

RAQAMLI TEXNOLOGIYALAR SOHASIDA FRANSUZ TILINING LEKSIK BIRLIKLARINI SHAKLLANTIRISH MODELLARI

Munisa Tolliboyeva Abdulla qizi

*Samarqand Davlat Chet Tillar Instituti Roman-German Tilari fakulteti,
fransuz filologiyasi kafedrası 4-bosqich talabasi*

Annotatsiya: *Ushbu ilmiy maqola zamonaviy raqamli texnologiyalar, xususan sun'iy intellekt, katta ma'lumotlar va lingvistik ontologiyalar kontekstida fransuz tilining leksik birliklarini (so'z va iboralarni) shakllantirish va tashkil etishning zamonaviy modellarini o'rganadi. Tadqiqotda fransuz raqamli leksikografiyasining an'anaviy lug'atlarni raqamlashtirishdan tortib, interfaol hamkorlikdagi loyihalar va sun'iy intellekt asosida ishlaydigan tizimlargacha bo'lgan evolyutsiyasi tahlil qilinadi. Ikkita asosiy yo'nalish – raqamli vositalar yordamida leksik boylikni yaratish (generativ model) va mavjud leksikani tizimlashtirish va qayta ishlash (tashkiliy model) – chuqur o'rganiladi. Maqolada fransuz tilining raqamli muhtidagi barqarorligi va rivojlanishi uchun muhim bo'lgan "Linked Lexical Open Data" (LLOD) va lingvistik ontologiyalar kabi tushunchalar tahlil etiladi. Tadqiqot natijalari shuni ko'rsatadiki, raqamli texnologiyalar nafaqat fransuz leksikasini saqlash va targ'ib qilishni, balki uning dinamik rivojlanishini ham shakllantirishda hal qiluvchi rol o'ynamoqda. Leksikani yaratish va tashkil etish jarayoni endi faqatgina tilshunoslarning ishi emas, balki keng hamjamiyatning ishtirok etadigan, katta ma'lumotlar va sun'iy intellekt vositalari bilan boyitilgan ochiq ekotizimga aylangan.*

Kalit so'zlar: *raqamli leksikografiya, fransuz tili, leksik birlik, ontologiya, katta ma'lumotlar, Linked Open Data (LLOD), hamkorlik loyihalari, sun'iy intellekt.*

Kirish. Har qanday tirik tilning asosiy xususiyati uning leksik boyligining doimiy o'zgarishi va yangi so'z, ibora va tushunchalar bilan to'lishidir. Zamonaviy raqamli davr bu jarayonga tubdan yangi dinamika kiritdi: leksikani shakllantirish endi nafaqat tabiiy til evolyutsiyasining natijasi, balki qasddan dasturlashtirilgan texnologik platformalar, algoritmlar va hamjamiyat faoliyatining mahsulidir. Fransuz tili, o'zining boy an'anaviy merosi va butun dunyo bo'ylab tarqalishi bilan, bu raqamli transformatsiyaning eng yorqin namunalaridan biridir. Raqamli texnologiyalar tilning leksik fondini faqatgina saqlash va ko'paytirish vositasi emas, balki uni tashkil etish, qayta ishlash va global miqyosda yetkazib berishning asosiy mexanizmidir. "Raqamli leksikografiya" tushunchasi an'anaviy lug'atchilikni kompyuterlashtirishdan hamda raqamli muhtida tug'ilgan va rivojlanadigan yangi leksik resurslar yaratishgacha bo'lgan kompleks hodisani anglatadi .

Ushbu maqolaning maqsadi – fransuz tilining leksik birliklarini raqamli sohada shakllantirishning asosiy modellarini belgilash, tahlil qilish va tizimlashtirishdir. Tadqiqot quyidagi savollarga javob berishga intiladi: Zamonaviy raqamli vositalar fransuz leksikasini qanday usullar bilan to'ldirmoqda va yangilamoqda? Katta ma'lumotlar va sun'iy intellekt leksikani qayta ishlash va izlash qobiliyatlarini qanday o'zgartiradi? Fransuz tilining leksik boyligini tashkil etish va saqlashda ontologik yondashuvlar qanday rol o'ynamoqda? Bu masalalarni o'rganish orqali tilning texnologiya bilan o'zaro ta'sirining yangi qonuniyatlari ochiladi va fransuz tilining raqamli kelajagi uchun istiqbollar belgilanadi.

