



IJTIMOIIY-GUMANITAR SOHADA ILMIY-INNOVATSION TADQIQOTLAR

ILMIY METODIK JURNALI

ISSN 3060-5059



VOL.3 № 4

2026

JADID MEROSINI KORPUSGA: ELBEK–FITRAT LUG‘ATCHILIGI, OMONIMIYA MEZONLARI VA DRAMALAR LEKSIKASINI SEMANTIK TEGGLASH MODELI

Rustamov Doniyorbek Avazbekovich
Toshkent amaliy fanlar universiteti, PhD

Annotatsiya

Maqola jadidlar merosini leksikografik va korpus lingvistik yondashuv asosida qayta talqin etadi. Unda Elbek, Abdurauf Fitrat va boshqa ma’rifatparvarlarning lug‘atshunoslik ishlari (EOL, Fitratning etimologik izohlari va boshqalar) asosida davr tilining fonetik–grafik, leksik va grammatik xususiyatlari tavsiflanadi. Sh. Bobojonova tahliliga tayangan holda EOL (416 birlik) va ROL (497 birlik) o‘rtasidagi tafovutlar, shuningdek, 1929-yilgi lotin yozuviga o‘tish va singarmonizm asosidagi imlo me’yorlari tufayli omonimlarni belgilash mezonlarini (talaffuz va yozuv birligi, ma’no tafovuti) qayta ko‘rib chiqish zarurligi asoslanadi. Fitratning turkiy leksemalarga bergan etimologik izohlari (masalan, yitika, mung‘ilamak, qub) jadid lug‘atchiligida manbalararo uyg‘un yondashuv mavjudligini ko‘rsatadi. Behbudiy, Avloniy, Qodiriy, Cho‘lpon dramalari asosida leksikani tematik guruhlariga (ijtimoiy-siyosiy, diniy-axloqiy, maishiy, ta’lim, matbuot, o‘lchov va boshqalar) ajratish, arxaizm va istorizm belgilarini qo‘llash hamda “jadid dramalari terminlari”ni semantik teglash taklif etiladi. Mualliflik korpusi uchun lemmatizatsiya, token–lemma–birikma–konkordans asosida qidiruv tizimi, operator va konstanta teglar, sinonim va antonim qatlamlari hamda razmetka modeli ishlab chiqiladi. Natijada jadid dramalari asosida izohli–kollokatsion hamda tematik–etimologik lug‘atlar majmuasini yaratish istiqbollari aniqlanadi.

Kalit so‘zlar: jadidlar, leksikografiya, Elbek, Abdurauf Fitrat, EOL, ROL, omonim, 1929 lotin yozuvi, singarmonizm, etimologiya, korpus lingvistika, semantik teglash, lemmatizatsiya, jadid dramaturgiyasi, arxaizm/istorizm, izohli–kollokatsion lug‘at.

К КОРПУСУ ДЖАДИДСКОГО НАСЛЕДИЯ: ЛЕКСИКОГРАФИЯ ЭЛЬБЕКА– ФИТРАТА, КРИТЕРИИ ОМОНИМИИ И МОДЕЛЬ СЕМАНТИЧЕСКОЙ РАЗМЕТКИ ЛЕКСИКИ ДРАМ

Рустамов Дониёрбек Авазбекович
Ташкентский университет прикладных наук, PhD

Аннотация

В статье переосмысливается наследие джадидов на основе лексикографического и корпусно-лингвистического подходов. На материале работ Эльбека, Абдурауфа Фитрата и других просветителей (EOL, этимологические комментарии Фитрата и др.) описываются фонетико-графические, лексические и грамматические особенности языка эпохи. С опорой на анализ Ш. Бобожиновой обосновывается необходимость пересмотра критериев выделения омонимов (тождество произношения и написания, различие значений) в связи с различиями между EOL (416 единиц) и ROL (497 единиц), а также влиянием латинизации 1929 года и орфографических норм, основанных на сингармонизме. Этимологические толкования тюркских лексем у Фитрата (например, yitika, mung‘ilamak, qub) демонстрируют межисточниковый синтез в джадидской лексикографии. На примере драм Бехбуди, Авлони, Кадыри и Чулпана предлагается тематическая классификация лексики (социально-политическая, религиозно-нравственная, бытовая, образовательная, газетно-публицистическая, измерительная и др.), использование помет архаизмов и историзмов, а также модель семантической разметки «терминов джадидской драматургии». Для авторского корпуса предлагаются инструменты лемматизации,

