



Distinctive aspects and structure of contextual dictionaries

Umida TURAEVA¹

National University of Uzbekistan named after Mirzo Ulugbek

ARTICLE INFO

Article history:

Received February 2025
Received in revised form
10 March 2025
Accepted 25 March 2025
Available online
25 April 2025

Keywords:

corpus,
contextological dictionary,
artificial intelligence,
corpus integration,
idiomatics,
machine translation.

ABSTRACT

This article provides an in-depth analysis of the role of contextual dictionaries in modern linguistic research, their distinctive lexicographic features, and principles of construction. It explores the differences between context-based dictionaries and traditional ones, particularly focusing on how the meanings of lexical units are revealed in speech environments, how polysemantic units are defined in context-dependent ways, and how synonyms are differentiated through semantic differential features. Furthermore, the article discusses the key methodological criteria that should be considered when compiling contextual dictionaries, including the effective utilization of corpus materials, the scientific foundations for context selection, and their communicative-functional potential. The findings of this research hold significant scientific and practical importance for developing innovative approaches in lexicography and advancing context-based linguistic disciplines.

2181-3701/© 2025 in Science LLC.

DOI: <https://doi.org/10.47689/2181-3701-vol3-iss4/S-pp382-387>

This is an open-access article under the Attribution 4.0 International (CC BY 4.0) license (<https://creativecommons.org/licenses/by/4.0/deed.ru>)

Kontekstologik lug'atlarning o'ziga xos jihatlari, tuzilishi

ANNOTATSIYA

Kalit so'zlar:

korpus,
kontekstologik lug'at,
sun'iy intellekt,
korpus integratsiyasi,
idiomatika,
mashina tarjimas.

Ushbu maqolada kontekstologik lug'atlarning zamonaviy lingvistik tadqiqotlardagi o'rnini, ularning leksikografik tabiatiga xos jihatlari hamda tuzilish prinsiplari chuqur tahlil etiladi. Kontekst asosida tuzilgan lug'atlarning an'anaviy lug'atlardan farqli jihatlari – xususan, leksik birliklarning ma'nosi nutqiy muhitda ochilishi, polisemantik birliklarning kontekstga bog'liq holda aniqlanishi, sinonimlarning esa semantik differensial belgilar orqali farqlanishi kabi masalalar yoritiladi. Shuningdek, maqolada kontekstologik lug'atlar tuzishda e'tiborga olinishi

¹ Master degree student, National University of Uzbekistan named after Mirzo Ulugbek. E-mail: t.umida@bestaudit.co

lozim bo'lgan asosiy metodologik mezonlar, jumladan, korpus materiallaridan samarali foydalanish, kontekst tanlashning ilmiy asoslari va ularning kommunikativ-funksional imkoniyatlari haqida fikr yuritiladi. Mazkur tadqiqot natijalari leksikografiya sohasida innovatsion yondashuvlarni shakllantirish hamda kontekstga asoslangan tilshunoslik yo'nalishlarini rivojlantirishda muhim ilmiy va amaliy ahamiyatga ega.

Особенности и структура контекстологических словарей

АННОТАЦИЯ

Ключевые слова:

корпус,
контекстологический
словарь,
искусственный интеллект,
интеграция корпусов,
идиоматика,
машинный перевод.

В данной статье проводится углублённый анализ роли контекстологических словарей в современной лингвистике, их лексикографических особенностей и принципов построения. Рассматриваются отличительные черты контекстуально ориентированных словарей по сравнению с традиционными моделями, в частности: раскрытие значений лексических единиц в речевой ситуации, выявление полисемии через контекст, а также дифференциация синонимов на основе семантических дифференциальных признаков. Кроме того, в статье освещаются основные методологические критерии составления контекстологических словарей, включая эффективное использование корпусных материалов, научно обоснованный выбор контекста и коммуникативно-функциональные возможности таких ресурсов. Результаты исследования представляют собой важную научную и практическую ценность для развития инновационных подходов в лексикографии и контекстуально ориентированных направлений лингвистики.

