

**O‘ZBEKISTON RESPUBLIKASI OLIY VA O‘RTA MAXSUS
TA‘LIM VAZIRLIGI**

**ALISHER NAVOIY NOMIDAGI TOSHKENT DAVLAT O‘ZBEK
TILI VA ADABIYOTI UNIVERSITETI**

**«O‘ZBEK TILINI DUNYO MIQYOSIDA KENG
TARG‘IB QILISH BO‘YICHA HAMKORLIK
ISTIQBOLLARI»
mavzusidagi
xalqaro ilmiy-amaliy anjuman materiallari**

2020-yil 19-20-oktabr

TOSHKENT - 2020

5. Суропов Б. Электрон таълим муҳитида ахборот коммуникация технологиялари фанини ўқитишнинг методик таъминотини такомиллаштириш. Дисс. PhD. – Тошкент, 2019. –150б.

6. Таълимда ахборот-коммуникация технологияларидан фойдаланиш. Педагогик дастурий воситалар. Маъруза матни. – // [motm.tma.uz/site /wp-content/.../09/3-2-1-МАЪРУЗА.pdf](http://motm.tma.uz/site/wp-content/.../09/3-2-1-МАЪРУЗА.pdf).

Lingvistik kompyuter dasturlari uchun ma'lumotlar bazasi va lingvistik protsessor

Database and linguistic processor for linguistic computer software

Abjalova ManzuraAbdurashetovna*

Annotation: *The creation of linguistic programs for natural language processing (NLP) is determined by its formalization for computer technology. A language base is needed to create natural language translation products. Also, the formal model of the language is reflected in the linguistic processor. This article deals with a database of linguistic programs and a linguistic processor using the example of a machine translation program.*

Key words: *state language, lexicographic sources, dictionary, module, linguistic processor.*

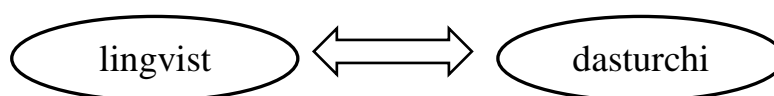
O'zbek tili Davlat tilimiz, davlatimizning milliy ramzi va milliy borligimizdir. Shu o'rinda izoh beraman: nega milliy boylik emas, borlik. Chunki til topib olingan, yig'ilgan va qo'ldan boy berish ehtimoli bo'lgan boylik emas, balki xalqimizning tili, madaniyati, xalqimiz o'zligi va o'zbekligini namoyon etuvchi ijtimoiy hodisa. Tilimizning ko'plab imkoniyatlariga nisbatan so'zlar xazinasini metaforasini qo'llashimiz mumkin, lekin butun bir tilga boylik sifatida qarashimiz tanqidiy mulohazani yuzaga keltiradi. O'zbek tilining mavqeyini mustahkamlash borasida fikr-mulohazalar azaldan davom etib kelyapti, lekin keyingi yillarda bu masala siyosiy islohot darajasiga ko'tarildi. Jumladan, Respublikamiz Prezidentining 2019-yil 21-oktabrdagi "O'zbek tilining davlat tili sifatidagi nufuzi va mavqeyini tubdan oshirish chora-tadbirlari to'g'risida"gi Farmoni [1] e'lon qilindi. Unga ko'ra Vazirlar Mahkamasi huzurida "Davlat tilini rivojlantirish departamenti" tashkil etildi va uning aniq vazifalari belgilab berildi. 2020-yil 10-aprelda O'zbekiston Respublikasi Qonuni bilan "21-oktabr – O'zbek tili bayrami kuni" etib belgilandi [2]. Davlat idoralari va tashkilotlari, jumladan, mahalliy hokimiyat organlarining davlat tilida ish yuritishi, davlat idoralarida davlat tili to'g'risidagi qonun hujjatlariga qat'iy rioya etilishi ustidan nazorat olib borishini monitoring qilish tizimi yaratildi.

