

ISSN 2181-922X

LANGUAGE & CULTURE

UZBEKISTAN O'ZBEKISTON

TIL VA MADANIYAT

KOMPYUTER
LINGVISTIKASI

2023 Vol. 2 (6)

www.compling.tsuull.uz

UZBEKISTAN

MUNDARIJA

Shahlo Hamroyeva, Noila Matyakubova

Mashina tarjimasida matnni moslashtirish usullari.....6

Zilola Xusainova

O'zbek tili milliy korpusi qidiruv tizimini optimallashtirishda
lemmatizatsiyadan foydalanish.....20

Shahlo Abdusalomova

O'zbek tilida pronominal anaforani hobbs yondashuvi
asosida hal etish modeli.....38

Botir Elov, Narzullo Alayev, Aziz Yuldashev

Svd va nmf metodlari orqali tematik modellashtirish.....55

Botir Elov, Madina Samatboyeva

Ner: o'zbek tilidagi matnlarda toponim(lar)ni
avtomatik aniqlash modellari.....67

Dilraxon Rustamova

Lingvistik atamalarining so'zligini shakllantirish hamda
terminlarni standartlashtirish asoslari.....85

NER: O'ZBEK TILIDAGI MATNLARDA TOPONIM(LAR)NI AVTOMATIK ANIQLASH MODELLARI

Botir Elov¹,
Madina Samatboyeva²

Annotatsiya

Ushbu maqolada toponimlarning tarkibi, ularning boshqa fanlar bilan aloqadorligi, grammatik tuzilishi tahlil qilindi. Shuningdek, toponimlarni etimologik va grammatik jihatdan tahlillari taqdim etildi. Sodda va qo'shma toponimlar to'g'risida fikr yuritilib, ularning toponimik modellari taqdim etildi. Tilimizda toponimlarning hosil bo'lishida ishtirok etadigan topoformant(qo'shimcha)lar tadqiq qilingan. Toponimlar ma'no jihatdan guruhlariga ajratilib, ular tarkibidagi topotermin (toponimga ishora qiluvchi ko'rsatkich so'zlar – indikator)lar ro'yxati taqdim etildi. Toponimlardagi shakldoshlik (omonimiya) hodisasi o'rganilib, joy nomlari mashinali o'qitish yordamida avtomatik aniqlanganda yuzaga keladigan muam molar keltirib o'tildi.

Kalit so'zlar: *NER, nom, nomlangan obyekt, joy nomi, toponomika, toponimiya, toponim, onomastika, lingvistika, etimologik tahlil, toponimik modeller, toponimik tip, topoformant, suffiks, affiks, topotermin, indikator.*

Kirish

Joy nomlari **geografik nomlar** yoki **toponimlar** deb ataladi. Toponimlarni **toponimika** fani o'rganadi. **Toponimika** yunoncha **topos – joy va onoma (yoki onima) - nom** so'zlaridan tarkib topgan.

Toponimlar til lug'at tarkibining bir qismi bo'lib, til qonuniyatlariga bo'ysunadi [Do'simov, Egamov, 1977: 5]. Toponimlar tilshunoslik (lingvistika)ning onomastika (yunoncha "**onomas-**

¹*Elov Botir Boltayevich* – texnika fanlari falsafa doktori (PhD), dotsent. Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti.

E-pochta: elov@navoiy-uni.uz

ORCID: 0000-0001-5032-6648

²*Samatboyeva Madina To'liqinjon qizi* –Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti Kompyuter lingvistikasi mutaxassisligi magistranti.

E-pochta: msamatboyeva@gmail.com

ORCID: 0000-0002-5146-4484

tike” – nomlash, nom qo‘yish san‘ati) [<https://qomus.info>] bo‘limida o‘rganiladi.

Toponimlar joy xususiyatini, uning tarixini, ushbu obyekt bilan bog‘liq voqea-hodisani anglatgan oddiy so‘z va nomning birikuvidan hosil bo‘lgan atoqli otdir. Toponimlar doirasida uchta tushunchani farqlash lozim: **“Toponimika”, “Toponimiya”, “Toponim”**.

Toponimika – joy nomlari, ularning tuzilishi, geografik nomlarning hosil bo‘lishi va rivojlanishi bilan shug‘ullanuvchi fan;

Toponimiya – joy nomlari yig‘indisi (masalan, Buxoro toponimiyasi – Buxoro joy nomlari yig‘indisi);

Toponim – bu bitta joyga tegishli *xususiy nom*.

Toponimlar tilshunoslik nuqtayi nazari bilan o‘rganilganda nomning paydo bo‘lishi, uning tarixi, yaratilishi, leksik-semantik tasnifi, lingvistik strukturasi, so‘z turkumi, uning atoqli otligi, qaysi tilga mansub leksika ekanligi, toponim bilan bog‘liq til qonuniyatlari o‘rganiladi. Ammo rivojlanib borayotgan asrimizda lingvistik qarashlar kompyuter tiliga moslashtirilib, zamonaviy elektron dasturlar yaratilayotgan bir davrda, joy nomlarini qo‘lda aniqlash va ular bilan bog‘liq statistik ilmiy tadqiqot ishlarini olib borish mushkul.

