

**ЎЗБЕКИСТОН РЕСПУБЛИКАСИ ВАЗИРЛАР МАҲКАМАСИ ҲУЗУРИДАГИ ДАВЛАТ  
ТИЛИНИ РИВОЖЛАНТИРИШ ДЕПАРТАМЕНТИ**

**ЎЗБЕКИСТОН РЕСПУБЛИКАСИ ОЛИЙ ВА ЎРТА МАХСУС ТАЪЛИМ ВАЗИРЛИГИ**

**ЎЗБЕКИСТОН РЕСПУБЛИКАСИ ХАЛҚ ТАЪЛИМИ ВАЗИРЛИГИ**

**ЎЗБЕКИСТОН РЕСПУБЛИКАСИ ИННОВАЦИОН РИВОЖЛАНИШ ВАЗИРЛИГИ**

**АЛИШЕР НАВОИЙ НОМИДАГИ  
ТОШКЕНТ ДАВЛАТ ЎЗБЕК ТИЛИ ВА АДАБИЁТИ УНИВЕРСИТЕТИ**

**“ЎЗБЕК МИЛЛИЙ ВА ТАЪЛИМИЙ  
КОРПУСЛАРИНИ ЯРАТИШНИНГ НАЗАРИЙ  
ВА АМАЛИЙ МАСАЛАЛАРИ”**

**мавзусидаги  
халқаро илмий-амалий конференция  
материаллари  
(2021 йил 7 май)**

## O‘ZBEK TILI MILLIY KORPUSIDA SINONIMAYZER YOXUD SINONIMIZATORNI YARATISH MASALASI

### CREATION OF A SYNONYMIZER OR SYNONYMIZER IN THE NATIONAL CORPUS OF THE UZBEK LANGUAGE

Абжалова Манзура Абдурашметовна\*

**Annotatsiya:** Bugungi kunda tilni o'rganish, uni o'rgatish masalasi global hisoblanmoqda. Turli lug'atlar, mobil ilovalar va kompyuter dasturlari bo'lishiga qaramasdan, mushtarak platformaga ehtiyoj sezilmoqda. Ana shu ehtiyojni qondirish maqsadida yaratilayotgan O'zbek tili Milliy korpusi ta'lim samaradorligini oshirish, lingvodidaktik qulay imkoniyatlarni berish, til o'rganishni so'z boyligini oshirishni osonlashtirish kabi masalalar yechimini beradi. Uning bazasida o'quv lug'atlarining mavjud bo'lishi korpusning mushtarak lingvo-axborot tizimi bo'lib xizmat qilishini ta'minlaydi. Bazada sinonimlar lug'ati ham kiritilgan bo'lib, bu korpusning sinonimayzer vazifani ham o'tashiga zamin yaratadi.

Mazkur maqolada sinonimayzer, uning ishlash prinsipi va O'zbek tili Milliy korpusida sinonimizatorning dastlabki jarayoni haqida so'z yuritildi.

**Kalit so'zlar:** O'zbek tili, Milliy korpus, sinonimayzer, sinonimizator, ma'nodosh so'zlar.

Maqola AM-FZ-201908172 raqamli “O'zbek tili ta'limiy korpusini yaratish” mavzusidagi amaliy loyiha doirasida bajarildi.

**Annotation:** Language learning is a global problem today. Despite the existence of various dictionaries, mobile applications and computer programs, there is a need for a common system. The National Corpus of the Uzbek Language, created to meet this need, deals with issues such as improving the effectiveness of education, providing convenient language and didactic opportunities, and promoting language learning and vocabulary development. The presence of dictionaries in its database ensures that the corpus serves as a common linguistic information system. The database also includes a synonym dictionary that allows the corpus to act as a synonymizer.

This article discusses the synonymizer, how it works, and the initial synonymizer process in the Uzbek National Corpus.

**Keywords:** Uzbek language, national corpus, synonymizer, synonymizer, synonyms.

The article was implemented within the framework of the practical project AM-FZ-201908172 "Creation of the educational corpus of the Uzbek language".

O'zbek tili so'zga boy til deyiladi. Shu o'rinda savol tug'iladi: tilning boyligi so'zlar miqdori bilan o'lchanadimi? Yoki uning tarixiyli bilanmi? Balki yana boshqa mezonni bordir. Albatta bor. Tilning boyligi har bir so'zning nutq uslublari (so'zlashuv/dialektal, badiiy, rasmiy, ilmiy, publitsistik)da muqobil ma'nodoshlari borligi bilan belgilanadi. Mana shundagina so'zlovchi o'z nutqini ham tushunarli, ham jozibali, ham rasmiy muhitga muvofiq, ham ilmiy jihatdan ravon, ham erkin mulohazali fikr yuritadi. Shuni e'tiborga olib, O'zbek tili Milliy korpusida sinonimayzer imkoniyati yaratildi.

