

ISSN 2181-922X

LANGUAGE & CULTURE

UZBEKISTON O'ZBEKISTON

TIL VA MADANIYAT

KOMPYUTER
LINGVISTIKASI

2024 Vol. 2 (6)

www.compling.tsuull.uz

ISSN 2181-922X

O‘ZBEKISTON

TIL VA MADANIYAT

KOMPYUTER LINGVISTIKASI

2024 Vol. 2 (6)

compling.tsuull.uz

Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti

Bosh muharrir:

Botir Elov

Bosh muharrir o'rinbosari:

Shahlo Hamroyeva

Mas'ul kotib:

Oqila Abdullayeva

Tahrir kengashi

Shuhrat Sirojiddinov (O'zbekiston), Eshref Adali (Turkiya), [Viktor Zaxarov] (Rossiya), Vladimir Benko (Slovakiya), Ayrat Gatiatullin (Tataristan), Rinat Gilmullin (Tataristan), Murat O'rxun (Turkiya), Suyun Karimov (O'zbekiston), Abduvali Qarshiyev (O'zbekiston), Muxammadjon Musayev (O'zbekiston), Kamoliddin Shukurov (O'zbekiston), O'tkir Hamdamov (O'zbekiston), Tal'at Zuparov (O'zbekiston), Bahodir Mo'minov (O'zbekiston), Faxriddin Nurullayev (O'zbekiston), Zulxumor Xolmanova (O'zbekiston), Muqaddas Abdurahmonova (O'zbekiston), Elova Dilrabo (O'zbekiston), Ruhillo Alayev (O'zbekiston), Rasuljon Atamuratov (O'zbekiston), Malika Abdullayeva (O'zbekiston), Mannon Ochilov (O'zbekiston), Xolisa Axmedova (O'zbekiston), Zilola Xusainova (O'zbekiston), Uldona Abdurahmonova (O'zbekiston).

Jurnal haqida ma'lumot

“O'zbekiston: til va madaniyat. Kompyuter lingvistikasi” seriyasi – Oliy attestatsiya komissiyasi ilmiy nashrlar ro'yxatidagi “O'zbekiston: til va madaniyat” akademik jurnalining ilovasi hisoblanib, unda professor-o'qituvchilar, doktorantlar, stajor-tadqiqotchilar, mustaqil izlanuvchilar, magistrantlarning kompyuter lingvistikasi, jumladan, tabiiy tilga ishlov berish (NLP), o'zbek tilining formal grammatikasi, korpus lingvistikasi, mashina tarjimasi, nutqni qayta ishlash tizimlari, intellektual tizimlar, kompyuter leksikografiyasi hamda lingvistik ontologiyalar kabi sohalarga oid tadqiqotlari nashr qilinadi.

Jurnal ilovasi bir yilda to'rt marta chop etiladi.

O'zbek, turk, rus va ingliz tillarida yozilgan maqolalar qabul qilinadi.

Jurnalda kitoblarga yozilgan taqrizlar, adabiyotlar sharhi, konferensiyalar hisobotlari va tadqiqot loyihalari natijalari ham e'lon qilinadi.

Mualliflar fikri tahririyat nuqtayi nazaridan farq qilishi mumkin.

“O'zbekiston: til va madaniyat. Kompyuter lingvistikasi” seriyasi 2023-yildan chiqa boshlagan.

Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti. O'zbekiston, Toshkent, Yakkasaroy tumani, Yusuf Xos Hojib ko'chasi, 103-uy.

E-mail: kompling@navoiy-uni.uz

Website: kompling.tsuull.uz

Alisher Navoi'i Tashkent State University of the Uzbek Language and Literature

Chief editor:	Botir Elov
Deputy editor-in-chief:	Shahlo Hamroyeva
Responsible secretary:	Oqila Abdullayeva