Adabiyotlar Tahlili va Metodologiyasi. Tadqiqotning nazariy asosi raqamli tilshunoslik, korpus lingvistikasi va bilim muhandisligi sohalaridagi ishlardir. Asosiy metodologik yondashuv sifatida tizimli tahlil va sintez, qiyosiy-metodologik tahlil va amaliy misollar orqali umumlashtirish usullari qo'llaniladi. Asosiy manba sifatida "French digital lexicography: Metamorphosis, Status Quo and Future Directions" (2023) nomli ilmiy maqola muhim ahamiyatga ega. Ushbu ishda fransuz raqamli leksikografiyasining tarixiy evolyutsiyasi, hozirgi holati va kelajak tendensiyalari keng yoritilgan. Maqola, an'anaviy nashrlarni raqamlashtirish (masalan, "Littré" yoki "Trésor de la langue française informatisé – TLFi"), hamda "Wiktionnaire" yoki "Dictionnaire des Francophones" (DDF) kabi tug'ma raqamli hamkorlik loyihalarining paydo bo'lishini asosiy bosqichlar sifatida ajratib ko'rsatadi. Bundan tashqari, leksik ma'lumotlarning ochiq va bog'langan ma'lumotlar sifatida (Linked Lexical Open Data – LLOD) ishlab chiqilishi fransuz leksik resurslarini global bilim tarmoqlariga integratsiya qilishning yangi modeli sifatida ta'kidlanadi. Bu manba fransuz tilidagi leksikani shakllantirishning tashkiliy modelini (mavjud boyligni raqamli muhtda tashkil etish va bog'lash) tushunish uchun asos yaratadi.

Boshqa bir muhim yo'nalish – katta ma'lumotlar va sun'iy intellektning leksik jarayonlarga ta'siridir. Tadqiqotda fransuz so'zlarini tez izlash va aniqlash uchun chuqur o'rganish (deep learning) algoritmlaridan foydalanishni o'rganuvchi ilmiy ishga tayaniladi. Ushbu tadqiqot, leksik birliklarni qidirish va olish samaradorligini oshirish uchun neyron tarmoqlar va "SGD synchronous update network" kabi algoritmlarni qo'llash orqali leksikani shakllantirishning generativ/kognitiv modelini namoyish etadi, ya'ni texnologiya nafaqat so'zlarni saqlaydi, balki ularni aqlli tarzda tushunadi va taqdim etadi. Shu bilan birga, "Corpora and AI for inductive learning" nomli maxsus nashr korpus lingvistikasi va yirik til modellarini (LLM) integratsiyalashgan holda til o'rganish va tahlil qilishda qo'llash imkoniyatlarini ko'rib chiqadi, bu esa leksikani qayta ishlashning yangi gibril usullarini ko'rsatadi. Fransuz tilining katta hajmdagi matnli korpuslari, masalan, Frantext, uning leksik boyligini tizimli o'rganish va yangi leksik modellarni qurish uchun muhim

asosdir. "Chapitre II. Traitement informatique – Corpus" nomli bob fransuz matnlarini morfologik va sintaktik teglash (tagging) jarayonlarini tushuntiradi. Ushbu avtomatik va yarim avtomatik annotatsiya leksik birliklarni ularning grammatik xususiyatlari va kontekstdagi ishlatilishiga ko'ra tizimlashtirish, shuningdek, yangi so'z birikmalari va frazeologik birliklarni aniqlash imkonini beradi. Bu yondashuv, masalan, "The Phraseology of Legal French..." maqolasida qonuniy fransuz tilida "n-gram"lar va leksikogrammatik naqshlarni aniqlash uchun qo'llanilgan, bu esa leksikani shakllantirishda korpusga asoslangan modelning ahamiyatini tasdiqlaydi.

