

TEZAURUSLAR VA ULARNI TAHLIL QILISH TAMOYILLARI

Oybekova Sarvinoz Oybekovna

O'zbekiston Milliy universiteti
Kompyuter lingvistikasi magistratura mutaxassisligi
2-bosqich magistranti.

Annotatsiya: Mazkur maqolada tezauruslar, ularning lug'atlar yaratishdagi ahamiyati hamda tasnif qilish tamoyillari haqida gap boradi. Tezaurus lug'atlarining afzalliklarini misollar yordamida izohlanadi.

Kalit so'zlar: tezaurus, lug'at, semantic bog'lanish, WordNet.

Abstract: This article discusses thesauruses, their importance in creating dictionaries, and principles of classification. The advantages of thesaurus dictionaries are explained with the help of examples.

Keywords: thesaurus, dictionary, semantic connection, WordNet.

Аннотация: В статье рассматриваются тезаурусы, их значение при создании словарей и принципы классификации. Преимущества тезаурусных словарей поясняются на примерах.

Ключевые слова: тезаурус, словарь, семантическая связь, WordNet.

Tezaurus lug'atlari, so'zlar va ularning o'zaro aloqalarini tahlil qilish uchun yaratilgan ma'lumotlar bazasi hisoblanadi. Ular bir-biriga yaqinligi bo'lgan so'zlar to'plamidir. Bu lug'atlarning asosiy maqsadi, bir so'zning ma'nosi bo'yicha boshqa so'zlarni topish va ularga o'xshash so'zlar topishdir. Bu matnlar yoki turli xil ma'lumotlar bazalarida ishlatilishi mumkin. Tezaurus lug'atlari, matnlarni tahlil qilishda va tarjimada yordam berishda keng ishlatiladi. Ular matnda o'xshash yoki o'zaro bog'liq so'zlar ro'yxatini ko'rsatadi. Bu matnda qidirayotgan so'zni topishga yordam beradi. Shuningdek, tezaurus lug'atlari, matnlarni taqsonlarga bo'lishda va taqsonlarnishakllantirishda ishlatiladi. Tezaurus lug'atlari, boshqa lug'atlarga nisbatan bir necha foydali imkoniyatlar taqdim etadi. Ular so'zlarning ma'nolarini tahlil qilish, ularning o'zaro aloqalarini aniqlash, so'zlar va ularning aloqadorligini tushuntirish va matnlarni taqsonlarga bo'lishda yordam beradi. Bu, matnlar va turli ma'lumotlar bazalaridagi ishlarimizda juda foydali bo'ladi. Tezaurus lug'atlari, dastlabki qarashda, lug'atlar bilan bir xil deb o'ylansa ham, o'zaro farqlanuvchi xususiyatlarga ega.

Tezauruslarni tasnif qilish muammosi bir necha o'n yillar davomida bir qator chet el tilshunoslari (V.V. Morkovkin, L.P.Stupin, K.Marello va va boshqalar.) tomonidan atroflicha o'rganilgan. Ushbu tasniflarning asosini quyidagi mezonlar tashkil etadi:

- 1. Orasida semantic bog'lanishlar mavjud bo'lgan lug'at birliklari.*
- 2. Lug'at hajmi.*
- 3. So'z boyligini umumlashtirish.*
- 4. Leksemalarning ma'nosini rivojlantirish.*
- 5. Grammatik-stilistik leksema malakasi.*
- 6. Taqdim etilgan tillar soni;*

7. Leksemaning semiotik vosita turi, semantizatsiya uchun ishlatilishi.

Tezaurus lugʻatlari, shu jumldan, bir soʻz yoki ifodaning oʻziga xos maʼnolarini va ulardan boshqa yuzlab oʻzgaruvchilarni koʻrsatadi. Bu yuzlab oʻzgaruvchilar, asosan, Giperonim, Giponim, Meronim, Partonim va Halonimlardir.