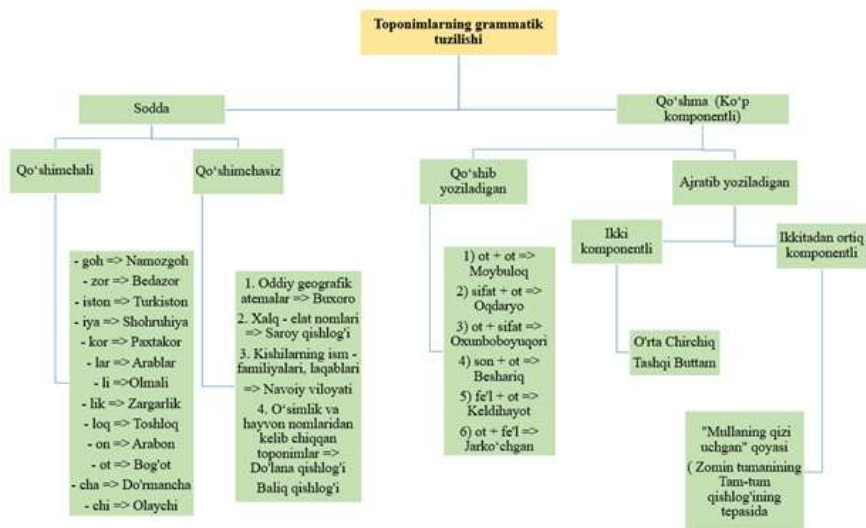
Ma‘lum bir hududdagi geografik nomlarning yig‘indisi shu hududda yashovchi xalqlarning asrlar davomida, nomlar yaratishdagi ijod mahsuli hisoblanadi. Dunyoda qancha geografik nom borligini hech kim aniq bilmaydi. Taxminiy hisoblarga ko‘ra, butun yer sharida yarim milliarddan ortiq geografik nom bor. Holbuki, ular qatoriga soy, jilg‘a, buloq, quduq, jar, qir, mahalla, guzar, ko‘cha kabi mikrotoponimlar kirmaydi. Agar yer yuzidagi barcha katta-kichik joy nomlarini hisobga olish imkoni bo‘lganda, son-sanoqsiz raqam hosil bo‘lar edi [Qilichev, 2023: 41].

Toponimlar tarkibi

Toponimlar geografiya, tarix, lingvistika sohalari doirasida o‘rganiladi. Toponimlarni lingvistik jihatdan o‘rganish ikkita usul yordamida amalga oshiriladi. Bular:

Etimologik tahlil - tarix fani bilan birgalikda amalga oshiriladi. Nomning paydo bo‘lishi tarixiy faktlarga asoslangan holda aniqlanadi.

Grammatik tuzulish jihatidan tahlil – grammatik tuzulish jihatidan tahlil jarayonida toponimlarning qaysi so‘z turkumida ekanligi, sodda, qo‘shma yoki birikmali ekanligi, so‘z tarkibidagi qo‘shimchalari tadqiq qilinadi.



1-rasm. Toponimlarning grammatik tuzilishi

Toponimlar *sodda* va *qo'shma* (ko'p komponentli) turga bo'linadi. Sodda tuzulishli toponimlar o'z tarkibida *qo'shimchali* va *qo'shimchasiz* turga ajratiladi.

Sodda toponimlar

Qo'shimchali toponimlar

Bir so'zdan iborat toponimlar, odatda, otlar va kamdan-kam hollardagina sifatlar shaklida boladi. Toponimlar muayyan qoliqlar shaklida hosil bo'ladi. Ana shunday qoliqlar – *toponimik modeller* deyiladi. Har bir til toponim o'z xususiyatlari va toponimik model-lariga ega. Masalan, o'zbek tilidagi toponimlar ko'p hollarda birdan ortiq unsurlardan iborat bo'ladi, qo'shimchalar o'zakdan keyin ke-ladi, sifat otdan oldin turadi. Shuningdek, toponimlarni o'rganishda toponimik modellardan tashqari *toponimik tip* ham muhim. Maxsus lingvistik adabiyotlarda toponimik tip deb ba'zan muayyan toponi-mik *yasalish modeli*, ayrim hollarda *formant* ham tushuniladi. Biroq toponimik tip deganda grammatik shakl bilan bir qatorda so'zning ma'nosini tushunish maqsadga muvofiq. Ana shunda har bir joy, har bir tarixiy davr va har bir til uchun o'zining toponimik tiplari aniq ko'rinadi [Qorayev, 2006: 33-34]

Toponimlarni yasashda shu til uchun xos bo'lgan lingvistik vositalar(modellar)dan foydalaniladi. Lekin shu bilan birgalikda har bir tilda *toponim yasovchi vositalar (suffiks, affiks)lar va so'zlar* yoki *toponimlar yaratishga ixtisoslashib qolgan unsurlar* mavjud bo'ladi.

Qo'shimchali toponimlar tarkibidagi qo'shimcha(suffiks, af-fiks)lar *topoformantlar* deb ataladi. *Topoformant* – (lotincha, for-mans - hosil qiluvchi) geografik nomlar hosil bo'lishida ishtirok

etadigan, lekin o'zi muayyan ma'noni anglatmaydigan qo'shimcha va suffikslar. Ushbu qo'shimchalar asosga qo'shilib yangi nom hosil qiladi. Masalan, *Chilonzor, Misrgarlik, Xizmatchi, Ohangaron, Biyxem, Satino* va boshqalardagi *-zor, -gar, -chi, -garon, -xem, -o* qo'shimchalari topoformantlardir [G'ulomov, Mirakmalov, 2005: 31].

Akademik A.Vostokov 1812-yilda yozgan "Etimologiya havaskorlariga mashq" nomli maqolasida geografik nomlarni old va oxirgi qo'shimchalarining o'xshashligiga qarab bir necha guruhga bo'lgan va bu hol tasodifiy emas, degan xulosaga kelgan. Shu sababdan mutaxassislar A.Vostokovni toponimik ilmiy tadqiqotlarda qo'llaniladigan topoformant usuli asoschisi deb e'tirof etishgan.

Nomlardagi elementlarning takrorlanishi ko'pincha ularning bir tipga xos ekanligidan dalolat beradi. Nomni (topo)formantlarga qarab tahlil etishning qulayligi shundan iboratki, *toponimlar tarkibida tez-tez takrorlanadigan qo'shimchalar ularning shakllanishi, etnik daxldorligi, tarqalish areali, ko'chma toponimlar haqida aniq ma'lumot beradi. Masalan, O'rta Osiyoda, xususan, O'zbekiston va Tojikistonda sermahsul formantlardan biri - iston qo'shimchasidir. Bu unsur mamlakat, o'lka nomini yasaydi yoki biron narsaning ko'p ekanligini bildiradi. Masalan, Afg'oniston - "Afg'onlar mamlakati", O'zbekiston - "O'zbeklar mamlakati" [Qorayev, 2006: 65].*