KIRISH

Korpus – til birliklarining imkoniyatlarini aniqlash borasida qidiruv dasturiga bevosita bo'ysundirilgan matnlar majmui maqsadida yaratilgan. Til korpuslari til bo'yicha tadqiqot va amaliy topshiriqlar yechimi uchun inkor etib bo'lmas ish quroli sanaladi. Korpus bir til doirasini yoki ko'p tillarni qamrab olishi mumkin. Korpuslardan foydalanishning qator afzalliklari mavjud. Masalan, korpuslar vaqt o'tishi bilan tildagi o'zgarishlarni tadrijiy o'rganish imkoniyatini beradi shu bilan birga foydalanuvchilar auditoriyasiga muvofiq bo'ladi. Korpusning ixtisosligiga qarab bir nechta turlari tasniflangan: informatsion, monolingval, billingval, multilingval, parallel, qiyosiy (comparable), diaxronik (diachronic) va shu kabi boshqa bir nechta tilni qayta ishlash uchun, tilni tahrir qilish uchun, tarjimani avtomatlashtirish hamda elektron lug'atga xos dasturlarni yaratish. Parallelizm mezoniga ko'ra, korpuslar bir tilli, ikki tilli va ko'p tilli toifalarga bo'linadi. Ikki tilli va ko'p tilli korpus bir xil tematik sohada ikki yoki undan ortiq tilda mustaqil ravishda yozilgan matnlarni birlashtiradi. Masalan, turli tillardagi konferensiyalar, ilmiy anjumanlar, siyosiy konferensiyalar va boshqa materiallar to'plami va boshqalar. Mazkur qiyosiy tadqiqotlar, tarjima nazariyasi va amaliyoti bo'yicha tadqiqotlar olib borish uchun muhim manba bo'lib xizmat qiladi.

ADABIYOTLAR TAHLILI VA METODOLOGIYA

Dunyo tajribasida parallel korpus yaratishning lingvistik, matematik va dasturiy tomonlari olimlar tomonidan tadqiqotlarda o'z ifodasini topgan. Chunonchi, rus va ingliz tillari bo'yicha korpus lingvistikasi turli sohalar kesimida V. Zaxarov, A. Sedov, A. Baranov, R. Potapova, V. Rikov, U. Frensis, N. Leontyeva, V. Martin, S. Kubler, A. Laurens, E. Etwell, S. Hunston, L. Boizou, McKenneri, J. Grafmiller, J. Grieve, N. Grum, S. Hansson, K. McAulif, M. Malberg, P. Milin, A. Murakami, R. Peych, A. Shembri, P. Tompson, B. Vinter, G. Lich kabi xorijiy olimlar tomonidan hamda Turkologiyada korpusshunoslik (korpus lingvistikasi) sohasi bo'yicha ilmiy tadqiqotlar olib borilgan. Turk tili korpusi bo'yicha Aksan, Deniz Zeyrek, Kemal Oflazer, Umut Özge Bular; uyg'ur tili bo'yicha Yusup Aibaidulla, Kim-Teng Lua; boshqird tili bo'yicha L.A. Buskunbaeva, Z. Sirazitdinov; hakas tili bo'yicha Sheymovich, tatar tili bo'yicha J. Suleymanov, A. Gatiatullin, O. Nevzorova, R. Gilmullin, B. Hakimov; qirimtatar tili bo'yicha L. Kubedinova hamda tuva tili bo'yicha Salchak kabi olimlarning ishlari diqqatga sazovor.