Yuqorida ko'rsatilgan Farmonda davlat tilining xalqaro miqyosdagi o'rni va nufuzi, uning boshqa tillar bilan aloqalari istiqbollarini belgilash, xorijda istiqomat qiluvchi vatandoshlar va o'zbek tilini o'rganish istagida bo'lgan chet el fuqarolari uchun o'zbek tili darsliklari va elektron dasturlarni ishlab chiqish va ularni keng miqyosda tarqatish; davlat tilining axborot va kommunikatsiya texnologiyalari, xususan, Internet jahon axborot tarmog'ida munosib o'rin

*Filologiya fanlari bo'yicha falsafa doktori (PhD), Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti. manzura_ok@mail.ru

egallashini ta'minlash, o'zbek tilining kompyuter dasturlarini yaratish vazifalari belgilab qo'yilgan bo'lib, mazkur vazifalar raqamli iqtisodiyotga jadal ravishda o'tilayotgan hozirgi davrda dolzarb masalalar sirasiga kiradi.

O'zbek tilining Davlat tili sifatida mavqeyini mustahkamlash va jahon tillari qatorida iste'molda bo'lishiga erishish maqsadi, shuningdek, kompyuter texnologiyalari ish faoliyatimiz, hattoki turmush tarzimizning ajralmas vositalariga aylangani o'zbek tilini formallashtirish, uning Milliy korpusini yaratish, o'zbek va o'zga tillar juftligi bo'yicha tarjimon dasturlarini yaratish, o'zbek tilini qayta ishlash tizimlari va dasturiy ta'minotlari, jumladan, nutqni tanish (распознавание речи) imkoniyati va nutq sintezatorini yaratish, avtomatik referatlash, morfoanalizator (spelling, lemmatization, stemming) va sintaktik tahlil (parsing)ni yaratishni taqozo etilmoqda. Ko'rsatib o'tilgan mazkur dastur, tizim va mobil ilovalarning yaratilishida mukammal lingvistik ta'minot va leksikografik ma'lumotlar bazasining shakllantirilishiga asoslaniladi. Tom ma'noda til imkoniyatlarini elektron hisoblash vositalarida namoyon etishda quyidagi bog'liklik mavjud bo'ladi:



Ikki soha mutaxassislarining maqsadli mushtarak harakati, ikki soha integratsiyasi mahsuli Kompyuter lingvistikasining rivojlanishiga olib keldi. Bugungi kunda Toshkent davlat o'zbek tili va adabiyoti universitetida mazkur ta'lim sohasida "Mashina tarjimasi", "Kompyuter lingvistikasi" va "Tabiiy tilni qayta ishlash" fanlari ikki modul: lingvistik bilimlar va dasturiy injiniring modullari asosida ta'lim yo'lga qo'yilgan.

Lingvistik bilimlar modulida fanning dolzarbligi, maqsad va vazifalari, obykti, predmeti; mashina tarjimasining yuzaga kelish tarixi; fan bo'yicha konseptual bilimlar; formallashtirish so'z turkumlari moduli; lingvistik modellash asoslari; lingvistik muammolar va ularning dasturiy yechimi, lingvistik ta'minotni shakllantirish kabi masalalar o'rganiladi.

Dasturiy injiniring modulida dasturlash va lingvistik ma'lumotlar bazasi asoslari; SQLda ma'lumotlar bazasini yaratish; kodlash va dekodlash jarayoni, tarjimon dasturlar algoritmini tuzish; mashina tarjimasi va avtomatik tarjima vositalarini yaratish texnologiyalari o'rganiladi.

Mashina tarjimasi til juftliklariga asoslangan jarayon bo'lib, boshqa dasturlarga nisbatan tillarni xarakterlovchi keng qamrovli ma'lumotlar va leksikografik elektron manbalar zarur hisoblanadi. Ular lingvistik protsessorida formal akslanadi. Quyida mashina tarjimasi dasturiy ta'minotlarning lingvistik ta'minoti va lingvistik protsessor tarkibi berildi.

Lingvistik ma'lumotlar bazasining tarkibi. Mashina tarjimasi tizimi uchun lingvistik ma'lumotlar bazasiga matnlar, kartotekalar, lug'atlar, formal grammatik qoidalar va boshqa lingvistik manbalar bilan to'plangan lingvistik ma'lumotlar kiradi. Lingvistik ma'lumotlar bazasining odatdagi tarkibi quyidagilarni tashkil etishi mumkin:

a) L1 va L2 tillarning **leksik-grammatik klassifikatori** (morfologik kodlangan tizimi). L1 til matnini tahlil qilishda undagi har bir so'z tegishli morfologik izohlarga ega bo'lishi kerak: turkumlik belgisi, shakli, kelishigi, tuslanishi, soni va boshqalar. Muhimi kodlash tizimi muayyan mashina tarjimasi tizimi uchun bir xil bo'lishi kerak.