Editorial board

Shukhrat Sirojiddinov (Uzbekiston), Eshref Adali (Turkiye), [Viktor Zakharov] (Russia), Vladimir Benko (Slovakia), Ayrat Gatiatullin (Tataristan), Rinat Gilmullin (Tataristan), Murat Orhun (Turkey), Suyun Karimov (Uzbekistan), Abduvali Karshiyev (Uzbekistan), Mukhammadjon Musayev (Uzbekistan), Kamoliddin Shukurov (Uzbekistan), O'tkir Hamdamov (Uzbekistan), Tal'at Zuparov (Uzbekistan), Bahadir Mo'minov (Uzbekistan), Fakhridin Nurullayev (Uzbekistan), Zulkhumor Kholmanova (Uzbekistan), Muqaddas Abdurakhmonova (Uzbekistan), Elova Dirlabo (Uzbekistan), Ruhillo Alayev (Uzbekistan), Rasuljon Atamuratov (Uzbekistan), Malika Abdullayeva (Uzbekistan), Mannon Ochilov (Uzbekistan), Kholisa Akhmedova (Uzbekistan), Zilola Khusainova (Uzbekistan), Uldona Abdurakhmonova (Uzbekistan).

Information about the magazine

"Uzbekistan: language and culture. "Computer Linguistics" series is an appendix of the academic journal "Uzbekistan: Language and Culture" in the list of scientific publications of the Higher Attestation Commission, in which computer linguistics, including natural language processing (NLP) of professors-teachers, doctoral students, intern-researchers, independent researchers, master's students, researches related to formal grammar of the Uzbek language, corpus linguistics, machine translation, speech processing systems, intelligent systems, computer lexicography and linguistic ontologies are published.

The magazine supplement is published four times a year.

Articles written in Uzbek, Turkish, Russian and English languages are accepted.

The journal also publishes book reviews, literature reviews, conference reports, and research project results.

The opinion of the authors may differ from the editorial point of view.

"Uzbekistan: language and culture. "Computer Linguistics" series has been published since 2023.

Tashkent State University of Uzbek Language and Literature named after Alisher Navoi. Yusuf Khos Hajib street, 103, Yakkasaray district, Tashkent, Uzbekistan.

E-mail: kompling@navoiy-uni.uz

Website: kompling.tsuull.uz

MUNDARIJA

Фархад Мирзаев Рамиз, Гюнель Новрузова Сиявуш

Компьютерное моделирование основные инструменты
исследования6

Xolisa Axmedova, Elbek Malikov

Bilinga asoslangan yondashuvlar asosida omonimiyani
farqlash.....15

Xolisa Axmedova, Shohnazar Sultonov

Statistik usullar yordamida polifunksional so'zlarni
semantik farqlash.....26

Oqila Abdullayeva, O'g'iloy Bozorqulova

Jahon tilshunosligida treebanklar tasnifi.....43

Zilola Xusainova, Surayyo Yangibayeva

Til korpusi turlari.....54

Shaxinabonu Mansurova

Son so'z turkumini grammatik pos teglashning
lingvistik modellari.....64

Botir Elov, Zilola Xusainova, Sarvinoz Qosimova

Katta til modellari.....78

Oqila Abdullayeva, Fotima O'tkirova

Jahon tilshunosligida Dependancy Parsingga oid
tadqiqotlar.....92

CONTENT

Farhad Mirzayev Ramiz, Gunel Novruzova Siyavush

Computer simulation basic research

tools.....13

Xolisa Axmedova, Elbek Malikov

Differentiating knowledge-based

Homonymy.....24

Xolisa Axmedova, Shohnazar Sultonov

Semantic differentiation of polyfunctional words using

statistikal methods.....40

Oqila Abdullayeva, O'g'iloy Bozorqulova

The classification of treebanks in world linguistics.....52

Zilola Xusainova, Surayyo Yangibayeva

Types of language corpus.....62

Shaxinabonu Mansurova

Linguistic methods of grammatical pos tagging

of the number word group.....76

Botir Elov, Zilola Xusainova, Sarvinoz Qosimova

Large language models.....90

Oqila Abdullayeva, Fotima O'tkirova

Dependancy Parsing in world

linguistics.....104

TIL KORPUSI TURLARI

Zilola Xusainova¹

Surayyo Yangibayeva²

Annotatsiya. Mazkur maqolada korpus tipologiyasi o'rganilib, ingliz va rus tilshunoslari tomonidan o'rganilgan tasniflash ishlari muhokama qilingan. Tilshunos olim V.P.Zaxarov ta'kidlaganidek, mavjud korpuslarning butun xilma-xilligi ular uchun yaratilgan tadqiqotlar va amaliy muammolarning ham xilma-xilligi bilan belgilanganligi bois bir qancha tasniflar qisqacha tahlilga tortilgan.