поисковые механизмы по схеме токен–лемма–сочетание–конкорданс, система операторов и константных тегов, слои синонимии и антонимии, а также модель разметки. В результате определяются направления создания комплекса толково-коллокационных и тематико-этимологических словарей джадидской драматургии.

Ключевые слова: джадиды, лексикография, Эльбек, Абдурауф Фитрат, EOL, ROL, омоним, латиница 1929 года, сингармонизм, этимология, корпусная лингвистика, семантическая разметка, лемматизация, джадидская драматургия, архаизм/историзм, толково-коллокационный словарь.

TO THE CORPUS OF JADID HERITAGE: ELBEK–FITRAT LEXICOGRAPHY, CRITERIA OF HOMONYMY, AND A MODEL FOR SEMANTIC TAGGING OF DRAMATIC LEXIS

Rustamov Doniyorbek Avazbekovich
Tashkent University of Applied Sciences, PhD

Abstract

The article reinterprets the heritage of the Jadids through lexicographic and corpus-linguistic approaches. Based on the works of Elbek, Abdurauf Fitrat, and other enlighteners (EOL, Fitrat's etymological explanations, etc.), the phonetic-graphic, lexical, and grammatical features of the language of the period are described. Drawing on Sh. Bobojonova's analysis, the need to reconsider the criteria for identifying homonyms (identity of pronunciation and spelling, difference in meaning) is substantiated due to discrepancies between EOL (416 units) and ROL (497 units), as well as the influence of the 1929 Latinization and orthographic norms based on vowel harmony. Fitrat's etymological interpretations of Turkic lexemes (e.g., yitika, mung'ilamak, qub) demonstrate a cross-source integrative approach in Jadid lexicography. Using the dramas of Behbudiy, Avloniy, Qodiriy, and Cho'lpon, the study proposes thematic classification of lexical units (socio-political, religious-ethical, everyday, educational, journalistic, measurement-related, etc.), the use of archaic and historical labels, and a model for semantic tagging of "Jadid drama terminology." For the authorial corpus, tools such as lemmatization, token–lemma–collocation–concordance search, operator and constant tags, synonym–antonym layers, and a tagging model are developed. The findings outline directions for creating a комплекс of explanatory-collocational and thematic-etymological dictionaries based on Jadid drama texts.

Keywords: Jadids, lexicography, Elbek, Abdurauf Fitrat, EOL, ROL, homonym, 1929 Latin script, vowel harmony, etymology, corpus linguistics, semantic tagging, lemmatization, Jadid drama, archaism/historicism, explanatory-collocational dictionary.

Jadidlar merosini o'rganish va ular bo'yicha ilmiy tadqiqotlar yaratish masalasida qator izlanishlar kuzatilgan. Xususan, jadidlar davrida Elbek, Abdurauf Fitrat kabi til fidoyilari tomonidan tuzilgan lug'atlar o'sha davr til holatini tasvirlash, zarur ma'lumotlarni yig'ish va umumlashtirish, shuningdek, bugungi kun bilan qiyosiy tahlil o'tkazish imkonini beradi. O'zbek leksikografiyasi shakllanishiga bag'ishlangan ishida S. E. Normatov jadidlar davri lug'atchiligi muammolarini keng ko'lamda tadqiq etib, Abdulla Qodiriy, Elbek, Abdurauf Fitrat, Ashurali Zohiriy, Ishoqxon Ibrat singari ma'rifatparvarlarning leksikografik qarashlarini yoritadi hamda davr til xususiyatlariga doir qimmatli dalillarni keltiradi. Shuningdek, M. M. Latipov Ishoqxon Ibrat asarlarining leksik-grammatik jihatlarini maxsus tahlil qiladi; mazkur tadqiqotda ham qator leksik birliklar mavzuviy guruhlariga ajratilib, etimologik nuqtayi nazardan izohlanadi. F. Bobojov esa o'zining o'zbek jadid dramalarining lisoniy xususiyatlariga bag'ishlangan dissertatsiyasida Abdulla Avloniy va Mahmudxo'ja Behbudiy dramalarini tahlil etadi. Ushbu tadqiqotlar natijalaridan kelib chiqib, jadid dramalari leksikasini semantik teglash mumkinligi ko'rsatib