b) **Ikki tilli morfologik lug'at.** Ushbu lug'atda L1 tildagi har bir so'z shaklining L2 tilning so'z shakllariga muvofiq so'zma-so'z o'rnatiladi.

v) **Qisqartmalar va abbreviatsiyalar lug'ati.** Lug'at dastlabki matnni so'zlar va jumalarga ajratish bosqichida qo'llaniladi. Qisqartmalarining imkon qadar yoyilmasi bo'lishi kerak, chunki ular gap bo'lagi vazifasida kelishi mumkin, shuning uchun ular sintaktik va semantik tahlil qilishda hisobga olinishi kerak.

d) **Iboralar lug‘ati.** Ushbu lug‘atdan sintaktik tahlilga qadar foydalaniladi, chunki iboralar bir butun semantis birlik hisoblanadi va u jumlada gap bo‘lagi vazifasida kelishi mumkin. Tarjima qilishda L1 tildagi iboralar tarjima qilingan tildagi bitta so‘zga to‘g‘ri kelishi mumkin.

e) **Sohalar bo‘yicha terminologik lug‘atlar.** Ixtisoslashgan matnlarni tarjima qilish uchun kerak bo‘lganda qo‘shimcha lug‘atlar ulanadi.

f) **Sintaktik lug‘at.** Ushbu lug‘atda jumla a‘zolarining asl tilida ham, tarjima qilingan tilda ham sintaktik muvofiqligi, shuningdek, tarjima uchun zarur bo‘lgan sintaktik yozishmalar to‘g‘risida ma‘lumotlar bo‘lishi kerak.

g) **Semantik lug‘at** (tezaurus, ontologiya). Ushbu komponent leksemalarning semantik muvofiqligi, leksikosemantik sohalar haqida ma‘lumotni o‘z ichiga oladi, semantik qurish bosqichida ishlatiladi.

h) **Parallel matnlarning korpusi.** Korpusda asl tilidagi matnlar va ularning boshqa tilga tarjimalari mavjud. Parallel matnlar korpusida jumla yoki uning qismi topilgan bo‘lsa, uning tarjima tilidagi yozishmalari tarjima matniga kiritiladi. Tarjima xotirasi texnologiyasi matn korpusidan foydalanishga asoslangan.

Mashina tarjimasida lingvistik protsessor. Lingvistik protsessor (LP) L1 tildagi matnni lingvistik tahlil qilish, shuningdek, L2 tilda matnni sintez qilishga mo‘ljallanadi. lingvistik protsessor Lingvistik protsessor tabiiy tilga qayta ishlashga mo‘ljallangan axborot tizimlarining asosiy elementi [Mushkova, elektron: <https>]. LP tizim va foydalanuvchi o‘rtasida ma‘lumot almashishni osonlashtiradi. LP uchun asosiy vazifa jumla yoki matndan kerakli ma‘noni ajratib olish va uni mashina tushunadigan rasmiy vositalar to‘plamidan foydalanib tahlil qilishdir. Tizim axborotni o‘zgartirishning uchta darajasini o‘z ichiga oladi: morfologik, sintaktik va semantik. Ularning har biri qoidalar to‘plami bilan belgilanadi. Protsessor asosini formal matn/til tashkil etadi [Bolshakova, 2011: 91]. Unda tabiiy til bilan bog‘liq operatsiyalar amalga oshiriladi va uning birlamchi manbasi til materialini tahlil qilish hisoblanadi va LPda tahlil hamda sintezni amalga oshirishni ta‘minlaydigan formal me‘yorlar va lug‘atlar o‘rin oladi. Jumladan, zamonaviy leksik tahlil mashina lug‘atlariga tayanadi [Marchuk, 2007: 70]. Yuqorida mashina tarjimasida lingvistik ma‘lumotlar bazasi tarkibi berilgan bo‘lib, unda har bir lug‘at imkoniyati ko‘rsatildi. Lingvistik protsessor quyidagi tarkibiy qismlarni o‘z ichiga oladi:

a) Matnni jumlar va so‘zlarga ajratish dasturi.

b) barqaror iboralarni tanib olish dasturi. Idiomalarni tahlil qilish va bo‘linmas bir butun sifatida tarjima qilish kerak.

