

**O'ZBEKISTON RESPUBLIKASI  
OLIV TA'LIM, FAN VA INNOVATSIYALAR VAZIRLIGI**

**TERMIZ DAVLAT UNIVERSITETI**

**Ro'yxatga olindi:**  
MD- 70230101 – 2.4.  
“ ” 2024-yil



**“TASDIQLAYMAN”**  
Termiz davlat universiteti  
O'quv ishlari bo'yicha prorektor  
R.To'rayev  
2024-yil

**KORPUS LINGVISTIKASI  
FANINING  
O'QUV DASTURI**

**Bilim sohasi:** 200000 – San'at va gumanitar fanlar  
**Ta'lim sohasi:** 230000 – Tillar  
**Ta'lim yo'nalishi:** 70230101 – Lingvistika (o'zbek tili)

Fan/modul kodi AT 304		O'quv yili 2024-2025	Semestr 3	ECTS – Kreditlar 4	
Fan/modul turi Tanlov		Ta'lim tili O'zbek		Haftadagi dars soatlari 4	
1.	Fanning nomi	Auditoriya mashg'ulotlari (soat)	Mustaqil ta'lim (soat)	Jami yuklama (soat)	
	KORPUS LINGVISTIKASI	60	60	120	

### Fanning mazmuni

**Fanni o'qitishdan maqsad** - talabalarni korpus lingvistikasining nazariy masalalari (tushunchalar bilan tanishish, sohaning amaliy masalalarini yechish usullari va boshqalar) bilan tanishtirish hamda korpus tuzishni o'rganish uchun poydevor yaratish, korpusdan foydalanish ko'nikmalarini shakllantirishdan iborat. Ayni paytda u talabalarni shu sohada mantiqiy fikrlashga, ilmiy xulosa chiqarishga, korpus lingvistikasi: korpus tuzish va foydalanish nazariy va amaliy muammolarini hal etish ko'nikmalarini shakllantirishga xizmat qiladi.

**Fanni o'qitishning vazifasi** - talabalarga korpus lingvistikasi nazariyasi va amaliyotiga oid bilim berish, nazariy bilimlarini amaliyotda qo'llay bilishga o'rgatish, kompyuter lingvistikasining o'zbek tilshunosligi sohalari bilan munosabatlarini bilish, ularga xos hodisalarning modellarini tuzish, tekshirish va tushuntira olishga o'rgatishdir.

### Asosiy nazariy qism (ma'ruza mashg'ulotlari)

Fan tarkibiga quyidagi mavzular kiradi: (3-semestr)

**1-mavzu.** Kirish. Korpus lingvistikasi fani obyekti, predmeti, maqsad va vazifalari. "Korpus lingvistikasi" fanining predmeti, maqsad, vazifalari. Korpus lingvistikasining tilshunoslik fanlari orasida tutgan o'mni. Korpus lingvistikasi obyekti, predmeti.

**2-mavzu.** Korpusning amaliy ahamiyati. Korpusning lingvodidaktika, lingvistikada foydalanish ahamiyati, tadqiqotning korpus metodi, matnni avtomatik qayta ishlash, tarjima dasturi. Korpusning leksikologiya, leksikografiya, morfologiya, sintaksis, xorijiy tilni o'rganishdagi ahamiyati.

**3-mavzu.** Korpus lingvistikasining asosiy tushunchalari. Lingvistik korpus, elektron kutubxona, korpus menedjeri, konkordans, korpus annotatsiyasi, korpus hajmi va klassifikatsiyasi.

**4-mavzu.** Til korpusi turlari va tipologiyasi. Nutq turiga ko'ra og'zaki, yozma, aralash korpus. Matn tiliga ko'ra ruscha, inglizcha va h. korpus. Parallelligiga ko'ra bir tilli, ikki tilli, ko'p tilli korpus. Matnning ixtisoslashuviga ko'ra badiiy, dialektal, suhbat, terminologik, aralash korpus. Korpusning izoh xususiyati, o'zgaruvchan va

5-turg'unligiga ko'ra, strukturasi, matn hajmi, xronologik jihatiga ko'ra turliari tavsifi.

**6-mavzu.** Jahon korpus lingvistikasining bugungi holati: milliy korpuslar va unga qo'yilgan talablar. Rus va ingliz tili korpuslari tarixi, korpus lingvistikasining bugungi holati, zamonaviy rus, ingliz, turk, tojik tili korpuslarining mushtarak, farqli tomonlari.

**7-mavzu.** Rus va ingliz korpus lingvistikasi: tarix va bugun. Rus korpus lingvistikasi tarixi. Elektron va annotatsiyalangan korpuslarning shakllanish bosqichlari. Rus tili korpuslarining yaratilishi. Rus tili milliy va Upsalla korpusi haqida.

