

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

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

**“FILOLOGIYA MASALALARI – YOSH
TADQIQOTCHILAR NIGOHIDA”**

mavzusidagi respublika ilmiy-amaliy anjumani materiallari

2019-yil 14-noyabr

TOSHKENT – 2019

~ 1 ~

<i>Shahnoza Sidiqova</i> O'zbek tilida va ingliz tilida rang-tus bildiruvchi sifatlarning lingvomadaniy xususiyatlari tahlili	75
<i>Dinoraxon To'g'onova</i> The role of metaphor and simile uzbek and english languages and comparison of them	78
<i>Guli Toirova, Aziza Zaripova</i> Nutqiy muloqotning yakuniy bosqichi haqida	81
<i>Solihabonu Toshbotirova</i> O'zbek tilida arab tilidan o'zlashgan so'zlarning o'rni	82
<i>Umida Yusubova</i> Noverbal ma'no ifodalashning milliy shakllari	84
<i>Mohinurbonu Xo'jamova</i> "Alpomish" dostonida dialektik-leksik birliklar tahlili	88
<i>Sayyora Abdurahmonova</i> Matnni morfologik darajada avtomatik tahrir qilish dasturining lingvistik moduli	91
<i>Gulchiroy Muhiddinova</i> O'zbek nutqiy etiketlarining lingvokulturologik tahliliga doir	97
<i>Oydin Nurullayeva</i> Dunyodagi mavjud til korpuslarining yaratilish tamoyillari	100
<i>Mahmud Rajabov</i> Badiiy o'xshatishlar va ularning leksik-grammatik xususiyatlari	103
<i>Rayxonoy Yangiboyeva</i> "Devonu lug'oti-turk va hozirgi o'g'uz shevasida qo'llanuvchi so'zlar talqini	109
<i>Zamira Sarieva</i> Semantic features of personal pronouns in indonesian language	112
<i>Mahfuzaxon Xomidova</i> Badiiy matn tushunilishida intertekstuallik muammosi	113
<i>Maftuna Ahmadjonova</i> "Devonu lug'otit-turk" boshqa xalqlar nigohida	118
<i>Anvarbek Turdaliyev</i> Kompyuter lingvistikasining amaliy masalalari	123
<i>Durdona Bultakova</i> O'zbek tilida argoning o'ziga xos jihatlari haqida	129

O'ZBEK TILI VA ADABIYOTINI O'QITISH METODIKASI

<i>Alisher Ashurov</i> Shaxsga yo'naltirilgan yondashuv asosida talabalarning kommunikativ kompetensiyalarini shakllantirish	134
<i>Odinaxon Saidaxmedova</i> Ona tili va chet tilini o'rganish bosqichlari	137
<i>Antonina Iplina, Indira Nabiyeva</i> Using newspapers and magazines at the english lesson	140
<i>Tog'aymurod Bekmurodov</i> Ideografik lug'atlar va ularning ona tili ta'limida tutgan o'rni	143
<i>Xasanov Sunnatbek, Ergasheva Maftuna</i> O'zbek tilini xorijiy til sifatida o'qitishda gapirishning ahamiyati	146

5. Qoshg'ariy M. Devonu lug'otit - turk (Turkiy so'zlar devoni). Uch tomlik. I, II, III tom(tarjimon va nashrga tayyorlovchi S. M. Mutallibov). – Toshkent: "O'zSSSR Fanlar akademiyasi". 1960 – 1963

KOMPYUTER LINGVISTIKASINING AMALIY MASALALARI

Anvarbek Turdialiyev

*Alisher Navoiy nomidagi ToshDO'TAU 2-kurs magistri
Ilmiy rahbar N. Abdurahmonova*

XX asrning 50-yillaridan boshlab tilshunoslikda "mashina tarjimasi", "mashina tilshunosligi" atamaları qo'llanila boshlandi. Bu asrning buyuk kashfiyoti bo'lgan kompyuter texnologiyalari tilshunoslikka ham kirib kelganining isboti edi. "Mashina tarjimasi yoki avtomatik tarjima deyilganda bir tildagi matni ikkinchi "bir tilga EHM (kompyuter) vositasida, tez vaqt ichida tarjima qilish nazarda tutiladi" [1. 32-36-betlar]. Mashina tarjimasining asoschilari muhandis va matematik olimlar bo'lib, keyinchalik bu ishda tilshunoslar ham faol qatnasha boshlagan. Shu tariqa mashina tarjimasi g'oyalari butun dunyoda nazariy va amaliy tilshunoslikning rivojlanishida katta ahamiyat kasb etdi.