1-jadval. Topoformantlar

№	Qo'shimcha	Misollar
1	-goh	Janggoh, Namozgoh, Sayilgoh, Shamolgoh, Korizgoh
2	-zor	Bedazor, Bodomzor, Kavarzor, Olmazor, Regzor, Sangzor
3	-iston	Bolujiston, Bog'iston, Sebiston, Turkiston, Sho'riston
4	-iya	Mingiya, Yuziya, Shohruhiya, Qodiriya
5	-kor	Lalmikor, Paxtakor, Sabzikor
6	-lar	Arablar, Beklar, Boylar, Zagarlar, Shayxlar, Shag'allar
7	-li	Alamli, Anjirli, Burganli, Olmali, Piyozli, Tolli, Chayonli
8	-lik	Attorlik, Zargarlik
9	-liq	Xo'jaliq, Soyliq, Tepalik
10	-loq	Sangloq, Toshloq, Qumloq
11	-lox	Gurlox
12	-on	Arabon, Mang'iton, Kosagaron, Mington, Namatgaron, Sayidon, Xudoyon
13	-ot	Bog'ot, Rabotot, Qishloqot
14	-cha	Buloqcha, Deycha, Rabotcha, Qo'rg'oncha, Bahrincha, Do'rmancha, Naymancha, Oqcha, Saricha, Yangicha
15	-chi	Zominchi, Farg'onachi, Olaychi

16	-kent(kand)/ kat (kass)	Toshkent, Binkat
----	----------------------------	------------------

Bulardan O'zbekistonda **-cha**, **-ak** ba'zan **-ik** affikslari *diminutive (kichraytirilgan shakldan toponim yasalishi)* shakldagi toponimlar yasaydi. Bunday toponimlar otdan (Hisorak, Xumdonak, Ro'dak), sifatdan (Ko'kcha, Oqcha, Qoracha, Saricha), etnonimlardan (Do'rmoncha, Naymancha) yasalishi mumkin, ba'zi hollarda esa biron toponimga taqlid qilib nom qo'yilganda ham diminutiv shakldagi toponimlar yasaladi. Masalan, Buxorcha, Samarqandak, Bog'dodik, Registonak va hokazo.

Arealli nomlar – bir xil qo'shimcha yoki topoterminlar ishtirokida yasalgan geografik nomlar joylashgan hududdagi toponimlar: kent, tepa, qo'rg'on va boshqa terminlar ishtirokida yasalgan nomlar tarqalgan hudud. Masalan, Toshkent viloyatida “kent” termini yordamida yasalgan nomlar ko'p: Alimkent, G'azalkent, Toshkent, Parkent, Zarkent va boshqalar [G'ulomov, Mirakmalov, 2005: 42-46].

Qo'shimchasiz toponimlar

Qo'shimchasiz toponimlar tarkibida hech qanday qo'shimcha mavjud bo'lmaydi. Bunday toponimlar 4 guruhga bo'linadi:

Oddiy geografik atamalar – Buxoro (viloyat), Osiyo (Buxoro viloyati, Qorako'l tumanidagi qishloq), Xovos shaharchasi (Xovos tumanida), Baxt shahri (Sirdaryo viloyatida).

Xalq-elat nomlari – Saroy (Namangan viloyatining Chortoq tumanidagi qishloq nomi), Qiyot (Qo'qon shahridagi qishloq).

Kishilarning ism-familiyalari, laqablari – Mirzo Ulug'bek tumani, Navoiy viloyati.

O'simlik va hayvon nomlaridan kelib chiqqan toponimlar – Tol (Sirdaryo viloyati Xovos tumanidagi qishloq), Tut (Samarqand viloyati Jomboy tumanidagi qishloq).

Qo'shma (ko'p komponentli) toponimlar

Qo'shma toponimlar ikki mustaqil asosdan tarkib topgan toponimlardir. Qo'shma (ko'p komponentli) toponimlar ham ikkiga bo'linadi: *qo'shib yoziladigan qo'shma toponimlar va ajratib yoziladigan qo'shma toponimlar*.

Qo'shib yoziladigan qo'shma toponimlar

Qo'shib yoziladigan qo'shma toponimlar o'z tarkibida ikkita asosdan iborat bo'ladi. Qo'shma toponimlar tarkibi umumiy grammatik tahlil jarayonida *topotermin(toponimikada – indikator)lar* yordamida tadqiq etiladi. *Topotermin* – joy nomlarini hosil qiluvchi terminlar (atamalar). Masalan, *daryo, soy, tepa, qo'rg'on, kent, obod...* [G'ulomov, Mirakmalov, 2005: 50] qo'shma toponimlar tarkibidagi

asoslar boshqa tildan o'zlashgan yoki istorik (istarizm) xarakteriga ega bo'lishi mumkin. Masalan, "Obihayot" toponimi ikki asosdan ("ob" va "hayot") tarkib topgan. Bunda birinchi asos "ob" so'zi fors tilida "suv" degani ammo o'zbek tilida "suv" ma'nosida qo'llanilmaydi. Ikkinchi asos "hayot" so'zi esa tilimizda mavjud.

Yoki yana bir misol "Toshsaqa" toponimi tarkibidagi birinchi asos "tosh" so'zi tilimizda mavjud, ammo ikkinchi "saqa" (kanalning bosh inshooti) so'zi tarixiy xarakterga ega. Ammo ushbu yuqoridagi ikki so'z ham qo'shma tarkibli bo'lganligi uchun qo'shma toponimlar tarkibida tadqiq qilinadi.

Toponimlarning *oykonim (aholi yashash manzillari)*lar turi hosil bo'lishida ishtirok etadigan leksikalar:

1. Suvga oid leksikalar.
2. Yer yuza relyeflariga oid leksikalar.
3. O'simlik, daraxtlarga oid leksikalar.
4. Hayvonlarga oid leksikalar.
5. Kasb-hunarga oid leksikalar.