s) qisqartirish va abbreviatsiyalarni dekodlash dasturi.

d) L1 tildagi matnni morfologik izohlash dasturi.

e) Parser (sintaktik tahlil) va tobelik daraxtini yaratish dasturi

f) semantik tahlil qilish va L1 matnidagi har bir jumlaning semantik grafini yaratish dasturi.

g) Ikki tilli lug‘at yoki parallel matnlar korpuslaridan tarjima muvofiqliklarini tanlash dasturi

h) tarjima qilingan tilda matnni semantik sintez qilish dasturi.

i) Gapning sintaktik tuzilishini qurishi va sintez qilingan gapdagi so‘zlarning tartibini aniqlash dasturi.

j) Tarjima matnida so‘z shakllarini morfologik sintez qilish dasturi

Lingvistik protsessorning sifati bir qator muhim omillar bilan belgilanadi. Masalan, axborot obyektlarini ajratish va ular o‘rtasidagi aloqalar, identifikatsiya qilish qoidalari va protseduralarining aniqligi, tabiiy til matnlarining uning istalgan korpusiga sozlash va murakkabligini bartaraf etish imkoniyatini yaratish, shuningdek, yangi axborot obyektlarini avtomatik tushunish, lingvistik protsessor ishlaydigan tezlikni oshirish, matnni tahlil qilish uchun zarur bo‘lgan vaqtni qisqartirish kabilar.

Xulosa o'rnida shuni aytish kerakki, lingvistik protsessorlarni ishlab chiqish sun'iy intellektni yaratishning yana bir qadamidir. Chunki tabiiy tildagi matnni aniq tahlil qilish uchun ideal algoritmnining yaratilishi kompyuterga til qurilishini bilish va tafakkur qilish imkoniyatini beradi.

Foydalanilgan adabiyotlar:

1. O'zbekiston Respublikasi Prezidentining "O'zbek tilining davlat tili sifatidagi nufuzi va mavqeyini tubdan oshirish chora-tadbirlari to'g'risida"gi PF-5850-son Farmoni. <https://lex.uz/docs/4561730>.
2. O'zbekiston Respublikasining 2020-yil 10-apreldagi "O'zbek tili bayrami kunini belgilash to'g'risida"gi O'RQ-615-son Qonuni. <https://lex.uz/docs/4784608>.
3. Мушкова В.В. Лингвистические процессоры и обработка текстов на естественных языках. <https://www.eduherald.ru/ru/article/view?id=18266>
4. Большакова Е.И., Клышинский Э.С. и др. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика: учеб.пособие. – Москва: МИЭМ. 2011. – Б. 91.
5. Марчук Ю.Н. Компьютерная лингвистика: учеб.пособие. – Москва: Восток – Запад. 2007. – Б. 70.

Родственные языки и их особенности при обучении английскому языку

*Вохидова Ситорабону Хурматуллоевна**

Annotation. This article is devoted to the questions of mutual influence and enrichments of English, German, Russian and Tajik languages. Result of these processes is learning relative languages with Venn diagram.

Key words. Venn diagram, symbolical logic, analyze, morphology, phonetics, word order, grammar, similarity, difference.

Английский язык считается самым легким языком во всем мире. Главная причина этого факта в том, что этот язык можно изучать даже за месяц! Существуют несколько методов и техник по которым можно проверить данный факт. В своей практике я часто использую метода диаграммы Венна.

Диаграмма Венна. Впервые прием описан английским ученым Джоном Венном в книге «Символическая логика». Это графический способ, который используется, когда нужно сравнить два или более понятия, явления, способа, предмета. «Кольца Венна» помогают выявить общее в двух или нескольких явлениях, подчеркнуть различия и обобщить знание по заявленной теме [1].

Диаграмму Венна применяю на практических уроках, с той целью чтоб участникам удалось легко и просто изучать новый язык. Создаю атмосферу на котором они чувствовали себя свободно и уверенно. И для этого мне пригодится знание нескольких языков, в особенности, родственных языков, таких как, английский, немецкий, русский и таджикский языки. Все эти языки относятся к индоевропейской языковой семье. К этой индоевропейской языковой семье ученые относят следующие группы нескольких языков:

* преподаватель кафедры «Методика преподавания языков» Бухарского РЦПиПКСНО.