O‘ZBEK TILIDA FE’L SO‘Z TURKUMINING MORFOLOGIK PARADIGMALARINI ELEKTRON BAZAGA AYLANTIRISH: MUAMMOLAR VA YECHIMLAR

Muslima Risqulova

Annotatsiya. *Ushbu maqolada o‘zbek tilidagi fe’l so‘z turkumining morfologik paradigmalarini to‘liq va tizimli elektron bazaga aylantirish masalasi yoritiladi. Fe’lning boy grammatik kategoriyalari tahlil qilinib, mavjud elektron resurslarning yetishmovchiliklari aniqlanadi. Tadqiqotda paradigmalarni elektron shaklga keltirishning asosiy muammolari ochib beriladi hamda ularni bosqichma-bosqich hal etish yo‘llari taklif etiladi. Maqola natijalari keyinchalik to‘liq morfologik leksikografik baza va annotatsiyalangan korpus yaratish uchun nazariy-amaliy asos bo‘lib xizmat qiladi.*

Kalit so‘zlar: *fe’l, morfologik paradigma, elektron baza, so‘z yasalishi, annotatsiya, NLP, o‘zbek tili.*

Abstract. *This article addresses the issue of converting the morphological paradigms of the Uzbek verb class into a complete and systematic electronic database. The rich grammatical categories of the verb are analyzed, and the shortcomings of existing electronic resources are identified. The study reveals the main problems in converting paradigms into electronic form and proposes step-by-step solutions for addressing them. The results of the article will serve as a theoretical and practical basis for the subsequent creation of a full morphological lexicographic database and an annotated corpus.*

Keywords: *verb, morphological paradigm, electronic database, word formation, annotation, NLP, Uzbek language.*

O‘zbek tili agglyutinativ tillar guruhiga mansub bo‘lib, unda so‘z shakllari asosga qo‘shimchalar ketma-ket qo‘shilishi orqali hosil qilinadi. So‘z turkumlari ichida fe’l morfologik jihatdan eng murakkab va boy tizimga ega. Birgina fe’ldan zamon, mayl, nisbat, shaxs-son, bo‘lishli-bo‘lishsizlik kabi kategoriyalar yordamida yuzlab shakllar yasalishi mumkin.

Hozirgi kunda sun’iy intellekt, mashina tarjimasi, avtomatik matn tahlili va boshqa NLP yo‘nalishlarining rivojlanishi tilning morfologik resurslarini raqamlashtirishni taqozo etmoqda. O‘zbek tili hali past resursli tillar qatorida qolayotganligi sababli, fe’l paradigmalarining to‘liq va tizimli elektron bazasi mavjud emas. Mavjud milliy korpus va morfologik analizatorlar umumiy xarakterga ega bo‘lib, fe’lga maxsus, to‘liq paradigmatik resurs sifatida ishlab chiqilmagan. Ushbu maqolaning maqsadi – o‘zbek tilida fe’l so‘z turkumining morfologik paradigmalarini elektron bazaga aylantirishdagi asosiy muammolarni aniqlash va ularni amaliy yechimlar bilan bog‘lab ko‘rsatishdan iborat.

1. O‘zbek tilida fe’l paradigmalarining hozirgi holati

Fe’l o‘zbek tili grammatikasining markaziy birligi hisoblanadi. Uning asosiy grammatik kategoriyalari quyidagilardan iborat:

zamon (hozirgi, o‘tgan, kelasi);

mayl (buyruq, shart, istak va boshqalar);

nisbat (aniq, majhul, o‘zlik, orttirma, birgalik);

shaxs-son;

bo‘lishli va bo‘lishsiz shakllar;

harakat nomi, sifatdosh, ravishdosh kabi noaniq shakllar.

Bu kategoriyalar o‘zaro birikib, juda keng paradigma hosil qiladi. Mavjud elektron resurslar (O‘zbek tili milliy korpusi, ayrim morfologik analizatorlar) fe’l shakllarini qisman qamrab olsa-da, quyidagi kamchiliklarga ega:

to‘liq paradigmalar tizimli tarzda berilmagan;

fe’lga maxsus leksikografik baza yaratilmagan;

annotatsiya sifati va qamrovi yetarli emas;

NLP tizimlari uchun tayyor, eksport qilinadigan formatda resurs mavjud

emas.