Kalit so'zlar: *korpus, korpus lingvistikasi, ekstralingvistik belgi, sinxron, diaxron, parallel korpus.*

Adabiyotlar sharhi

So'nggi yillarda korpus lingvistikasi sohasida mamlakatimizda ilmiy va amaliy tadqiqotlar olib borilmoqda. Korpus lingvistikasi lingvistik tadqiqotlar uchun mo'ljallangan lingvistik ma'lumotlarning bazasini (lingvistik korpuslar) yaratish va ulardan foydalanishning nazariy asoslari va amaliy ishlanmalarni o'z ichiga oladi. Til korpusi, shuningdek, dasturiy ta'minotni ishlab chiqishda qo'llaniladigan turli tabiiy tillardagi ma'lumotlar bazalarini yaratish uchun ishlatiladi. Jumladan, *imlo tekshiruvi, grammatikani tuzatish, matn/nutqni tanib olish tizimlari va mashina tarjimasini tizimlari* kabi NLP ilovalarini ishlab chiqishga asos bo'lib xizmat qiladi. Bugungi kunda, *korpusning tuzilishi, maqsadi, vazifasi, o'zgaruvchanligi* va boshqa jihatlarga ko'ra bir qancha turlarga bo'linishi mutaxassislar tomonidan turlicha tasniflanganligini ko'rishimiz mumkin [Murphy, 2015]. Xususan, Brona Murphy [Murphy, 2015], Elaine Riordan, Natalia Levshina [Levshina, 2022], Randi Reppen, Rita Simpson-Vlach [Reppen, Simpson-Vlach, 2019], V.P.Zaxarov [Zaxarov, 2015],

¹Xusainova Zilola Yuldashevna – filologiya fanlari falsafa doktori (PhD). Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti o'qituvchisi.

E-pochta: xusainovazilola@navoiy-uni.uz

ORCID: 0000-0003-4357-7515

²Yangibayeva Surayyo – ToshDO'TAU, Kompyuter lingvistikasi mutaxassisligi I bosqich magistranti.

E-pochta: surayyoyangibayeva4@gmail.com

M.V.Kopotev [Копотев, 2014], A.B.Kutuzov kabi kabi ko'plab chet el olimlari korpus turlarini bir necha xil usullarda, bir qancha tamoyillarga asosan tasniflashgan.

Olib borilgan ilmiy tadqiqotlarda ingliz tili korpuslari tasnifi nisbatan sodda shakllarda izohlangan [Murphy, Riordan, 2016]. Ushbu tadqiqotlarda, ingliz tilshunoslari korpuslarni *umumlashtirilgan, ixtisoslashtirilgan, o'rganuvchi, pedagogik, tarixiy, parallel, taqqoslanadigan va monitoring* kabi sakkiz turga ajratgan holda tasniflashgan. Korpusning muayyan turidan foydalanish esa NLP vazifasining yoki olib borilayotgan ilmiy tadqiqotning maqsadiga bog'liq bo'lib, asosan, ingliz tilshunoslari korpus yondashuvini qo'llash uchun quyidagi to'rtta turdagi korpuslarni asosiy korpuslar deya izohlashadi [Bennett, 2010].

1-jadval. Ingliz tili korpusi turlariga namuna

Nº	Korpus turi	Til korpusi
1.	<i>Umumiy korpuslar (Generalized Corpara)</i>	The British National Corpus (BNC), the American National Corpus (ANC)
2.	<i>Xususiy korpuslar (Specialized Corpara)</i>	Michigan Corpus of Academic Spoken English (MICASE), the CHILDES Corpus (MacWhinney, 1992)
3.	<i>O'rganuvchi korpuslar (Learner Corpara)</i>	International Corpus of Learner English (ICLE) (Granger, 2003)
4.	<i>Pedagogik korpuslar (Pedagogic Corpara)</i>	All the material (Student Book and Workbook) of the Funway series - the old coursebooks

Shuningdek, boshqa manbalarda ingliz tilshunoslari korpuslarni *til, vaqt, hozirgilik va boshqa ko'proq xususiyatlari* bo'yicha turlarga ajratgan.