beriladi. Elbekning lugʻatshunoslik faoliyatini oʻrgangan Sh. Bobojonova lingvistik terminlar, omonimlar va koʻp maʼnoli birliklarning Elbek talqinidagi shakllarini tahlil qiladi. Unga koʻra, “EOL (Elbekning Omonimlar Lugʻati)”da 416 ta, “ROL (Shavkat Rahmatullayev Omonimlar Lugʻati)”da esa 497 ta shakldosh birlik qayd etilgan. EOLDagi 178 ta omonim ROL tarkibida ham mavjud, qolgan 238 tasi esa quyidagi omillar sabab kiritilmagan: ROL omonimni belgilashda talaffuzdagi bir xillik, imlo jihatidan moslik va maʼnolarning keskin farqlanishi kabi uch mezonni bir vaqtda eʼtiborga oladi; EOLda esa asosan shakliy tenglik bosh mezon sifatida olingan. 1929-yilda lotin yozuviga oʻtish jarayonida singarmonistik tillargagina xos boʻlgan 9 unli asos sifatida qabul qilinib, yoʻgʻon-ingichka oppozitsiyalar boʻyicha imlo qoidasiga amal qilingan: yoʻgʻon soʻzlarda ingichka unilarga xos harflar (k, g, a), ingichka soʻzlarda esa yoʻgʻon unilarga xos grafemalar (q, gʻ, o) yozilmagan. Shu bois talaffuzda biroz farq qiladigan leksemalar yozuvda bir xil koʻrinishda berilgan. Masalan, bugungi adabiy tilimizda talaffuz va yozuvda farqlanuvchi olmos–olmas, boyroq–bayroq, qozi–qazi, chalgı–cholgʻu, chal–chol, chay–choy, barmoq–bormoq, olaman–olomon kabi juftliklar oʻsha davr imlosida bir xil koʻrinishda yozilgan va bu holat oʻsha zamonning imlo meʼyorlari uchun norma hisoblangan. Demak, Elbekning omonimlar lugʻati hozirgi oʻzbek adabiy tilidan fonetik, leksik va grammatik koʻrsatkichlari bilan farq qiladi; bu farqlar jadidlar yaratgan imlo konsepsiyalarini yangilashga doir urinishlarning mahsuli sifatida davr adabiy tilidan muayyan tafovutlarni yuzaga keltirganini koʻrsatadi.

Abdurauf Fitratning lugʻatshunoslik borasida alohida izlanishlari mavjud boʻlib, u turkiy leksemalarni izohlashda etimologik yondashuvga tayanadi. Masalan: Yatika – yitika, yetti yulduz; arabcha *banāt al-naʼsh* (“banotunnash”), tojikcha *haft dodaron* — yulduzlar majmuasini bildiradi (45). *Mungʻilamak (yamangʻilamak)* — “miya yemoq” maʼnosida: qadim turklar ziyofat paytida qoʻyning miyasini eng eʼzozli mehmoniga tortishgani bois, keyinchalik bu soʻz “mazali taom yemoq” maʼnosida ishlatilgan. Shuningdek, *sav* – sur; *qazirmoq* – qayirmoq, qaytarmoq; *qub* – taʼkid yuklovchisi (“juda ham”); *bulmoq* – topmoq; *kuvazlik* – kattalik, takabburlik, shodlik kabi birliklar ham izohlanadi. Bu leksemalar Mahmud Koshgʻariyning “Devonu lugʻotit turk”idagi birliklar boʻlib, Fitrat ularni turli manbalar va jonli xalq tili misollari bilan aniqlashtiradi. Fitratning tilshunoslik qarashlari M. Qurbonovanning tadqiqotida batafsil sharhlangan boʻlib, bu izlanishlar Fitratning tilshunos sifatidagi faoliyatini keng yoritadi. Shu bilan birga, Fitrat dramaturg sifatida ham sermahsul ijod qilgan; uning tarixiy dramalari tili va undagi leksemalarning semantik qirralari chuqur tahlilni talab qiladi.

