

MALOHAT BOBOZARIPOVA

Navoiy davlat universiteti tayanch doktoranti¹

WORDNET – INGLIZ TILIDAGI ZAMONAVIY TEZAURUS LUG‘ATI NAMUNASI SIFATIDA

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. WordNet – 1985-yilda Prinston 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 birgalikda erkin qidiriladigan va 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 qatoridir. 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 mum-

¹ E-mail: malohatbobozaripova9@gmail.com

kin. Unda “sinonimiyada soʻz shakllar oʻrtasidagi leksik munosabatlar hisoblanib, sinonimiyada soʻzlarning oʻrtasidagi asosiy farqi uchun {} belgisi, boshqa qoʻshimcha leksik aloqalar uchun [] belgilari bilan belgilangan.

Asosiy qism. WordNet tizimida ingliz tilidagi barcha soʻz turkumlarining semantik maʼnolari ifodalangan va u semantik aloqalar orqali tashkillashtiradi. Ot soʻz turkumi misolida ushbu lugʻatning xususiyatlari tahlilga tortilar ekan, uning hajmi 80000 dan ziyod ot soʻz turkumiga toʻgʻri kelishi va bu koʻrsatkich muayyan leksemani nutqiy birlik sifatida voqealanishida soʻz birikmasi holatida qoʻllanish imkoniyati qamrab olingan. WordNet kompyuter leksikografiyasida erishilgan innovatsion yutuqlardan biridir. Chunki unga ilova etilgan soʻzlar mashina oʻqishi uchun mumkin boʻlgan holatga moslashtirilgan.

Odatiy lugʻatlarda kiritilgan soʻzliklarning talaffuzi, grammatik shakllanishi, yasalishi, etimologiyasi, izoh hamda sinonim, antonim kabi yana bir qator lingvistik xususiyatlari singdirilgan boʻladi. Bunday belgilar mashina oʻqishi mumkin boʻlgan imkoniyati mavjud boʻlmaganligi bois aksariyat jihatlari tushurib qoldiriladi, xususan, WordNet yuqorida qayd etilgan talaffuz, etimologiyasi tavsif va shunga oʻxshash maʼlumotlarni oʻz ichiga qamrab olmaydi. Ushbu lugʻatda oʻrin olgan maqolalarning qoʻllanishida ifoda maʼnolarining semantik xususiyatlari ochib berish bilan bogʻliq afzalliklarni yaratish, qolaversa, toʻgʻri tarjimaga erishishda soʻz maydonlarini mantiqiy loyihalash birlamchi mezon boʻlib hisoblanadi.

Shu jihatdan WordNet soʻzlarning semantik aloqalarini chuqur oʻrganish sinonimiya hodisasi muhim deb qaraladi. WordNetda sinonimlar muayyan bloklarga ajratilib, semantik jihatdan tasniflangan. Sparck Jones (1964-1986) oʻzining semantik tasnif nazariyasida asos solgan tadqiqotchidir. U oʻzining bu boradagi izlanishlarida matndan olingan muayyan soʻz shakllari bilan birga boʻlishi mumkin boʻlgan barcha sinonimlar tizimi yaratadi [2]. Masalan, u ingliz tilidagi qurol-yarogʻ bilan bogʻliq tushuncha **pellet** hamda **injection** (inyeksiya) soʻzlarining kontekstdagi maʼnodoshi **shot** leksemasi bilan bogʻlaydi, bu esa shartli ravishda muayyan holat uchun sinonimik qatorni hosil qiladi. Sinonimlar tizimi (synset) oʻzaro semantik bogʻlansa-da har ikki tizimning strukturasi har xil tipga tegishli boʻladi: {*shot, pellet*} va {*shot, injection*}. Bu ikki toʻplam oʻzaro hech qanday maʼno jihatdan aloqaga kirishmaydi, yaʼni ularni faqat “shot” leksemasigina bogʻlab turadi. Aksariyat synsetda oddiy lugʻatlardagi izohlardan foydalaniladi. Polisemantik soʻzlarda bir necha koʻchma maʼnolar ifodalangan glossemalar mavjud boʻlsa, synsetda faqat yagona glossema mavjud boʻladi. WordNetdagi synset orqali soʻzliklarning leksikallashuv konsepti shakllanadi.

