasdiqlash to‘g‘risida”gi 2024-yil 14-oktabrdagi PQ-358-son qarori. https://lex.uz/docs/-7158604
4. Элтазаров Ж.Д. Ўзбек тилида сўз туркумларининг ўзаро алоқаси ва кўчиши: Филол. фан. докт. ... дисс. – Самарқанд, 2005. – 284 б.
5. Исақов З. Ўзбек тилида сўз туркумларининг ўзаро муносабати ва унда юкламаларнинг ўрни: монография. – Қўқон, 2011. – 83 б.
6. Nurmonov A., Shahobiddinova Sh., Iskandarova Sh., Nabiyeva D. O‘zbek tilining nazariy grammatikasi. Morfologiya. – Toshkent: Yangi asr avlodi, 2001. – 80 b.
7. Шоабдураҳмонов Ш., Асқарова М., Ҳожиев А., Расулов И., Дониёров Х. Ҳозирги ўзбек адабий тили. I қисм. – Тошкент: Ўқитувчи, 1980. – 446 б.
8. Quirk R., Greenbaum S., Leech G., Svartvik J. A Comprehensive Grammar of the English Language. – London–New York: Longman, 1985. – 1779 p.
9. Croft W. Radical Construction Grammar: Syntactic Theory in Typological Perspective. – Oxford: Oxford University Press, 2001. – 416 p.
10. Abduraxmonova N.Z. Korpus lingvistikasi. – Toshkent, 2023. – 357 b.
11. Sharipov M., Kuriyozov E., Vičič J. UzbekPOS: A multi-domain dataset for Uzbek part-of-speech tagging // Data in Brief. – 2026. – Vol. 66. – Art. 112640. https://doi.org/10.1016/j.dib.2026.112640
12. Bobojonova L., Akhundjanova A., Ostheimer P.S., Fellenz S. BBPOS: BERT-based Part-of-Speech Tagging for Uzbek // Proceedings of the First Workshop on Language Models for Low-Resource Languages. – Abu Dhabi, 2025.
13. Zarmasov Sh.Sh. O‘zbek tili so‘z turkumlarini modellashtirish tamoyillari (ot so‘z turkumi misolida): Filol. fan. bo‘yicha fals. dokt. (PhD) diss. avtoref. – Qo‘qon, 2025. – 58 b.