SUPERFRAZALI TURG‘UN BIRLIKLARNI RAQAMLASHTIRISH MUAMMOLARI

Kabulova Umida Sayidmaxamadovna
ADCHTI Umumiy va qiyosiy tilshunoslik kafedrasi professori
Email: umidakabulova@mail.ru

Annotatsiya: Mazkur maqolada biz matnli birliklarni kompyuter lingvistikasida kodlashtirish xususida so‘z yuritamiz.Biroq elektron lug‘at yaratishning eng poydevor bosqichlari texnologik va konseptual jarayonlar hisoblanib, ular mutaxassisdan alohida ilmiy yondashuvni talab etadi. Aks holda, zamonaviy elektron leksikografiya mezonlariga mos kelmaydigan mahsulot yuzaga kelishi mumkin.

Kalit so‘zlar: *superfrazali birliklar, kompyuter leksikografiyasi, kodli format, raqamli ko‘rinish, teglash, elektron lug‘at, interpretatsiya*

Metaleksikografik modellashtirishdagi diskursiv muammolar. Kompyuter lug‘atshunosligida SFBlarni tizimga kiritishning asosiy qiyinchiligi ularning diskursiv tabiati bilan bog‘liqdir. Standart leksema yoki tayyor iboralardan farqli o‘laroq, SFB tarkibidagi mazmun faqatgina komponentlar yig‘indisidan iborat bo‘lmay, balki kontekstual yaxlitlik, ya’ni koheziya va koherensiya qonuniyatlari asosida yuzaga chiqadi.

Elektron lug‘atlar platformasi uchun SFB strukturasini shakllantirish quyidagi tizimli ilmiy-metodik bosqichlar asosida amalga oshiriladi:

1. Formal-struktural koordinatsiyalash va modellashtirish. SFB tarkibiga kiruvchi gaplararo bog‘lanish usullari (zanjirsimon aloqa va parallel munosabatlar) raqamli indekslar va operatorlar korpusiga o‘tkaziladi:

2. Anaforik munosabatlar identifikatsiyasi: matndagi olmoshlar, kontekstual sinonimlar va koreferentlik zanjirlarini kompyuter orqali avtomatlashtirilgan aniqlash va teglash tizimi yaratiladi.

3. Konnektorlar taksonomiyasi: *"shuning uchun", "binobarin", "xullas"* singari matn shakllantiruvchi va mantiqiy aloqani ta'minlovchi vositalarning semantik-logik vazifalari raqamli algoritm ko'rinishida kodlanadi.

Elektron lug'at bazalarida har bir superfrazali birlik muayyan konseptual soha yoki mikromavzu doirasida semantik jihatdan klassifikatsiyalanadi. Mazkur jarayon ontologik iyerarxiya tamoyillariga tayangan holda ikki bosqichda amalga oshiriladi:

1) denotativ sath; 2) konnotativ sath.

Denotativ sathda matnning bevosita nima haqida ekanligi, undagi asosiy axborot va referensiyalar aniqlansa, konnotativ sathda esa muallifning matnga yuklagan yashirin ma'nosi, jumladan, emotsional-ekspressiv bo'yoqlar, kesatiq (ironiya) hamda pragmatik yo'naltirilganlik darajasi baholanadi.

Algoritmik tahlil va kodlashtirish tamoyillari. Kompyuter lingvistikasida superfrazali birlikni identifikatsiyalash — eng avvalo, uning sintaktik va semantik chegaralarini aniq belgilashni anglatadi. Ushbu maqsadlar uchun maxsus makro va mikrostruktura modellari ishlab chiqilib, XML maketlari orqali formatlanadi. XML formatidagi teglashga misol:

Elektron lug'at platformalarida superfrazali birliklar uchun shakllantiriladigan maxsus matnlar muayyan funksional komponentlar majmuasidan iborat bo'lishi shart.