Jadid maʼrifatparvarlaridan Mahmudxoʻja Behbudiy asarlarining til va uslub xususiyatlari D. Rahmatova tomonidan oʻrganilgan. Tadqiqotda asosan Behbudiyning publitsistik matnlari tili tahlil qilingan. Biroq Behbudiy dramalari ham oʻzbek dramaturgiyasining debochasi sifatida oʻziga xos til-uslubiy xususiyatlarga ega. Dramaturgiyadagi soʻzlarni mavzuviy guruhlariga ajratish va korpus formatida berishning ahamiyati shundaki, jadidlar yashab ijod qilgan davrda til va imlo masalalarida jiddiy islohotlar boshlandi; tilni yot unsurlardan asrash, adabiy tilni shakllantirish, jonli xalq tilining sinkretik tabiati aynan pyesalarda yorqin namoyon boʻladi. Shu bois jadid dramalarida ishlatilgan leksik birliklarning izohini alohida lugʻat tartibida tizimlash hamda ularni “jadid dramalari terminlari” sifatida semantik teglash maqsadga muvofiqdir.

Oʻzbek tili korpusi qidiruv tizimiga qoʻyiladigan asosiy talablar quyidagicha:

- soʻz va soʻz birikmalarini ularning grammatik, semantik va boshqa belgilariga koʻra qidirish;
- matn (tugal fikr ifodalovchi parcha) va birliklar orasidagi masofani (soʻzlararo masofani) inobatga olish;
- metamatnli maʼlumot boʻyicha qidiruvni qoʻllash;
- rivojlangan mantiqiy bogʻlovchilar, qavslar va matn operatorlarini qamrab olish;
- indeksatsiyalash samaradorligini taʼminlash;
- murakkab savollarga yetarli tezlikda javob qaytarish;
- keng koʻlamlilik, yaʼni yuz millionlab soʻzdan iborat resurslar bilan ishlay olish.

Korpusda izlash to‘rt yo‘l bilan amalga oshiriladi: token, lemma, birikma va konkordans bo‘yicha. Qidiruv jarayonida tanlangan usul bilan birga uslub va davr parametrini ham belgilash zarur. Token bo‘yicha izlash so‘z shakllarini topadi, shu sababli izlovga leksemaning aynan ishlatilgan shaklini kiritish kerak. Chunki Abdurauf Fitrat asarlarida (umuman, jadid adabiyotida) grammatik qo‘shimchalar hozirgi adabiy me‘yorlardan sezilarli farq qiladi. Masalan, *ot* so‘zini token bo‘yicha qidirish istagida foydalanuvchi jadid davrida uchragan *otg‘a* shaklini aynan shunday kiritib yuborsa, natija doirasi torayadi; bu birlik o‘sha davrda *otqa* tarzida ham qo‘llangan (“Қараш белгиси бир отқа қўшилғач, ундан кейин келган бир отға учинчи турли ‘олмош излари’ қўшиладир” [Abdurauf Fitrat, 2006:171]).