Kompyuter lingvistikasining amaliy masalalari bu bir qancha og'riqli nuqta hisoblanadi, sababi amaliy masalalarni joriy etish uchun sharoit va maxsus laboratoriyalar kerak bo'ladi. Xorijda va O'zbekistonda bu yo'nalishlar bo'yicha bir qancha izlanishlar olib borilgan. Mashina tarjimasi – bir tabiiy tildan ikkinchi tilga (mukammal yozma, og'zaki) matni maxsus kompyuter programmalari yordamida tarjima qilish. Bundan tashqari bu kabi tizimlar ustida ilmiy ishlar olib boruvchi Kompyuter lingvistikasining bir yo'nalishi hisoblanadi.

Mashina tarjimasi (MT) nazariyasi quyidagi ikki bosqichdagi ishlarni o'z ichiga oladi:

- a) bir tildan ikkinchi tilga matni tarjima qilish masalalarini aniqlash;
 - b) tarjima qilingan matni avtomatik tahrir qilish tamoyillarini o'rganish.
- **Lingvistik ta'minot** (matni tahlil va sintez qilish uchun grammatik hamda semantik ma'lumotlarni beruvchi avtomatik lug'atlar);
 - **Matematik-algoritmik ta'minot** (matnga ishlov berish algoritmlari va ma'lumotlarni qayd qilish uchun qoidalar tizimi);
 - **Dasturiy ta'minot** (tarjima algoritmini amalga oshiruvchi va qo'shimcha dasturlar);
 - **Informatsion ta'minot** (tegishli fan sohalariga doir ma'lumotlar bazasi);

Mantiqiy ta'minot (matni fan sohasiga doir ma'lumotlar bilan taqqoslash uchun mantiqiy xulosa usullari).

Mashina tarjimasi keng tushunchani anglatgani bois, hozirda uning to'rt turi mavjud:

Ø mashina yordamidagi inson tarjimasi (machine-assisted human translation –MAHT)-inson tomonidan matnni tarjima qilishda kompyuter elektron lug'atlaridan foydalanish va unga kiritilgan ko'rsatmalaridan foydalanish;

Ø kompyuter yordamidagi tarjima (computer assisted-translation –CAT)-matnni kompyuter vositasida tarjima qilish va unda anglashilgan asosiy tushunchalarni ajratishga xizmat qiladi;

Ø inson yordamidagi mashina tarjimasi (human-assisted machine translation –HAMT)-kompyuter dasturiy ta'minoti orqali amalga oshiriladigan tarjimaga inson resursi tomonidan tahrir qilish vazifasini yuklash;

Ø to'liq avtomatik mashina tarjimasi (fully automatic machine translation –FAMT)-tarjimon dasturi orqali matnni ham tarjima, ham tahrir qilish vazifalarini amalga oshirish.

- Amaldagi mtlari qoidaga ko'ra aniq ikki tilga (masalan, fr-ru yoki yap-ing) yo'naltirilgan bo'lib, xoh yuzaki, xoh ikki tildagi ba'zi qoidalar asosidagi mashina tarjimasi bo'lsin, uning sifati lug'atlar va ma'lumotlar hajmiga, qayd etilgan leksik birliklarga, algoritmning analiz va sintez jarayoni puxta va xatosiz ishlashiga hamda dasturiy ta'minotning samaradorligiga bog'liq. Zamonaviy uskunalar va dasturiy vositalar mukammal Grammatik ma'lumotlarni o'zida jamlagan katta hajmdagi lug'atlardan foydalanish imkonini beradi. Ma'lumotlar deklarativ (tavsiflangan) va protsedura (hisoblangan) shakllarda bo'lishi mumkin.

