

## FE'L ANALIZI VA SINTEZI DASTURI: KOMPYUTER LINGVISTIKASIDA ILMIY-AMALIY YONDASHUV

Ibragimova Gulchexra Madaminjonovna.

Andijon davlat chet tillari inistituti

f.f.f. (PhD), dotsenti

Nabiqulov Haqqulbek Poziljon o'g'li

Andijon davlat chet tillari inistituti magistranti

**Annotatsiya.** Ushbu maqolada kompyuter lingvistikasi doirasida o'zbek tilidagi fe'llarning morfologik tahlili va sintezini amalga oshiruvchi dasturiy ta'minotning loyihaviy asoslari, algoritmik strukturalari va amaliy samaradorligi yoritiladi. Fe'l morfologiyasining murakkabligi, uning zamon, shaxs, mayl, nisbat va boshqa grammatik kategoriyalar bilan bog'liqligi ko'rib chiqiladi. Tadqiqot davomida mavjud nazariy manbalar tahlil qilinib, qoidaviy yondashuv asosida modul ishlab chiqildi. Natijalar test korpusida sinovdan o'tkazilib, samaradorlik darajasi baholandi.

**Kalit so'zlar:** fe'l, morfologik tahlil, sintez, kompyuter lingvistika, qoidaviy yondashuv, o'zbek tili.

**Kirish.** Kompyuter lingvistikasi - bu tabiiy tillarni avtomatlashtirilgan tarzda qayta ishlashga qaratilgan fanlararo yo'nalish bo'lib, u lingvistika, informatika, sun'iy intellekt, va statistik modellashtirish kabi sohalarining kesishgan nuqtasida joylashgan. Mazkur yo'nalishning markaziy tadqiqot ob'ektlaridan biri morfologik tahlil hisoblanadi. Morfologik tahlil orqali so'zlarning ichki tuzilishi, ularning affiksial tizimi, leksik va grammatik kategoriyalariga ajratilishi amalga oshiriladi. Ayniqsa, fe'l kabi harakat va holatni ifodalovchi markaziy leksik-grammatik birlikni avtomatik tarzda tahlil qilish kompyuter lingvistikasida muhim o'rinni egallaydi.

O'zbek tilida fe'l shakllarining morfologik tizimi juda murakkab va ko'p variantli bo'lib, vaqt, shaxs, son, mayl, nisbat, inkor, taxmin, kechish kabi grammatik kategoriyalarni o'z ichiga oladi. Shuning uchun ham o'zbek tili uchun morfologik analizator va sintezatorlar yaratish NLP (Natural Language Processing - tabiiy tilni

qayta ishlash) sohasida dolzarb vazifa hisoblanadi. Ushbu murakkablik, bir tomondan, lingvistik bilimlarga asoslangan modellarni talab qilsa, ikkinchi tomondan, statistik va neyron tarmoqlarga asoslangan zamonaviy yondashuvlar uchun sinov maydoniga aylanadi.

Fe'l morfologiyasini kompyuter vositalarida modellashtirishga oid bir qancha ilmiy tadqiqotlar mavjud. Jumladan, S. Jorayev o'zining tadqiqotlarida o'zbek tilidagi fe'llarning morfologik tahlilini formal grammatik modellar yordamida ifodalashga harakat qilgan. Uning ishlari, ayniqsa, affiksial tizimni algoritmik shaklda modellashtirish va shakl hosil qilish jarayonini avtomatlashtirish borasida muhim amaliy ahamiyatga ega<sup>1</sup>.

Xalqaro miqyosda esa Jurafsky & Martin tomonidan taklif etilgan morfologik tahlil va sintez tizimlari hozirgi NLP texnologiyalarining poydevorini tashkil etadi. Ular tomonidan ishlab chiqilgan Finite State Transducer (FST – Nihoyalangan Holatlar O'zgartirgichi) asosidagi morfologik analizatorlar turli tillarda muvaffaqiyatli sinovdan o'tgan va ular morfemalarning ketma-ketligiga asoslanib, so'zlarning grammatik tuzilishini aniqlay oladi<sup>2</sup>.

Ushbu maqolada aynan o'zbek tilidagi fe'llarning morfologik shakllarini avtomatik tarzda aniqlovchi va hosil qiluvchi (analiz va sintez) dasturiy vosita yaratish va uning ilmiy-lingvistik asoslarini yoritish maqsad qilingan. Maqola doirasida mavjud nazariy yondashuvlar, ularning texnik ko'rinishdagi modellari, shuningdek, dasturiy vositaning ishlash prinsipi va amaliy qiymati tahlil qilinadi.

**Metodologiya.** Maqolada qoidaviy yondashuv (rule-based approach) asosida dasturiy modul ishlab chiqildi. Quyida bu yondashuvning komponentlari ko'rib chiqiladi:

1. *Fe'l analiz algoritmi.* Fe'llar korpusi yaratilgan bo'lib, u 5000+ fe'l namunalarini o'z ichiga oladi. Har bir fe'l quyidagi parametrlarga ajratiladi:

<sup>1</sup> Jo'rayev S. *O'zbek tilining morfologik tuzilishi*. Toshkent: Fan, 2020, 134 b. 43-bet.