Bu mulohazalarga qo‘shimcha ravishda ta’kidlash lozimki, jadid matnlarida, ayniqsa dramalarda bir so‘zning turli morfologik allomorflari uchraydi: masalan, Abdulla Avloniy pyesalaridagi *borurg‘a* formasi Niyoziyning dramalarida *borarg‘a* shaklida uchrashi mumkin; shuningdek, *ishqig‘a–ishqiga* kabi parallel shakllanishlar mavjud. Tahlillar shuni ko‘rsatadiki, jadid dramalari leksemalari bir adabiy davrga mansub bo‘lsa-da, adabiy til va sheva nuqtayi nazaridan farqlanadi; morfologik ko‘rinishda adabiy me‘yor, sheva, fonetik variatsiya va kontekst birgalikda ta’sir ko‘rsatadi. Bu holat korpuslarda token bo‘yicha qidiruvni murakkablashtiradi. Shu sababdan jadid dramalari leksik fondini korpusga kiritishda sinonimik qatorlarni ham alohida belgilash muhim. Lemma bo‘yicha qidiruvda so‘zning lug‘aviy (asl) shakli topiladi; birikma bo‘yicha qidiruv esa Abdurauf Fitrat matnlarida uchraydigan tilshunoslik terminlarining kollokativ ko‘rinishlarini aniqlashga yordam beradi. Konkordans esa leksemaning qo‘llanish muhiti, ya’ni kollokatsion o‘rab turuvchi birliklarni aniq ko‘rsatadi. Jadid dramalarini mualliflik korpusi doirasida o‘rganish davriy, tilshunoslik va mazmun nuqtayi nazaridan katta ahamiyatga ega, biroq qidiruv tizimida ayrim murakkabliklarni tug‘diradi: jadid davri tili bilan hozirgi davr tili to‘liq mos kelmaydi, natijada foydalanuvchi kutilgan natijani olmasligi mumkin. Mualliflik korpusi sifatida ishlatilganda esa qidiruv aynan o‘sha davr til me‘yorlari asosida quriladi. Bundan tashqari, mazmun tomondan ham jadid dramalarini ajratib, alohida korpus sifatida tadqiq etish zarur. Yuqorida sanab o‘tilgan ilmiy ishlar esa jadidlar mualliflik korpuslarini yaratishda lingvistik ta’minot vazifasini bajaradi. Masalan, jadid dramalarida uchraydigan lavozim va unvon bildiruvchi so‘zlarni quyidagicha tartiblash mumkin:

– jadid dramaturgiyasiga oid asarlar tanlanadi (Behbudiy, Fitrat, Cho‘lpon, Avloniy va b.);

– matnlar elektron ko‘rinishga keltiriladi (faqat qog‘ozda bo‘lsa, raqamlashtiriladi);
– leksik birliklar ajratilib, tahlil qilinadi.

Keyingi bosqich:

– so‘z formalarini belgilash (lemmatizatsiya — so‘zning lug‘at shakliga keltirish);
– sintaktik va semantik belgilash (*ot*, *fe‘l*, sifati va h.k.).

Korpusga integratsiya:

– O‘zbek tili milliy korpusi ma’lumotlar bazasiga yuklash;
– qo‘shimcha belgilash va tasniflash (mavzu, davr, uslubiy belgilarga ko‘ra).

Misollarni jadval bilan ko‘rsatish:

Asar nomi	Muallif	Tarixi	Leksik birlik	Ma’nosi / zamonaviy shakli	Sintaktik xususiyat
Padarkush	Behbudiy	1913	isloh	isloh, tuzatish	ot
O‘zbek yoshlarining g‘ami	Fitrat	1915	maorif	ta’lim, bilim	ot
Halima	Cho‘lpon	1920	farzand tarbiyasi	bola tarbiyasi	ot birikmasi
Topshirgan bola	Avloniy	1912	shuuriy	ongli	sifat
Baxtsiz kuyov	S. Nosiriy	1911	ro‘zg‘or	oilaviy hayot	ot

Jadvalda jadid dramalaridan kiritilishi kerak bo‘lgan leksik birliklar, ularning zamonaviy ma’nosi va grammatik tavsifi aks etgan. Bu birliklar O‘zbek tili milliy korpusiga kiritilgach, ularning tarixiy hamda zamonaviy qo‘llanilish xususiyatlari bo‘yicha qiyosiy kuzatuvlar o‘tkazish mumkin bo‘ladi.