2-jadval. Turli mezonlarga ko'ra tili korpuslari tasnifi

Nº	Mezon	Korpus turi	Til korpusi
1.	Til xususiyatlariga ko'ra	Monolingual korpus	The Czech National Corpus
		Parallel korpus	Croatian-English parallel corpus hrenWaC 2.0
		Ko'p tilli korpus	Aarthus corpus

		Taqqoslanadigan korpus	Reuter's Corpus in English
2.	Vaqt xususiyatlariga ko'ra	Sinxron korpus	Synchronic English Web Corpus
		Diaxron korpus	Diachronic English Web Corpus
		Statistik korpus	Brown Corpus, COCA
3.	Tilning hozirgi holatlariga ko'ra	Monitor korpus	English Trends corpus
		O'quvchi korpusi	Longman Learners' Corpus, the Cambridge Learner Corpus
4.	Ko'proq xususiyatlariga ko'ra	Xato bilan izohlangan korpus	Error-Annotated Corpus of Latvian
		Ixtisoslashgan korpus	Air Traffic Control (ATC) Corpus
		Multimediya korpusi	The Multimedia Russian Corpus (MURCO)

Ushbu tasniflar orqali ingliz korpus lingvistikasida korpuslar asosan, *maqsadi* hamda *ifodalaydigan xususiyatlariga* ko'ra tasniflanishini qayd etishimiz mumkin.

Korpusni ma'lum bir toifaga tasniflash, ular qo'llanilishi mumkin bo'lgan *maqsadga* yoki *xususiyatlarga* asoslanadi. Til korpuslari mavjud xususiyatlaridan kelib chiqqan holda rus tilshunoslari tomonidan ham tasniflangan. Rus tilshunosligida korpus turlarini tasniflash ishlariga dastlab bir necha olimlar tadqiqotlar olib borishgan. Jumladan, V.P.Zaxarov, V.V.Rikov va M.V.Kopotev kabi olimlar tasniflash borasida bir qancha ilmiy ishlarni olib borishgan bo'lib, ushbu tasniflar rus korpusshunosligi uchun muhim tasniflar hisoblanadi. Bu tasniflarning bir-biridan farqli va o'xshash tomonlari mavjud. Xususan, V.P.Zaxarov korpuslarni bajaradigan *vazifalaridan, tilidan, oldiga qo'yilgan maqsadlaridan* va *qo'llanilish imkoniyatlaridan* kelib chiqqan holda tasniflagan [Захаров, Богданова, 2022]. Har bir turdagi korpus biz sanab o'tgan ushbu xususiyatlardan biri orqali keng ko'lamda yoritiladi. Masalan, rus tili milliy korpusi tarkibidan joy olgan dialektal korpusni olaylik. Ushbu korpus **599 258** dialektal so'zlarni o'zida jamlagan. Ya'ni bu korpus qo'llanilish imkoniyatidan kelib chiqib, obyektidan chetga chiqmagan holda faqat dialektal birliklarni izohlaydi. Ammo rus tilshunoslari bu bilan cheklanib qolmaydi.

3-jadval. V.P.Zaxarovning til korpuslari tasnifi

saqlash shakliga ko'ra	matn tiliga ko'ra	Janriga ko'ra	korpusga kirish imkoniyatiga ko'ra	maqsadiga ko'ra	dinamik-ligiga ko'ra	qo'shimcha axborotga egaligiga ko'ra
Ovozli yozma aralash	bir tilli ko'p tilli	adabiy dialektal og'zaki publisistik aralash	Erkin tijorat yopiq	tadqiqiy illyustrativ	dinamik (monitor) turg'un	Teglangan Teglanmagan

Tasniflash borasida M.V.Kopotev ham keng qamrovli ishlarni olib borgan. Ilmiy adabiyotlarni qiyosiy tahlil qilish yagona tasnifni sintez qilish imkonini berdi. Mana shu jihatni inobatga olgan holda olim korpus turlarini nisbatan keng shaklda izohlaydi [Копотев, 2014].