**8-mavzu.** Rus tili og'zaki nutq korpusi. Rus tili maxsus mantlar korpusi.

**9-mavzu.** Rus tili og'zaki nutq korpusi. Tuzilishi va interfeysi, o'ziga xos xususiyati vazifasi va lingvistik, lingvodidaktik, ijtimoiy ahamiyati. Korpus tarkibi, reprezentativligi, razmetkaning o'ziga xos turi. Rus tili maxsus mantlar korpusi. Tuzilishi va interfeysi, o'ziga xos xususiyati vazifasi va lingvistik, lingvodidaktik, ijtimoiy ahamiyati. Korpus tarkibi, materiali, reprezentativligi, razmetkaning o'ziga xos turi.

**10- mavzu.** Turkiy korpus lingvistikasi tadriji va bugungi holati. Turkiy tillardayaratilgan elektron lug'at, korpus, elektron atlas va platformalar. Korpus ishlab chiqish bo'yicha amalga oshirilgan loyihalar. Turkiy til korpus lingvistikasini rivojlantirish yo'lidagi harakatlar.

### **Amaliy mashg'ulotlar bo'yicha ko'rsatma va tavsiyalar Amaliy mashg'ulotlar uchun quyidagi mavzular tavsiya etiladi:**

1. Internetdan korpus saytlarini izlash. Korpus interfeyslari bilan tanishish.
2. Milliy korpuslar qidiruv funksiyasi ustida ishlash.
3. Turkiy tillardagi korpuslardan ma'lumot olish.
4. Ekstralingvistik va lingvistik teg/parametrlar asosidagi qidiruv.
5. Grammatik shakllar qidiruvi.
6. Bir xil semantik tavsifga ega so'zlarni topish.
7. Matnni muallif, janr, yaratilish joyi va vaqtiga ko'ra qidirish.
8. Korpusda nasriy va nazmiy matnlar qidiruvi.
9. Muayyan faoliyat sohasining nobadiiy matnlarini tanlash.

### **Mustaqil ta'lim va mustaqil ishlar**

Fan bo'yicha mustaqil ta'limni tashkil etishda quyidagi shakllardan foydalanish tavsiya etiladi:

- ✓ darslik va o'quv qo'llanmalar bo'yicha fan boblari va mavzularini o'rganish;
- ✓ tarqatma materiallar bo'yicha ma'ruzalar qismini o'zlashtirish;
- ✓ maxsus adabiyotlar bo'yicha fanlar bo'limlari yoki mavzulari ustida ishlash;
- ✓ talabning o'quv-ilmiy-tadqiqot ishlarini bajarish bilan bog'liq bo'lgan fanlar



- bo'limlari va mavzularni chuqur o'rganish;
- ✓ o'rganilmagan, muammoli mavzular bo'yicha tezis, maqolalar tayyorlash;
- ✓ badiiy asarlarni o'qish va mutolaa qilish;
- ✓ faol va muammoli o'qitish uslubidan foydalaniladigan o'quv mashg'ulotlari.
- ✓ masofaviy (distsion) ta'lim.

### **Mustaqil ta'lim uchun tavsiya etiladigan topshiriqlar:**

1. Korpus interfeyslarini tahlil qilish. Milliy korpuslar qidiruv funksiyasini qiyoslash: umumiy, farqli, afzal tomonlarni topish.
2. Turkiy tillardagi korpuslarning imkoniyatlari.
3. Ekstralingvistik va lingvistik teglar.
4. Lingvistik ta'minot tayyorlash

### **I. Fan tarkibiga quyidagi mavzular kiradi: (4-semestr)**

**1-mavzu.** Parallel korpuslar. Parallel korpuslarning tipologiya, tarjimonlik nazariyasi va xorijiy tilni o'rganishdagi ahamiyati. Parallel korpusning tuzilishi, interfeysi. Korpus tarkibi, materiali, reprezentativligi, razmetkaning o'ziga xos turi. Parallellik xususiyati, tillarning morfologik mutanosiblik darajasi.

**2-mavzu.** Mualliflik korpuslari. Mualliflik korpusi tuzishning lingvistik asoslari. Mualliflik korpuslarining mushtarak va farqli jihatlarini. Mualliflik korpusining tuzilishi, tarkibi, hosilasi, rus va ingliz tilida mavjud mualliflik korpuslarining o'xshash hamda farqli jihatlarini.

**3-mavzu.** Korpus tuzish: bosqichlar va reprezentativlik. loyihalash va tuzish jarayonlari. Matn yig'ish va unga avtomatik ishlov berish. Konversiyalash, grafematik tahlil, noleksik elementlarni belgilash. Til korpusida reprezentativlik. Korpus birliklarining janriy va miqdoriy ko'rsatkichi.