Jadid dramalari leksikasini milliy korpusga kiritishda quyidagi pometa va tasniflar qo‘llanadi:

1) Pometa (belgilar):

[hist] – tarixiy birlik; [archaic] – eskirgan; [dial] – shevaviy; [lit] – badiiy; [neol] – yangi ma’no; [tech] – ilmiy/texnik; [top] – toponim; [fraz] – frazeologizm; [maq] – maqol; [mat] – matal.

2) Semantik guruhlar:

- ijtimoiy-siyosiy atamalar (islohot, millat, hurriyat);
- ta’lim-maorif leksikasi (maorif, madrasa, maktab);
- oilaviy va axloqiy birliklar (tarbiya, odob, farzand);
- diniy-axloqiy terminlar (imom, shariat, tavba);
- kasb-hunar va iqtisod (dehqon, tijorat, sanoat);
- madaniyat va san’at (she’r, sahna, adabiyot).

3) Semantik to‘dalar (guruh ichidagi tor toifalar):

Semantik guruh	Semantik to‘da	Misol
Ijtimoiy-siyosiy	Milliy harakat	hurriyat, istiqlol
	Siyosiy tizim	majlis, sho‘ro
Maorif-ta’lim	O‘qitish jarayoni	dars, maktab
	Ilmiy unvonlar	muallim, ulamo
Oilaviy hayot	Nikoh va oila	kuyov, kelin
	Tarbiya	odob, axloq
Diniy-axloqiy	Din asoslari	namoz, shariat
	Ma’naviyat	tavba, axloq
Kasb-iqtisod	Tijorat	bozor, savdo
	Dehqonchilik	ekin, hosil
Madaniyat-san’at	Adabiyot	she’r, qalam
	Teatr	sahna, aktyor

Korpusda joylashtirishning ishchi shakli (namuna):

T/r	So‘z	Pometa	Semantik guruh	Semantik to‘da	Izoh
1	Vasiqa	[hist]	hujjat	rasmiy shartnoma	1917-yilgacha yer oldi-sotdi va shu kabilarga oid shartnoma
2	Tanob	[archaic]	o‘lchov	maydon birligi	Markaziy Osiyoda qo‘llangan; ~39,9 m ² yoki hududiy me‘yor bo‘yicha talqin
3	Mirzo	[archaic]	lavozim	unvon	Temuriylarda kotiblik, jadidlarda “kotib/yordamchi”
4	Yag‘ir	[]	belgi	jarohat	Og‘ir yukdan hayvon yelkasida paydo bo‘ladigan yiringli yara; ko‘chma ma’nolari ham bor
5	Xatib	[din]	lavozim	diniy unvon	Xutba o‘quvchi; juma xutbasini ado etuvchi
6	Ustuxon	[archaic]	tana a‘zosi	qism	Tojikcha “suyak”; qiyoslarda qo‘llanadi

Korpus bo‘yicha yirik tadqiqotlar muallifi Sh. Hamrayeva mualliflik korpuslarini yaratishda quyidagilarni uqtiradi: razmetkalash (belgilash) avtomatik yoki yarim avtomatik rejimda bajariladi; muallif qo‘llagan neologizmlar va boshqa alifboda yozilgan so‘z shakllarining

lemmatizatsiyasiga alohida e'tibor qaratiladi. Korpus tayyor bo'lgach, tabiiy tarzda muallif asarlari bo'yicha chastotali lug'at shakllanadi, chunki lemmatizatsiya jarayonida so'z shakllari va leksik birliklar (lemmalar) soni aniqlanadi. Masalan, Chexov korpusi 36 153 lemma (taxminan 1 381 000 qo'llanish holati, 120 000 so'z shakli)dan iborat bo'lib, gapning o'rtacha uzunligini ham hisoblash imkonini beradi. Qo'llanish chastotasi asar yili, janri, gap uzunligi bo'yicha ham hisoblanishi mumkin, bu esa hozirgi adabiy til leksikasi bilan muallif davri leksikasini qiyoslashga, xulosalar chiqarishga yo'l ochadi. Bunday qiyosiy tahlillar milliy korpus bazasida ham bajarilishi mumkin.