Quyidagi jadvalda ushbu leksikalarning ba'zilarini keltirib o'tamiz:

2-jadval. Oykoni mlarni hosil qiluvchi topotermin(indikator)lar

Toponimlar		
Tarkibida topoterminlar ishtirok etgan – Oykoni mlar (aholi yashash manzillari)		
№	Suvga oid leksikalar	Hosil bo'lgan joy nomlari
1	Anhor (nahr)	Anhor, Ulug'nor
2	Arna	Arna, Anhor, Ulug'nor,
3	Ariq	Ariqboshi, Beshariq, Jarariq
4	Buloq	Jarbuloq, Oybuloq, Oqbuloq
5	Daryo	Daryobo'yi, Daryoliq, Sirdaryo, Qoradaryo, To'polondaryo, Sangardakdaryo
№	Yer yuza relyeflariga oid leksikalar	Hosil bo'lgan joy nomlari
1	Adir	Bo'lakadir, Adirmoq
2	Bel	Beltov, Beltepa, Belariq
3	Dara	Omondara, Palandara
4	Dahana	Oqdahana, Qoradahana
5	Do'ng	Do'ng, Do'ngariq, Do'ngovul
№	O'simlik, daraxtlarga oid leksikalar	Hosil bo'lgan joy nomlari
1	Paxta	Paxta, Paxtakor, Paxtakon, Paxtakesh,

2	Ajriq	Ajriqti
3	Arpa	Arpa, Arpako'l, Arpaliko'l, Arpapoya
4	Archa	Archazor, Yakkaarcha
5	Beda	Bedazor, Bedapoya
№	Hayvonlarga oid leksikalar	Hosil bo'lgan joy nomlari
1	Ayg'ir	Ayg'ir, Ayg'irtepa
2	Baliq	Baliqli, Baliqliko'l
3	Burga	Burgali, Burgalik
4	Bo'ri	Bo'risoy, Bo'ritepa,
5	Do'lta	Do'lta, Do'ltali
№	Kasb-hunarga oid leksikalar	Hosil bo'lgan joy nomlari
1	Aravachi	Aravakashlar, Aravachilar
2	Baxshi	Baxshilar, Baxshiyon, Baxshitepa
3	Baliqchi	Baliqchi, Baliqchilar
4	Bo'zchi	Bo'zachi, Bo'zchi
5	Bo'yra	Bo'yraf, Bo'yraf, Bo'yraf, Bo'yraf, Bo'yraf

Yuqoridagi indikatorlardan tashqari hech bir guruhga kirmaydigan, ammo faol qo'llaniladigan "xona", "obod" kabi so'zlar ham mavjud. Ushbu indikatorlar so'zga qo'shilib yoziladi; toponimlarning oykonim turini hosil qiladi. Masalan, *Dehqonobod, Bekobod, Mirobod, Paxtaobod, Yunusobod, Chinobod...*; *Yomxona, Qalandarxona, Qulonxona, Jayraxona, Chinnixona...*

3-jadval. *Gidronimlarni hosil qiluvchi topotermin(indikator)lar*

Tarkibida topoterminlar ishtirok etgan - Hidronimlar (suv inshootlari)		
№	Gidronimlarga oid leksikalar	Hosil bo'lgan joy nomlari
1	Dengiz	Dengizko'l
2	Daryo	Sirdaryo, Qashqadaryo, Qoradaryo
3	Ariq	Yangiariq, Toshariq, Moyariq, Ariqbo'yi, Beshariq, Bo'zariq, Damariq, Javharariq
4	Ko'l	Ko'liyon, Ko'lob, Labiko'l, Luqliko'l, Oqko'l, Oy-dinko'l, Porsonko'l, Qo'shko'l, Xubbonko'l
5	Rud	Shahrud (shahr i rud), Bolorud, Labirud (lab i rud), Ro'dasoy (rud-ak)
6	Nahr (nar, nor)	Nahr ash-Shosh, Ulug'nor, Norin, Minor,
7	Quduq	Sho'rquduq, Uchquduq, Jalolquduq, Oqquduq, Otquduq, Oyoqquduq, Qo'shquduq, Toshquduq
8	Suv	Oqsuv, Yettisuv, Sho'rsuv, Qorasuv, Bo'zsuv, Egrisuv, Iyrisuv

9	Hovuz	Labihovuz
10	Soy	Arnasoy, Badal soy, Bodoqsoy, Dahanasoy, Daryosoy, Kattasoy, Kosonsoy, Moylisoy, Oqsoy
11	Ob (Forscha – suv)	Obi Farg'ona (Farg'ona daryosi), Obirahmat, Obi Xo'jand, Panjob (panj ob), Ponob, Qarnob
12	Buloq	Tovoqbuloq, Toshbuloq, Yo'rg'abuloq,

Shuningdek, gap tarkibida quyidagi gidronim obyektlarni bildiruvchi so'zlar qo'llanilsa, gap tarkibidagi bosh harfda yozilgan so'z katta ehtimol bilan toponim bo'ladi: *Akvatoriya, akveduk, anhor, arashon, ariq, arna, artezion, afdoq, badoq, balchiq, band, botqoq, bo-shot, bo'g'at, bo'g'oz, guzar, dayravot, dam...*

4-jadval. Oronimlarni hosil qiluvchi topotermin(indikator)lar

Tarkibida topoterminlar ishtirok etgan – Oronimlar (yer yuza relyeflari)		
№	Oronimlarga oid leksikalar	Hosil bo'lgan joy nomlari
1	Tog' (tov)	Olatog', Bobotog', Qo'ngirtog', Tog'tosh, Shertog', Chimqortog'
2	Tosh	Teshiktosh, Tog'tosh, Toshariq,
3	Bobo	Bobotog', Boboqum
4	Ota	Nurota, Yog'ochota, Zangiota
5	Ranglar (Oq, qora, sari(q), qo'ng'ir...)	Qo'ng'iroq, Qo'ng'irobod
6	Bel – Baxmal, Zomin tumanlarida “dovon”	O'rtabel, Ayribel, Beltov, Oqbel, Qumbel
7	Dovon – “tog'dan oshaverishga yaqin dam olinadigan tekisroq maydoncha”	Kandirdovon
9	Egar (Navoiyda va qurama shevasida – “iyar”)	Egarchi, Iyartosh, Egarbozor
10	Zov (zovlin)	Zovurak, Avezov, Surxzov
11	Qir (qirang, qirpang)	Qaqir, Qirguli, Qoraqir, Uzunqir

