Sifatli ta'lim-yorqin kelajak
ISSN: 3060-4761

TerDPI ILMIY XABARNOMASI
SCIENTIFIC BULLETIN OF THE TerSPI
НАУЧНЫЙ ВЕСТНИК ТерГПИ

www.xabarnoma.terdpi.uz

UDK: *811.512.133'367'322*
ORCID: 0009-0007-8690-2442

TIL KORPUSLARI ASOSIDA O'ZBEK TILINING LEKSIK-SEMANTIK TAHLILI

Eshqulova Nilufar O'ktamovna
Tadqiqotchi
E-mail: n.eshqulova93@gmail.com.

Annotatsiya. O'zbek tili milliy korpusi va boshqa lingvistik resurslar asosida so'zlarning ma'no doirasi, semantik guruhlanishi va ularning turli nutqiy vaziyatlarda qo'llanishi tahlil qilinadi. Natijalar o'zbek tilshunosligi uchun yangi imkoniyatlarni ochib berishi bilan birga, ta'lim jarayonida ham qo'llanishi mumkinligini namoyon etadi.

Tadqiqot maqsadi. O'zbek tilining leksik-semantik tizimini korpus lingvistika yondashuvlari asosida tahlil qilish, so'zlarning semantik guruhlanishi, ma'no doiralari va ularning nutqiy kontekstdagi qo'llanish qonuniyatlarini aniqlashdan iborat.

Tadqiqot metodlari: Tadqiqot jarayonida korpus lingvistik tahlil, semantik tahlil va qiyosiy tahlil metodlaridan kompleks foydalanildi. Xususan, O'zbek tili milliy korpusi asosida leksik birliklarning qo'llanish chastotasi va semantik guruhlanishi kuzatildi, ularning lug'aviy ma'nosi, polisemiya hamda omonimiya hodisalari kontekstual tahlil qilindi. Shuningdek, o'zbek tilidagi semantik hodisalar boshqa tillar korpuslari bilan qiyosiy o'rganilib, umumiy va o'ziga xos jihatlar aniqlab berildi.

Tadqiqot natijalari: Tadqiqot jarayonida o'zbek tili milliy korpusi asosida leksik birliklarning qo'llanish chastotasi, semantik guruhlanishi, polisemiya va omonimiya hodisalari hamda kollokatsion birliklari tizimli tahlil qilindi. Natijada so'zlarning faol-passiv qatlamlari, semantik maydonlari va kontekstda ma'no kengayishi yoki torayishi aniqlandi. Korpus asosida olingan statistik dalillar izohli lug'atlar tuzish, ta'lim jarayonida real nutqiy materiallardan foydalanish va avtomatlashtirilgan til texnologiyalarini rivojlantirishda amaliy ahamiyat kasb etishi belgilandi.

Xulosa qilib aytadigan bo'lsak mazkur maqolada o'zbek tilining leksik-semantik tizimini korpus lingvistika yondashuvi asosida o'rganish maqsad qilib olindi. Tadqiqot jarayonida korpus lingvistik tahlil, semantik va kollokatsion tahlil, kontekstual yondashuv hamda statistik metodlardan foydalanildi. Natijada leksik birliklarning qo'llanish chastotasi, semantik guruhlari, polisemiya va omonimiya hodisalari, shuningdek, so'zlarning real kontekstdagi qo'llanish xususiyatlari aniqlandi.

Kalit so'zlar: Leksik-semantik tahlil, milliy korpus, semantika, tilshunoslik, matnshunoslik, semantik guruh, innovatsion metodlar.

ЛЕКСИКО-СЕМАНТИЧЕСКИЙ АНАЛИЗ УЗБЕКСКОГО ЯЗЫКА НА ОСНОВЕ ЯЗЫКОВЫХ КОРПОРАЦИЙ

Аннотация. На основе Национального корпуса узбекского языка и других лингвистических ресурсов анализируется объём значений слов, их семантическая группировка и их употребление в различных речевых ситуациях. Полученные результаты не только открывают новые возможности для узбекской лингвистики, но и демонстрируют их применимость в образовательном процессе.

Цель исследования. Цель исследования – проанализировать лексико-семантический строй узбекского языка на основе подходов корпусной лингвистики, выявить семантические

группировки слов, смысловые группы и закономерности их употребления в речевых контекстах.