Jadid dramalari uchun alohida mualliflik korpuslarini barpo etishda semantik teglash jarayoni hal qiluvchi ahamiyatga ega. Sh. Hamrayeva, Sh. G'ulomova, D. Ahmedova, A. Eshmo'minov ishlarida semantik teglash masalalari yoritiladi. D. Ahmedovanning o'zbek tili korpuslari uchun atov birliklarini leksik-semantik teglashga bag'ishlangan tadqiqotida operator va konstanta teglar farqlanadi: korpus birliklari semantik maydon, to'da va guruhga mansubligiga ko'ra teglanadi; to'da/guruh — operator teg, maydon esa konstanta teg bilan belgilanadi. Geografik terminlarni korpusga kiritish masalasini ko'rib chiqqan I. Islomov o'zbek tili manbalarida tilning bir necha mavzuviy guruhlariga ajratib kelinishi operator teglar uchun tabiiy asos bo'la olishini ta'kidlaydi. Demak, umumiy mavzuviy guruhlar operator teglarni, nisbatan turg'un semantik maydonlar esa konstanta teglarni anglatadi.

Korpusning semantik razmetkasi uchun lingvistik ta'minot sifatida quyidagilar ko'rsatiladi:

- so'zlik (lemma ro'yxati);
- semantik lug'at yoki semantik ma'lumotlar ombori;
- semantik razmetkaning lingvistik modellari;
- ko'p ma'nolilik va omonimlikni farqlash mezonlari.

Masalan, ko'p ma'nolilikni quyidagicha rasmiylashtirish mumkin:

So'z	Semantik teg	Izoh 1	Izoh 2	Izoh 3	Gap namunasi
Yag'ir	belgi	Hayvon yelkasida og'ir yukdan paydo bo'ladigan yiringli yara	Ko'chma: inson yelkasidagi og'ir yukdan qolgan jarohat	Ko'chma: juda iflos, isqirt (ust-bosh)	“Unga tekkanimdan buyon yeganim zahar-zaqqum, yeng boshim yag'ir.” (<i>Advokatlik osonmi</i>)

Korpus lingvistikasida leksik birliklarning semantik darajalarda joylashuvi *taksonomiya* atamasi bilan belgilanadi. Jadid dramalarida yuqorida ko'rsatilganidek, turli semantik maydonlar mavjud; omonim, sinonim, antonim va ko'p ma'noli birliklarni berishda O'TIL tajribasiga o'xshash tarzda alohida semalarni kiritish maqsadga muvofiq. A. Eshmo'minov o'zbek tilida taksonimik va meronimik munosabatli leksemalarni korpusda teglash bo'yicha tadqiqot olib borib, O'TILdagi so'zlar va ularning statistikasi asosida JavaScript kodlari orqali tegga solish mexanizmini ishlab chiqadi; qidiruv parametrlariga tayanib taksonimik/meronimik birliklarni tanlashning lingvistik bazasi va statistikasi taklif etiladi. Jadid dramalari korpusini ham mavzuviy guruhlariga ajratish mumkin. *Ruscorpora.ru* bazasida taksonomik va meronimik qidiruv tizimlari mavjud bo'lib, tanlangan mavzular doirasida kerakli teglar bo'yicha izlashni qo'llab-quvvatlaydi.

Biz avval jadid dramalarini alohida mavzuviy guruhlariga ajratib chiqdik; jadid dramalari leksikasi lug'ati tuzilib bo'lgach, lug'atda so'zlarning semantik guruhlari ham ajratiladi. Ketma-ketlik quyidagicha bo'lishi mumkin: **Semantik guruh** → **so'z** → **sinonim** → **antonim** → **izoh**.