Tarjimada elektron lug'atlar mashina tarjimasidan foydalanish jarayonining tez va samarador bo'lishiga xizmat qilish bilan birga, yangi ilmiy muammolarni o'rganishga bo'lgan ehtiyojni ham keltirib chiqarmoqda. Mashina tarjima tizimlari, avtomatlashtirilgan lug'atlar, terminologik ma'lumot banklarini yaratish parallel matnlarni tadqiq va tahlil qilishni talab qiladi. Buning samarasi sifatida dunyo bo'ylab parallel matndan foydalanish parallel korpuslarni tadqiq qilish va natijalarni mukammallashtirish imkoniyatini yaratadi. Shu o'rinda parallel korpus bo'yicha quyidagi ta'riflarni keltirish o'rinli. Parallel korpus – tarjima matnlarining parallel elektron analogi hisoblanib, asosan ikki tilli tarjima matnlardan iborat. Korpusdagi elektron matnlar asl matnning o'zi yoki uning bir qismi ham bo'lishi mumkin. Bir qancha olimlar, jumladan, D. Dobrovolskiy, Yu. Tao, V. Zaxarov, A. Kokoreva (11, 53) hamda E. Sosninalar parallel korpusning tuzilishi, tarkibi va imkoniyatlari haqida o'zlarining tadqiqot ishlarida alohida e'tirof etishgan. Parallel korpusda bajariladigan amallarni takomillashtirish yuzasidan ham ko'plab olimlar ilmiy izlanishlar olib borishgan. Ulardan M. Kay, M. Roscheisen, W. Gale, K. Church, P. Resnik, P. Brawn, J. Lai, R. Merserlar parallel korpus yaratish bosqichlari, yo'llari xususida o'z fikr-mulohazalarini bildirishgan. Ushbu olimlar orasidan Key va Roscheisenning fikricha, tarjimada gaplar bir-biriga mos kelishi uchun, avvalo, so'zlar muqobilligiga erishish muhim o'rin tutadi, lekin ikki tilda matn so'zlarining muqobilligiga erishish oson emas va bu muqobillik adekvat bo'lmasligi mumkin. Shunday bo'lsa, parallel korpusdagi izlangan so'zning bir qancha matndagi ma'nolarining berilishi orqali tadqiqotning o'rganish jarayoni osonlashadi.

Sun'iy intellekt (AI) tadqiqotlari axborot-kommunikatsiya texnologiyalarining eng muhim sohasi bo'lib, natijalari hamma foydalanishi mumkin bo'lgan ilovalar ko'rinishida paydo bo'ldi. Sun'iy intellektning eng muhim yo'nalishlari qatoriga mashina tarjimasini ham kiradi, tadqiqotchilar tilning nozik tomonlarini shubha ostiga qo'yib, mashinani matnlarni bir tildan ikkinchi tilga tarjima qilish qobiliyatiga ega qilish uchun turli yondashuv va metodologiyalarni taklif qilmoqdalar. Shu sababli, tarjima vositalari texnologiyaning so'nggi rivojlanishi bilan barqaror rivojlandi; Ko'p yillar oldin tarjimonlar o'zlarining tarjimalarini yozish uchun tahrirlash dasturidan foydalanganlar va hozirda ko'plab tarjima vositalari mavjud. Ikki soha, ya'ni mashina tarjimasini (MT) va kompyuter yordamida tarjima (CAT) ular orasidagi farqni ko'rsatib, qanday o'zaro bog'liqligini ko'rishning muhimligi ta'kidlangan. Mashina tarjimasini (MT) asosiy vazifani avtomatlashtirish uchun mo'ljallangan, ya'ni manba matnning tarjimasini sifatida

hisoblangan soʻzlar qatorini ishlab chiqarish. Boshqalar, kompyuter yordamida tarjima (CAT) vositalari sarlavhasi ostida guruhlangan, asosiy vazifani inson tarjimoniga qoldirib, samaradorlikni oshirish uchun moʻljallangan. Biroq MT va CAT vositalari, xususan, onlayn lugʻatlar asta-sekin birlashtirildi, yaʼni tarjima mexanizmlari kabi xususiyatlar ushbu onlayn platformalarga qoʻshildi.