<sup>2</sup> Jurafsky D., Martin J.H. *Speech and Language Processing* (3rd ed.). Stanford University, 2021, 75-78-bet.

Asos (leksik ildiz)

Zamon affikslari (di, moqda, edi, etc.)

Shaxs-son affikslari (man, san, lar, etc.)

Mayl affikslari (sin, ing, etc.)

Analiz algoritmi suffix tree va lookup table texnikasiga tayangan. Hamma affikslar tartib bilan qoidaga ko'ra ajratiladi.

2. *Fe'l sintezi algoritmi*. Foydalanuvchi tomonidan belgilangan grammatik parametrlar asosida quyidagi formula asosida yangi fe'l shakli generatsiya qilinadi:

Asos + Zamon + Mayl + Shaxs

Masalan:

Asos: "bor"

Zamon: hozirgi zamon

Shaxs: 2-shaxs, ko'plik

Natija: borasizlar

3. *Dasturiy muhit*

Tillar: Python 3.11

Kutubxonalar: NLTK, Pandas, Tkinter (GUI uchun)

Platforma: Desktop (Windows/Linux)

**Natijalar.** Quyidagi jadvalda asosiy funksiyalar bo'yicha tafsilotlar keltirilgan<sup>3</sup>:

| Funktsiya    | Aniqlik<br>(%) | Test<br>namunalari<br>soni | Xatolik<br>sabablari                          | Izohlar                                                     |
|--------------|----------------|----------------------------|-----------------------------------------------|-------------------------------------------------------------|
| Fe'l tahlili | 93.4%          | 4000                       | Noaniq<br>affikslar,<br>noturg'un<br>shakllar | Asosan standart<br>fe'llarda yuqori<br>aniqlikka erishildi. |

<sup>3</sup> Matyakubov M. *Kompyuter lingvistikasi asoslari*. Toshkent: TDPU nashriyoti. 2018, 56-60-bet.

| Funktsiya                      | Aniqlik (%) | Test namunalari soni | Xatolik sabablari                              | Izohlar                                                             |
|--------------------------------|-------------|----------------------|------------------------------------------------|---------------------------------------------------------------------|
| Fe'l sintezi                   | 95.2%       | 2500                 | Affikslar tartibi xatolari                     | Qoidaviy modelga asoslangan shakllarda yuqori aniqlik.              |
| Murakkab shakl ajratish        | 88.1%       | 1500                 | Affikslar zanjirining chigal shaklda birikishi | "Olmoqda edi", "borayotgan edilar" kabi shakllarda xatoliklar.      |
| Dialektal variantlar           | 76.5%       | 1000                 | Rasmiy leksikaga mos kelmasligi                | "Borayapman", "kelayotirmidi" kabi og'zaki variantlarda xatoliklar. |
| Sintaktik bog'liqlikda aniqlik | 81.9%       | 1200                 | Kontekst yetishmasligi                         | Gapda fe'lining subyektga bog'liqligi har doim to'g'ri aniqlanmadi. |

#### Qo'shimcha Kuzatishlar:

*Aniqlik eng yuqori:* "Kelganman", "borasiz", "olmayapti" kabi aniq qoidaviy shakllarda.

*Aniqlik eng past:* Dialektal nutqdagi "qilarmidi", "bormoqchi edim", "kelolarmidi" kabi shakllarda.

*Xatoliklar soni eng ko'p:* Kontekstda semantik yondashuv talab etiladigan holatlarda (masalan, "bor" - fe'l yoki ot<sup>4</sup>).

<sup>4</sup> Islomov X. "Tabiiy tilni qayta ishlash va zamonaviy texnologiyalar." *Filologiya va AT*, 1(1), 21-29. 2022, 23-25-bet.

*Yechim takliflari:*

Dialektal variantlar uchun sinonimik bazani kengaytirish.

Affikslarni statistik modellar asosida dinamik tartiblash.

Sintaktik bog‘liqlikda BERT asosidagi modelni integratsiyalash.

Olingan natijalar o‘zbek tilining agglutinatив tuzilmasi qoidaviy model asosida samarali ishlashini ko‘rsatadi. Qoidaviy yondashuv yordamida ko‘pchilik standart va sodda fe‘l shakllarini aniqlash va yaratish muvaffaqiyatli amalga oshirildi. Biroq, istisno holatlar, noan‘anaviy shakllar va kontekstga bog‘liq fe‘l birliklari uchun yanada murakkab yondashuvlar talab qilinadi. Ayniqsa, gap ichidagi semantik kontekstga bog‘liq holda aniqlikni oshirish zarur.

Goral ta’kidlaganidek, morfologik tahlilda neyron modellar va statistika asosidagi yondashuvlar - ayniqsa mashinali o‘rganish yondashuvi - yuqori darajadagi aniqlik va o‘zgaruvchanlikni ta’minlaydi<sup>5</sup>. Bu fikrlar asosida, kelgusida fe‘l analiz va sintezi modulini BERT yoki GPT kabi yirik til modellariga integratsiyalash g‘oyasi ilgari surildi.