Shuningdek, gap tarkibida quyidagi oronim obyektlarni bildiruvchi so'zlar qo'llanilsa, gap tarkibidagi bosh harfda yozilgan so'z katta ehtimol bilan toponim bo'ladi: *tekislik, tog', tog' tizmasi, pasttekislik, soylik, adir, jarlik, tepalik, jaylov(yaylov), jar, nov, nova nura, buloq, gaza, dara, aqba (ovg'a)...*

5-jadval. Etnotoponimlarni hosil qiluvchi topotermin(indikator)lar

Tarkibida topoterminlar ishtirok etgan – Etnotoponimlar (urug'-qabila nomlari)		
№	Etnotoponimlarga oid leksikalar	Hosil bo'lgan joy nomlari
1	O'zbek	O'zbekon, Yuqori O'zbekiston, O'zbekon

2	Arab	Zarafshon vodiysida: Arab (12), Arabla(r) (4), Arabon (14), Arabxona (40), Arabtepa, Qo'ldosharab, Milkiarab, Chilonguarab
3	Baxrin	Baxrin, Bahrinbayot, Bahrincha, Bahrinboz
4	Do'rman (Do'rmon)	Do'rman, Do'rmanon, Do'rmantepa
5	Qalmoq	Qalmoqon, Ravoti qalmoq, Qalmoqkendi, Qalmoqbevezanon
6	Qirg'iz	Qirg'iz, Qirg'izon, Qirg'izsaroy
7	Qiyot	Qiyot, Qiyoton
8	Qozoq	Qozoq, Qozoqon, Maydaqozoq, Qozoqqoldikalono'rta, Qozoqqoldikalonyuqori, Beshqozoq,
9	Qipchoq	Qipchoq, Qipchoqon, Qipchoqxo'ja, Qipchoqbolo
10	Qitoy	Qitoy, Qoraqitoy, Qitoyon
11	Mang'it	Mang'iton, Mang'itpoyon
12	Mo'g'ul	Mo'g'ulon, Mo'g'ulkenti
13	Ming	Mingla(r), Mingon
14	Saroy	Saroy, Uyg'ursaroy, Saroypoyon, Saroybolo
15	Tojik	Tojikon, Tojiklar
16	Tatar	Tataron, Tatarcha, Tatarmahalla
17	Turk	Turkon, Turkman
18	Urganji	Urganjiyon, Urganjilar

Bundan tashqari yuqoridagi indikatorlardan tashqari anatomik (inson tanasi a'zolari nomi) topoterminlar ishtirok etgan qo'shma olmoshlar ham mavjud. Ba'zi adabiyotlarda anatomik topoterminlar – *metafora atamalar(terminlar)* atamasi bilan ham yuritiladi [Qorayev, 2006: 95].

6-jadval. *Tarkibida anatomik topoterminlar ishtirok etgan toponimlar*

Tarkibida anatomik (inson tanasi a'zolari nomi) ishtirok etgan toponimlar			
No	Termin	Geografik ma'nosi	Misol
	Barmoq	Tik qoya	Beshbarmoq,
	Bel	Dovon, oshuv, ko'tal	Oltinbel, Beltepa, Belariq
	Bosh	Daryoning boshlanish qismi, tog' cho'qqisi	Ko'prikboshi, Ariqboshi, Oqbosh
	Burun(murun)	Quruqlikning dengiz ichiga kirib borgan qismi	Xo'ja Axchaburun, Muruntov

	Bo'yin(mo'yin)	Quruqliklar, tog'larni (relyefning musbat shakllarini) tutashtirib turuvchi joy	Qoramoyin, Tuyamoyin
	Bo'g'iz	Dengizlarni, ko'llarni va ularning qismlarini tutashtirib turuvchi suv hududi, akvatoriya	Tuyabo'g'iz
	Kindik	Tabiiy-hududiy birliklarning o'rta, markaziy qismi	Kindiktepa
	Ko'z	Buloq ko'zi – buloq chiqqan joy	Qorako'z
	Oyoq	Etak, quyi qism, odoq	Qiziloyoq
	Og'iz	Daryolarning quyilish joyi (ko'rfaz), tog' darasining eng tor joyi	Jarog'iz
	Til	Ko'rfaz oldi qum orollar	Ko'haktil
	Tirsak	Daryoning burilish joyi	Tirsak
	Tumshuq	Tog', tepalik va orollarning burni, uch tomoni	Tumshuqariq, Oqtumshuq
	Qovoq	Daryo va ko'llarning baland qirg'og'i, jarlik; Baland terrasa	Oqqovoq (Aqqabaq)
	Qosh (qash)	Uzun cho'zilgan tepalik, marza	Yettiqashqa, Qashqalar

Ajratib yoziladigan toponimlar

Ajratib yoziladigan toponimlar *ikki komponent(qism)li* va *ikkitanadan ortiq komponent(qism)li* shaklda bo'ladi.

Ikki komponentli toponimlar ikki so'zdan tashkil topib, bitta mustaqil joy nomini bildiradi. Odatda, ikki komponentli toponimlarda birinchi qism sifat so'z turkumiga oid bo'lib, u o'zidan keyingi so'zdan ajratib yoziladi. Masalan: *Kichik Aravon, Katta Beshkapa, Pastki Oyoqchi, Quyi Chirchiq, Eski Xo'jaobod, O'rta Sho'rchi, Yuqori Olot, Yangi Baliqchi...*

Ajratib yoziladigan toponimlar tarkibi ikkitadan ham ortiq bo'lishi mumkin. Bunday toponimlar ko'proq kishi ismlari bilan ataluvchi toponimlarda, ko'chalar nomlarida, yer yuza relyeflari nomlarida va yo'nalishni bildiruvchi toponimlarda uchraydi. Masalan, *Buyuk Ipak yo'li bekati, O'zbekiston ovozi gazetasi ko'chasi, Hamza Hakimzoda Niyoziy ko'chasi, Zahiriddin Muhammad Bobur mahallasi, Mullaning qizi uchgan qoyasi...*

Toponim(lar)ni avtomatik aniqlashdagi muammolar

Toponimlarda shakldoshlik (omonimiya) hodisasi va ularni aniqlashdagi muammolar

Toponimlar bir qarashda faqatgina joy nomlari emas. Toponim – ushbu joyning tarixi, o'tmishi, qadriyati, madaniy boyligi, ma'naviy merosi bo'libgina qolmay, unda tilimizning tarixiy xususiyatlari aks etadi. Joyga nom qo'yish jarayonida kimgadir (yoki

Ner: o'zbek tilidagi matnlarda toponim(lar)ni avtomatik aniqlash modellari nimagadir), biror voqea-hodisaga nisbat berilib, kishi ismlarini yoki biror muhim kunni abadiylashtirish maqsad qilinadi.