WordNet-dagi asosiy munosabat, berkitish va yopish yoki mashina va avtomobil soʻzlarni orasidagi sinonimlikdir. Sinonimlar – bir xil tushunchani bildiruvchi va koʻp kontekstda bir-birining oʻrnini bosadigan soʻzlar boʻlib, ular tartiblanmagan toʻplamlarga (sinsets) birlashtirilgan. WordNetning 117000 ta sinsetining har biri “konseptual munosabatlar” yordamida boshqa sinsetlar bilan bogʻlangan. Bundan tashqari, sinset qisqacha taʼrifga (gloss) ega boʻlib, aksariyat hollarda bir yoki bir nechta qisqa jumlar bilan izohlanadi. Bir necha maʼnoli soʻz shakllari bir necha xil sinsetlarda kuzatilishi mumkin. Shu sababli, WordNetdagi har bir shakl-maʼno juftligi oʻziga xosdir.

Sinsetlar orasida eng koʻp qoʻllangan 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*) bogʻlaydi. Shu sababli, WordNetda ushbu 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 oʻtimli boʻlib kreslo stulning bir turi boʻlsa va stul – bu mebelning bir turi boʻlsa, unda kreslo – bu mebelning bir turi deb qaraladi. WordNet turdosh otlarni va atoqli otlarni (maʼlum bir shaxslar, mamlakatlar

va geografik hududlar) ajratib turadi. Albert Eynshteyn esa olim soʻziga misol boʻla oladi. Atoqli otlar oʻz iyeararxiyalarda eng soʻnggi oʻrinda yoziladi.

Sinonimlar oʻrtasida maʼno yaqinligi mavjud boʻlsa-da, semantik tizim bundan ancha kengroq tushunchani qamrab oladi. Ayniqsa, otlarning leksikallashuv jarayonida ular oʻrtasidagi iyerarxik bogʻlanishlar muhim oʻrin tutadi. Bunday bogʻlanishlar subordinatsiya asosida shakllanadi va tasnifiy guruhlanishni yuzaga keltiradi. Ushbu tizimda torroq maʼnoni bildiruvchi birliklar giponim, umumiyroq maʼnoni 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, maʼlum bir qush turi *“qush”* kategoriyasiga, u esa *“hayvon”*, undan keyin *“tirik mavjudot”* tushunchasiga ulanadi. Bu jarayonda har bir pogʻona oldingi birlikni yanada umumiyroq semantik maydonga olib chiqadi. Belgilash tizimida giperonimdan giponimga yoki aksincha yoʻnalishlar maxsus belgilar orqali ifodalanadi va bu maʼno kengayishi yoki torayishini koʻrsatadi. Natijada synsetlar oʻrtasida doimiy oʻzaro almashinuv – yaʼni biri uchun umumiy, boshqasi uchun xususiy boʻlish holati yuz beradi.

Bunday tarmoqli tuzilma anʼanaviy lugʻatlardagi chiziqli berilishdan farq qiladi. Tarmoqlanish modeli soʻzlarni aylana yoki ketma-ket izohlash oʻrniga ularni oʻzaro bogʻliq tugunlar sifatida koʻrsatadi. Kompyuter texnologiyalarida aynan shu usul samarali hisoblanadi, chunki har bir birlik maxsus identifikator orqali boshqa birliklar bilan bogʻlanadi. Katta hajmdagi maʼlumotlar bazasini shakllantirishda iyerarxik tizim joyni tejaydi va maʼlumotlarni tezkor izlash imkonini beradi [4]. Shu bois bu model *“vorislik tizimi”* deb ham ataladi — yaʼni maʼno yuqoridan pastga yoki pastdan yuqoriga meros tarzida uzatiladi.

Semantik aloqalar sinonimiya bilan cheklanmaydi. Antonimlar qarama-qarshi maʼnoni, meronimlar esa butun va qism munosabatini ifodalaydi. Masalan, *“gʻildirak”* va *“velosiped”* oʻrtasidagi aloqa qism-butun munosabatiga asoslanadi. Bundan tashqari, bir xil soʻz turkumi doirasidagi munosabatlar bilan birga turkumlararo bogʻlanishlar ham mavjud. Masalan, feʼl, ot va sifat oʻrtasidagi morfosemantik aloqalar bir oʻzakdan yasalgan birliklar orqali namoyon boʻladi: *observe – observant – observation* kabi. Feʼllar tizimida troponimik munosabatlar ham uchraydi. Bu holda bir feʼl boshqa feʼlning bajarilish usulini aniqlashtiradi. Masalan, *“vaysamoq”* feʼli *“gapirmoq”* harakatining maʼlum usulini bildiradi. Shuningdek, feʼllar oʻzaro antonimik munosabatga ham kirishadi: *“turmush qurmoq”* va *“ajrashmoq”* kabi.

