

O‘ZBEK TILIDA AVTOMATIK FONETIK TRANSKRIPSIYA TIZIMINI YARATISHDAGI MUAMMOLAR

Maxmudjonova Gulshaxnoz Ulug‘bek qizi,
Tayanch doktorant (PhD)
gulshaxnozmahmudjonova@gmail.com
ToshDO‘TAU

Annotatsiya: Maqolada o‘zbek tilida avtomatik fonetik transkripsiya tizimini yaratishdagi asosiy lingvistik va amaliy muammolar tahlil qilinadi. Adabiy og‘zaki nutqni sheva va kundalik so‘zlashuv nutqidan farqlash, yozuv va talaffuz o‘rtasidagi tafovutlar, grafema-fonema nomuvofiqligi, urg‘u, o‘zlashma so‘zlar hamda dialektal variantlarning transkripsiya tizimiga ta’siri yoritiladi. Shuningdek, avtomatik transkripsiya oddiy harfma-harf o‘g‘irish emas, balki orfoepik me’yor, fonologik qoidalar va kontekstga tayangan holda qurilishi zarurligi asoslanadi.

Kalit so‘zlar: *avtomatik fonetik transkripsiya, o‘zbek tili, orfoepik me’yor, grafema-fonema munosabati, urg‘u, o‘zlashma so‘zlar, sheva, fonologik modellashtirish.*

Abstract: This article analyzes the main linguistic and practical challenges in developing an automatic phonetic transcription system for the Uzbek language. It examines the distinction between standard spoken Uzbek and dialectal and everyday colloquial speech, as well as discrepancies between orthography and pronunciation, grapheme-phoneme inconsistencies, stress patterns, borrowed words, and the influence of dialectal variants on the transcription system. Furthermore, the study argues that automatic transcription should not be based on simple letter-by-letter conversion, but must be constructed on the basis of orthoepic norms, phonological rules, and contextual factors.

Key words: *automatic phonetic transcription, Uzbek language, orthoepic norm, grapheme–phoneme correspondence, stress, borrowed words, dialect, phonological modeling.*



Fonetik transkripsiya nutq tovushlarini ilmiy asosda ifodalashning eng muhim vositasi bo'lib, u oddiy yozuv tizimining cheklovlarini bartaraf etadi. Xalqaro fonetik alfavit (IPA) yordamida nutqdagi nozik artikulyatsion va akustik farqlar aniq qayd etilishi mumkin. Shu bilan birga, transkripsiya nutqning to'liq nusxasi emas, balki uning abstrakt modeli hisoblanadi, chunki real nutq dinamik va variativ xususiyatga ega. Fonetik transkripsiya, nafaqat, lingvistik tadqiqotlar, balki nutq texnologiyalari, xususan, G2P, TTS va ASR tizimlari uchun ham fundamental ahamiyatga ega [1].

Chomskiy va Halle nazariyasida fonetik transkripsiya nutq signalining oddiy grafikli nusxasi emas, balki grammatikaning fonologik komponenti tomonidan hosil qilinadigan yakuniy representatsiya sifatida talqin qilinadi. Transkripsiya fizik signalning barcha tomonlarini emas, balki til tizimi uchun muhim bo'lgan belgilarni ifodalaydi. Leksik birliklarni bevosita fonetik transkripsiya ko'rinishida berish yetarli emas, chunki bitta birlik turli kontekstlarda turli fonetik shakllarda namoyon bo'ladi. Urg'u joylashuvini misol sifatida keltirish mumkin. Shuning uchun fonetik transkripsiya abstrakt leksik representatsiya va uni yuzaga chiqaruvchi fonologik qoidalar tizimi bilan birgalikda tushunilishi lozim. Demak, ilmiy nuqtayi nazardan fonetik transkripsiya – bu “eshitilganini yozib qo'yish” emas, balki nutqning fonologik tahlil asosida strukturaviy va tizimli ifodalanishidir. Avtomatik transkripsiya tizimi ham aynan shu sababli harfma-harf ko'chirishga emas, balki fonologik qoidalar, kontekst va me'yoriy talaffuzga tayangan holda qurilishi kerak [2; 3]. Fonetik transkripsiyani oddiy yozib olish amaliy emas, balki metodologik va empirik jihatdan asoslanishi lozim bo'lgan ilmiy jarayondir.