Tadqiqot jarayonida ushbu turli manbalardan olingan ma'lumotlar tizimlashtirilgan tahlil usuli yordamida umumiy modellar va tendensiyalarga bo'linadi. Muhokama va natijalar leksikani shakllantirishning asosiy yo'nalishlari bo'yicha: yaratish/generatsiya, tashkil etish/strukturizatsiya va integratsiya/aloqa jihatlaridan tashkil etiladi.

Muhokama va Natijalar. Tadqiqot natijalari asosida raqamli texnologiyalar sohasida fransuz tilining leksik birliklarini shakllantirish modellari quyidagi uchta asosiy yo'nalishda namoyon bo'lishini ko'rish mumkin. Bu model leksik boylikni faol ravishda kengaytirish va yangilash mexanizmlarini o'z ichiga oladi. U ikki shaklda namoyon bo'ladi: Bu model an'anaviy yopiq akademik va nashriyot jarayonlarini demokratiklashtiradi. Bunga eng yorqin misol – "Dictionnaire des Francophones" (DDF) loyihasidir. Bu loyiha butun dunyo bo'ylab fransuz tilida so'zlashuvchi hamjamiyatga fransuz tilining barcha lahja va variantlarini (Afrika, Karib havzasi, Okeaniya va boshqalar) kiritish uchun loyiha hisoblanadi. Bu yerda har bir ishtirokchi nafaqat leksikani iste'mol qiluvchi, balki uni faol ravishda yaratuvchiga aylanadi. Xuddi shunday, "Wiktionnaire" (fransuzcha Vikilug'at) ham o'ziga xos me'yoriy doiralardan tashqarida rivojlanib, zamonaviy sleng, texnologik terminlar va internet-memlarni tezda o'z ichiga oladi.

Katta til modellari (LLM) va maxsus algoritmlar leksikani qayta ishlash va hatto yaratishda yangi imkoniyatlar ochmoqda. Tadqiqotlar shuni ko'rsatadiki, chuqur o'rganish algoritmlari fransuz tilidagi nutqiy va matnli leksikani aniqligi va tezligi yuqori bo'lgan holda qidirish va tahlil qilish imkonini beradi. Bundan tashqari, Chat GPT kabi vositalar korpus tahlili bilan integratsiyalashgan holda, yangi so'z birikmalarini generatsiya qilish yoki mavjud leksikani turli kontekstlarda namoyish etish uchun ishlatilishi mumkin. Bu jarayon leksikani shakllantirishda inson va mashina hamkorligining yangi modelini yaratadi. Yaratilgan va to'plangan leksik boylikni tizimli, qidirish va tahlil qilish uchun qulay shaklga keltirish muhim vazifadir. Bu yerda quyidagi modellar ajralib turadi: Bu model leksik birliklarni nafaqat alifbo tartibida, balki ular orasidagi semantik munosabatlar (sinonimiya, giperonim/giperonim, meronimiya va boshqalar) tarmog'i sifatida tashkil etadi. Fransuz

tilida bunga “WordNet” ga o‘xshash tizimlar (masalan, “WOLF – Wordnet Libre du Français”) yoki “Dictionnaire des synonymes du CRISCO” kabi resurslar misol bo‘la oladi. Ushbu model kompyuter tizimlari tomonidan ma’noni tushunish va mantig‘iy xulosalar chiqarish uchun asos yaratadi. O‘zbek tilidagi ontologik tadqiqotlar ham aynan shu tamoyillar – semantik munosabatlarni formalizatsiya qilishga qaratilganligini ta’kidlash lozim.