Методы исследования: В процессе исследования использовались методы корпусного лингвистического анализа, семантического анализа и сопоставительного анализа. В частности, на основе Национального корпуса узбекского языка исследовались частотность употребления и семантическая группировка лексических единиц, контекстуально анализировались их лексическое значение, явления полисемии и омонимии. Также изучались семантические явления в узбекском языке в сравнении с корпусами других языков, выявлялись общие и специфические черты.

Результаты исследования: В ходе исследования на основе национального корпуса узбекского языка был проведен системный анализ частотности употребления лексических единиц, семантической группировки, явлений полисемии и омонимии, а также устойчивых словосочетаний. В результате были определены активно-пассивные слои слов, семантические поля и расширение или сужение значения в контексте. Было установлено, что полученные на основе корпуса статистические данные имеют практическое значение при создании толковых словарей, использовании реальных речевых материалов в учебном процессе и разработке автоматизированных языковых технологий.

В заключение следует отметить, что данная статья направлена на исследование лексико-семантической системы узбекского языка на основе подхода корпусной лингвистики. В ходе исследования были использованы корпусный лингвистический анализ, семантический и устойчивый анализ, контекстуальный подход и статистические методы. В результате были определены частотность употребления лексических единиц, семантических групп, явлений полисемии и омонимии, а также особенности употребления слов в реальных контекстах.

Ключевые слова: Лексико-семантический анализ, национальный корпус, семантика, лингвистика, текстология, семантическая группа, инновационные методы.

LEXICAL-SEMANTIC ANALYSIS OF THE UZBEKISTAN LANGUAGE BASED ON LANGUAGE CORPORATIONS

Abstract. The scope of meaning of words, semantic grouping and their use in various speech situations are analyzed based on the National Corpus of the Uzbek language and other linguistic resources. The results not only open up new opportunities for Uzbek linguistics, but also demonstrate their applicability in the educational process.

Research objective. The aim of the research is to analyze the lexical-semantic system of the Uzbek language based on corpus linguistics approaches, to identify the semantic grouping of words, meaning groups and the patterns of their use in speech contexts.

Research methods: The methods of corpus linguistic analysis, semantic analysis and comparative analysis were used in the research process. In particular, the frequency of use and semantic grouping of lexical units were observed based on the National Corpus of the Uzbek language, their lexical meaning, polysemy and homonymy phenomena were contextually analyzed. Also, semantic phenomena in the Uzbek language were studied comparatively with corpora of other languages, and common and specific aspects were identified.

Research results: In the course of the research, the frequency of use of lexical units, semantic grouping, phenomena of polysemy and homonymy, and collocational units were systematically analyzed based on the national corpus of the Uzbek language. As a result, the active-passive layers of words, semantic fields, and expansion or narrowing of meaning in context were

determined. It was determined that the statistical evidence obtained on the basis of the corpus is of practical importance in the creation of explanatory dictionaries, the use of real speech materials in the educational process, and the development of automated language technologies.

In conclusion, this article aims to study the lexical-semantic system of the Uzbek language based on the corpus linguistics approach. In the course of the research, corpus linguistic analysis, semantic and collocational analysis, contextual approach, and statistical methods were used. As a result, the frequency of use of lexical units, semantic groups, phenomena of polysemy and homonymy, as well as the features of the use of words in real contexts were determined.

Key words: Lexical-semantic analysis, national corpus, semantics, linguistics, textual studies, semantic group, innovative methods.

KIRISH. Tilshunoslikda korpus lingvistika yo‘nalishi zamonaviy ilmiy tadqiqotlarning muhim sohasi sifatida shakllanib bormoqda. Bugungi kunda o‘zbek tilshunosligi ham ushbu yo‘nalishda jadal rivojlanmoqda. Xususan, o‘zbek tili milliy korpusining yaratilishi, uning lingvistik va texnologik imkoniyatlari, shu jumladan, leksik-semantik tizimni chuqur tahlil qilishga xizmat qilmoqda. Maqolaning maqsadi – o‘zbek tilining leksik-semantik tizimini korpus asosida ilmiy jihatdan o‘rganish, mavjud imkoniyatlar va muammolarni aniqlash hamda istiqbolli yo‘nalishlarni belgilab berishdan iborat[1].

ADABIYOTLAR TAHLILI VA METODLAR. O‘zbek tili milliy korpusi leksik birliklarning qo‘llanish chastotasi, semantik guruhlanishi va ma’no o‘zgarishlarini o‘rganishda muhim manba bo‘lib xizmat qiladi. Korpus lingvistika usullari yordamida so‘zlarning real nutqiy vaziyatlardagi ishlatilishi kuzatilib, ularning ma’no doirasi kengayishi yoki torayishi mumkinligi aniqlanadi. Bu jarayon semantik tahlilning asosiy yo‘nalishlaridan biri bo‘lib, lingvistik tadqiqotlarda an’anaviy metodlarni to‘ldiradi[2].