Giperonim, bir soʻzning umumiy yoki katta maʼnolarini ifodalovchi soʻzdir. Misol uchun, "hayvon" soʻzi "it", "sigir", "qoʻy" va boshqa hayvonlar bilan bogʻliq soʻzlar uchun giperonim boʻlib hisoblanadi.

Giponim, esa, bir soʻzning kichik yoki toʻgʻri maʼnolarini ifodalovchi soʻzdir. Misol uchun, "it" soʻzi "dalmachiq", "pudralik", "pishiq" va boshqa kichik hayvonlar bilan bogʻliq soʻzlar uchun giponim boʻlib hisoblanadi.

Meronim, bir obyektning tarkibiy qismlarini ifodalovchi soʻzdir. Misol uchun, "avto" soʻzi "motor", "shassi", "faralar" va boshqa avtomobilning qismlari bilan bogʻliq soʻzlar uchun meronim boʻlib hisoblanadi.

Holonim, esa, bir obyektning tarkib qiladigan butunliklarni ifodalovchi soʻzdir. Misol uchun, "uy" soʻzi "xona", "oʻzi", "qavat" va boshqa uyning qismlari bilan bogʻliq soʻzlar uchun holonim boʻlib hisoblanadi.

Partonim, esa, bir obyektning tarkib qiladigan qismlarni ifodalovchi soʻzdir. Misol uchun, "kitob" soʻzi "sahifa", "qogʻoz", "qop" va boshqa kitobning qismlari bilan bogʻliq soʻzlar uchun partonim hisoblanadi. Halonim, esa, bir soʻzning oʻzining oʻzini tushunish uchun zarur boʻlgan soʻzdir. Misol uchun, "qalam" soʻzi "qogʻoz", "yozuv", "toʻqima" va boshqa qalamni tushunish uchun zarur boʻlgan soʻzlar uchun halonim boʻlib hisoblanadi.

Tezaurus lugʻatlari, soʻzlar orasidagi yaqin munosabatlarni aniqlashda va matnlarni taxmin qilishda yordam beradi. Ularning asosiy manbalari WordNet va Roget's Thesaurus kabi kompyuter dasturlari hisoblanadi. Tezaurus lugʻatlari, matnlar tarjimasida ham ishlatiladi, chunki ular soʻzlar orasidagi munosabatlarni aniqlashda yordam beradi. [5, 78-79]

Tezaurus lugʻatlari shunday afzalliklarga egaki, ular:

1. Soʻzlar orasidagi munosabatlarni aniqlashga yordam beradi: Tezaurus lugʻatlari, belgilangan soʻzning boshqa soʻzlar bilan boʻlgan yaqin munosabatlarini aniqlashda va ularni taxmin qilishda yordam beradi. Bunday qoʻllanmalar matn tuzish, tarjima va tafsilotlar yozish uchun juda foydali boʻladi.

2. Matnlarning koʻrsatilgan maʼnolarini kengaytirishga imkon beradi: Tezaurus lugʻatlari, berilgan soʻzning koʻrsatilgan maʼnolarini kengaytirish uchun qoʻshimcha soʻzlar va ularning sinonimlari va antonimlari bilan birga koʻrsatuvlar beradi.

3. Ishtirokchilar uchun maʼnolarni aniqlashda va tarjima qilishda yordam beradi:

4. Tezaurus lugʻatlari, matnlar tarjimasida yoki ishtirokchilar oʻzining soʻzlari bilan gaplashishda yordam beradi. Ular soʻzlarning koʻrsatilgan maʼnolarini aniqlash va ularning sinonimlari va antonimlari bilan birga koʻrsatuvlar beradi.

5. Matndagi soʻzlar orasidagi munosabatlarni tahlil qilishga imkon beradi: Tezaurus lugʻatlari, matndagi soʻzlar orasidagi munosabatlarni tahlil qilishga yordam beradi. Ular misol uchun, bir soʻzning giperonimlari, giponimlari, meronimlari, holonimlari, partonimlari va halonimlari bilan birga koʻrsatuvlar beradi.

