



UDK: 81'373.46:811.512.133

Baxodir SUYUNOV,

Chirchiq davlat pedagogika universiteti dotsenti, filologiya fanlari doktori (DSc)

E-mail: Bahodirsuyunov66@gmail.com

Toshkent davlat o'zbek tili va adabiyoti universiteti professori, f.f.d. S.Muhamedova taqrizi asosida

THESAURUS IS A FACTOR OF LINGUISTIC DEVELOPMENT

Annotation

This article discusses the thesaurus and its types, as well as the application of computer technologies to the dictionary. The author purposefully used scientific sources in Uzbek and Russian languages, as well as various dictionaries and scientific research works to cover the topic. Instead, referring to examples and proofs that confirm certain concepts and events, he drew relevant scientific conclusions and gave practical recommendations.

Key words: thesaurus, homonym, synonym, paronym, lexicon, semantics, hyponym, hyperonym, adequate, system, computer, linguistics, corpus, engineering.

ТЕЗАУРУС – ФАКТОР ЛИНГВИСТИЧЕСКОГО РАЗВИТИЯ

Аннотация

В данной статье ведется рассуждение о тезаурусе и его видах, а также о внедрении в словароведение компьютерных технологий. При освещении темы автор целенаправленно использует научные источники в узбекском и русском языках, а также различные словари и научно-исследовательские работы. Уместно обращаясь к примерам и фактам, подтверждающим определенные понятия и явления, сделал соответствующие научные выводы по теме и дал практические рекомендации.

Ключевые слова: тезаурус, омоним, синоним, антоним, пароним, лексика, семантика, гипоним, гипероним, адекват, систем, компьютер, лингвистика, корпус, инженеринг.

TEZAURUS – LINGVISTIK TARAQQIYOT OMILI

Annotatsiya

Mazkur maqolada, tezaurus va uning turlari, shuningdek lug'atchilikka kompyuter texnologiyalarini tatbiq etish xususida fikr-mulohaza yuritilgan. Muallif mavzuni yoritishda o'zbek va rus tilidagi ilmiy manbalar hamda turli xildagi lug'atlar va ilmiy tadqiqot ishlaridan maqsadli foydalangan. O'rni bilan ma'lum tushuncha va hodisalarni tadqiqlovchi misollar va dalillarga murojaat qilib, mavzu yuzasidan tegishli ilmiy xulosalar chiqargan hamda amaliy tavsiyalar bergan.

Kalit so'zlar: tezaurus, omonim, sinonim, antonim, paronim, leksika, semantika, giponim, giperonim, adekvat, sistem, kompyuter, lingvistika, korpus, injenering.

Kirish. O'zbekiston Respublikasi Prezidenti Shavkat Mirziyoyevning "O'zbek tilining davlat tili sifatidagi nufuzi va mavqei tubdan oshirish chora-tadbirlari to'g'risida" 2019 yil 21 oktabrdagi PF-5850-son Farmoni ona tilimizning dunyo miqyosidagi nufuzini, jamoatchilikning davlat tiliga bo'lgan e'tiborni yanada oshirdi [1]. Bu esa o'z navbatida ona tilimizga kompyuter texnologiyalarini tatbiq etish, kompyuter lug'atlari (tezaurus)ni yaratish vazifalarini qo'yadi.

Tezaurus – hozirgi kundagi zamonaviy tilshunoslikda leksik birliklar o'rtasidagi semantik aloqalar – omonim, sinonim, antonim, paronim, giponim, giperonim va boshqa tushunchalarni ko'rsatadigan umumiy yoki maxsus lug'at, shunday lug'atlarning maxsus turi hisoblanadi.

Nazariy jihatdan tezaurus leksik-semantik tizimning ehtimoliy modellardan biridir. Amalda undan individual lug'atni boyitish va tezkor qidiruv vositasi sifatida foydalaniladi.

Tezaurus termini ikki xil ma'noda ishlatiladi:

1. Muayyan leksik birlik semantikasiga oid lingvistik tahlillarga asoslangan to'liq ma'lumot yoritilgan lug'atlar;

2. Bilimlarning muayyan sohasiga taalluqli tizimlashtirilgan axborotlar majmui.

Tezaurus milliy til, qayta ishlangan tabiiy til, axborot uslubi sifatida shakllangan sun'iy til semantikasini belgilaydi. Dunyoning rivojlangan tillaridan hisoblangan ingliz, rus tillaridagi tezaurslar milliy korpuslarning tarkibiy qismi

sifatida ma'lumotlar bazasi va tadqiqotlar uchun material bo'lib xizmat qilmoqda.