Ilmiy kuzatishlar shuni ko'rsatadiki, SFBlarni raqamlashtirishda ulardagi transformatsion o'zgarishlar va dinamika eng murakkab bosqich hisoblanadi. Turg'un iboralar hamda frazeologik birliklar SFB muhitiga kirganda, ko'pincha, semantik

modifikatsiyaga uchraydi (yangi ma’no kasb etadi). Kelajakda ushbu muammolarni bartaraf etish uchun quyidagi istiqbolli texnologiyalar integratsiya qilinadi:

1. Neyrotarmoqli modellar. Bunda SFB strukturalarini shunchaki bazada saqlab qolmay, balki kontekstual o‘xshash muqobil variantlarni avtomatik generatsiya qilish tizimini yo‘lga qo‘yiladi.

2. Interaktiv korpus integratsiyasi (Linguistic linked data). Unda lug‘at tizimini real o‘quv va amaliy korpuslar bilan o‘zaro bog‘liq holda dinamik yangilab boriladi.

3. Ontologik ma’lumotlar bazasi. Ushbu bazada SFBlarni umumiy lisoniy olam manzarasi (konseptual xaritalar) bilan uzviy aloqadorlikda tizimlashtiriladi.

Superfrazali birliklarni raqamli lug‘atlarga kiritish va ularni kodlashtirish — an’anaviy, statik leksikografiyadan dinamik xarakterga ega bo‘lgan kompyuter semantikasiga o‘tishni ifodalaydi. Bu kabi murakkab vazifalarni bajarish zamonaviy tilshunosdan klassik sintaktik va paremiologik fundamental bilimlaridan tashqari, ma’lumotlar bazasini boshqarish (DBMS) hamda tarkibiy tilshunoslikning eng yangi metodologiyalarini mukammal darajada egallagan bo‘lishni talab qiladi.

Bizning ilmiy-tadqiqot ishimiz doirasida milliy tilshunoslik paradigmasiga oid turg‘un birliklarning raqamli platformasini loyihalashda mazkur kodlashtirish qoidalari bazaviy metodologik asos bo‘lib xizmat qiladi. Leksikografik manbalarni raqamlashtirish va kodlashtirishning konseptual jihatlari quyidagi tizimli bosqichlarni qamrab oladi:

II. Leksikografik manbalarni raqamlashtirish bosqichlari:

1.Leksikografik materiallarni unifikatsiyalash: mavjud matnli ma’lumotlarni dastlabki kompyuterbop (raqamli) ko‘rinishga keltirish.

2.Abstrakt modellashtirish. Lug‘atning umumiy axborot-funksional tizimini va uning ishlash algoritmlarini (biznes-jarayonlarini) loyihalash.

3.Tizimli strukturalash. Lug‘at maqolasidagi so‘z shakli, morfologik ko‘rsatkichlar, semantika, illyustrativ misollar va sinonimik qatorlarni alohida-alohida ma’lumotlar maydoniga (fields) joylashtirish.

III. Dasturiy muhandislik (Programming). *Eslatma:* dasturlash bosqichi — tayyorlangan raqamli lug‘at materiallarini vizual interfeys ko‘rinishiga keltirish va foydalanuvchi uchun moslashuvchan qidiruv mexanizmlarini ta’minlovchi funksional dasturiy kodni ishlab chiqish jarayonidir.

IV. Kodlashtirish amaliyotining afzalliklari:

1.Tezkor va intellektual qidiruv. Maxsus optimallashtirilgan algoritmlar yordamida muayyan so‘z yoki frazeologik birikmalarni soniyaning ulushlarida topish imkoniyati.

2.Multimediali integratsiya. Matnli ma’lumotlar bilan bir qatorda, audio (to‘g‘ri talaffuz), vizual tasvirlar hamda animatsion komponentlarni yagona kod ostida sinxronizatsiya qilish imkoniyati.

Yelena Karpilovkaning yondashuviga ko‘ra[1], raqamli (kompyuter) lug‘atlardagi maqola strukturasi — bu unda ifodalangan til birliklari to‘g‘risidagi ma’lumotlarni tizimlashtirish, joylashtirish va vizual-grafik aks ettirishning o‘ziga xos modelidir. Til materiallarini dasturiy ta’minot orqali joylashtirish moslashuvchan mexanizmga asoslanadi; bunda har bir axborot toifasi uchun alohida zona ajratiladi va ushbu qismlarning mavjudligi funksional ahamiyatga ega. Bunday raqamli formatni

[1] Карпиловка Е.А. Введение в прикладную лингвистику. – М.: Наука. – С. 49.

shakllantirish ma’lumotlar bazasini yaratish va leksikografik protsessorni loyihalash bilan bir qatorda, elektron lug‘atlarni ishlab chiqish jarayonining fundamental bosqichi hisoblanadi.