Toponimlarni aniqlashdagi muammolardan biri ulardagi omonimiya (shakldoshlik) hodisasi. Ya'ni bir nom ostida bir nechta obyekt ifodalanishi mumkin. Bunday hodisa toponimlarning *antro-toponimlar* (kishi ismlari, familiyasi, laqabi, taxallusi bilan ataluvchi joy nomlari) guruhi doirasida ko'p kuzatiladi. Bu o'rinda kishi ismlari, familiyasi, laqabi, taxallusi *eponim* (kingadir, qayergadir yoki biror narsaga nomi berilgan yoki shunday deb taxmin qilinadigan *shaxs, joy yoki narsa* nomi) bo'lib xizmat qiladi.

1989-yildan 2003-yil oktabr oyigacha Toshkentdagi mustaqil nomlarga ega bo'lgan ko'chalardan 1000ga yaqin ko'cha qayta nomlandi. Nomi o'zgartirilgan ko'chalarning qariyb uchdan bir qismiga Ahmad Yassaviy, Bahouddin Naqshband, Imom at-Termiziy, Mahmud Torobiy, Jaloliddin Manguberdi, Amir Temur, Bobur, Husayn Boyqaro, Temur Malik kabi buyuk shaxslarning nomlari berilib poytaxt toponimiyasida abadiylashtirildi. Shu bilan birga Muxtor Avezov, Betxoven, Maxatma Gandi, Lohutiy, Neru, Firdavsiy, Lev Tolstoy, Chexov kabi jahon klassiklarining nomlari Toshkent ko'chalarini bezab turibdi [Qorayev, 2006:145].

7-jadval. Kishi ismlari bilan ataluvchi NER obyektlari

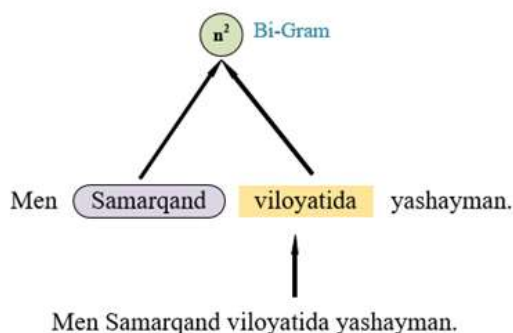
No	Ibn Sino
1	Ibn Sino – shaxs
2	Ibn Sino nomidagi ixtisoslashtirilgan maktab
3	Ibn Sino nomidagi jamoat salomatligi texnikumi
4	Ibn Sino nomidagi kollej
5	Ibn Sino nomidagi davlat stipendiyasi
6	Ibn Sino ko'chasi
7	Ibn Sino nomidagi Buxora Davlat Tibbiyot Instituti
8	Ibn Sino nomidagi tibbiy oliy va o'rta bilim yurtlari (Tojikiston)
9	Ibn Sino haykali (Buxoro shahrida, Afshona qishlog'ida va Belgiyaning Kortreyk shahrida)
10	Ibn Sino muzeyi (Afshonada)
11	Ibn Sino nomidagi nashriyot
12	Ibn Sino nomidagi sanatoriya
13	Ibn Sino nomidagi shifoxona
14	Ibn Sino nomidagi kutubxona
15	Ibn Sino nomidagi respublika davlat mukofoti (Tojikiston)
16	Ibn Sino xalqaro jamg'armasi (O'zbekiston. 1999)
17	“Ibn Sino“ va „Sino“ xalqaro jurnallar

18	“Tbn Sino” portreti (S. Marfin)
19	“Abu Ali ibn Sino” spektakl-film (1984)

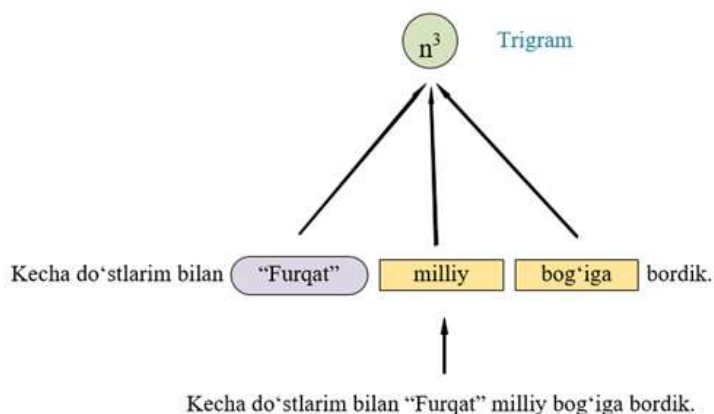
№	Alisher Navoiy
1	Alisher Navoiy – shaxs
2	Alisher Navoiy mahallasi (Samarqand viloyati Jomboy tumani va Sirdaryo viloyati Boyovut tumani)
3	Alisher Navoiy ko‘chasi (Yangiyer shahri, Mirzaobod, Sardoba tumani)
4	“Alisher Navoiy” romani
5	“Alisher Navoiy” fondi
6	Alisher Navoiy nomidagi milliy kutubxona
7	“Alisher Navoiy” istirohat bog‘i
8	Navoiy viloyati
9	Alisher Navoiy nomidagi davlat mukofoti
10	“Alisher Navoiy” filmi
11	O‘zRFA Alisher Navoiy nomidagi Til va adabiyot instituti
12	Alisher Navoiy nomidagi opera va balet akademik teatri
13	Alisher Navoiy nomidagi Toshkent davlat o‘zbek tili va adabiyoti universiteti