O'zbek tilida avtomatik fonetik transkripsiya tizimini yaratishda eng muhim metodologik talab adabiy og'zaki nutqni sheva va kundalik so'zlashuv nutqidan ajratishdir. Sababi har bir so'zning talaffuzi asosida transkripsiyani belgilashda orfoepik standartga tayaniladi. *Adabiy og'zaki nutq* umumxalq uchun tushunarli



bo'lgan, orfoepik me'yorlarga tayanuvchi va ijtimoiy jihatdan qabul qilingan talaffuz shaklidir. *Sheva* esa muayyan hududga xos fonetik, leksik va grammatik xususiyatlari bilan ajralib turuvchi hamda adabiy norma bilan har doim ham mos kelmaydigan shakl. *Kundalik so'zlashuv nutqi* bo'lsa spontanlik, tezkorlik va iqtisodiylikka moyilligi sababli tovush tushishi, so'z qisqarishi, assimilyatsiya kabi tovush hodisalar bilan xarakterlanadi, biroq bu belgilar doimo adabiy talaffuz me'yorlari sifatida qabul qilinmaydi. Masalan, *olib keldi* so'zi og'zaki adabiy nutqda [*olipkeldi*], kundalik so'zlashuvda [*obkeldi*], shevada esa [*apkelli*] tarzida talaffuz qilinishi mumkin. So'zning shakli tamoman boshqa bo'lishi ham mumkin: narvon (adabiy talaffuzda), shoti (shevada). Adabiy og'zaki nutq, sheva va kundalik so'zlashuv nutqi bir-biridan me'yorliq darajasi, qo'llanish doirasi va fonetik barqarorligi orqali ajratiladi.

Adabiy talaffuz normasi dinamik xarakterga ega, chunki u til bilan birga rivojlanadi, jamiyat nutqiga moslashib o'zgarib boradi, ayrim til birliklari esa vaqt o'tishi bilan eskirib, iste'moldan chiqadi. Talaffuz standartini belgilashda faqat qoidalarga tayanish yetarli emas, balki jonli nutqiy muloqot ham hisobga olinishi lozim. Standartlashtirish barcha mavjud talaffuz variantlari ichidan bitta standart variantni tanlash va uni norma darajasiga ko'tarish jarayonidir. Bu jarayonda fonologik izchillik, ijtimoiy qabul qilinganlik va funksional barqarorlik mezonlari asos bo'lib xizmat qiladi[4]. Tizimning maqsadi “har qanday eshitilgan variantni yozish” emas, balki adabiy talaffuzni me'yorlar asosida modellashtirish, zarur holatlarda esa uning yonida so'zlashuv yoki dialektal variantlarni qo'shimcha qatlam sifatida ko'rsatishdan iborat bo'lishi lozim.

Orfoepik manbalarda ham o'zbek tilining ko'p shevali ekani, shu sababli adabiy talaffuzni alohida me'yor sifatida belgilash zarurligi, ayniqsa rus tili orqali kirgan o'zlashma so'zlarning talaffuzini me'yorlashtirish lozimligi ta'kidlanadi. Amaliy kuzatuvlarda yozuv bilan talaffuz orasidagi tafovutlar bir necha ko'rinishda



namoyon bo‘ladi. Ba’zi tovushlar yozuvda saqlansa-da, talaffuzda juda sustlashadi yoki qo‘shni tovush ta’sirida assimilyatsiyaga uchraydi: *adabiyotchi* shaklining [adabiyochchi], *tuhmatchi* shaklining [tuxmachchi] tarzida aytilishi shunday hodisalardandir. Boshqa hollarda tovush tushishi kuzatiladi: *murojaat* → [murojat], *komissiya* → [komisiya], *abstrakt* → [abstrak]. Yana bir turdagi nomuvofiqlikda yozuvdagi fonema talaffuzda boshqa shaklga o‘tadi: *kitob* → [kitop], *ahd* → [aht][5]. Bu misollar o‘zbek tili uchun transkripsiya tizimi oddiy “harfma-harf o‘g‘irish” emas, balki kontekstga sezgir fonologik modellashtirish bo‘lishi kerakligini ko‘rsatadi.