1. Til korpuslari va leksik-semantik tahlilning nazariy asoslari

Til korpusi – bu raqamlashtirilgan matnlar majmuasi bo‘lib, u maxsus lingvistik tamoyillar asosida yig‘iladi va tizimlashtiriladi. Hozirgi davrda korpus lingvistikasining asosiy yo‘nalishlaridan biri “leksik-semantik tahlil” hisoblanadi. Chunki korpus nafaqat til birliklarini yig‘ish va saqlash vositasi, balki ularni mazmuniy qatlamda o‘rganish, tilning semantik imkoniyatlarini ochish imkonini beradi. Leksik-semantik tahlil deganda so‘zning lug‘aviy ma’nosi, uning kontekstdagi qo‘llanish xususiyati, sinonimik qatorlar, omonimiya va polisemiya jarayonlarini tizimli tahlil qilish tushuniladi.

O‘zbek tili uchun yaratilayotgan milliy korpus aynan shu maqsadlarga xizmat qilishi, tilning lug‘aviy boyligi va ma’no imkoniyatlarini raqamli asosda qayta ko‘rsatib berishi mumkin. Bu esa tilshunoslikning an’anaviy usullaridan farqli ravishda, keng qamrovli statistik tahlil, kompyuter lingvistikasi vositalaridan foydalanish imkoniyatini beradi[3].

2. Leksik birliklarning korpus asosida tadqiq etilishi

O‘zbek tilida mavjud so‘z boyligini o‘rganishda korpus yondashuvi yangi ufqlarni ochadi. Masalan, bir so‘zning qo‘llanish chastotasi, turli janrdagi matnlarda uchrash tezligi, dialektal yoki uslubiy farqlanishi korpus orqali aniq statistik dalillar bilan asoslanadi.

Korpus lingvistikasida keng qo‘llaniladigan “kollokatsion tahlil” usuli so‘zlarning qanday birliklarda qo‘llanishini, qaysi so‘zlar bilan ma’nodosh yoki antonimik munosabatda ekanini aniqlash imkonini beradi. Bu, o‘z navbatida, leksik birliklarning semantik tarmoqlarini qurish, ularning ma’no maydonlarini belgilash uchun muhim ahamiyatga ega[4].

Masalan, o‘zbek tilidagi “yorug‘” leksemasi turli kontekstlarda “nur”, “chiroq”, “quvonchli”, “porloq” kabi turli semantik o‘zgarishlarda namoyon bo‘ladi. Korpus materiallari asosida bunday semantik o‘zgarishlar ko‘lamini tizimli tarzda aniqlash mumkin.

3. Semantik tahlilda kontekstning roli

Korpus lingvistikasida so‘zning ma’nosini izohlashda “kontekst” alohida ahamiyat kasb etadi. Chunki ko‘p hollarda so‘zning asosiy lug‘aviy ma’nosi kontekstdagi vazifasiga qarab kengayadi yoki torayadi. Shu bois semantik tahlil nafaqat so‘zning lug‘aviy izohi, balki u qo‘llangan matn birliklari, sintaktik qurilmalar va uslubiy vaziyatlar bilan birga ko‘rib chiqiladi.

O‘zbek tilida polisemiya hodisasi keng tarqalgan. Masalan, “ko‘z” so‘zi bir vaqtning o‘zida inson a’zosi, igna yoki buloqning “ko‘zi”, mevaning “ko‘zi” kabi turli ma’nolarni ifodalaydi. Korpus materiallari asosida bu turli semantik qatlamlarning kontekst bo‘yicha farqlanishi, qaysi janrda yoki hududda qaysi ma’no ko‘proq ishlatilishi haqida aniq ilmiy xulosalar chiqarish mumkin[5].

4. Til korpuslari va o‘zbek tilining semantik maydonlari. Korpus tahlili asosida o‘zbek tilidagi semantik maydonlarni aniqlash imkoniyati mavjud. Semantik maydon – bu ma’no jihatdan o‘zaro bog‘liq so‘zlarning majmuasidir. Masalan, “inson” tushunchasi atrofida “erkak”, “ayol”, “bola”, “ota”, “ona” kabi ko‘plab leksik birliklar shakllanadi. Korpusda bu birliklarning qo‘llanish chastotasi va o‘zaro munosabatlari tahlil qilinib, ular semantik maydon doirasida qay tarzda tizimga tushishi ilmiy asosda ko‘rsatiladi.