**4-mavzu.** Tabiiy tilni qayta ishlash. Razmetka texnologiyalari. Tokenizatsiya, lemmatizatsiya, stemming, parsing. Teglashning korpus tuzishdagi ahamiyati. Teg va uning turlari. Matnning avtomatik razmetkasi: tagging va parsing.

**5-mavzu.** Annotatsiya turlari: lingvistik va ekstralingvistik razmetka. Teg lingvistik vosita sifatida. SGML, XML tili. Lingvistik va ekstralingvistik razmetka. Lingvistik razmetka turlari.

**6-mavzu.** Korpus qidiruv tizimi sifatida. Korpus menejeri va interfeys. Korpus menejerining qidiruv, saralash, filtrlash vazifasi va uning BONITO, XAIRA, SARA, SQR, DDS va boshqa turlari haqida ma'lumot.

**7-mavzu.** Korpus menejeri va interfeys. Konkordans va interfeys, uning turlari haqida ma'lumot berish.

**8-mavzu.** Korpus asoslanuvchi leksikografik, grammatik va diskurs tadqiqotlar. Milliy korpusga asoslangan tarjima. Korpus asoslanuvchi leksikografik tadqiqotlar haqida ma'lumot. Leksikografik tadqiqotlar ko'lami, korpus vositasida leksikografik tadqiqotlarning aniqligini Leksikografik tadqiqotlardagi dolzarb muammolar.

III. Amaliy mashg'ulotlar bo'yicha ko'rsatma va tavsiyalar Amaliy mashg'ulotlar uchun quyidagi mavzular tavsiya etiladi:

1. Rus tili og'zaki nutq korpusi ustida ishlash.
2. Parallel korpus ustida ishlash.

3. Poetik korpus ustida ishlash.
4. Milliy korpus asosida ilmiy tadqiqot uchun material yig'ish.
5. Mualliflik korpusi ustida ishlash.
6. Mualliflik korpusi asosida manba yig'ish.
7. Maqsadli shaxsiy korpus yaratish.
8. AntConc dasturi yordamida korpus tuzish

#### IV. Mustaqil ta'lim va mustaqil ishlar oshirish.

Fan bo'yicha mustaqil ta'limni tashkil etishda quyidagi shakllardan foydalanish tavsiya etiladi:

- ❖ darslik va o'quv qo'llanmalar bo'yicha fan boblari va mavzularini tarqatma materiallar bo'yicha ma'ruzalar qismini o'zlashtirish, o'rganish;
- ❖ maxsus adabiyotlar bo'yicha fanlar bo'limlari yoki mavzulari ustida ishlash;
- ❖ talabani o'quv-ilmiy-tadqiqot ishlarini bajarish bilan bog'liq bo'lgan fanlar bo'limlari va mavzularni chuqur o'rganish;
- ❖ o'rganilmagan, muammoli mavzular bo'yicha tezis, maqolalar tayyorlash;
- ❖ badiiy asarlarni o'qish va mutolaa qilish;
- ❖ faol va muammoli o'qitish uslubidan foydalaniladigan o'quv mashg'ulotlari.

masofaviy (distantion) ta'lim.

1. Korpus interfeysi loyihasi ishlan chiqish.
2. Til korpuslari tahlili, tavsifi yoki korpus tuzish muammolari bo'yicha tahliliy maqola tayyorlash.
3. Mavjud korpus menejeri yordamida tanlangan soha (mavzu) korpusini tuzish
4. Lingvistik ta'minot tayyorlash.

#### Fanni o'zlashtirish natijasida talaba:

- 1) "Korpus lingvistikasi"ning fan sifatidagi o'rni, uning asosiy tushunchalarini umumlashirish, til korpuslarining jahon tilshunosligi va o'zbek tilshunosligida o'rganilish holati, korpus tuzishning nazariy asoslari to'g'risida tasavvur va bilimga ega bo'lishi;
  - 2) til korpusi bilan ishlash ko'nikmalariga ega bo'lishi;
  - 3) til korpusidan ma'lumot qidirish, materilalami ajratib olish, milliy, parallel, mualliflik korpuslari bilan ishlash malakalariga ega bo'lishi kerak.
- 1) ma'ruza;

#### Ta'lim texnologiyalari va metodlari:

- 2) interfaol keys-stadi;
- 3) seminar (mantiqiy fiklash, tezkor savol-javob);
- 4) guruhda ishlash;
- 5) taqdimot qilish;
- 6) individual loyihalar;
- 7) jamoa bo'lib ishlash va himoya qilish uchun loyihalar.

#### VII. Kreditlarni olish uchun talablar:

Fanga oid nazariy va uslubiy tushunchalarni to'la o'zlashtirish, tahlil natijalarini to'g'ri aks ettira olish, o'rganilayotgan jarayonlar haqida mustaqil mushohada yuritish va nazorat shakllarida berilgan vazifa hamda topshiriqlarni bajarish, yakuniy nazorat bo'yicha yozma ish topshirish.