Bundan tashqari, semantik tahlilni qo‘shish orqali ayrim leksik homonimlarning ajratilishi (masalan, “bor” — fe‘l yoki ot yanada aniqroq bo‘lishi mumkin. Nutq tanish tizimlari, avtomatik tarjima va matn tuzuvchilari uchun bu modulni asosiy komponent sifatida ishlatish mumkinligi e’tiborga loyiq.

Xalqaro tajriba shuni ko‘rsatadiki, turkiy tillar (turk, qozoq, qirg‘iz) uchun ishlab chiqilgan morfologik analizatorlar (masalan, TRmorph yoki Zemberek) o‘xshash strukturalarga ega bo‘lib, ushbu modelni o‘zbek tiliga moslashtirish imkonini beradi. Bu modul ularning texnologik tajribasiga tayangan holda mahalliyashtirilgan yechim hisoblanadi.

Boshqa bir jihatdan, model o‘zbek tilining ta’lim tizimida, ayniqsa chet elliklar uchun mo‘ljallangan avtomatik mashg‘ulot platformalarida grammatika mashqlarini

---

<sup>5</sup> Goral, P. "Morphological analysis of agglutinative languages: A hybrid model." *Language Modelling Journal*, 8(2), 112-125. (2019). 119-bet

yaratish, o'quvchilar xatolarini avtomatik tahlil qilish va fe'l shakllarini real vaqtda ishlab chiqarishda keng qo'llanilishi mumkin. Munosabatan, ushbu loyiha kompyuter lingvistikasida maxsus fe'l tahlil modullarini yaratishda zamonaviy va an'anaviy yondashuvlar sintezini o'zida mujassam etadi. Bu esa o'zbek tilini raqamli muhitda keng ko'lamda qayta ishlash imkonini beradi.

**Xulosa.** Tadqiqot natijalari o'zbek tilidagi fe'l shakllarining morfologik tahlil va sintezi uchun qoidaviy yondashuv asosida yaratilgan dasturiy modul yuqori samaradorlikka ega ekanligini ko'rsatdi. Fe'l tahlil va sintezining algoritmik modeli foydalanuvchilarga grammatik shakllarni avtomatik tahlil qilish va hosil qilish imkonini beradi.

Dasturiy vositaning til o'rgatish, matn tahlili, tarjima va nutq texnologiyalarida qo'llanilishi kelajakda bu sohada yangi yutuqlar sari yo'l ochadi. Fe'l strukturasi bo'yicha aniqlik darajasi 90% dan yuqori bo'lib, bu NLP sohasidagi boshqa turkiy tillarga oid dasturlar bilan raqobatlasha oladi.

Kelgusida sintez algoritmiga semantik va sintaktik tahlil komponentlarini qo'shish, neyron tarmoqlar bilan integratsiya qilish va istisno holatlarni o'rganuvchi mashinali o'rganish modullarini joriy etish ko'zda tutilgan. Bu esa tizimning ko'p tilli platformalar uchun universal yechim bo'lishiga xizmat qiladi. Shuningdek, fonetik o'zgaruvchanlikni ham hisobga oluvchi modellar (masalan, talaffuzga bog'liq qisqartmalar) qo'shilishi taklif etiladi.

Ushbu loyiha, nafaqat lingvistik jihatdan, balki axborot texnologiyalari sohasida ham katta ilmiy va amaliy ahamiyat kasb etadi. Uning zamonaviy AI vositalari bilan integratsiyalashgan shakli o'zbek tilining raqamli manbalarda to'liq aks ettirilishiga xizmat qiladi.

### FOYDALANILGAN ADABIYOTLAR

1. Jo'rayev S. *O'zbek tilining morfologik tuzilishi*. Toshkent: Fan. 2020.
2. Jurafsky D., Martin J.H. *Speech and Language Processing* (3rd ed.). Stanford University, 2021.

3. Goral, P. "Morphological analysis of agglutinative languages: A hybrid model." *Language Modelling Journal*, 2019.
4. Matyakubov M. *Kompyuter lingvistikasi asoslari*. Toshkent: TDPU nashriyoti, 2018.
5. Islomov X. "Tabiiy tilni qayta ishlash va zamonaviy texnologiyalar." *Filologiya va AT*, 2022.
6. Ibragimova. G. Tilshunoslikda paranteza xodisasining tadqiqi xususida. Редакционная коллегия, 260, 2024.
7. Ibragimova, N. ISSUES OF TEACHING METHODS OF UZBEK LITERATURE ON THE EXAMPLE OF THE LIFE AND WORK OF ABDULLA ORIPOV. *International Journal of Artificial Intelligence*, 1(2), 1824-1828, 2025.
8. Ibragimova, G. M. Study of the Parentheses by World Linguistics. *Excellencia: International Multi-Disciplinary Journal of Education* (2994-9521), 2(3), 5-7. <https://doi.org/10.5281/> 2024.