

SHEVALAR VA MEDIA: O‘ZBEK TILINING RAQAMLI RIVOJIDA XORAZM SHEVASINING O‘RNI

Musayeva Mahina Maqsudbek qizi,
magistrant
musayevamahina41@gmail.com
Urganch davlat universiteti

Annotatsiya. So‘nggi yillarda o‘zbek tilining raqamli shakllanishi va sun’iy intellekt tizimlarida qo‘llanilishi tez sur‘atlarda rivojlanmoqda. Ammo bu jarayonda hududiy shevalarning, xususan, Xorazm shevasining o‘rni va ifodalanish darajasi hali ham yetarli emas. Ushbu maqolada O‘zbek tilining shevalari, ayniqsa Xorazm shevasi, zamonaviy media platformalarida (YouTube, TikTok, Telegram va h.k.) qanday aks etayotgani tahlil qilinadi. Shuningdek, Xorazm shevasida yaratilgan kontentlarning hajmi, sifati va ularni til korpuslari yoki AI modellar uchun potensial ma’lumot manbai sifatida foydalanish imkoniyatlari ko‘rib chiqiladi. Shevalarning raqamli kontent orqali saqlanishi va rivojlanishi til siyosati va madaniy meros nuqtai nazaridan ham dolzarb ekani asoslab beriladi.

Abstract. In recent years, the Uzbek language has been rapidly developing in the digital sphere and increasingly integrated into artificial intelligence systems. However, regional dialects-especially the Khorezm dialect-remain underrepresented in this process. This paper analyzes how Uzbek dialects, particularly the Khorezm dialect, are reflected on modern media platforms such as YouTube, TikTok, and Telegram. It explores the volume, quality, and characteristics of content in the Khorezm dialect and its potential as a data source for language corpora and AI models. The importance of preserving and promoting dialects through digital media is emphasized from the perspectives of language policy and cultural heritage.

Аннотация. В последние годы узбекский язык активно развивается в цифровом пространстве и внедряется в системы искусственного интеллекта. Однако региональные диалекты, в частности хоразмский говор, остаются недостаточно представленными в этих процессах. В данной статье анализируется присутствие узбекских диалектов, особенно хоразмского, на современных медиа-платформах (YouTube, TikTok, Telegram и др.). Рассматриваются объем, качество и особенности контента на хоразмском говоре, а также его потенциал как источника данных для лингвистических корпусов и моделей ИИ. Подчеркивается значение сохранения и развития диалектов через цифровой контент с точки зрения языковой политики и культурного наследия.

Kalit so‘zlar: *Xorazm shevasi, media korpus, o‘zbek tili, raqamli tilshunoslik, shevalar, sun’iy intellekt, til korpusi.*

Raqamli texnologiyalarning tezkor rivojlanishi tilshunoslikda yangi yo‘nalishlarning shakllanishiga sabab bo‘lmoqda. Xususan, sun‘iy intellekt (SI) tizimlarida tabiiy tilni qayta ishlash (Natural Language Processing – NLP) metodlari asosida til modellarini yaratish uchun katta hajmdagi matnli korpuslar talab etiladi. O‘zbek tilida bunday korpuslar shakllantirila boshlagan bo‘lsa-da, bu jarayonda asosan adabiy til asosiy manba sifatida ishlatilmoqda. Hududiy shevalar, ayniqsa Xorazm shevasi, bu jarayonda yetarlicha e‘tibordan chetda qolmoqda.

Shevalar har bir tilning ichki boyligini, tarixiy qatlamlarini va madaniy xilma-xilligini ifodalovchi muhim til resursidir. Agar ularni raqamli formatda saqlab qolish va tizimlashtirish amalga oshirilmasa, bu so‘z boyligi va madaniy qatlamlar vaqt o‘tishi bilan yo‘qolishi mumkin. Shu bois Xorazm shevasining media kontentlar orqali qay darajada aks etayotgani, bu kontentlarni korpus yaratishda qanday foydalanish mumkinligi muhim tadqiqot yo‘nalishidir.