Avtomatik transkripsiya tizimi, asosan, sistematik transkripsiya tayanishi lozim, chunki u tilning fonologik tizimi va adabiy talaffuz normasini aks ettiradi. Avtomatik transkripsiyani ishlab chiqishdan maqsadlardan biri elektron lug‘atlarga so‘zlarning to‘g‘ri talaffuzi uchun transkripsiyani qo‘shishdir. Longman [6] va Oxford [7] lug‘atlarida har bir so‘zning talaffuz varianti audio shaklda, shuningdek, to‘g‘ri talaffuzni ko‘rsatish uchun transkripsiya shaklida beriladi. Bu hol ingliz tilini o‘rganuvchilarga so‘zlarni to‘g‘ri o‘zlashtirishga va talaffuzda xatoga yo‘l qo‘ymaslikka yordam beradi. O‘zbek tili lug‘atlarida esa aksariyat hollarda na transkripsiya, na audio shakldagi talaffuz namunasi mavjud.

O‘zbek tiliga xos asosiy muammolar

Birinchi katta muammo – grafema va fonema o‘rtasidagi notekis munosabat. O‘zbek yozuvida ayrim digraflar bitta fonemani bildiradi: *ch, sh, ng* kabi birliklar shunga misol. Shuningdek, *g‘* va *o‘* kabi birliklarda apostrof belgisi oddiy ajratuvchi emas, grafemaning tarkibiy qismi sifatida ishlaydi. Bu belgi elektron manbalarda turli shakllari foydalaniladi Bu avtomatik segmentatsiya va tokenizatsiya jarayonini murakkablashtiradi.

Ikkinchi muammo – urg‘u masalasi. O‘zbek tilida ko‘pincha urg‘u so‘nggi bo‘g‘inga tushadi, degan umumiy tasavvur mavjud bo‘lsa-da, zamonaviy fonetik



tavsiflar kishi ismlari, o'zlashma so'zlar, so'z turkumlarida turli bo'g'inlarga tushadi[6]. Demak, urg'uni umuman e'tibordan chetda qoldirish ham, uni mutlaq so'nggi bo'g'inga bog'lash ham xato bo'ladi.

Uchinchi muammo – o'zlashma so'zlar va ikkilangan talaffuz. Ruscha-internatsional qatlam, arabcha-forscha tarixiy qatlam va hozirgi inglizcha o'zlashmalar bir xil prinsiplarda ishlamaydi. Amaliy transkripsiya va transliteratsiya bo'yicha zamonaviy tadqiqotlar rasmiy hujjatlarda ham o'zbekcha ism va birliklar ko'pincha rus tili vositasida o'zgartirilgan shaklda berilganini, bu esa yagona me'yorlashuvga to'sqinlik qilayotganini ko'rsatadi.

To'rtinchi muammo – sheva, yosh va bilingual muhit ta'siri. O'zbek adabiy talaffuzida keksa va yosh avlod nutqi orasida, shuningdek, shahar va hududiy shevalar orasida sezilarli tafovutlar uchraydi. Ayniqsa, ba'zi unlilar va o'zlashma birliklarning talaffuzi hududga qarab farqlanadi. Avtomatik tizim uchun bu shuni anglatadiki, dastlabki bosqichda dialektal xilma-xillikni emas, standart qatlamni markazga olish, sheva variantlarini esa keyingi kengaytma sifatida ko'rish maqsadga muvofiq.

Xulosa o'rnida, o'zbek tilida avtomatik fonetik transkripsiya tizimini yaratish uchun, avvalo, “nima transkripsiya qilinmoqda?” degan savolga aniq javob berish kerak. Adabiy talaffuzmi, so'zlashuv variantimi yoki dialektal shaklmi? Ilmiy jihatdan eng to'g'ri birinchi qadam, adabiy og'zaki normani markazga olgan, orfoepik lug'atga tayanuvchi, qoida hamda ma'lumotga asoslangan usullarni birlashtiruvchi gibridd tizimni qurishdir. Shunda transkripsiya tizimi nafaqat TTS va ASR uchun, balki o'zbek tilining raqamlashuvi, elektron lug'atlar va fonetik korpuslar uchun ham mustahkam poydevor vazifasini bajaradi.

Foydalanilgan adabiyotlar ro'yxati

1. Ladefoged Peter, and Keith Johnson. A Course in Phonetics. 6th ed. Boston, 2011.



2. Chomsky N., Halle M. The Sound Pattern of English. New York, 1968
3. Cucchiarini C. Phonetic Transcription: A Methodological and Empirical Study. PhD thesis, Radboud University Nijmegen, The Netherlands, 1993.
4. Begmatov E. and Mamatov A. Adabiy norma nazariyasi. 1-qism. Toshkent, 1997.
5. O'zbek adabiy talaffuzi lug'ati. – Toshkent: Fan nashriyoti, 1984.
6. <https://www.ldoceonline.com>
7. <https://www.oxfordlearnersdictionaries.com>