4-jadval.M.V. Kopotevning til korpuslari tasnifi

№	Tasniflash mezonlari	Korpus turlari
1.	<i>Ma'lumotlarni taqdim etish formati</i>	Monomodal / Multimodal (multimediya)
2.	<i>Ma'lumotlarning hajmi va to'liqligi</i>	Umumiy (vakillik, milliy) / Maxsus
3.	<i>Tillar soni</i>	Bir tilli / ikki tilli / ko'p tilli
4.	<i>Bitta matndagi tillar soni (ikki va ko'p tilli korpuslar uchun)</i>	Parallel
5.	<i>Turli tillardagi matnlarni tekislash (ikki va ko'p tilli korpuslar uchun)</i>	mutanosib yoki parallel / o'lchovsiz yoki parallel bo'lmagan
6.	<i>Til materialini tanlashning yagona tamoyillarining mavjudligi</i>	Taqqoslanadigan / solishtirilmaydigan
7.	<i>To'plangan til ma'lumotlari tegishli bo'lgan vaqt davri</i>	Sinxron / diaxronik
8.	<i>Belgilarning mavjudligi</i>	Belgilangan (annotatsiya qilingan) / Belgilanmagan (annotatsiya qilinmagan)
9.	<i>Belgilash turi (belgilangan korpuslar uchun)</i>	Metatmatn belgisi / Lingvistik belgi (morfologik, sintaktik, semantik) / Ekstralingvistik belgilar
10.	<i>Mavjudligi</i>	Erkin mavjud / Qisman mavjud / Yopiq (tijorat)
11.	<i>Mualliflarning lingvistik malaka darajasi</i>	Haqiqiy matnlar korpusi / Ikkinchi/chet tilini o'rganuvchilar tomonidan yaratilgan matnlar korpusi / Bolalar nutqi korpusi
12.	<i>Til materialini ifodalash</i>	To'liq matn / Fragmentli (jumladan, n-gramm korpus va muvofiqlik)

Tasnifmuallifikorpuslarninghartomonlama imkoniyatlaridan va xususiyatlaridan kelib chiqqan holda har bir korpusni alohida turlarga ajratib chiqqan. M.V.Kopotev til korpuslarini tasniflash jarayonida “parallel” soʻzini ikki oʻrinda ishlatgan boʻlib, parallel korpuslar tasnifini belgilash uchun, yaʼni turli xil tillardagi matnlarning oʻxshash qismlarini taqqoslashni bildiradi. Tilshunos olimlardan yana biri V.V.Rikov ham korpuslarni tasniflash ishlarini olib borgan. V.V.Rikov ham V.P.Zaxarov, M.V.Kopotev kabi korpuslarning yuqorida biz izohlagan xususiyatlarini inobatga olgan boʻlsa ham, tasniflashga boshqacha tamoyil bilan yondashganini qayd etish mumkin [Корпусная лингвистика, 2008].

5-jadval. V.V.Rikovning til korpuslari tasnifi

Tashkil qilish darajasi va qurilishiga koʻra	Xronologik belgisiga koʻra	Koʻrsatkichning mavjudligiga koʻra	Tiliga koʻra	Qoʻllash, foydalanish usuliga koʻra	Mavjudlik usuliga koʻra
elektron arxiv elektron kutubxona matnlar korpusi subkorpus	sinxron monitor diaxron	Oddiy Annotatsiyalangan	bir tilli ikki tilli koʻp tilli	tadqiqot illyustrativ parallel	dinamik turgʻun

Yuqoridagi tasnifga koʻra V.V.Rikovning tasnifi nisbatan tor doirada qisqartirilgan koʻrinishdagi tasnif shaklida koʻrinishi mumkin. Rikov korpuslarning qanday yaratilishiga, xronologiyasiga, tiliga hamda usullarini inobatga olgan holda tasniflagan.

Yuqorida keltirilgan rus olimlarining tasnifi orqali korpuslarning juda koʻplab imkoniyatlari mavjudligini qayd etish mumkin. Lekin, tasniflashda turlicha yondashuvlar boʻlganligi sababli, ushbu tasniflarni quyidagi shaklda umumlashtirish mumkin.