Bu lug'at boshqa lug'atlardan farqli o'laroq, biror bir ilm-fan sohasiga oid lug'aviy birliklar yoki biror mavzu tarkibida joylashtirilgan ana shunday birliklar o'rtasidagi semantik munosabatlarni aks ettirilgan ideografik lug'atdir.

Tezaurus umumiy ma'noda – maxsus terminologiya, ya'ni lug'at, ma'lumotlar yig'indisi, korpus yoki jamlanma, ma'lum sohadagi bilimlar yoki faoliyat sohasidagi tushunchalar, ta'riflar va terminlarni to'liq o'z ichiga oladi.

Muayyan yozuv yodgorliklaridan leksik birliklarni yoppasiga terib olishga asoslangan (masalan, yunon, lotin tillariga tuzilgan) lug'atlar tezaurus hisoblanadi. So'z tanlash tamoyillariga ko'ra yozuvchilar yoki ularning ijodiga mansub biror asar tili bo'yicha tuzilgan lug'atlar ham tezaurus bo'ladi.

U leksik-semantik, korporativ kommunikatsiya (bir fan yoki kasb orqali o'zaro bog'liq bo'lgan shaxslarning muloqotda bir-birlarini tushunishlari) uchun xizmat qiladi. Tezaurslar ma'lum bir fanni talqin qilishda muhim vositalardan biridir.

Lug'at va tezaurus tushunchalarini qiyoslasak, unda har ikkisi so'z va uning ma'nolarini bilish uchun xizmat qilsa-da, biroq ularning so'z ma'nolari haqida ma'lumot berish usullari har xil, ya'ni o'zaro farqlanadi.

Lug'at va tezaurus so'zlari lug'aviy ma'nosiga ko'ra, ot hisoblanadi. Lug'at ko'proq til o'rganuvchilar tomonidan so'zlarning ma'nolari, talaffuzi va orfografiyasini aniqlash

uchun ishlatilsa, tezaurus tadqiqotchilar tomonidan so'zning sinonim, antonim, omonim va boshqa ma'no shakllarini topish uchun ishlatiladi.

Tezaurus – so'z yoki terminning tegishli tushunchalar guruhidagi ro'yxati. U murojaat qilinayotgan so'z bilan bir xil bo'lgan boshqa so'zlar to'g'risida ham keng qamrovli ma'lumotlar beradi.

Aksariyat tadqiqotchilar sinonim, antonim, paronim va boshqa shu kabilarni bilish uchun tezaurusdan foydalanishadi. Odatda, tezaurus tilda so'zlarning kelib chiqishiga unchalik bog'liq bo'lmaydi, xususan, so'zlarning etimologiyasi haqida qo'shimcha ma'lumotlar bermaydi. Ammo, tezaurus so'zlarning boshqa shakllari, masalan, nominal shakllari, sifatlar va adverbial shakllar haqida juda ko'plab ma'lumotlar beradi.

Lug'at – bu tildagi so'zlarni (odatda alifbo tartibida) ro'yxatga oladigan va ularning ma'nosini izohlab beradigan yoki shunday so'zlarni boshqa tilda beradigan, ko'pincha to'g'ri talaffuz haqida ma'lumot beradigan kitob yoki elektron manba hisoblanadi. Boshqacha qilib ifodalasak, lug'at tildagi so'zlar haqida bilim beradigan etimologik va grammatik vositadir.

Lug'atda so'zning jinsi va nutqda ishlatilishi haqida qo'shimcha ma'lumotlar mavjud. U tildagi birorta so'z bilan bir qatorda, unga qo'shimcha ravishda ma'lumotlar ham berishi mumkin. Lug'at bir nechta aniq tillarda tuzilishi mumkin, masalan, ingliz, fransuz, rus va boshqa tillarda. Bir tilda tuzilgan lug'atda boshqa tillardagi shu so'zning ekvivalent ma'nolari ham beriladi. Odatda, bu sifatlar tezaurusga tegishli emas.