Rus tilidagi WordNet modeli — **RuWordNet** — ham shunday tamoyillar asosida qurilgan boʻlib, unda ot, sifat va feʼllar synsetlar koʻrinishida beriladi [5]. Masalan, *“bermoq”* maʼnosini anglatuvchi feʼl bir nechta sinonimik shakllar bilan bir synsetga birlashtiriladi, undan yuqori bosqichda umumiyroq maʼnodagi giperonimlar, pastki bosqichda esa xususiyoq giponimlar joylashadi. Bu esa maʼno tasnifini aniq koʻrsatishga xizmat qiladi.

Lugʻat bazasida uyadosh va semantik bogʻliq soʻzlar ham alohida ahamiyatga ega. Uyadosh birliklar umumiy mavzuviy guruh asosida birlashadi. Masalan, *“ananas”* va *“oʻrik”* sinonim emas, biroq *“meva”* kategoriyasiga mansub boʻlgani sababli bir paradigmatic maydonni tashkil etadi. Semantik bogʻliq soʻzlar esa koʻproq nutq jarayonida bir-birini talab qilishi bilan xarakterlanadi. Masalan, *“paxta”* soʻzi *“yigʻmoq”*, *“terim”*, *“yumshoq”* kabi birliklar bilan tez-tez qoʻllanib, asotsiativ bogʻlanish hosil qiladi.

Semantik maydon yondashuvi ayniqsa feʼl, ot va sifat turkumlarida yaqqol namoyon boʻladi. Masalan, *“ishla”* leksemi markaziy birlik sifatida olinib, uning atrofida sinonimlar (*mehnat qilmoq*, *ter toʻkmoq*), leksik-semantik guruh (*boshqarmoq*, *qurmoq*, *yaratmoq*), semantik toʻda (*zavod*, *korxona*, *ishchi*), shuningdek giponim, xolonim va meronim munosabatlari shakllanadi.

Bu fe'ning grammatik imkoniyatlari ham keng: u turli zamon, nisbat va mayl shakllarini qabul qiladi, o'timli va o'timsiz ko'rinishda qo'llanadi. Sintaktik jihatdan esa "zavodda ishlamoq", "qattiq ishlamoq" kabi birikmalarda namoyon bo'ladi va ishchi, dehqon, korxona kabi birliklar bilan assotsiativ aloqaga kirishadi.

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 paradigmatic 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.

Sintaktik aloqalar va assotsiativ bog'lanishlar esa leksemaning real nutq jarayonidagi funksional imkoniyatlarini namoyon etadi. "Zavodda ishlamoq", "qattiq ishlamoq" kabi birikmalar fe'ning kontekstual aniqligini belgilasa, ishchi, dehqon, korxona kabi birliklar bilan yuzaga keladigan assotsiatsiyalar uning konseptual maydonini kengaytiradi.

Umuman olganda, keltirilgan tahlil shuni ko'rsatadiki, semantik maydon yondashuvi til birliklarini izolyatsiyada emas, balki tizimli va o'zaro bog'liq holda o'rganishga imkon beradi. Uyadoshlik, semantik bog'liqlik, iyerarxik munosabatlar hamda morfologik va sintaktik xususiyatlarning birgalikda tahlil qilinishi lug'atshunoslikda mukammal, ko'p qatlamli lingvistik model yaratishga xizmat qiladi. Shu asosda shakllangan lug'aviy tizim esa nafaqat nazariy tadqiqotlar, balki amaliy tilshunoslik, tarjima va kompyuter lingvistikasi sohalarida ham samarali qo'llanilishi mumkin.

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.
5. Loukachevitch N., Lashevich G. *RuWordNet: Russian WordNet Development and Application* // *Computational Linguistics and Intellectual Technologies*. – 2016.
6. Fellbaum C. (Ed.). *WordNet: An Electronic Lexical Database*. – Cambridge, MA: MIT Press, 1998.
7. Jurafsky D., Martin J. H. *Speech and Language Processing*. – 3rd ed. – Pearson, 2021.
8. Levin B. *English Verb Classes and Alternations*. – Chicago: University of Chicago Press, 1993.