- Tarjima amaliyoti va axborot texnologiyalarida mtning ikkita asosiy yo'nalishi mavjud. Bir tomondan, mt natijalari orqali notanish tildagi ma'lumot mazmuni bilan yuzaki bo'lsada tanishib chiqish mumkin. Bu holda tarjima ishoraviy ma'lumot sifatida qaralib, puxta tahrirni talab etmaydi. Boshqa tomondan esa mt oddiy inson tomonidan qilingan tarjima o'rniga foydalaniladi va u puxta tahrirni hamda muayyan fan sohasiga mt ni sozlashni talab etadi. Bunda lug'at boyligi, mazmun va til birliklarining tarjima matniga yo'nalganligi, leksik ko'p ma'noviylik masalalarini unumli hal qilinishi, gr ma'lumotlarni ajratishda algoritmning samarali ishlashi, tarjimaning bir-biriga mosligi muhim ahamiyatga ega. Bu turdagi tarjima amaliyoti tarjima hajmi yetarlicha katta bo'lganida (yiliga bir necha o'n ming sahifadan kam bo'lmaganda), matnlar o'xshash, lug'at tizimlari to'liq, keyingi kengaytirishga yo'l qo'yib, dasturiy ta'minoti tarjimadan keying tahrirga qulay bo'lganda iqtisodiy tejamkor hisoblanadi. Mtdagi shuncha ixtirolarga qaramay, barcha tarjima tizimlari faqat yuzaki sintaktik darajadagi tarjimani amalga oshira oladi. Chunki hali barcha ko'rinishdagi fikrning formal modeli samarali ishlab chiqilmagan. Mutaxassislar mtning mos tushuvchi tarjima qurilishini sun'iy intellekt bilan bog'laydilar: mashina bir tildan boshqa tilga insondek fikrlashni o'rgangandagina tarjima qila oladi.

- Zamonaviy mtt tanqidchilari taxminicha, amaliyotda janrlar bo'yicha chegaralash (mashinaga avval juda oson maxsus tanlangan matnlarni o'rgatish) o'ratilganligi shunga olib

keldiki, tabiiy tilni modellashtirish masalasi haqiqatdan ham alohida chegaralangan (ilgarigi) tilosti bilim sohalarini modellashtirishga ham yo'l ochib berdi. Bu yo'lda kanadaning mashhur TAUM-METEO tizimi ing- frcha ob-havo ma'lumoti tarjimasini a'lo darajada bajarib eng yaxshi natija ko'rsatdi. Bu turdagi oddiy ko'rinishdagi tizimlar turistlar uchun foydalanuvchiga ikki yoki bir nechta tilda birmuncha har turli "menyu" bo'yicha standart savol-javoblarni tavsiya etuvchi avtomatik so'zlashgich hisoblanadi.

- Tm- bu bajarilgan tarjimalar saqlanuvchi ma'lumotlar ombori. Tm texnologiyasi jang'arib borish prinsipi bo'yicha ishlaydi, ya'ni tarjima jarayonida dastlabki ma'lumotlar va uning tarjimasini tmda saqlanib qoladi. Tarjimaga berilgan yangi matn bilan ishlash jarayonida tizim har bir gapni bazadagi ma'lumotlar bilan taqqoslaydi. Agar aynan bir xil yoki o'xshash birliklar topilsa, bu birliklar tarjimasini va tasodiflar ko'rsatkichi foizlarda tarjima bilan birgalikda ko'rsatiladi. Har bir o'zgarish yoki yangi tarjima tmda saqlanadi. Buning natijasida bitta gapni ikki marta tarjima qilishga hojat qolmaydi.

- Avtomatik tarjima tizimi (CAT, computer-aided/assisted translation) dasturlari tm bilan ta'minlanganligiga qaramay, ularni mt dasturlari (machine translation) bn adashtirmaslik lozim: tm hech nimani o'zi tarjima qilmaydi, mt esa berilgan matnning grammatik tahlili natijalariga ko'ra tarjimani keltirib chiqarishga asoslangan.

- Mt texnologiyasi bosqichma-bosqich takomillashib bormoqda. Agar ilk tarjima tizimlari sifati va tezligi qoniqarli darajada bo'lmagan bo'lsa, hozirda kompyuter haqiqatdan ham bir tildagi matnni boshqa tilga mazmunini kelishtirib tarjima qila oladi. Yanada zamonaviy eng yuqori sifatli dasturlar esa bir daqiqada bir sahifani tarjima qilib berishi mumkin. Bu esa vaqtni ham, mablag'ni ham tejaydi.

- Ushbu ma'lumotlarni hisobga olgan holda, mashina tarjimasini dasturlari ko'pincha mahalliy tilda gapiradigan kishi uchun boshqa tilda yozilgan ma'lumotlarning taxminiy ma'nosini bilib olish uchun etarlicha yaxshi ishlaydi. Qiyinchiligi muayyan usulni qo'llab - quvvatlash uchun to'g'ri turdagi ma'lumotlarga ega bo'lish kerak. Masalan, statistik metodlar uchun zarur bo'lgan katta ko'p tilli ma'lumotlar korpusining grammatikaga asoslangan usullari uchun zaruriyat yo'q. Ammo ular ishlatadigan grammatikani diqqat bilan loyihalash uchun malakali tilshunos kerak. Bir-biriga yaqin bo'lgan tillarni tarjima qilish uchun qoidaga asoslangan MT deb nomlanadigan texnik usulni qo'llash mumkin.