Mavzuga oid adabiyotlar tahlili. Birinchi zamonaviy ingliz tili tezaurusi Piter Mark Roger tomonidan 1805 yilda yaratilgan. Mazkur tezaurus 1852 yilda Angliyada nashr etilgan, taxminan 15 mingta tushunchani o'z ichiga oladi va shu vaqtdan beri foydalanib kelinadi. Bunday lug'atlarning asosiy maqsadi matn yozishda tegishli so'zlarni tanlashga yordam berish, mavzu doirasida atama yoki termin o'rtasidagi munosabatlarni tavsiflovchi ma'lumotlarni qidirib topishdan iborat. Bunday lug'atlar ma'lum bir soha mutaxassislari tomonidan yaratilgan va mavzu doirasidagi ma'lumotni qidirib topish uchun mo'ljallangan.

1870 yillarda tezauruslardan ma'lumot izlash ishlarida faol foydalanila boshlandi. Shunday lug'atlardan yana biri Vikilug'at deb ataladi. U ko'p tilli bepul yangilanadigan lug'at va tezaurus bo'lib, "Wikimediya" fondi loyihasi doirasida yaratilgan va 2004 yildan hozirgi kungacha ishlatilib kelinadi. Unda 250 mingta tushuncha va 67 mingta semantik munosabatlar o'z aksini topgan.

Rus tilidagi xuddi shunday tezauruslardan biri "RuTez" deb nomlangan. Bu lug'at 1997 yilda Axborot tadqiqotlar markazi tomonidan avtomatik indeksatsiya vositasi sifatida yaratilgan. U hozirgi kunga qadar ishlab chiqilgan 45 mingta tushuncha, 107 mingta so'z va iboralar, 177 mingta semantik munosabatlarni qamrab oladi.

O'zbekiston Milliy Ensiklopediyasida tezaurus terminiga quyidagicha izoh berilgan: "Tezaurus – muayyan tildagi barcha so'zlarni qamrab oladigan, ularning matnda qo'llanish holatlarini to'la-to'kis aks ettiradigan lug'at" [2].

Tezaurus – bu lug'atlar tushunchasi va uning birliklari o'rtasidagi sobit semantik aloqalar bo'lib, har ikkala mustaqil xususiyat ham tushuncha aniqlash uchun muhim ahamiyat kasb etadi.

"Tezaurus termini, avvalambor tilning lug'at fondini maksimal darajada to'liqlik bilan ifodalaydigan xazina sifatida talqin etiladi. Matnlarning adekvat talqini va ma'lumotlari uning sistem ma'nolarini o'zida mujassam etgan va soha vakillari tomonidan e'tirof etilgan tezaurusda aniqlanadi" [3].

Hozirgi vaqtda ushbu tushunchaga kompyuter texnologiyalarida so'zlarni qidirishda axborot

texnologiyalarining tarqalishi, Internet va mashinalarning tarjimai bilan bog'liq lug'atlar kiradi. Psixologiyada shaxsning tezaurusi ma'lumotni idrok etish va tushunish bilan xarakterlanadi. Bunda aloqa nazariyasi va uning elementlari o'zaro ta'sir ko'rsatadigan murakkab tizimning umumiy tushuncha-tezislari ko'rib chiqiladi.

Kompyuter lingvistikasining tarkibiy qismi bo'lgan tezauruslar leksemalarning o'z va ko'chma ma'nolari, kommunikativ xususiyatlari, emotsional-ekspressiv funksiyalari haqida keng ma'lumot berishi bilan qimmatlidir. Shu o'rinda ta'kidlash joizki, jahon tilshunosligida kompyuter lingvistikasi, korpus tushunchasiga doir ilk ma'lumotlar XX asrning qirinchii yillarida yuzaga kelganligi ilmiy manbalarda qayd etilgan [4]. O'zbek tilshunosligida korpus lingvistikasi masalalari keyingi yillarda ilmiy tadqiqot ishlari sifatida keng o'rganila boshladi.

Xususan, Sh.M. Hamroyevaning "O'zbek tili mualliflik korpusini tuzishning lingvistik asoslari" nomli dissertatsiyasida o'zbek tilshunosligida birinchi marta korpus, uning o'ziga xos xususiyatlari, nazariy asoslari, til korpusining lingvistik hamda nazariy va amaliy ahamiyati yoritib berilgan [5].

