RAQAMLI TILSHUNOSLIK VA ELEKTRON LUG'ATLAR

Bobozaripova Malohat Ergash qizi
Navoiy davlat universiteti tayanch doktoranti
E-mail: malohatbobozaripova9@gmail.com

Annotatsiya. *Mazkur maqolada ingliz tilidagi WordNet leksik ma'lumotlar bazasi mukammal tezaurus namunasi sifatida tahlil qilinadi. Unda WordNet'ning yaratilish tarixi, tuzilishi, asosiy birliklari – synsetlar, shuningdek sinonimiya, polisemiya, giperonimiya, giponimiya, meronimiya va antonimiya kabi semantik munosabatlar yoritiladi. Leksik iyerarxiya va tarmoqlanish tamoyillari asosida so'zlarning konseptual tizimda tashkil etilishi ko'rsatib beriladi. Shuningdek, turkumlararo morfosemantik bog'lanishlar, troponimik munosabatlar hamda RuWordNet modeli misolida ko'p tilli leksik bazalar masalasi tahlil qilinadi. Maqolada semantik maydon yondashuvi asosida "ishla" leksemasining ko'p qatlamli strukturasi ochib berilib, uning leksik, grammatik va sintaktik xususiyatlari yoritiladi.*

Kalit so'zlar: *WordNet, tezaurus, synset, semantik munosabatlar, giperonimiya, giponimiya, meronimiya, antonimiya, semantik maydon, kompyuter lingvistikasi.*

Abstract. *This article analyzes WordNet as a comprehensive thesaurus model in English lexical studies. It discusses the history of WordNet's development, its structural organization, and its core unit – the synset. Special attention is given to semantic relations such as synonymy, polysemy, hypernymy, hyponymy, meronymy, and antonymy. The study highlights the hierarchical and network-based structure of lexical organization and examines cross-part-of-speech morphological-semantic relations and troponymy. The RuWordNet model is also considered as an example of multilingual lexical database development. Furthermore, the semantic field approach is applied to*

the analysis of the verb "ishla," demonstrating its multi-layered lexical, grammatical, and syntactic structure.

Keywords: *WordNet, thesaurus, synset, semantic relations, hypernymy, hyponymy, meronymy, antonymy, semantic field, computational linguistics.*

Kirish (mavzuning dolzarbligi, maqsad va vazifalar). WordNet – 1985- yilda Princetin Universitetida ishlab chiqilgan. WordNet so‘zlar o‘rtasidagi semantik va leksik munosabatlarni o‘z ichiga olgan ma’lumotlar bazasidan iborat [1]. Bular psixolingvistik bilimlarga muvofiq ishlab chiqilgan, chunki WordNet dastlab kompyuter uchun tabiiy matnni tushunarli qilish uchun ishlatilgan. Dasturiy ta’minot bilan burgalikda erkin qidiriladigan ve bepul mavjud bo‘lgan ma’lumotlar bazasi ham boshqa maqsadlarda foydalaniladi.

WordNet so‘zlar o‘rtasidagi semantik va leksik munosabatlarni o‘z ichiga olgan ma’lumotlar bazasidan iborat. Lug‘at nutqning asosiy muhim qismlari uchun to‘rtta tarmoqdan iborat: ***ot, fe’llar, sifatlar va qo‘shimchalar.***

WordNet-dagi asosiy lug‘at birligi bitta so‘z emas, balki o‘xshash ma’noga ega tarmoq so‘zlarni semantik tarmoq tuguniga birlashtiradigan “senset” deb nomlanagan sinonimik qatordir. Muhim jihatlaridan biri sinonimiya va polisemiya tarjimada bir qator murakkabliklarni yuzaga keltiradi. Masalan, bir so‘zning bir nechta ma’nolari va ularning sinonimlari bo‘lishi mumkin. Unda “sinonimiyada so‘z shakllar o‘rtasidagi leksik munosabatlar hisoblanib, sinonimiyada so‘zlarning o‘rtasidagi asosiy farqi uchun {} belgisi, boshqa qo‘shimcha leksik aloqalar uchun [] belgilari bilan belgilangan.

Asosiy qism. WordNet tizimida ingliz tilidagi barcha so‘z turkumlarining semantik ma’nolari ifodalangan va u semantik aloqalar orqali tashkillashtiradi. Ot so‘z turkumi misolida ushbu lug‘atning xususiyatlari tahlilga tortilar ekan, uning hajmi 80000 dan ziyod ot so‘z turkumiga to‘g‘ri

kelishi va bu ko‘rsatgich muayyan leksemani nutqiy birlik sifatida voqealanishida so‘z birikmasi holatida qo‘llanish imkoniyati qamrab olingan. WordNet kompyuter leksikografiyasida erishilgan innovatsion yutuqlardan biridir. Chunki unga ilova etilgan so‘zlar mashina o‘qishi uchun mumkin bo‘lgan holatga moslashtirilgan.