Katta hajmdagi elektron matn korpuslari (masalan, Frantext, korpus Le Monde) fransuz leksikasining haqiqiy qo‘llanilishini o‘rganish uchun asosdir. Avtomatik teglash va annotatsiya vositalari har bir leksik birlikning grammatik xususiyatlarini, chastotasini va boshqa so‘zlar bilan birikish qonuniyatlarini (kollokatsiyalar) aniqlash imkonini beradi. Masalan, yuridik fransuz tilining frazeologiyasini o‘rganishda korpus usullari an’anaviy iboralarni (“aux termes des présentes”) va ularning soddalashtirilgan variantlarini (“dans ce contrat”) aniqlashga yordam bergan. Bu model leksikani statik lug‘at elementlari sifatida emas, balki dinamik matnlar tarkibida ishlaydigan birliklar sifatida ko‘rishga imkon beradi. Ushbu model yaratilgan va tizimlashtirilgan leksik ma’lumotlarni foydalanuvchilarga samarali yetkazish va boshqa tizimlar bilan bog‘lash mexanizmlarini belgilaydi. Bu zamonaviy raqamli leksikografiyaning eng muhim tendensiyalaridan biridir. Maqsad – turli lug‘atlar va lingvistik ma’lumotlar bazalarini bir-biri bilan va tashqi bilim manbalari (masalan, Vikipediya) bilan semantik bog‘lash. Masalan, “Dictionnaire des Francophones” dagi ma’lumotlar LLOD tamoyillari asosida ishlab chiqilgan bo‘lib, bu uning ma’lumotlarini boshqa ilmiy yoki ta’lim resurslari bilan avtomatik integratsiya qilish imkonini beradi. Bu model leksikani yopiq ilovalardan global bilim tarmog‘ining ochiq qismiga aylantiradi. Raqamli texnologiyalar fransuz tilini o‘rganish vositalariga (masalan, Duolingo, AI muloqot botlari) chuqur integratsiyalashgan. Bu platformalar nafaqat lug‘atni taklif qiladi, balki kontekstda leksikani o‘rganish, interfaol mashqlar orqali mustahkamlash va shaxsiylashtirilgan tavsiyalar berish imkoniyatini yaratadi. Shu bilan birga, kengaytirilgan raqamli nashrlar (augmented digital publishing) matn ichida leksik ma’lumotlarni darhol ko‘rsatish, audio talaffuzni eshitish yoki vizual materiallar bilan bog‘lash imkonini beradi.

Fransuz tilining raqamli leksikani shakllantirish modellari

Model Nomi	Asosiy Xususiyati	Asosiy Texnologiyalar	Namuna Loyihalar/Misol
Kollaborativ	Keng hamjamiyat ishtirokida lug‘at	Viki-platformalar, bulutli	Dictionnaire des Francophones

Model Nomi	Asosiy Xususiyati	Asosiy Texnologiyalar	Namuna Loyihalar/Misol
Yaratish Modeli	yaratish va to‘ldirish.	hammualliflik.	(DDF), Wiktionnaire .
AI Generativ Modeli	Leksikani qayta ishlash, tahlil qilish va yangi kontentni yaratishda AI dan foydalanish.	Katta Til Modellari (LLM), Chuqur O‘rganish.	ChatGPT integratsiyasi bilan korpus vositalari, nutqni aniqlash tizimlari .
Ontologik Tizimlashtirish Modeli	Leksik birliklarni semantik munosabatlar tarmog‘i sifatida tashkil etish.	Ontologiyalar, tezauruslar, semantik veb tamoyillari.	WOLF (Wordnet Libre du Français), CRISCO sinonimlar lug‘ati .
Korpusga Asoslangan Tahlil Modeli	Haqiqiy matnlar korpusidan leksik qonuniyatlarni chiqarish.	Avtomatik teglash, n-gramma tahlili, konkordanslar.	Frantext, Le Monde korpusi, yuridik frazeologiya tadqiqoti .
Ochiq Bog‘langan Ma’lumotlar Modeli	Turli lingvistik resurslarni global bilim tarmog‘iga ulash.	Linked Open Data (LOD), RDF formatlari, FAIR tamoyillari.	DDF ning LLOD sifatida ishlab chiqilishi, boshqa ma’lumotlar bazalari bilan integratsiya .