Shuningdek, korpus lingvistikasining shakllanish tarixi, taraqqiyot yo'li, o'ziga xos xususiyatlari va bugungi holati masalalari atroflicha talqin qilingan.

Tadqiqot metodologiyasi. O'zbekiston Respublikasi Prezidentining "O'zbekiston Respublikasining "Davlat tili haqida"gi Qonuni qabul qilinganligining 30 yilligini keng nishonlash to'g'risida"gi PQ-4479-son qarori (2019 yil 4 oktabr), "Mamlakatimizda o'zbek tilini yanada rivojlantirish va til siyosatini takomillashtirish chora-tadbirlari to'g'risida"gi PF-6084-son Farmoni (2020 yil 20 oktabr) hamda tilshunoslik sohasiga tegishli boshqa me'yoriy-huquqiy hujjatlarda davlat tilini yanada rivojlantirish, uning xalqaro nufuzini oshirish, maishiy xizmat, harbiy, tibbiyot va boshqa sohalarda keng qo'llanilayotgan xorijiy so'z va terminlarning o'zbekcha muqobillari va ularning lug'atini yaratish, o'zbek tiliga ilg'or texnologiyalarni tatbiq etish va bu boradagi ilmiy tadqiqot ishlarni qo'llab-quvvatlash singari dolzarb masalalar ko'zda tutilgan.

Yuqoridagi Farmon va qarorlar ushbu tadqiqotning metodologik asosini tashkil etadi.

Tahlil va natijalar. O'zbek tilini jahon tillari qatorida rivojlantirish, uni dunyoviy tillar sirasiga kiritish va til o'rganish hamda o'rgatishda kompyuter lingvistikasi fani dolzarb ahamiyat kasb etmoqda. Texnika taraqqiyoti mahsuli bo'lgan kompyuter tizimi barcha sohalarda qulayliklar yaratadi, ma'lumotlarning tezkor yetkazib berilishini, tarjima, tahrir jarayonlarining mashina yordamida qisqa muddatlarda amalga oshirilishini ta'minlaydi.

Kompyuter lingvistikasi amaliy tilshunoslik yo'nalishlaridan biridir. "Amaliy lingvistika" termini esa ko'p ma'noli termin bo'lib, bu termin umumtilshunoslikda har xil talqin qilinadi. G'arb tilshunosligida applied linguistics, angewandte linguistik termini, birinchi navbatda, o'qitish metodikasi, grammatik xususiyatlarni qamrab olgan holda, ona tili va chet tillarini o'qitish amaliyoti bilan bog'lanadi.

Aslida mazkur tushunchalar kompyuter texnologiyalarining ishlab chiqilishi va axborotni qayta ishlash tizimi (boshqaruvning avtomatlashgan tizimi, informatsion qidiruv tizimi, matnni qayta ishlashning avtomatik tizimi) shakllanishi jarayonida yuzaga kelgan.

Xorijiy tillar, jumladan rus tilida yaratilgan tezauruslar o'rganilganda ularning tuzilishi, modellari jihatidan har xillikni kuzatdik [6]. YA'ni tezaurus tuzishda bir xil prinsip yoki kriteriy mavjud emas, shuningdek axborot qidiruv jarayonida foydalanuvchilarga qulaylik yaratishga asosiy e'tibor qaratilganini ko'ramiz.

Ayrim rus tilidagi adabiyotlarda, ko'p hollarda, "kompyuter lingvistikasi" ("computational linguistics"), "hisoblash lingvistikasi", "avtomatik lingvistika", "injener lingvistikasi" tarzida qo'llaniladi [7]. Amaliy lingvistikaning mazkur nomlar bilan yuritilishi uning faoliyati xarakterini belgilaydi. Ushbu terminlarda amaliy lingvistikaning integratsiyaga asoslanishi, avtomatlashish xususiyati o'z ifodasini topgan.

Mavjud tezaurus maqolalarining modellari (O.Baranov, S.YE.Bikova, N.R.Valitova, YE.V.Smirnova va boshqalar izlanishlarida qayd etilgan) va boshqa tadqiqotlar xarakteridan kelib chiqib [8], tezaurus maqolasining quyidagi umumiy modelini keltirish mumkin:

termin; 2) aniqllovchi; 3) sinonimlari; 4) antonimlari; 5) golonimlari; 6) o'z ichida komponent, belgi, funksiya, xususiyat, aspekt munosabatlarini qamrab oluvchi meronimlari; 7) giperonimlari; 8) giponimlari.