## O'QUV ADABIYOTLARI VA AXBOROT MANBALARI

### Asosiy adabiyotdar:

1. Zaxarov V., Mengliyev B., Hamroyeva Sh. Korpus lingvistikasi. O'quv qo'llanma. GlobeEdit, 2021.-198 b.
2. Пулатов А. Компьютер лингвистикаси. Тошкент: Akademnashr, 2011.
3. Хамроева Ш.М. Корпус лингвистикаси атамаларининг кискача изоҳли лугати. - Тошкент: Камалак, 2018. - 96 б.

### Qo'shimcha adabiyotlar:

4. Ўзбекистон Республикаси Президентининг –Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси" тўғрисидаги Фармони ("Халқ сўзи", 2017 йил, 8 февраль).

5. Абдурахмонова Н.З. Инглизча матиларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти. Монография. Тошкент, 2018. - 152б.

6. Абжалова М.А. Ўзбек тилидаги матиларни тахрир ва таҳлил қилувчи дастурнинг лингвистик модуллари (Расмий ва илмий услублаги матнлар тахрири дастури учун). Филол.фан бўйича фалсафа доктори (PhD) диссер. - Фарғона, 2019.- 164 б.

7. Информационные технологии в лингвистике: учеб. Пособие / Л.Ю.Щипицина. -М.: ФЛИНТА: Наука, 2013.-128 с.

8. Захаров В.П., Богданова С.Ю. Корпусная лингвистика. Иркутск ИГЛУ, 2020. - 161 с.

9. Структурно-функциональная компьютерная модель татарских морфем. Казань: Фэн, 2003. - 220 с.

10. Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол.фан.бўйича фалсафа доктори (PhD) диссер. Карши, 2018. - 250 б.

11. Хамроева Ш. М. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари. Монография. Globe edit: ISBN-13: 978-620-0-51507- 0; ISBN-10: 6200515077.-2020.-260 б.

12. Хамроева Ш. М. Ўзбек тили морфологик анализаторининг лингвистик таъминоти. Монография. Globe edit: 2021. - 250 б.

13. Эшмунинов А.А. Ўзбек тили миллий корпусининг синоним сўзлар базаси: Филол.фан бўйича фалсафа доктори (PhD) диссер. Карши, 2019. 140 б

### Internet manbalari:

14. <http://khnt.hit.uib.no/icame/manuals/lobman/>
15. [www.ruscorpora.ru](http://www.ruscorpora.ru)
16. <http://www.ucl.ac.uk/english-usage>
17. <http://www.collins.co.uk/Corpus/CorpusSearch.aspx>
18. <http://www.pearsonlongman.com/dictionaries/corpus/index.html>
19. <http://www.natcorp.ox.ac.uk/>
20. <http://www.ucl.ac.uk/english-usage/ice/> <http://www.americannationalcorpus.org/>

21. <http://www ldc upenn edu>
23. <http://corpora phil spbu ru>
24. [www slaviska uu se/korpus htm](http://www slaviska uu se/korpus htm)
25. [www sfb441 uni-tuebingen de/b1/rus/korpora html](http://www sfb441 uni-tuebingen de/b1/rus/korpora html)
26. [www philol msu ru/~lex/corpus](http://www philol msu ru/~lex/corpus)
27. <http://www philol msu ru/~lex/corpus>
24. <http://www ruscorpora ru> HKPS
27. <http://ruscorpora ru>
28. <http://ruscorpora ru/search-murco html>
29. <http://www ling helsinki fi/projects/hanco/>
30. <http://corpus leeds ac uk/>
40. <http://corpus leeds ac uk/ruscorpora html>
41. <http://aot ru/search1 html>

**Termiz davlat universiteti tomonidan ishlab chiqilgan va tasdiqlangan.**

O'quv dasturi Termiz davlat universiteti o'quv-uslubiy Kengashining 2024-yil "26" - 06 dagi 11 -sonli yig'ilish bayoni bilan tasdiqlangan.

O'zbek filologiyasi fakultetining 2024-yil "12" - 06 dagi 10 -sonli yig'ilish bayoni bilan ma'qullangan.

"O'zbek tilshunosligi" kafedrasining 2024-yil "7" - 06 dagi 20 -sonli yig'ilish bayoni bilan ma'qullangan.

**Fan/modul uchun mas'ullar:**

M.Xolova – Termiz davlat universiteti "O'zbek tilshunosligi" kafedra katta o'qituvchisi, f.f.f.d (PhD)

**Taqrizchilar:**

**Sh.M.Hamroyeva** - Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti dotsenti, filologiya fanlari doktori (DSc)

**Sh.Gulyamova** - Turkiston yangi innovatsiyalar universiteti dotsenti, filologiya fanlari doktori (DSc)