Bu jarayon, bir tomondan, o‘zbek tilining lug‘at boyligini yanada chuqurroq yoritishga xizmat qilsa, ikkinchi tomondan, kompyuter lingvistikasi, avtomatik tarjima va sun’iy intellekt tizimlarida o‘zbek tilini qo‘llash imkoniyatlarini kengaytiradi[6].

5. Leksik-semantik tahlilning amaliy ahamiyati. O‘zbek tili milliy korpusi asosida olib boriladigan leksik-semantik tadqiqotlarning natijalari:

Leksikografiyada yangi avlod izohli lug‘atlarini yaratishga, lug‘aviy birliklarning kontekstual ma’nolarini ko‘rsatishga xizmat qiladi. Til o‘qitish metodikasida o‘quvchilarga real nutqiy material asosida so‘z boyligini o‘rgatish imkoniyatini beradi.

Avtomatlashtirilgan tizimlarda (avtomatik tarjima, matn tahlili, nutqni tanish dasturlari) o‘zbek tilining imkoniyatlarini kengaytiradi.

Tilshunoslik nazariyasida semasiologiya, leksikologiya va stilistika bo‘yicha yangi ilmiy xulosalar chiqarishga asos bo‘ladi[7].

Shuningdek, raqamli texnologiyalar asosida yaratilgan korpus ma’lumotlari o‘zbek tilidagi leksik birliklarning sinonimiyasi, antonimiyasi, polisemiyasi va boshqa semantik hodisalarni keng miqyosda o‘rganishga imkon beradi.

Shunday qilib, korpus lingvistikasi o‘zbek tilini chuqur leksik-semantik jihatdan tahlil qilish, uning imkoniyatlarini ilmiy asosda yoritish hamda amaliy qo‘llashni rivojlantirishda strategik ahamiyat kasb etadi.

NATIJALAR VA MUHOKAMA

O‘zbek tilining milliy korpusi va unga oid ilmiy manbalar tahlili natijasida bir qator muhim ilmiy xulosalarga kelindi. Avvalo, milliy korpusning yaratilishi o‘zbek tilshunosligida zamonaviy til resurslari asosida olib boriladigan tadqiqotlarning sifat jihatidan yangi bosqichga ko‘tarilganini ko‘rsatadi. Ushbu korpus o‘zbek tilining leksik, grammatik va uslubiy imkoniyatlarini kompleks o‘rganish, so‘z birliklarining real qo‘llanish chastotasi, kollokatsion xususiyatlari hamda semantik maydonlarini aniqlashda asosiy manba sifatida xizmat qilmoqda.

Milliy korpus nafaqat ilmiy-tadqiqot ishlari, balki amaliy sohalarda ham keng

qo‘llanish imkoniyatiga ega. Jumladan, ta’lim tizimida o‘quv materiallarini samarali tanlash, tarjima amaliyotida adekvat muqobillarni topish, axborot texnologiyalari sohasida esa tabiiy tilni qayta ishlash tizimlarini rivojlantirishda asosiy manba bo‘lib xizmat qilmoqda. Xalqaro miqyosdagi tajribalar bilan solishtirilganda, o‘zbek til korpusi hali boshlang‘ich bosqichda bo‘lsa-da, mavjud materiallar kelajakda tilshunoslik, leksikografiya va kompyuter lingvistikasi bo‘yicha fundamental izlanishlar uchun poydevor yaratadi.

Shu bilan birga, o‘rganilgan natijalar shuni ko‘rsatadiki, mavjud korpus materiallari hajmi va ko‘lami jihatidan hali to‘liq mukammal emas. Xususan, matnlarning janrlar bo‘yicha notekis taqsimlanishi, ayrim dialekt va lahjalarning yetarlicha qamrab olinmagani, pragmatik nuqtai nazardan qo‘llanish imkoniyatlarining cheklangani kabi masalalar dolzarb bo‘lib qolmoqda. Biroq bu kamchiliklar ilmiy hamjamiyat e’tiborini yanada ko‘proq jalb qilmoqda va istiqbolda o‘zbek til korpusini kengaytirish hamda boyitish jarayonlarini jadallashtirish zarurligini ko‘rsatmoqda.