Xulosa va takliflar. Tezaurus – bu umumiy ma'noda maxsus bilim sohasi yoki faoliyat sohasining tushunchalari, ta'riflari, atama yoki terminlarini to'liq qamrab oladigan maxsus terminologiya, yanada qat'iy va obyektiv ravishda lug'at, ma'lumotlar to'plami, korpus yoki kod, demakdir.

Tilda tezauruslar elektron formatda, alohida fan sohalarini tavsiflashning samarali vositalaridan biri bo'lib, biror fan yoki kasb-hunar bilan bog'liq bo'lgan munosabatlar va shu munosabatlardagi tushunchalarni rivojlantirishga xizmat qiladi.

Tezaurusning milliy tilshunoslikka tatbiq etilishi axborotlarning tezkorlik bilan o'zlashtirilishi, shuningdek bajariladigan amallarning aniqligini ta'minlaydi.

Kompyuter lingvistikasi har bir tilning o'ziga xos tabiatini hisobga olib, muayyan tilning fonetik, leksik, grammatik sathlariga oid masalalarni mashina yordamida hal qilish vazifasini qo'ygandagina lingvistik taraqqiyot omili bo'la oladi.

ADABIYOTLAR

1. Ўзбекистон Республикаси Президентининг "Ўзбек тилининг давлат тили сифатидаги нуфузи ва мавқеини тубдан ошириш чора-тадбирлари тўғрисида"ги Фармони. "Халқ сўзи" газетаси // 2019 йил, 22 октябрь, 218-сон.
2. Ўзбекистон Миллий Энциклопедияси. – Тошкент, 2004
3. Валитова Н.Р. Формирование профессионального тезауруса у студентов физкультурного вуза. Автореф. на соиск. уч. степени канд. пед. наук. – Омск, 2012
4. Кутузов А.Б. Курс "Корпусная лингвистика". Лицензия Creative commons Attribution Share-Alike 3.0 Unported - // lab314.brsu.by/kmp-lite/kmp-video/ CL/ CorporeLingva.pdf.
5. Хамроева Ш.М. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари. Филол. ф. бўйича ф.д.(PhD) дисс. автореф. – Қарши, 2018
6. Лукова В.А. таҳрири остида "Тезаурусный анализ мировой культуры". Сборник научных трудов. // – М.: 2013., Крюкова М. И. Экфрастический тезаурус в прозе А.С.Грина. Дисс., на соиск уч. степени к.ф.н. // – Новосибирск: 2018 ва б.
7. Баранов А.Н.таҳрири остидаги "Словарь тезаурус современной русской идиоматики". – М.: "Мир энциклопедии", 2007
8. Баранов О. Русский тезаурус (Электронное издание). – М.: 2014
9. Быкова С.Е. Тезаурусное моделирование юридической терминологии (на материале французского языка в сопоставление с русским). Автореф. на соиск. уч. степени канд. фил. наук. – Екатеринбург, 2006
10. Валитова Н.Р. Формирование профессионального тезауруса у студентов физкультурного вуза. Автореф. на соиск. уч. степени канд. пед. наук. – Омск, 2012
11. Смирнова Е.В. Структурно-семантические и лексикографический аспекты медицинской терминологической (на примере кардиологической лексики). Автореф. дис. на соискание уч. степени канд. филол. наук. – Ярославль, 2012. – 48 с.
12. Суюнов Б.Т. Тиббий терминлар семантикаси ва тезауруси. Монография. – Тошкент: 2022
13. Суюнов Б.Т. Тезаурус – ҳозирги замонавий лугат тури. IX Global science and innovations 2020: Central Asia. International scientific practical conference. Kazakhstan, Nur-Sultan (Astana).
14. Суюнов Б.Т. Tezaurus va uni turkiy tillarga tatbiq etish. "Turkiy filologiyaning dolzarb masalalari, nazariya va innavtsion ta'lim texnologiyalari" Xalqaro ilmiy-amaliy konferensiya materiallari to'plami / Ajiniyoz nomidagi Nukus davlat pedagogika instituti. – Нукус: 2021