Sinonimayzer – bu matndagi so'z birikmalari yoki so'zlarni sinonimlar (dastur bazasida joylashgan) bilan almashtiradigan dastur [1]. Ayrim manbalarda sinonimizator, sinomizator nomlari bilan yuritiladi. Ushbu atama matnlarni qayta yozishni qulaylashtirishi mumkin bo'lgan dasturni tavsiflash uchun birinchi marta SEO-forumda taxallusi Kevindark bo'lgan kopirayter tomonidan kiritilgan.

Sinonimayzer matnni unikal (betakror) qilib o'zgartirishga xizmat qiladi. Sinonimizator dasturi yoki Internetdagi veb-sayt bo'lishi mumkin. Jahon kompyuter lingvistikasida sinonimizatorlar, asosan, veb-saytlar uchun o'ziga xos tarkib yaratish uchun qayta yozish (rerayting) va kopiraytinglar (reklama matnlarini yozish sohasi)da qo'llaniladi.

\*Филология фанлари бўйича фалсафа доктори ( PhD), Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университети, [manzura\\_ok@mail.ru](mailto:manzura_ok@mail.ru)

### Sinonimayzerlarning ishlash prinsipi.

Sinonimayzerlar avtomatik va yarimavtomat bo‘linishi mumkin. Yarimavtomat sinonimayzerlar foydalanuvchiga u belgilab olgan so‘zga mos sinonimlar ro‘yxatidan muvofiq‘ini tanlab olishni taklif qiladi. To‘liq avtomatik sinonimayzerlar matndagi so‘zlarni foydalanuvchi ishtirokisiz ularning ma‘nodoshlari bilan avtomatik tarzda almashtiradi.

Avtomatik sinonimizatorlarga xos matn almashtirishlarning ikki turi mavjud: nomorfologik (qatorli) va morfologik (lug‘atli) [2].

Nomorfologik (qatorli) almashtirish matndagi belgilangan so‘zning dastur backendi, ya‘ni ma‘lumotlar bazasida ikkala tomonida harfiy simvollar hisoblanmagan belgilar bo‘lgan so‘zga uning o‘sha shakliga mos ma‘nodoshiga almashtirish jarayoni e‘tiborga [3] olinadi. Aniqroq tushuntirilganda, bu jarayonda faqat sinonimlarning izohli lug‘ati [4, 6] va yordamchi lug‘atlar [5,7,8]ga tayaniladi. So‘zshakli e‘tiborga olinmaydi. Masalan, agar so‘ziga uning mabodo [f-t], magar [f-t], basharti [f-t+a], bordiyu kabi ma‘nodoshlari taqdim etiladi (1-rasm).



Bosh sahifa

Lug‘atlar

Kirish

agar

So‘z tarkibi: [agar]

Antonim(lar): mavjud emas.

Sinonim(lar): mabodo [f-t], magar [f-t], basharti [f-t+a], bordiyu

- ☐ Izohli lug‘at
- ☐ Omonimlar lug‘ati
- ☒ Sinonimlar lug‘ati
- ☒ Antonimlar lug‘ati
- ☐ Paronimlar lug‘ati

#### 1-rasm

Ammo o‘rin-payt kelishigi affiksini qo‘shib agarda tarzida so‘zshakli qidiruv satriga yozib, qidiruvga berilsa ma‘lumot taqdim etilmaydi (2-rasm)



Bosh sahifa

Lug‘atlar

Kirish

agarda

So‘z tarkibi: [agarda]

Antonim(lar): mavjud emas.

Sinonim(lar): mavjud emas.

- ☐ Izohli lug‘at
- ☐ Omonimlar lug‘ati
- ☒ Sinonimlar lug‘ati
- ☒ Antonimlar lug‘ati
- ☐ Paronimlar lug‘ati

#### 2-rasm

*Morfologik* (lug‘atli) almashtirishlar so‘zning grammatik shaklini ham hisobga oladi va har qanday so‘zshakli o‘zining muvofiq shaklida ma‘nodoshi bilan almashtiriladi. Ba‘zi dasturlar nafaqat so‘zlarni, balki tarkibi o‘nta so‘zgacha bo‘lgan so‘z birikmalari yoxud iboralarni ham o‘zgartirishi mumkin.

Ba‘zi sinonimizatorlar almashtirishning ikkala turini ham amalga oshirishi mumkin. Agar dastur morfologiyani tushunish imkoniyatiga ega bo‘lmasa, unda ma‘lumotlar bazasi qo‘lda yoki turli xil yordamchi dastur (utilit)lar yordamida har xil so‘zshakllarini lemmalash yoki stemlash imkoniyati natijasida so‘zshakllarga ham sinonimlarni taqdim etish yoxud ma‘nodoshlariga almashtirish imkoniyati yuzaga keladi. Bunday dasturlar “turlovchi/tuslovchilar” deb ham nomlanadi.