Avvalo, onlayn lugʻatlar elektron yoki hisoblash leksikografiyasi deb ataladigan narsaga kiradi, yaʼni “elektron lugʻatlarni (ED) loyihalash, ishlatish va qoʻllash” degan atama boʻlib, ular oʻz navbatida asosan insonga yoʻnaltirilgan elektron toʻplamlar sifatida belgilanadi. Xususan, onlayn lugʻatlar 1990-yillarda texnologiyaning jadal rivojlanishi bilan paydo boʻldi. Oʻsha paytda ular dizayni va funksiyasi jihatidan sodda edi, yaʼni CD-ROM versiyalari va qoʻl qurilmalari oʻrnini bosadigan veb-saytdagi raqamli yoki ikki tilli yozuvlardan iborat boʻlgan. Biroq, eski bosma nashrlarning elektron nusxalari zamonaviy lugʻatlar tomonidan takomillashtirilgan korpus integratsiyasi kabi innovatsiyalar koʻp funksiyali lugʻatlarga aylandi. Ganger “Elektron leksikografiya” ning kirish muharriri maqolasida “gibridizatsiya” atamasidan ikki yoki undan ortiq til resurslarini bir mahsulotda, yaʼni onlayn lugʻatda integratsiyalashuv tendentsiyasini ifodalash uchun foydalanadi. Soʻnggi paytlarda onlayn lugʻatlarga qoʻshilgan eng keng tarqalgan funksiya va vositalar tarjima mexanizmlari koʻrinishidagi mashina tarjimasi (MT), korpusning integratsiyasi, yaʼni lugʻatlar yozuvlari korpusga asoslangan kontent bilan almashtiriladi va hamkorlikda kiritish (foydalanuvchi kiritishi) imkonini beradi. Myuller-Spitzer, Lew & Mickiewicz tomonidan nashr etilgan “Onlayn lugʻatlardan foydalanish” nomli kitobning sharh maqolasida “foydalanish boʻyicha nashr etilgan empirik tadqiqotlar” dan onlayn elektron lugʻatlardan foydalanish boʻyicha olib borilgan adabiyotlarni oʻrganish mumkin.

Parallel korpuslar, ikki yoki undan koʻp tillar boʻylab muqobil maʼlumotlarni oʻz ichiga olgan maʼlumotlar bazasi hisoblanadi. Kontekstologik lugʻatlar esa bu maʼlumotlarni yaxshi foydalanish va oʻrganish uchun juda foydali boʻladi.

Kontekstologik lugʻatlar, parallel korpus maʼlumotlaridan foydalanib soʻzlarning turli kontekstlar boʻyicha maʼnosini aniqlashda yordam beradi. Bu, til oʻrganuvchilar uchun va tarjimonlar uchun juda muhimdir, chunki ular oʻzaro almashtirilgan matnlarda soʻzlar yoki iboralar turli vaziyatlarda qanday ishlatilganligini koʻrsatadi.

Parallel korpus maʼlumotlari, kontekstologik lugʻatlarga koʻp toʻplangan matnlar orqali asoslangan boʻlib, bu esa soʻzlarning turli vaziyatlarda qanday maʼnoni ifodalay olishi va ularning turli kontekstlar boʻyicha qanday ishlatilishi haqida fikr qilishga yordam beradi. Bunday lugʻatlar til oʻrganish, tarjima qilish va matn tahlil qilishda juda muhim boʻlib, maʼlumotlar bazasini oʻrganish va aniqlashda yordam beradi.

Bu lugʻatlar soʻzlarning turli jumalardagi qoʻllanishi, sinonimlari va antonimlari, hamda ularning registri (rasmiy, norasmiy) haqida maʼlumot beradi.

Kontekstologik lugʻatlarning imkoniyatlari:

1. Turli kontekstlarda soʻzlarni qanday ishlatishni koʻrsatib berib, til oʻrganishni osonlashtiradi.

2. Noaniq yoki koʻp maʼnoli soʻzlarning kontekstdagi aniq maʼnosini tushunishga yordam beradi.

3. Soʻzlarning toʻgʻri qoʻllanishini oʻrganishda yordam beradi.

Masalan, “run” soʻzi kontekstga qarab turli maʼnolarni anglatishi mumkin:

– *Asosiy maʼno*: “yugurmoq” (He can run fast).

– *Kontekstual maʼno*: “boshqarish” (She runs a successful business).

– *Idiomatika*: “Run into” (U kutilmaganda kim bilandir toʻqnashdi).

Контекстологик lug'atlar bu turli ma'nolarni va ularning qo'llanilishiga oid misollarni ko'rsatadi.

Tavsiya etiladigan kontekstologik lug'atlar:

1. Oxford English Dictionary (OED) – ingliz tilidagi so'zlar uchun keng qamrovli kontekstual ma'nolarni beradi.

2. Merriam-Webster – ingliz tili uchun yaxshi manba.

3. Collins Dictionary – ingliz va boshqa tillar uchun kontekstual ma'lumotlarni taqdim etadi.

O'zbek va turk tillari uchun maxsus kontekstologik lug'atlar kamroq uchrasa-da, umumiy lug'atlardan va onlayn manbalardan foydalanib, so'zlarning turli kontekstlardagi ma'nolarini o'rganish mumkin.