№	Jaloliddin Manguberdi
1	Jaloliddin Manguberdi shaxs nomi
2	Jaloliddin Manguberdi dahasi
3	Jaloliddin Manguberdi istirohat bog‘i (Qo‘qonda)
4	Jaloliddin Manguberdi ko‘chasi
5	Jaloliddin Manguberdi ordeni
6	“Sulton Jaloliddin Manguberdi” romani (Baxtiyor Abdug‘afur)
7	“Jaloliddin Manguberdi” dramasi (Maqsud Shayxzoda)
8	Xalq og‘zaki ijodida “Jaloliddin Manguberdi” dostoni
9	“Jaloliddin Manguberdi” pyesasi
10	“Jaloliddin Manguberdi” videofilm
11	Jaloliddin Manguberdi haykali

Yuqoridagi jadvallarda qora bilan belgilangan so‘zlar bu – *bigram/trigram* metodlaridan foydalanib toponimlarni aniqlashga yordam beruvchi kalit so‘zlardir. Bigram/Trigram – N-gram tahliliy metodi tarkibi bo‘lib, NLP (Natural language Processing – Tabiiy tilni qayta ishlash)da keng qo‘llaniladi. N-gramma matn yoki nutqning berilgan namunasidagi “n” ta elementning ketma-ketligidir. Ushbu metod matnda qaysi so‘zlarning ketma-ket kelishini ko‘rsatadi (bigram – ikkita so‘z ketma-ketligi; trigram – uchta so‘z ketma-ketligi). Masalan:



2-rasm. Bigram metodida toponimning aniqlanishi



3-rasm. Trigram metodida toponimning aniqlanishi

Bu o'rinda bigram metodi yordamida "viloyat" so'zi aniqlanadi va uning yonidagi bosh harfda kelgan so'z – "Samarqand" so'zi "joy nomi" – toponim sifatida qabul qilinadi. Xususan, bu o'rinda "viloyat" so'zi joy nomiga ishora qiluvchi indikator bo'lsa-da, uning tarkibida egalik –i qo'shimchasi bo'lishi zarur. Chunki modelga tarkibida –i egalik qo'shimchasi bo'lmagan "viloyat" so'zi olinsa, model faqat NER obyektlarini umumiy aniqlashga ishlaydi. Ammo aniqlangan obyekt toifasini to'g'ri aniqlamaydi. Masalan:

Kecha *Alisher viloyatdan* qaytdi;

Kecha *Alisher Samarqand viloyatidan* qaytdi.

Ushbu ikki gapga e'tibor qaratisa, ikkala gapda ham "viloyat" so'zi ishtirok etgan. Ammo birinchi gapdagi "viloyat" so'zi tarkibida –i egalik qo'shimchasi mavjud emas, ikkinchi "viloyat" so'zida esa i- qo'shimchasi mavjud. Birinchi gapda bigram mavjud – "Alisher viloyatdan"; ikkinchi gapda bigram – "Samarqand viloyatidan". Agar model qoidasiga ko'ra "viloyat" indikatorini qatnashgan so'z va bigram holatida uning yonidagi bosh harfda kelgan so'z toponim deb aniqlansa, yuqorida birinchi gapdagi "Alisher" so'zini antroponim (kishi ismi) sifatida emas, aksincha, toponim deb xato aniqlaydi.

Modelda bunday xatolik bo'lmisligi uchun kompyuter buyrug'iga "viloyat" so'zi tarkibida –i egalik qo'shimchasi bor holatida ("viloyati") kiritiladi. Demak, tarkibida egalik qo'shimchasi mavjud indikator kelsa, uning bigram holatida yonidagi bosh harfli so'z – toponim; agar egalik indikator so'z qo'shimchasiz kelsa, unda bigram bo'lib kelayotgan so'z toponim emas.

8-jadval. Toponim(joy nomi)larning antroponim (inson ismlari) bilan shakldosh holatda kelishi

№	Shaxs nomlari	Joy nomi
1	Abdusamad	Farg'ona viloyati Dang'ara tumanidagi qishloq.
2	Abu Muslim	Buxoro viloyati hududlaridagi kanal va tepalikning nomi
3	Bektemir	Toshkent shahridagi tuman, Qashqadaryo viloyatining Kitob tumani
4	Feruza	Samarqand shahri devoriga janubi-g'arb tomondan o'rnatilgan darvoza
5	Furqat	Farg'ona viloyati tarkibidagi tuman nomi
6	Karim (Karimobod)	Farg'ona viloyati Bag'dod tumanidagi qishloq va mahalla nomi
7	Madaminov	Farg'ona viloyati Yozyovon tumanidagi qishloq nomi
8	Marhamat	Andijon viloyati tarkibidagi shahar va tuman nomi
9	M a ' m u r (Ma'murobod)	Andijon viloyatining Qo'rg'ontepa tumanidagi qishloq
10	Mirza(o) (ariq,)	Samarqand viloyatining Jomboy tumanidagi ariq; Zarafshon daryosidan chiqib, Bulung'ir kanaliga quyadigan kanal
11	Mirza (cho'l)	Jizzax viloyati tarkibidagi tuman nomi; O'zbekiston va qo'shni davlatlar (Qozog'iston, Tojikiston) hududidagi tekislik nomi
12	Mo'min(obod) – Mo'min ato ismli shaxs nomiga	Buxoro viloyati Vobkent tumanidagi qishloq
13	Nasriddin(o-bod) – Qo'qon xoni Xudoyorxonning katta o'g'lining ismi	Andijon viloyati Bo'z tumani markazining eski nomi.
14	Nodirqo'rchi – Qo'qon xonligida qurol-aslah boshlig'i bo'lgan shaxs	Buxoro shahri guzari

15	Norbo'ta – Qo'qon xoni Norbo'tabiy (1769-1801) tomonidan asos solingan	Andijon viloyatining Qo'rg'ontepa tumanidagi qishloq
----	---	--

Yuqoridagi nomlar birgina antroponim (shaxs nomi) bo'lgan holatida boshqa obyekt(lar) nomi bo'lib ham keladi. Masalan: Navoiy – shaxs nomi (antroponim), Navoiy viloyati – joy nomi (toponim). Shuningdek, shunday toponimlar mavjudki, ular birgina toponim bo'lgan holatida ham turli toponimik obyektlar nomini bildiradi va o'z ichida shakldoshlikni yuzaga keltiradi.