Xulosa. Xulosa qilib aytish mumkinki, raqamli texnologiyalar fransuz tilining leksik birliklarini shakllantirish jarayonini tubdan o‘zgartirdi. An’anaviy, yuqoridan pastga qarab (top-down) va markazlashtirilgan model o‘rniga, bugungi kunda ko‘p qirrali, dinamik va ochiq ekotizim shakllandi. Bu ekotizim quyidagi asosiy xususiyatlar bilan tavsiflanadi: Leksikani yaratish huquqi faqat akademik muassasalar yoki nashriyotlarga emas, balki butun dunyo bo‘ylab fransuz tilida so‘zlashuvchi keng hamjamiyatga o‘tdi. “Dictionnaire des Francophones” kabi loyihalar tilning farqliligi va boyligini aks ettirishda hal qiluvchi rol o‘ynamoqda. Aqlli va Adaptiv: Sun’iy intellekt va katta ma’lumotlar tahlili leksikani nafaqat saqlash, balki uni tushunish, tahlil qilish va foydalanuvchi ehtiyojlariga moslashtirilgan tarzda taqdim etish imkonini beradi. Nutqni aniqlash, kontekstli tavsiyalar va interfaol o‘rganish platformalari buning aniq misolidir. Leksik birliklar endi alohida lug‘atlar sahifalarida emas, balki semantik munosabatlar tarmog‘ida (ontologiyalar) va global ochiq ma’lumotlar fazosida (LLOD) mavjud. Bu nafaqat qidiruv samaradorligini oshiradi, balki tilni kompyuter tizimlari tomonidan yanada chuqurroq qayta ishlash imkoniyatini ochadi. Dinamik va Doimiy Yangilanuvchi: Raqamli muhtda lug‘atlar statik nashrlar emas, balki doimiy rivojlanayotgan, real vaqt rejimida yangilanadigan resurslardir. Bu fransuz tilining zamonaviy texnologik, madaniy va ijtimoiy o‘zgarishlarga tez moslasha olishini ta’minlaydi. Fransuz tilining raqamli sohadagi muvaffaqiyati uning an’anaviy maqomini mustahkamlagan holda, uni kelajakka moslashtirishga xizmat qilmoqda. Ushbu tajriba, shu jumladan O‘zbekiston kabi boshqa mamlakatlar uchun ham qimmatli bo‘lib, tilning raqamli sovg‘asini yaratishda hamjamiyat, texnologiya va tizimli lingvistik yondashuvlarni uyg‘unlashtirish zarurligini ko‘rsatadi.

Foydalanilgan Adabiyotlar

1. Chen, L. (2023). French digital lexicography: Metamorphosis, Status Quo and Future Directions. Akademik maqola. https://www.academia.edu/125242287/French_digital_lexicography_Metamorphosis_Status_Quo_and_Future_Directions
2. ACORP jurnali maxsus soni. (2025). Corpora and AI for inductive learning: Theory and practice. Applied Corpus Linguistics. <https://www.sciencedirect.com/science/article/abs/pii/S2666799125000474>
3. Li, S., & boshq. (2022). Deep Learning Models for Fast Retrieval and Extraction of French Speech Vocabulary. PMC, Computational Intelligence and Neuroscience. <https://pmc.ncbi.nlm.nih.gov/articles/PMC9287002/>

4. Habert, B., & boshq. (1997). Chapitre II. Traitement informatique – Corpus. Openedition Books. <https://books.openedition.org/pupvd/42518?lang=en>
5. Bouyé, M., & Gledhill, C. (2024). The Phraseology of Legal French and Legal Popularisation in France and Canada: A Corpus-Assisted Analysis. Languages, 9(3), 107. <https://www.mdpi.com/2226-471X/9/3/107>
6. Ishibie, C. I. (2025). The Impact of Digital Technologies on French Language Learning. Cascades Journal, 3(1), 33–39. <https://cascadesjournal.com/index.php/cascades/article/view/73>
7. StudySmarter. (2024). Restructuration Textuelle: Techniques & Définition. <https://www.studysmarter.fr/resumes/traduction/traduction-adaptative/restructuration-textuelle/>
8. Abjalova, M. (2021). O‘ZBEK TILI ONTOLOGIYASI: YARATISH VA TEKNOLOGIYASI. Monografiya. https://www.academia.edu/96998515/O%CA%BBZBEK_TILI_ONTOLOGIYASI_YARATISH_TEXNOLOGIYASI_VA_KONSEPSIYASI_Monografiya_DSc_Abjalova_M_qayta_nashr