Namuna:

№	Semantik guruh	So'z	Sinonim	Antonim	Turkumi
1	Ta'lim va fan	Ma'rifat	–	–	ot
2	Ta'lim va fan	Usuli jadidiy	–	–	izofa
3	Diniy	Xurofot	–	–	ot

4	Ijtimoiy	Yo'qsul	kambag'al	boy	sifat
---	----------	---------	-----------	-----	-------

Jadid dramalari uchun lingvistik bazani yaratishning yana bir yo'li — mualliflik lug'atlarini tuzishdir: bunda so'z, so'z shakli, semantik guruh va maydon, sinonimik qator, antonimlar, ko'p ma'nolilikning kontekstual farqlari tizimli qayd etiladi. Jahonda muayyan adib ijodini qamrab oluvchi maxsus mualliflik korpuslari keng qo'llanmoqda. Mualliflik leksikografiyasi odatda ikki bo'lakdan iborat: asos (korpus yoki konkordans) va mahsulot (lug'at). So'nggi yillarda, ayniqsa korpus lingvistikasida, "konkordans" tushunchasi keng tarqaldi: u ma'lum til vositasining kontekstdagi barcha qo'llanishlarini manbaga havola bilan beruvchi ro'yxatni anglatadi; shuningdek, biror asarning kalit so'zlari bo'yicha kontekstsiz, alifbo tartibidagi ko'rsatkich ma'nosida ham ishlatiladi. P. Chexovning barcha tugallangan badiiy matnlari kiritilgan korpusda matnlar yozuvchining to'liq asarlar to'plamiga tayangan holda tuzilgan.

Jadidlar asarlari bo'yicha mualliflik korpuslarini yaratishda dramatik matnlarga doir ma'lumotlardan lingvistik ta'minot sifatida foydalanish mumkin. Masalan, Behbudiy korpusi uchun belgilashning quyidagi bosqichlari taklif etiladi:

1. so'z shakli;
2. so'zning boshlang'ich (lemma) shakli;
3. maxsus qo'llanishlar;
4. ekstralingvistik
 - asar nomi (masalan, drama "Padarkush");
 - sarlavha/bag'ishlov;
 - yozilgan yil;
 - mavzular guruhi;
 - asosiy mavzu;
 - semantik guruh (ilm-ma'rifat, diniy, ijtimoiy-siyosiy, ijtimoiy-maishiy, ilmiy, badiiy).

belgilash:

Lug'atning grammatik razmetkasida har bir qo'llanish holati quyidagicha belgilab boriladi: leksema (lug'at shakli), leksemaning grammatik belgisi (so'z turkumi, aniq–mavhum), so'z shaklining grammatik ko'rsatkichlari (son, kelishik, tuslanish, zamon). Kompyuter "tushunmagan" joylar qo'lda qayta razmetkalanadi va davr tilining o'ziga xosliklari aniqlashtiriladi. Belgilangan korpus bazasida lug'atning turli ko'rinishlari — konkordans, chastotali, teskari (*reverse*), grammatik, nutq bo'laklari bo'yicha, matn va personajlar kesimidagi maxsus lug'atlar — yaratish imkoniyati vujudga keladi.

FOYDALANILGAN ADABIYOTLAR RO'YXATI

1. Nuritdinov A. S. Jadid davri adabiy muhitiga doir asarlardan korpusda foydalanish // Kompyuter lingvistikasi: muammolar, yechim, istiqbollor. — 2024. — Vol. 1, № 1.
2. Xamroyeva Sh. M. O'zbek tili mualliflik korpusini tuzishning lingvistik asoslari: filol. fan. bo'yicha falsafa doktori (PhD) dis. — Buxoro, 2018. — 250 b.
3. Axmedova D. B. Atov birliklarini o'zbek tili korpuslari uchun leksik-semantik teglashning lingvistik asos va modellari: filol. fan. bo'yicha falsafa doktori (PhD) dis. avtoref. — Buxoro, 2020. — 54 b.
4. Isломov I. O'zbek tilining geografik terminologiyasi: tizimi, genezisi, semantik strukturasi va leksikografik talqini: filol. fan. dok. dis. avtoref. — Qarshi, 2021.