Namunalar:

O'zbek tilida:

– So'z: "Qo'l"

– Kontekst 1: U menga qo'l berdi.

– Kontekst 2: Uning rejasi qo'l kelmadi.

– Kontekst 3: Bu ishchi qo'l ostimda ishlaydi.

1-matndagi "qo'l berdi" birikmasidan salomlashish ma'nosida foydalanilgan.

2-matndagi "qo'l kelmadi" birikmasi muvaffaqiyatsizlik ma'nosini bildiradi.

3-matndagi "qo'l ostimda" birikmasi esa mening nazoratim ostida, yonimda kabi ma'nolarni ifodalaydi. Ya'ni birikmalarning qanday ma'no ifodalayotganini faqat matn mazmuni orqali bilish mumkin.

Xuddi shu matnlarning turk tilidagi namunlari:

– Sözcük: "El"

– Kontekst 1: Elini bana uzattı. (Selamlaşmak)

– Kontekst 2: Planı el vermedi. (Başarısız oldu)

– Kontekst 3: Bu işçi benim elimde çalışıyor. (kontrol altında, birlikte)

Kontekstologik lug'atlar shunday ko'plab misollarni jamlab, so'zlarning turli kontekstlardagi ma'nolarini aniq ko'rsatib beradi.

Kontekstologik lug'atlar quyidagi jihatlarga e'tibor beradi:

• Siyosiy kontekst: So'z yoki ifoda siyosiy matnlarda qanday ishlatilishini tushuntiradi.

• Ijtimoiy kontekst: So'zning ijtimoiy vaziyatlarda qanday ma'noga ega bo'lishi mumkinligini ko'rsatadi.

• Madaniy kontekst: So'zning madaniy ahamiyati va madaniy kontekstdagi ishlatilishi.

• Texnik kontekst: So'zning muayyan texnik sohada qanday ishlatilishini tushuntiradi.

• Shaxsiy kontekst: Shaxsiy tajriba yoki hissiyot bilan bog'liq bo'lgan ma'nolar.

Demak, kontekstologik lug'atlar va parallel korpuslar til o'rgatish va tarjima jarayonlarida o'zaro bog'liq bo'lgan kuchli vositalar hisoblanadi. Ular so'zlarning turli kontekstlarda qanday ishlatilishini tushunishga va ularning ma'nolarini to'g'ri aniqlashga yordam beradi. Bu vositalar, til o'rgatuvchilar va tarjimonlar uchun so'zlarning aniq va samarali qo'llanilishiga imkon yaratadi va til o'rgatish jarayonini yanada samarali qilishga yordam beradi.

FOYDALANILGAN ADABIYOTLAR RO'YXATI:

1. Abdurakhmonova N. Kompyuter lingvistikasi. Darslik. – London: Globe edit publishing, 2020. 395 b.
2. Abduraxmonova N. O'zbek tili elektron korpusining kompyuter modellari. Monografiya. – T.; 2021.7-8-b.
3. Раҳматуллаев Ш. Тил қурилишининг асосий бирликлари // – Т.: Университет. 2002. – 662 б.
4. Сафаров Ш. Когнитив тилшунослик. 2006. –91 б.
5. Ataboev N. Chomsky's. Critics to Corpus and the Problem of Expert Speakers' Linguistic Intuition // Filologiya masalalari. 2015. – №4. – B. 108-111.
6. Ataboyev N.B. Korpus tilshunosligi shakllanishining diaxronik tadqiqi // Berdaq nomidagi Qoraqalpoq davlat universitetining Axborotnomasi. 2018. –№4. – 147-150-b.
7. Elov B. Tabiiy tilni qayta ishlash (NLP)da Spacy modulidan foydalanish // Илм-фан ва инновацион ривожланиш. 2022.
8. Ro'ziyev H., Xodjayeva N. Matnlarning parallel korpuslarini tuzish muammolari // Academic Research in Educational Sciences. 2021.
9. Abdurakhmonova N. The bases of automatic morphological analysis for machine translation // Izvestiya Kirgizskogo gosudarstvennogo tekhnicheskogo universiteta. 2016. – №2 (38). – B. 12–17.