Globalizatsiya va raqamli texnologiyalar rivojlanishi natijasida tillarning ijtimoiy hayotdagi, ayniqsa internet va media platformalaridagi roli keskin o‘zgarib bormoqda. Shu nuqtai nazardan qaralganda, o‘zbek tilining raqamli makondagi o‘rni, uning ichki ko‘rinishlari, ya‘ni lahja va shevalarining taqdiri ham dolzarb masalaga aylanmoqda. Chunki raqamli axborot almashinuvi faqat adabiy til orqali emas, balki jonli xalq og‘zaki nutqi - shevalar orqali ham amalga oshmoqda. O‘zbek tilining hududiy shevalari ichida Xorazm shevasi o‘zining boy lug‘aviy zaxirasi, qadimiylilik bilan uyg‘unlashgan zamonaviy ifoda uslublari va raqamli kommunikatsiyada faol qo‘llanilishi bilan ajralib turadi. Ushbu maqolada Xorazm shevasining raqamli kommunikatsiyadagi roli, media vositalaridagi ifodalanish shakllari, hamda umumtil taraqqiyotiga qo‘shayotgan hissassi tahlil qilinadi. Shu bilan birga, raqamli tilshunoslik doirasida ushbu shevani o‘rganish va undan zamonaviy til texnologiyalarida foydalanish imkoniyatlari ko‘rib chiqiladi.

Xorazm shevasi fonetik, morfologik va leksik jihatdan o‘ziga xos bo‘lib, o‘zbek tilining boshqa shevalaridan farqlanuvchi ko‘plab xususiyatlarga ega. Ushbu sheva nafaqat muloqot vositasi, balki Xorazm viloyatining madaniy identitetini ifodalovchi element hamdir. Biroq bugungi kungacha bu sheva asosida yaratilgan hech qanday katta korpus yoki lingvistik resurs mavjud emas.

Zamonaviy media platformalarida ayniqsa, YouTube, Instagram, TikTok va Telegram kabi platformalar, xorazm shevasining tabiiy kontekstdagi namunalarini saqlab qolish va tahlil qilish uchun yangi imkoniyatlar yaratmoqda. Ayniqsa, Xorazmda faol bloggerlar, kontent yaratuvchilar tomonidan yaratilgan vloglar, sahnalashtirilgan skitlar va suhbat janridagi videolar shevaning zamonaviy shaklini to‘g‘ridan-to‘g‘ri aks ettiradi. Ushbu materiallar og‘zaki tilning real va kontekstli misollari sifatida juda katta ahamiyatga ega.

Mazkur media kontentlar raqamli korpus yaratishda foydali manba bo'lishi mumkin. Biroq, bu jarayonda bir nechta muammo yuzaga chiqadi:

1. *Standartlashtirish muammosi* - Xorazm shevasida yozilgan yoki talaffuz qilingan iboralarni adabiy tilga yoki fonetik yozuvga keltirish qiyin.
2. *Kontent sifati va barqarorligi* - Media platformalarda joylashtirilgan materiallar vaqtinchalik bo'lishi, yo'qolib ketishi yoki o'chirilishi mumkin.
3. *Annotatsiya muammolari* - Dialektal so'zlar, ijtimoiy jargon va kontekstual iboralarni aniqlash va ularni standart shaklda belgilash uchun lingvistik bilim va metodik yondashuv talab etiladi.

Shunga qaramay, xorazm shevasida mavjud kontentlarni tizimli yig'ish, ularni transkripsiya qilish, matnga aylantirish va lingvistik belgilash orqali yirik media korpus yaratish mumkin. Bu korpus keyinchalik SI modellarini shevaga moslashtirish, avtomatik tarjima, ovozdan matnga konversiya, va boshqa NLP vazifalarda foydalanish imkonini beradi.

Xorazm shevasi o'zbek tilining qadimiy va mustahkam mintaqaviy shevalaridan biri bo'lib, u fonetik, leksik va grammatik jihatdan ancha farq qiluvchi jihatlarga ega. Bu shevada qadimgi o'zbek tili unsurlarining saqlanib qolganligi, shuningdek, turkiy tillarning umumiy xususiyatlarini aks ettiruvchi shakl va so'zlar uchrashi kuzatiladi. Masalan, “bormoqchiman” o'rniga “barjaqman”, “yaxshimisiz?” o'rniga “nichiksiz?”, “nima qilay” o'rniga “nishatin” kabi shakllar fonetik o'zgarishlarning yorqin namunalaridandir. Xorazm shevasining leksikasida ham ko'plab o'ziga xos atamalar mavjud: “sovuq” o'rniga “suvug’”, “uyqu” o'rniga “uxi”, “katta” o'rniga “ulli” va hokazo. Ushbu leksik birliklar nafaqat tilshunoslik, balki madaniyatshunoslik uchun ham katta ilmiy qiziqish uyg'otadi. Bugungi kunda Xorazm shevasi nafaqat og'zaki nutqda, balki ijtimoiy tarmoqlar, audio-video mahsulotlarda ham faol ishlatilmoqda.