9-jadval. *Toponim(joy nomi)larning shakldosh holatda kelishi (qizil rangda berilgan obyektlar toponim emas, ammo NER obyekti)*

№	Joy nomlari	Shakldoshlik holati
1	Zarafshon	1. Vodiy nomi (Samarqand va Qashqadaryo hududlari) 2. Daryo nomi 3. Shahar nomi (Navoiy viloyati) 4. Ko'cha nomi (Navoiy viloyati) 5. Tog' tizmasi nomi (Pomir-Oloy tizmalari)
2	Sirdaryo	1. Viloyat nomi 2. Daryo nomi 3. Telekanal nomi
3	Andijon	1. Viloyat nomi 2. Shahar nomi 3. Suv ombori 4. Mehmonxona nomi 5. Telekanal nomi 6. Muzqaymoq nomi 7. Polka nomi

4	Samarqand	<ol style="list-style-type: none"> 1. Viloyat nomi 2. Shahar nomi 3. Do'kon nomi 4. Mehmonxona nomi 5. Kafe nomi 6. Restoran nomi 7. "Samarqand" hududgaz ta'minoti MChJ 8. "Samarqand" asari (prof. Z. M. Akramov) 9. Gazeta nomi 10. Telekanal nomi 11. "Samarqand" non
5	Q a s h q a - daryo	<ol style="list-style-type: none"> 1. Viloyat nomi 2. Daryo nomi 3. Ko'prik nomi 4. Gazeta nomi 5. Telekanal nomi
6	Farg'ona	<ol style="list-style-type: none"> 1. Viloyat nomi 2. Shahar nomi 3. Vodiy nomi 4. Tuman nomi 5. Mehmonxona nomi; 6. Kanal nomi; 7. Telekanal nomi;

Shuningdek, ba'zi adabiyotlarda nomlarning ko'chish hodi-sasi "*birlamchi va ikkilamchi nomlar*" atamasi ostida ham yuritiladi. Bunda birlamchi nomlar asl nomlar; ikkilamchi nomlar nom ko'chishidan hosil bo'ladi. Masalan: *Chirchiq daryosi, Chirchiq shahri*. Nom daryo nomidan shaharga ko'chgan [G'ulomov, Mirakmalov, 2005: 6].

Xulosa

Xulosa o'rnida shuni aytish mumkinki, matnlar tarkibida joy nomlarini aniqlashda toponimlarning tuzilish jihatidan tarkibiga

Ner: o'zbek tilidagi matnlarda toponim(lar)ni avtomatik aniqlash modellari ko'ra modellar yaratish maqsadga muvofiq. Chunki NER obyektlarini aniqlashga yo'naltirilgan modellar toponimlarning tarkibiga ko'ra aniqlash tizimiga qulay. Ushbu maqolada toponimlarni aniqlash metodlari, modellari va avtomatik aniqlashga yo'naltirilgan tizimlar taqdim etildi va misollar bilan isbotlandi.

Foydalanilgan adabiyotlar

Do'simov Z., Egamov X. Joy nomlarining qisqacha izohli lug'ati. -T., O'qituvchi, 1977.

G'ulomov P, Mirakmalov M. Toponimika va geografik terminshunoslik. -Toshkent, 2005.

Hakimov Q. Toponimika. Toshkent: Mumtoz so'z, 2016.

<https://qomus.info/oz/encyclopedia/o/onomastika/>

Qilichov B. Onomastika. O'quv qo'llanma. - Buxoro, 2023.

Qorayev. S. Toponimika. -Toshkent: O'zbekiston faylasuflari milliy jamiyati, 2006.

NER: MODELS FOR AUTOMATIC DETECTION OF TOPONYM(S) IN UZBEK LANGUAGE TEXTS

Botir Elov¹

Madina Samatboyeva²

Abstract

This article analyzed the composition of toponyms, their relationship with other disciplines, and their grammatical structure. Etymological and grammatical analyzes of toponyms were also presented. Simple and compound toponyms were discussed and their toponymic models were presented. Topoformants (supplements) involved in the formation of toponyms in our language have been studied. Toponyms were divided into semantic groups, and a list of topoterms (indicative words referring to the toponym - indicator) was presented. The phenomenon of homonymy in toponyms was studied, and the problems that arise when place names are automatically determined with the help of machine learning were brought up.

Keywords: NER, name, named object, place name, toponymy, toponymy, toponym, onomastics, linguistics, etymological analysis, toponymic modeller, toponymic type, topoformant, suffix, affix, topoterm, indicator.

References

- Do'simov Z., Egamov X. Joy nomlarining qisqacha izohli lug'ati. -T., O'qituvchi, 1977.
- G'ulomov P, Mirakmalov M. Toponimika va geografik terminshunoslik. -Toshkent, 2005.
- Hakimov Q. Toponimika. Toshkent: Mumtoz so'z, 2016.
- <https://qomus.info/oz/encyclopedia/o/onomastika/>
- Qilichov B. Onamastika. O'quv qo'llanma. - Buxoro, 2023.
- Qorayev. S. Toponimika. -Toshkent: O'zbekiston faylasuflari milliy jamiyati, 2006.

¹*Elov Botir Boltayevich* – doctor of philosophy of technical sciences (PhD), associate professor. Tashkent State University of Uzbek Language and Literature named after Alisher Navoi.

E-mail: elov@navoiy-uni.uz

ORCID: 0000-0001-5032-6648

²*Samatboyeva Madina To'lqinjon qizi* – Graduate student of Computer Linguistics. Tashkent State University of Uzbek Language and Literature named after Alisher Navoi.

E-mail: msamatboyeva@gmail.com

ORCID: 0000-0002-5146-4484