Odatiy lug‘atlarda kiritilgan so‘zliklarning talaffuzi, grammatik shakllanishi, yasalishi, etimologiyasi, izoh hamda sinonim, antonim kabi yana bir qator lingvistik xususiyatlari singdirilgan bo‘ladi. Bunday belgilar mashina o‘qishi mumkin bo‘lgan imkoniyati mavjud bo‘lmaganligi bois aksariyat jihatlari tushurib qoldiriladi, xususan, WordNet yuqorida qayd etilgan talaffuz, etimologiyasi tavsif va shunga o‘xshash ma’lumotlarni o‘z ichiga qamrab olmaydi. Ushbu lug‘atda o‘rin olgan maqolalarning qo‘llanishida ifoda ma’nolarining semantik xususiyatlari ochib berish bilan bog‘liq afzalliklarni yaratish, qolaversa, to‘g‘ri tarjimaga erishishda so‘z maydolarini mantiqiy loyihalash birlamchi mezon bo‘lib hisoblanadi.

Shu jihatdan WordNet so‘zlarning semantik aloqalarini chuqur o‘rganish sinonimiya hodisasi muhim deb qaraladi. WordNetda sinonimlar muayyan bloklarga ajratilib, semantik jihatdan tasniflangan. Sparck Jones (1964-1986) o‘zining semantik tasnif nazariyasida asos solgan tadqiqotchidir. U o‘zining bu boradagi izlanishlarida matndan olingan muayyan so‘z shakllari bilan birga bo‘lishi mumkin bo‘lgan barcha sinonimlar tizimi yaratadi [2]. Masalan, u ingliz tilidagi qurol-yarog‘ bilan bog‘liq tushuncha **pellet** hamda **injection** (inyeksiya) so‘zlarining kontekstdagi ma’nodoshi **shot** leksemasi bilan bog‘laydi, bu esa shartli ravishda muayyan holat uchun sinonimik qatorni hosil qiladi. Sinonimlar tizimi (synset) o‘zaro semantik bog‘lansa-da har ikki tizimning strukturasi har xil tipga tegishli bo‘ladi: {*shot, pellet*} va {*shot, injection*}. Bu ikki to‘plam o‘zaro hech qanday ma’no jihatdan aloqaga kirishmaydi, ya’ni ularni faqat “shot” leksemasigina bog‘lab turadi. Aksariyat synsetda oddiy lug‘atlardagi izohlardan foydalaniladi. Polisemantik so‘zlarda

bir necha ko‘chma ma’nolar ifodalangan glossemalar mavjud bo‘lsa, synsetda faqat yagona glossema mavjud bo‘ladi. WordNetdagi synset orqali so‘zliklarning leksikallashuv konsepti shakllanadi.

WordNet-dagi asosiy munosabat, berkitish va yopish yoki mashina va avtomobil so‘zlarni orasidagi sinonimlikdir. Sinonimlar – bir xil tushunchani bildiruvchi va ko‘p konteksda bir-birining o‘rnini bosadigan so‘zlar bo‘lib, ular tartiblanmagan to‘plamlarga (sinsets) birlashtirilgan. WordNetning 117000 ta sinsetining har biri “konseptual munosabatlar” yordamida boshqa sinsetlar bilan bog‘langan. Bundan tashqari, sinset qisqacha ta’rifga(gloss) ega bo‘lib, aksariyat hollarda bir yoki bir nechta qisqa jumlalar bilan izohlanadi. Bir necha ma’noli so‘z shakllari bir necha xil sinsetlarda kuzatilishi mumkin. Shu sababli, WordNetdagi har bir shakl-ma’no juftligi o‘ziga xosdir.

Sinsetlar orasida eng ko‘p qo‘llangan munosabatlar ikkinchi darajali munosabatlar (shuningdek, giperonomiya, giponimiya yoki iyerarxik munosabatlar) deb ataladi [3]. Ular umumiy sinsetlarni (*mebel, uy jihozi*) tobora aniqroq shakllangan sinsetlarga (*krovat, yotoq*) bog‘laydi. Shu sababli, WordNetda ushbe mebellar kategoriyasiga krovat va yotoq kiradi: yoki aksincha, krovat va yotoq kabi tushunchalar mebellar kategoriyasini tashkil qiladi, deb qabul qiladi. Barcha ot iyerarxiyalari oxir-oqibat ildiz tuguniga borib taqaladi. Giponimiya munosabati o‘timli bo‘lib kreslo stulning bir turi bo‘lsa va stul – bu mebelning bir turi bo‘lsa, unda kreslo – bu mebelning bir turi deb qaraladi. WordNet turdosh otlarni va atoqli otlarni (ma’lum bir shaxslar, mamlakatlar va geografik hududlar) ajratib turadi. Albert Eynshteyn esa olim so‘ziga misol bo‘la oladi. Atoqli otlar o‘z iyeararxiyalarda eng so‘nggi o‘rinda yoziladi.