6. Oʻzlashtirilgan soʻzlar va maxsus terminologiyani qamrab oladi: Tezaurus lugʻatlari, oʻzlashtirilgan soʻzlar va maxsus terminologiyani koʻrsatishda ham yordam beradi. Bunday qoʻllanmalar, maxsus terminologiyaga ega soʻzlarni aniqlash uchun juda foydali boʻladi. Tezaurus

lugʻatlaridan biri WordNet soʻzlarning leksik- semantik, antonimlik maʼnolarini aniqlashda eng muhim vosita hisoblanadi. [2, 45-48]

WordNet - bu Prinston universitetida ishlab chiqilgan inglizcha soʻzlarning semantik tarmogʻini oʻz ichiga olgan elektron lugʻat va tezaurus tizimi. WordNet ingliz tilidagi soʻzlarning semantik tuzilishini tavsiflash uchun ishlatiladigan juda koʻp turli atamalar (sinonim, antonim, gipernim, giponim, meronim, holonim va boshqalar) roʻyxatini oʻz ichiga oladi. Ushbu atamalar soʻzning maʼnosini yaxshiroq tushunish va soʻzlar guruhlari oʻrtasidagi munosabatlarni koʻrsatish uchun ishlatilishi mumkin [4, 69]. WordNet, ayniqsa, tabiiy tillarni qayta ishlash, kompyuter lingvistikasi va sunʼiy intellekt sohalarida keng qoʻllaniladi. WordNet soʻzning maʼnosini yaxshiroq tushunish va iboralar oʻrtasidagi munosabatlarni koʻrsatish uchun ishlatilishi mumkin, bu iboralar orasidagi munosabatlarni koʻrsatish uchun ishlatiladigan daraxt tuzilishi tufayli.

Masalan, “mushuk” soʻzi “uy hayvonlari” iborasining kichik toʻplami va “sutemizuvchilar” iborasining yuqori toʻplamidir. WordNet soʻz maʼnolari va ishlatilishini aniqroq va izchil aniqlashga yordam beradi, shuning uchun uni matn tasnifi, maʼlumotlarni qidirish va mashina tarjimasini kabi koʻplab sohalarida qoʻllash mumkin. WordNet - bu ingliz tilidagi soʻzlar uchun asosiy leksik tizim hisoblanadi. Bu tizimda soʻzlar taqsonlarni shakllantirish uchun ishlatilgan. Bu degani, bitta soʻzning oʻzining oʻrniga, uni ifodalaydigan boshqa soʻzlar ham mavjud. Misol uchun, "cat" soʻzi "feline" yoki "animal" deb ataladigan taqsonlarga tegishli boʻlib, "claw" soʻzi esa "cat" soʻzining meronimi (qismini ifodalaydi), yoki "tool" soʻzining holonimi (uchlashuvchisi) boʻlib ifodalaydi. WordNet oʻzida koʻp til uchun ishlatiladi va bu tizimning turli dasturlarga integratsiyasi mavjud.

Tezaurus - soʻzlarning maʼnolari va soʻz turkumlarining bir-biri bilan munosabatlari tizimli ravishda tartibga solingan lugʻat turi. Tezaurus til oʻrganish, matnlarni tasniflash, sinonimlarni topish, kontseptsiyani xaritalash va axborot qidirish kabi sohalarida keng qoʻllaniladi. Tezaurus tarkibida soʻzning sinonimlari (sinonimlari) va antonimlari (antonimlari) boʻlishi mumkin. Shuningdek, ularda soʻzning qaysi kontekstda qoʻllanishi va uning boshqa soʻz turkumlari bilan aloqasi koʻrsatilgan. Shu tarzda, ular soʻzning maʼnosi va ishlatilishini yaxshiroq tushunish uchun ishlatilishi mumkin.