N.Sivakova elektron lug‘atlarning gipermatnli tabiati an’anaviy “kompozitsiya” tushunchasini tubdan qayta ko‘rib chiqishni talab qilishini ta’kidlaydi[2].

Raqamli lug‘at makonida kompozitsiya deganda ko‘pincha dastur menyusi yoki uning tashqi interfeysi tushuniladi. Biroq, bu faqat yuzaki (tashqi) tasnifdir. Zamonaviy elektron lug‘atlarda foydalanuvchi o‘zining shaxsiy so‘rovlari orqali muayyan axborot zonalarini faollashtirishi va shu tariqa interfeysning individual kompozitsion yechimini mustaqil ravishda shakllantirishi mumkin. I.Kudashevning ta’kidlashicha[3], lug‘at tarkibini tizimlashtirish dastlabki bosqichda mikrostatlarni — lug‘at birliklari va ular to‘g‘risidagi ma’lumotlar bazasini to‘g‘ri modellashtirishdan boshlanishi maqsadga muvofiqdir. Bunda leksikograf mikrostatlar tasnifini shakllantiradi, ularning semantik-struktur tuzilishini loyihalashtiradi va o‘zaro muvofiqlashtiradi. Ko‘p tilli lug‘atlarda tarjima muqobillariga ega bo‘lgan leksik birliklarning mikrostatlari lug‘at maqolasining serqirra modellarini yuzaga keltirishi bilan hisoblanadi. Shuningdek, tilshunos olim V.Dubichinskiy ham lug‘at maqolasini lug‘at tizimining eng markaziy va asosiy bo‘g‘ini sifatida baholaydi.

Xulosa va takliflarimiz quyidagicha, elektron leksikografik manbalarni ishlab chiqish ichki va tashqi o‘zaro aloqalar tizimidan iborat murakkab strukturani

[2] Сивакова Н.А. Лексикографическое описание английских и русских фитонимов в электронном словаре: Автореф. дисс. к.ф.н. –Тюмень, 2004. – С. 15
[3] Кудашев И.С. Проектирование переводных словарей специальных словарей. – Хельсинки, 2007. – С. 364

modellashtirish va uning komponentlarini kompozitsion jihatdan tashkil etish jarayonidir. Bugungi kunda elektron lug‘atlarni loyihalash bosqichlarini tartibga soluvchi yagona va universal tizim shakllanmagan. Biroq elektron lug‘at yaratishning eng poydevor bosqichlari texnologik va konseptual jarayonlar hisoblanib, ular mutaxassisdan alohida ilmiy yondashuvni talab etadi. Aks holda, zamonaviy elektron leksikografiya mezonlariga mos kelmaydigan mahsulot yuzaga kelishi mumkin. Mazkur muammo, ayniqsa, ta’lim platformalari uchun mo‘ljallangan o‘quv lug‘atlarini yaratishda yaqqol ko‘zga tashlanadi. Tahliliy va konseptual bosqichda yo‘l qo‘yilgan har qanday metodologik xatolik yakunda raqamli ta’lim resursining didaktik prinsiplariga putur yetkazadi.

Foydalanilgan adabiyotlar:

1. Кудашев И.С. Проектирование переводных словарей специальных словарей. – Хельсинки, 2007. – 364 с.
2. Карпиловка Е.А. Введение в прикладную лингвистику. – М.: Наука. – С. 49.
3. Лингвистико-информатические студии: Прази украинского Мовноинформатического Фонда национальной академии наук Украины: 3-jild. Лексикография. – Тлумачна. Книга 1. – Киев, 2018.
4. Сивакова Н.А. Лексикографическое описание английских и русских фитонимов в электронном словаре: Автореф. дисс. к.ф.н. –Тюмень, 2004. – С. 15.