Sinonimlar o‘rtasida ma’no yaqinligi mavjud bo‘lsa-da, semantik tizim bundan ancha kengroq tushunchani qamrab oladi. Ayniqsa, otlarning leksikallashuv jarayonida ular o‘rtasidagi iyerarxik bog‘lanishlar muhim o‘rin tutadi. Bunday bog‘lanishlar subordinatsiya asosida shakllanadi va tasnifiy

guruhlanishni yuzaga keltiradi. Ushbu tizimda torroq ma’noni bildiruvchi birliklar giponim, umumiyroq ma’noni ifodalovchi birliklar esa giperonim sifatida talqin qilinadi. Masalan, “*burgut*” tushunchasi “*qush*”ning bir turi hisoblanadi, demak u giponim; “*qush*” esa kengroq tushuncha sifatida giperonim vazifasini bajaradi.

Leksik iyerarxiya bosqichma-bosqich yuqoriga qarab umumlashib boradi. Masalan, ma’lum bir qush turi “*qush*” kategoriyasiga, u esa “*hayvon*”, undan keyin “*tirik mavjudot*” tushunchasiga ulanadi. Bu jarayonda har bir pog‘ona oldingi birlikni yanada umumiyroq semantik maydonga olib chiqadi. Belgilash tizimida giperonimdan giponimga yoki aksincha yo‘nalishlar maxsus belgilar orqali ifodalanadi va bu ma’no kengayishi yoki torayishini ko‘rsatadi. Natijada synsetlar o‘rtasida doimiy o‘zaro almashinuv – ya’ni biri uchun umumiy, boshqasi uchun xususiy bo‘lish holati yuz beradi.

Xulosa va amaliy tavsiyalar. Ko’rinadiki, lug‘at tuzishda faqat sinonimik munosabatlarni emas, balki uyadoshlik va semantik bog‘liqlikni ham hisobga olish muhim ahamiyat kasb etadi. Uyadosh so‘zlar bir paradigmatik maydon doirasida umumiy semantik belgi asosida birlashadi, semantik bog‘liq birliklar esa ko‘proq sintagmatik munosabat asosida, ya’ni nutq jarayonida bir-birini talab qilishi orqali shakllanadi. Shu jihatdan lug‘atning semantik bazasi nafaqat ma’no o‘xshashligi, balki ma’no yaqinligi va funksional qo‘llanish chastotasini ham qamrab oladi. “Ishla” leksemasining semantik maydonini tahlil qilish esa bitta fe’l atrofida qanday keng va ko‘p qatlamli mazmun tizimi shakllanishini ko‘rsatadi. Ushbu leksema markaziy birlik sifatida turib, uning atrofida sinonimlar, giponimlar, xolonim va meronim munosabatlar, graduonimik darajalanish hamda funksional ma’no kengayishlari yuzaga keladi. Natijada oddiy ko‘rinadigan bir fe’l murakkab semantik struktura markaziga aylanadi. Shuningdek, morfologik imkoniyatlarning kengligi leksemaning nutqdagi faolligini oshiradi. O‘timli va o‘timsiz shakllarda qo‘llanishi, nisbat, mayl va zamon qo‘shimchalarini qabul qilishi uning

grammatik moslashuvchanligini ta’minlaydi. Bu esa semantik maydonning nafaqat leksik, balki grammatik sath bilan ham uzviy bog‘liqligini ko‘rsatadi.

FOYDALANILGAN ADABIYOTLAR RO‘YXATI

1. Miller G. A. WordNet: A Lexical Database for English // Communications of the ACM. – 1995. – Vol. 38, No. 11. – P. 39–41.
2. Sparck Jones K. Synonymy and Semantic Classification. – Edinburgh: Edinburgh University Press, 1986.
3. Amsler R. A. The Structure of the Merriam-Webster Pocket Dictionary. – Austin: University of Texas, 1980.
4. Touretzky D. S. The Mathematics of Inheritance Systems. – Los Altos: Morgan Kaufmann Publishers, 1986.