6-jadval. Til korpuslari tasnifini qiyoslash

№	Korpus turlari	V.P.Zaxarov	M.V.Kopotev	V.V.Rikov
1.	Saqlash shakliga koʻra: ovozli, yozma, aralash	+	+	
2.	Tashkil qilish darajasi va qurilishiga koʻra: elektron arxiv, elektron kutubxona, matnlar korpusi, subkorpus			+
3.	Matn tiliga koʻra: bir tilli, ikki tilli, koʻp tilli	+	+	+
4.	Janriga koʻra: adabiy, dialektal, ogʻzaki, publisistik, aralash	+		
5.	Maʼlumotlarni taqdim etish formati: Monomodal / Multimodal (multimediya)		+	

6.	Xronologik belgisiga ko'ra: sinxron, diaxron		+	+
7.	Qo'shimcha axborotga egaligiga ko'ra: teglangan, teglanmagan	+	+	+
8.	Ma'lumotlarning hajmi va to'liqligi: umumiy (vakillik, milliy) / maxsus		+	
9.	Korpusga kirish imkoniyatiga ko'ra: erkin, tijorat, yopiq	+	+	
10.	Maqsadiga ko'ra: tadqiqiy, illyustrativ	+		+
11.	Mavjudlik usuliga ko'ra: dinamik, turg'un	+	+	+
12.	Belgilash turi (belgilangan korpuslar uchun): Metatmatn belgisi / Lingvistik belgi (morfologik, sintaktik, semantik) / Ekstralingvistik belgilar		+	
13.	Mualliflarning lingvistik malaka darajasi: Haqiqiy matnlar korpusi / Ikkinchi/ chet tilini o'rganuvchilar tomonidan yaratilgan matnlar korpusi / Bolalar nutqi korpusi		+	
14.	Til materialini ifodalash: To'liq matn / Fragmentli (jumladan, n-gramm korpus va muvofiqlik)		+	
15.	Bitta matndagi tillar soni (ikki va ko'p tilli korpuslar uchun): parallel/ aralash		+	
16.	Turli tillardagi matnlarni tekislash (ikki va ko'p tilli korpuslar uchun): mutanosib yoki parallel / o'lchovsiz yoki parallel bo'lmagan		+	

Yuqoridagi 6-jadval orqali uchala olim ham tasniflashda korpuslarning *saqlash shakliga, tiliga, mavjud usuliga ko'ra* umumiy yondashuvni ko'ramiz. V.P.Zaxarov va M.V.Kopotevlar tasniflari nisbatan bir-biriga yaqin bo'lsa ham, M.V.Kopotevning tasniflari ko'p tilli korpuslarga alohida e'tibor qaratganligi bilan farqlanadi. V.V.Rikov tasnifidagi eng katta farq bu korpuslarning xronologik jihatdan tasniflanmaganligidir. Chunki, xronologik jihatdan tasniflanishi korpuslarning muhim tomoni hisoblanadi.

Umuman olganda, barcha tabiiy tillar uchun **til korpusni loyihalash va joriy etishning** umumiy metodikasi mavjud emas. Shunday bo'lsa-da, bu borada V.V.Rikov til korpuslari uchun quyidagi asosiy talablarni keltirgan:

1. *Korpus foydalanuvchilari talablari.*
2. *Korpusning mantiqiy maqsadlari.*
3. *Korpus tuzishda ishlanadigan ma'lumotlar bazasi hajmi.*

4. *Mantdan foydalanish usullari.*

Shuningdek, ushbu tasniflar orqali til korpusiga matn tanlashda quyidagi maqsadlarga asoslanadi:

1. *Matnni skanerlash.*
2. *Ma'lumotlarni tahlil qilish.*
3. *Korpusni shakllantirish.*

Olib borilgan tadqiqotlar asosida, eng avvalo, *yozma, og'zaki manbaga* asoslanishiga ko'ra til korpuslari 2 turga ajratiladi. Dastlabki til korpuslari faqat yozma matnlar asosida shakllantirilgan. Chunki, yozma matnlarni yig'ish ancha qulayroq hisoblanib, ularni ma'lumotlar bazasiga saqlashning 3 usuli mavjud:

1. *Matnlarni tahrirlash.*
2. *Elektron shaklda mavjud matnlardan foydalanish.*
3. *Bosma matnni skanerlash.*

Tajribalar shuni ko'rsatadiki, korpus lingvistikasi kamida 2 tipdagi obyekt ustida amaliyot bajaradi. Korpuslarning birinchi tipi **universal** bo'lib, ular o'zida nutqiy faoliyatning barcha xilma-xilliklarini aks ettiradi. Ikkinchi turdagi korpus jamiyatda nutqiy amaliyotning ma'lum bir lingvistik yoki madaniy fenomenning odatga aylanishi (yoki aylanganligi)ni aks ettirib, ular ma'lum maqsadni ko'zlab tuziladi. Bunday korpus sifatida *maqollar korpusi*, *OAV* korpusini sanab o'tish mumkin [Zaxarov, Mengliyev, Xamroyeva, 2021].

Xulosa

Ushbu maqolada keltirilgan rus tilshunosligidagi tasniflar o'zbek tilshunosligi uchun muhim manba hisoblanadi. Bugungi kunda, ushbu tasniflarga asoslangan holda o'zbek tilshunosligida ham bir qancha korpuslar yaratilgan. Umuman olganda, ushbu maqolada korpuslarning *tuzilishi, maqsadi, turg'un, o'zgaruvchanligi* va boshqa jihatlariga ko'ra bir qancha turlarga ega ekanligining va bu NLP mutaxassislari tomonidan turlicha tasniflanganligi keltirildi. Bu esa, o'zbek korpus lingvistikasida bajarilishi lozim bo'lgan ilmiy tadqiqotlarga asos bo'ladi. Xususan, bugungi kunda o'zbek tilidagi til korpuslari bo'yicha amaliy ishlar deyarli bajarilmagan. Shu sababli, bu sohada bajarilishi lozim bo'lgan ishlar juda ko'p deb hisoblaymiz.

Foydalanilgan adabiyotlar

Gena R.Bennett. An introduction to corpus linguistics. Part 1. Michigan ELT, 2010.

<https://catalog.ldc.upenn.edu/LDC94S14A>

<https://cc.au.dk/en/research/knowledge-communication-lab/products/corpus/eucorpus>
<https://ruscorpora.ru/corpus/dialect>
<https://www.sketchengine.eu/corpora-and-languages/corpus-types/>

Levshina, N. (2022). Corpus-based typology: Applications, challenges and some solutions. *Linguistic Typology*, 26(1), 129-160.

Reppen, R., & Simpson-Vlach, R. (2019). Corpus linguistics. In *An introduction to applied linguistics* (pp. 91-108). Routledge.

V.Zaxarov, B.Mengliyev, Sh.Xamroyeva "Korpus lingvistikasi" O'quv qo'llanma. Toshkent, 2021

Захаров В.П. Корпусная лингвистика. Учебно-методическое пособие. – Санкт-Петербург, 2005. – 48 с

Копотев М. Введение в корпусную лингвистику: Учебное пособие для студентов филологических и лингвистических специальностей университетов. – Прага: Animedia, 2014. – 230 с.

Копотев, М. (2014). *Введение в корпусную лингвистику: Учебное пособие для студентов филологических и лингвистических специальностей университетов. Корпусная лингвистика. Филология*. Animedia Company.

Корпусная лингвистика. Институт лингвистических исследований РАН, 2008 (электрон ресурс).

Корпусная лингвистика. Статья в интернет-энциклопедии Википедия. <http://wikipedia.org/wiki/>

Курс "Корпусная лингвистика" (А.Б. Кутузов) Лицензия Creative commons Attribution Share-Alike 3.0 Unported (Электрон ресурс)

TYPES OF LANGUAGE CORPUS

Zilola Xusainova¹,
Surayyo Yangibayeva²

Abstract. This article explores the typology of corpora and examines the classification efforts studied by English and Russian linguists. As noted by linguist V.P. Zakharov, the entire diversity of existing corpora is determined by the variety of research and practical issues for which they are created. Therefore, several classifications have been briefly analyzed.