Xalqaro tajriba shuni ko'rsatmoqdaki, ko'plab tillar uchun dialektal korpuslar yaratilib, ular lingvistik tadqiqotlarda ham, sun'iy intellekt modellarida ham muvaffaqiyatli qo'llanilmoqda. Masalan, arab tili, italyan tili, nemis tili dialektlari bo'yicha yaratilgan korpuslar tilshunoslikda muhim o'rin egallaydi. O'zbek tilida bu yo'nalishda hali dastlabki bosqichlar boshlanmoqda, va Xorazm shevasi shu jarayonga kiritilishi zarur.

Xulosa qilib aytganda, shevalar, xususan, Xorazm shevasi o'zbek tilining muhim tarkibiy qismi sifatida nafaqat madaniy merosni saqlashda, balki zamonaviy kommunikatsiya va raqamli til muhitida ham alohida o'rin egallamoqda. Raqamli platformalarning keng tarqalishi bu shevaning ommalashuvi va yoshlar orasida qayta qadrlanishiga sabab bo'lmoqda. Ayniqsa, media kontentlarda Xorazm shevasining faol qo'llanishi uning ijtimoiy qadriyat, estetik vosita va identitet belgisi

sifatida mustahkamlanishiga xizmat qilmoqda. Xorazm shevasi o'zbek tilining boyligidir. Uni raqamli asr talablariga mos tarzda saqlab qolish, o'rganish va rivojlantirish – nafaqat tilshunoslarning, balki jamiyatning har bir a'zosining ham burchidir.

Xorazm shevasining media kontentlarda aks etishi va ularni raqamli til resurslariga aylantirish imkoniyati o'zbek tilining raqamli rivoji uchun muhim ahamiyat kasb etadi. Bu jarayon nafaqat tilni sun'iy intellektga moslashtirish, balki madaniy boylikni saqlash vazifasini ham bajaradi. Bugun boshlangan harakatlar, ertaga xorazm shevasining yo'qolish emas, balki rivojlanishiga xizmat qiladi. Shu sababli, xorazm shevasidagi media kontentlar tizimli o'rganilishi, saqlanishi va korpuslashgan shaklga keltirilishi zarur.

Foydalanilgan adabiyotlar:

1. M. Abdul-Mageed, D. Kuebler, and M. Diab, “SAMAR: A System for Arabic Dialect Annotation of Social Media Data,” *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 2, pp. 250–255, 2018.
2. E. Kuriyozov, D. Vilares, and C. Gómez-Rodríguez, ‘BERTbek: A pretrained language model for Uzbek’, in *Proceedings of the 3rd Annual Meeting of the Special Interest Group on Under-resourced Languages@ LREC-COLING 2024*, 2024, pp. 33–44.
3. P. Soulos et al., "Structural Biases for Improving Transformers on Translation into Morphologically Rich Languages," *arXiv preprint arXiv:2208.06061*, 2022.
4. R. G. Yuldashovna, S. Sukhrob, K. Barno, P. Mukhayyo, and D. Shokhida, ‘Linguistic Considerations on the Functional Semantics of English Reduplicative Elements’, *Remittances Review*, vol. 8, 2023.
5. A. Vaswani et al., "Attention is All You Need," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017, pp. 5998–6008.
6. B. Mansurov and A. Mansurov, "UzBERT: Pretraining a BERT Model for Uzbek," *arXiv preprint arXiv:2108.09814*, 2021.
7. B. Allaberdiev, G. Matlatipov, E. Kuriyozov, and Z. Rakhmonov, ‘Parallel texts dataset for Uzbek-Kazakh machine translation’, *Data in Brief*, vol. 53, p. 110194, 2024.