Umuman olganda, olib borilgan tahlillar milliy korpusning o‘zbek tilshunosligi taraqqiyotida yangi metodologik yo‘nalishlarni belgilash, tilning real qo‘llanish qonuniyatlarini aniqlash va uni zamonaviy lingvistik tadqiqotlarga moslashtirishda muhim vosita ekanligini isbotladi. Korpus lingvistikasi yondashuvi orqali o‘zbek tilining semantik, grammatik va uslubiy imkoniyatlari keng qamrovda tadqiq etilishi, natijada tilning ilmiy-nazariy asoslarini chuqurlashtirish bilan bir qatorda, amaliy qo‘llanish sohalarida ham samarali yechimlar taklif qilish imkoniyati yuzaga kelmoqda.

XULOSA: Xulosa qilib aytganda, til korpuslari asosida o‘zbek tilining leksik-semantik tahlili milliy tilshunoslikda yangi bosqichni boshlab berdi. Korpus lingvistika nafaqat tilning nazariy jihatlarini tadqiq qilishda, balki ta’lim jarayonida, til o‘qitish metodikasida ham samarali qo‘llanilishi mumkin. Korpusga asoslangan tadqiqotlar orqali o‘zbek tilining boy semantik imkoniyatlari tizimli va ilmiy asosda yoritiladi.

O‘zbek tili milliy korpusi asosida olib borilgan tadqiqot natijalari shuni ko‘rsatadiki, korpus lingvistikasiga tayangan holda til birliklarini tahlil qilish nafaqat leksik-semantik qatlamni chuqurroq ochib berish, balki ularning qo‘llanish chastotasi, kontekstdagi o‘zgaruvchanligi va semantik taraqqiyotini aniqroq aniqlash imkonini beradi. Ushbu yondashuv orqali so‘zlarning ma’no tizimi, polisemiya va omonimiya hodisalarining tabiati, shuningdek, ularning nutqiy jarayonda tutgan o‘rni izchil yoritildi.

Tadqiqot jarayonida qiyosiy tahlil metodidan foydalanish orqali o‘zbek tilidagi semantik hodisalar boshqa tillar bilan taqqoslab o‘rganildi hamda ularning umumiylik va milliy xususiyatlari ilmiy asoslangan holda aniqlab berildi. Bu esa o‘zbek tilining leksik-semantik tizimi o‘ziga xosliklarini belgilash, shu bilan birga, uni jahon lingvistik tadqiqotlari kontekstida baholash imkonini yaratdi.

Umuman olganda, mazkur tadqiqot o‘zbek tilining zamonaviy lingvistik tahlil usullari asosida o‘rganilishiga hissa qo‘shadi, korpus lingvistikasining amaliy qo‘llanish salohiyatini ko‘rsatadi va kelgusida tilshunoslikning boshqa yo‘nalishlarida ham keng tatbiq etilishi mumkinligini isbotlaydi. Shunday qilib, tadqiqot o‘zbek tilshunosligi uchun nazariy hamda amaliy ahamiyatga ega bo‘lib, milliy korpus imkoniyatlarini yanada kengroq tadqiq etishga ilmiy zamin yaratadi.

FOYDALANILGAN ADABIYOTLAR:

1. McEnery, Tony; Hardie, Andrew. “Corpus Linguistics: Method, Theory and Practice”. Cambridge: Cambridge University Press, 2012. – 294 p. ISBN 978-0-521-84876-9.

2. Eshmo‘minov, A. (2019). “O‘zbek tili milliy korpusining sinonim so‘zlar bazasi”: Filologiya fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati. Qarshi.

3. Hamroyeva, Sh. M. “Korpus lingvistikasi atamalarining qisqacha izohli lug‘ati”. – Toshkent: Kamalak nashriyoti, 2019. – 120 b.

4. Hamroyeva, Sh. M. “O‘zbek tili mualliflik korpusini tuzishning lingvistik asoslari”: Filologiya fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati. – Qarshi, 2018. – 48 b.

5. O‘zbekiston Respublikasi Fanlar Akademiyasi Til va adabiyot instituti materiallari. – Toshkent: O‘zR FA nashrlari, yil va sahifalar ko‘rsatiladi.

6. O‘zbek tilining milliy korpusi. (2020). Toshkent: O‘zbekiston Respublikasi Fanlar akademiyasi, Til va adabiyot instituti.

7. Primova, M. (2025). “O‘zbek tili milliy korpusining tuzilishi, asosiy komponentlari va axborot turlari”