Keywords: *corpus, corpus linguistics, extralinguistic feature, synchronic, diachronic, parallel corpus.*

References

- Gena R.Bennett. An introduction to corpus linguistics. Part 1. Michigan ELT, 2010.
<https://catalog.ldc.upenn.edu/LDC94S14A>
<https://cc.au.dk/en/research/knowledge-communication-lab/products/corpus/eucorpus>
<https://ruscorpora.ru/corpus/dialect>
<https://www.sketchengine.eu/corpora-and-languages/corpus-types/>
- Levshina, N. (2022). Corpus-based typology: Applications, challenges and some solutions. *Linguistic Typology*, 26(1), 129-160.
- Reppen, R., & Simpson-Vlach, R. (2019). Corpus linguistics. In *An introduction to applied linguistics* (pp. 91-108). Routledge.
- V.Zaxarov, B.Mengliyev, Sh.Xamroyeva "Korpus lingvistikasi" O'quv qo'llanma. Toshkent, 2021.
- Zaxarov V.P. Korpusnaya lingvistika. Uchebno-metodicheskoe posobie. – Sankt-Peterburg, 2005. – 48 s
- Kopotev M. Vvedenie v korpusnuyu lingvistiku: Uchebnoe posobie dlya studentov filologicheskix i lingvisticheskix spetsialnostey universitetov. – Praga: Animedia, 2014. – 230 s.
- Kopotev, M. (2014). *Vvedenie v korpusnuyu lingvistiku: Uchebnoe*

¹Xusainova Zilola Yuldashevna – doctor of philosophy in philology, teacher. Alisher Navo'i Tashkent State University of Uzbek Language and Literature.

E-pochta: xusainovazilola@navoiy-uni.uz

ORCID: 0000-0003-4357-7515

²Yangibayeva Surayyo – Master of degree. Alisher Navo'i Tashkent State University of Uzbek Language and Literature.

E-pochta: surayyoyangibayeva4@gmail.com

*posobie dlya studentov filologicheskix i lingvisticheskix
spesialnostey universitetov. Korpusnaya lingvistika. Filologiya.*
Animedia Company.

Korpusnaya lingvistika. Institut lingvisticheskix issledovaniy RAN,
2008 (elektron resurs).

Korpusnaya lingvistika. Statya v interenet-ensiklopedii Vikipediya.
<http://wikipedia.org/wiki/>

Kurs "Korpusnaya lingvistika" (A.B. Kutuzov) Litsenziya Creative
commons Attribution Share-Alike 3.0 Unported (Elektron
resurs).

Jurnal 2017-yil 26-oktyabrda O'zbekiston Respublikasi Matbuot
va axborot agentligi tomonidan 0936-raqam bilan ro'yxatdan o'tgan.

Jurnal O'zbekiston Respublikasi Oliy Attestatsiya Komissiyasi tomonidan
filologiya fanlari bo'yicha falsafa doktori (PhD) va fan doktori (DSc)
dissertatsiyalari asosiy ilmiy natijalari chop etilishi lozim bo'lgan ro'yxatga
kiritilgan (30.10.2021. № 308/6).

Tahririyatga kelgan maqolalar mualliflarga qaytarilmaydi.

Manzil: Toshkent shahri, Yakkasaroy tumani, Yusuf Xos
Hojib ko'chasi 103-uy.
Telefonlar: +99871 281-45-11, +99871 281-41-93.
Website: compling.tsuull.uz
E-mail: kompling@navoiy-uni.uz

Bosishga **.**.****-yilda ruxsat etildi.
Bichimi 70x100 1/16, Ofset bosma. "Cambria" garniturasida.
Shartli b.t. 7,51. Nashr b.t. 7,62.

"O'zbekiston: til va madaniyat" jurnali tahririyatida
tayyorlandi va sahifalandi.
"YASHNOBOD NASHR" bosmaxonasida chop etildi.
Adadi 300 nusxa. Buyurtma №2.
Bosmaxona manzili: Toshkent shahar Yashnobod tumani,
58-a harbiy shaharcha.