O‘zbek tili Milliy korpusida sinonimayzerning nomorfologik turi ishga tushirilgan (3-rasm) bo‘lib, bu o‘zbek tili uchun yaratilgan 1-sinonimayzer hisoblanadi. Kelajakda uning morfologik turi bo‘yicha so‘zshakllarni ularning mos shakliga almashtirish imkoniyatini yaratish uchun ma‘lumotlar bazasini mukammallashtirilishi ustida ish olib borilmoqda. 3-rasmda O‘zbek tili Milliy korpusining “Lug‘atlar” rukni qidiruv satriga kiritilgan katta so‘ziga korpus sinonimayzeri tomonidan uning ma‘nodoshlari va, hatto, zid ma‘nolari (antonimlari) taqdim etildi.

katta

Izlash

- ☐ Izohli lug'at
- ☐ Omonimlar lug'ati
- ☒ Sinonimlar lug'ati
- ☒ Antonimlar lug'ati
- ☐ Paronimlar lug'ati

So'z tarkibi: [katta]

Antonim(lar): chog', kichik, kichkina, mayda, yosh

Sinonim(lar): azamat [a], azim [a], bahaybat [f-t+a], buyuk, gigant [r<grek],  
haybatli, ulkan, ulug', yirik, boshliq, zo'r

3-rasm.

**Sinonimayzelardagi muammo.** Asosiy muammolar quyidagilardan iborat:

- Sinonimlarni tanlash kontekstga bog'liq bo'lishi kerak.
- Gapning inversiyaga uchrashi sintaktik analizator ishini qiyinlashtiradi va natijada so'z birikmalarini ularning ma'nodosh birikmalariga almashtirish imkoniyati yo'qoladi.
- Tilga kirib kelgan yangi so'z (termin)lar bazaga kiritilmagan bo'lsa, neologizm ma'nodoshlar foydalanuvchiga taqdim etilmay qoladi.

Sinonimayzer tomonidan qayta ishlangan matn uslubi ravon bo'lishi uchun dasturda Zipfa qonuni hisobga olinishi kerak, ya'ni faol so'zlar faol ma'nodoshlari bilan, kam uchraydigan so'zlar - kam uchraydigan sinonimlari bilan almashtirilishi kerak.

Sinonimizator yordamida qayta ishlangan matn boshqa lingvistik kompyuter dasturlarida grammatik, sintaktik va semantik qo'shimcha ishlov berishni talab qiladi. Qayta ishlash jarayonida matnlar bir-biri bilan taqqoslanadi va keyin eng muvofiq varianti tanlanadi.

#### ADABIYOTLAR:

1. Abjalova M.A. O'zbek tili milliy korpusida so'zshaklni leksikografik baza asosida qidiruv imkoniyatlari. // Kompyuter lingvistikasi: muammolar, yechim, istiqbollar/ Respublika ilmiy-texnik konferensiya to'plami. Elektron nashr /ebook. – Toshkent: ToshDO'TAU, 23.04.2021. – B.12-17.
2. Hojiyev A. O'zbek tili sinonimlarining izohli lug'ati. – Toshkent: O'qituvchi, 1974. – 308 b.
3. O'zbek tilining izohli lug'ati: 80 000 dan ortiq so'z va so'z birikmasi (A. Madvaliyev tahriri ostida). 5 jildli. – Toshkent: O'zbekiston milliy ensiklopediyasi, 2006.
4. Shukurov O., Boymatova B. O'zbek tilining ma'nodosh so'zlar o'quv lug'ati: Maktab o'quvchilari uchun. – Toshkent: Yangi asr avlodi, 2007. – 49 b.
5. Bobojonov Sh., Islomov I. O'zbek tilining so'zlar darajalanishi o'quv lug'ati: Maktab o'quvchilari uchun. – Toshkent: Yangi asr avlodi, 2007. – 38 b.
6. Mengliev B., Shahabitdinova Sh., Khamroeva Sh., Gulyamova Sh., Botirova A. The morphological analysis and synthesis of word forms in the linguistic analyzer // Linguistica Antverpiensia, 2021, 2021(1), стр. 703-712.
7. Khamroeva Sh. Morphotactic rules in the morphological analyzer of the uzbek language // “MIDDLE EUROPEAN SCIENTIFIC BULLETIN”. VOLUME 6, NOVEMBER 2020. – ISSN 2694-9970 – P. 45-49.
8. Хамроева Ш. Корпус лингвистикаси атамаларининг изоҳли луғати. Globe Edit, 2020. – 96 b.
9. Qurbonova M. Abjalova M. va boshq. O'zbek tili o'zlashma so'zlarining urg'uli lug'ati. [Matn]: o'quv-uslubiy lug'at / M.Qurbonova, M.Abjalova, N.Axmedova, R.To'laboyeva. – Toshkent: Nodirabegim, 2021. – 988 b.
10. <https://xn----7sbbaqh1km9ah9aiq.net/news-new/sinonimizator-ru.html>
11. <https://animatika.ru/info/gloss/synonymizer.html#:~:text>