Tezaurus mualliflar, muharrirlar va tarjimonlarga yordam berishi mumkin. Biroq, namoyishlardan foydalanish ehtiyotkorlik bilan va toʻgʻri bajarilishi kerak. Soʻzlarning maʼnolari va qoʻllanishlarini toʻliq tushunmaslik ularning notoʻgʻri ishlatilishiga olib kelishi mumkin. Shuning uchun, koʻrinishlarni ishlatishdan oldin, soʻzning maʼnosi va ishlatilishini toʻliq tushunish muhimdir. Tezaurus til oʻrganish, matn tasnifi, sinonimlarni topish, kontseptsiyani xaritalash va maʼlumot olish kabi sohalarida muhim oʻrin tutadi. Tezaurus soʻz maʼnolari va ishlatilishini yaxshiroq tushunish uchun ishlatilishi mumkin. Xususan, yozuvchilar, muharrirlar va tarjimonlar soʻz tanlash va imloni tekshirish uchun aforizmlardan foydalanishlari mumkin.

Bundan tashqari, tezislarda maʼlumot olishda muhim rol oʻynaydi. Xususan, qidiruv tizimlari va axborotni boshqarish tizimlari foydalanuvchilarga tezaurus yordamida oʻzlari qidirayotgan maʼlumotlarga tezroq va aniqroq kirish imkonini beradi. Tezaurus soʻz maʼnolari va qoʻllanishlarini aniqroq va izchil aniqlashga yordam beradi, shu bilan maʼlumot olish samaradorligini oshiradi. Natijada, dissertatsiya til oʻrganish, matnlarni tasniflash, axborotni boshqarish va maʼlumot qidirish kabi koʻplab sohalarida muhim oʻrin tutadi. Tezaurusdan foydalanish soʻzlarning maʼnolari va qoʻllanishini yaxshiroq tushunish, toʻgʻri va izchil yozish,

ma'lumotlarga kirish samaradorligini oshirish uchun juda foydali. Tezaurus lug'atlaring tarjima sohasida tutgan o'rni beqiyosdir.

Foydalanilgan adabiyotlar ro'yxati:

1. Abdurakhmonova N, Tuliyeu U. Morphological analysis by finite state transducer for Uzbek-English machine translation/Foreign Philology: Language. Literature, Education. 2018(3):68.
2. Abdurakhmonova N, Urdishev K. Corpus based teaching Uzbek as a foreign language. Journal of Foreign Language Teaching and Applied Linguistics (JFLTAL).
3. Abdurakhmonova N. The bases of automatic morphological analysis for machine translation. Izvestiya Kyrgyzskogo gosudarstvennogo tekhnicheskogo universiteta. 2016;2 (38):12-7.
4. Abduraxmonova, N. Z. "Linguistic support of the program for translating English texts into Uzbek (on the example of simple sentences): Doctor of Philosophy (PhD) il dis. aforef." (2018).
5. Aripov, M., Sharipbay, A., Abdurakhmonova, N., Razakhova B.: Ontology of grammar rules as example of noun of Uzbek and Kazakh languages. In: Abstract of the VI International Conference "Modern Problems of Applied Mathematics and Information Technology - Al-Khorezmiy 2018", pp. 37–38, Tashkent, Uzbekistan (2018)
6. Atkins, B. T. S. and Beth Levin, 1991. Admitting impediments. In Uri Zernik (ed.), Lexical Acquisition: Exploiting On-Line Resources to Build aLexicon. Hillsdale,
7. Atkins, Sue, 1993. Tools for computer-aided corpus lexicography: the Hector project. Acta Linguistica Hungarica, 41:5–72.
8. Baeza-Yates, Ricardo and Berthier Ribeiro-Neto, 1999. Modern Information Retrieval. ACM Press and Addison Wesley.
9. Buitelaar, Paul, 1998. CORELEX: Systematic Polysemy and Underspecification. Ph.D. thesis, Brandeis University.
10. CIDE, 1995. Cambridge International Dictionary of English. CUP, Cambridge, England. New Jersey: Lawrence Erlbaum, pages 233–262.