2. Elektron bazaga aylantirishdagi asosiy muammolar

Fe’l paradigmalarini elektron bazaga aylantirish jarayonida bir qator jiddiy muammolar yuzaga keladi:

Birinchidan, paradigmalarning miqdoriy ko‘pligi. Birgina fe’ldan yuzlab, ba’zi hollarda undan ham ko‘p shakl hosil bo‘lishi mumkin. Bu esa ma’lumotlarni to‘plash, saqlash va boshqarishni qiyinlashtiradi.

Ikkinchidan, morfofonologik o‘zgarishlar va qo‘shimchalar tartibi masalasi. Unli uyg‘unligi, undosh assimilatsiyasi, ba’zi qo‘shimchalarning shakl o‘zgarishi kabi holatlar avtomatik tahlilni murakkablashtiradi.

Uchinchidan, omonimiya va noaniqlik. Ayrim shakllar bir nechta ma’no yoki kategoriya ifodalashi mumkin, bu esa annotatsiya jarayonida qiyinchilik tug‘diradi.

To‘rtinchidan, yagona annotatsiya standartining yo‘qligi. Hozircha o‘zbek tili uchun fe’l paradigmalariga mos keladigan umumqabul qilingan teglash tizimi to‘liq ishlab chiqilmagan.

Beshinchidan, texnik jihatlar. Ma’lumotlarni qulay saqlash, qidiruv va eksport qilish imkoniyatini yaratish dastlabki bosqichda ham e’tibor talab qiladi. Biroq bu jarayonni murakkab dasturlashsiz, oddiy va amaliy vositalar yordamida boshlash mumkin.

3. Yechimlar va taklif etilayotgan yondashuv

Yuqoridagi muammolarni bosqichma-bosqich hal etish mumkin. Dastlabki bosqichda quyidagi yondashuv tavsiya etiladi:

Eng faol va yuqori chastotali fe’llarni tanlab, ularning to‘liq paradigmalarini tuzish. Ma’lumotlarni dastlab Excel yoki CSV formatida tizimlashtirish. Bu usul dasturlash bilimi past bo‘lgan tadqiqotchilar uchun ham qulay. Annotatsiya qoidalarini ishlab chiqish: har bir shakl uchun ildiz, qo‘shimchalar ketma-ketligi, grammatik kategoriyalar va lemma ko‘rsatiladi. Paradigmalarni ikki asosiy qismga bo‘lish: fleksion (shakl o‘zgartiruvchi) paradigmalar; derivatsion (yasovchi qo‘shimchalar) tizimi. Keyingi bosqichlarda bazani kengaytirib, annotatsiyalangan

korpus va oddiy qidiruv interfeysiga o‘tish. Bunday yondashuv lingvistik chuqurlikni saqlagan holda amaliy resurs yaratish imkonini beradi. Natijada hosil bo‘ladigan baza keyinchalik morfologik analizatorlar, generatorlar va mashina o‘rganish modellari uchun asos bo‘lib xizmat qilishi mumkin.

O‘zbek tilida fe’l so‘z turkumining morfologik paradigmalarini elektron bazaga aylantirish hozirgi tilshunoslik va kompyuter lingvistikasi uchun dolzarb vazifadir. Mavjud resurslarning yetishmovchiliklari ushbu ishning zarurligini tasdiqlaydi. Maqolada ko‘rsatilgan muammolar va yechimlar keyingi bosqichda to‘liq leksikografik baza hamda annotatsiyalangan korpus yaratish uchun mustahkam asos bo‘lib xizmat qiladi.

Foydalanilgan adabiyotlar ro‘yxati:

1. Abduraxmonova N. va boshq. O‘zbek tili kompyuter lingvistikasi masalalari.
2. Hojiyev A. O‘zbek tili morfologiyasi, morfemikasi va so‘z yasalishi. – Toshkent, 2010.
3. Rahmatullayev Sh. Hozirgi o‘zbek adabiy tili. Morfologiya. – Toshkent, 2006.
4. O‘zbek tili milliy korpusi materiallari (uzbekcorpus.uz).