Xulosa o'rinda shuni aytish mumkinki, barcha tarjimonlar uchun umumiy bo'lgan tx uchun yagona ma'lumotlar bazasini tuzish lozim. Aks holda hamma alohida bir gapni qayta-qayta tarjima qilishi, bir ishni takroran bajarishi mumkin. Buni esa shu yo'l bilan hal qilish mumkin. Yana bir boshqa saytda esa tarjimonlar tomonidan hali tarjima qilinmagan (ishlanmagan) narsa (ish) lar berib borilishi kerak.

Ma'lumki, leksikografiya tilshunoslik fanining amaliy sohalaridan biri hisoblanadi. U yunoncha «lexikos» - «so'z», «lug'at», «grapho» - yozaman degan ma'nolarni anglatadi. Leksikografiya bo'limida lug'atlar, ularni tuzish tamoyillari o'rganiladi. Avtomatizatsiya ijtimoiy hayotning barcha jabhalariga kirib kelishi natijasida lug'atlar ham elektron holatda yaratilish imkoniyati yuzaga keldi. Kompyuter yordamida lug'atlar bilan ishlashning optimallasuvi natijasida kompyuter leksikografiyasi yo'nalishi shakllandi. "Kompyuter leksikografiyasi amaliy tilshunoslikning muhim tarkibiy qismi bo'lib, unda lingvistik va dasturiy ta'minot asosida ishlaydigan, kodlash va dekodlash prinsipi asosida yaratilgan kompyuter lug'atlari, ularni tuzish dasturlari, algoritmlari o'rganiladi. Mazkur sohaning rivojlanishi natijasida turli nomlarda elektron lug'atlar yaratilmoqda. Eng mashhur elektron lug'atlar sirasiga CONTEXT, ABBY LINGVO, MULTI-TRAN, POLYGLOSSUM, MULTILEKS kabilari kiradi" [2. 75-77-betlar].

Leksikografiya sohasiga kompyuter texnologiyalarining tatbiq etilishi lug'at tuzish ishlarini ancha yengillashtiradi. Lug'at tuzish jarayonining naqadar murakkabligini bir tilshunos olim hazilomuz tarzda quyidagicha izohlagan: *«Agar biror kishi og'ir jinoyat sodir qilsa, uni katorgaga surgun qilish shart emas. U lug'at tuzsin, shunda u katorganing barcha azoblarini, qiyinchiliklarini tortgan bo'ladi»*. Axborot manbalarini yig'ish bo'yicha maxsus dasturlar (Database Software) lug'at asosini tashkil etuvchi barcha ma'lumot va misollarni jamlash va sistematik tarzda ishlov berishga qulaylik tug'diradi. Bundan tashqari, boshqa maxsus dasturlar lug'atni tahrir qilish va chop etish borasidagi ancha mashaqqatli mehnatni bir qancha osonlashtiradi.

Elektron lug'atlar ham, o'z navbatida, an'anaviy lug'atlarga qaraganda birmuncha afzalliklarga ega. Bugungi kunda, masalan, ABBY LINGVO kompaniyasining lug'at va tarjima dasturlari ulardan foydalanuvchi har bir xaridorga o'z lug'atini tuzish yoki mavjud lug'at so'z boyligini to'ldirib borish imkonini beradi. Ochiq turdagi Internet lug'atlari (online lug'atlar) ham ko'p hollarda barcha foydalanuvchilarga ushbu lug'atlarni to'ldirish imkonini beradi. Shuningdek, elektron lug'atlarning hajm jihatidan ixchamligi, kompaktiligi, boshqa manbalarga (Internet tarmog'i orqali, giper-murojaatlar yordamida) ulanish imkoniyati, mavjud matnlar korpuslari yordamida ilustrativ misollar tuza olish imkoniyati, multimediali misollar bilan boyitilganligi, so'zlarning sinonimik variantlari, omonimlik xususiyatlari, grammatik ma'lumotlar bilan ta'minlanganligi va ularning tezlik bilan foydalanuvchiga havola etilishi kabi jihatlari elektron lug'atlarning optimalligini ta'minlaydi [3. 180-bet].

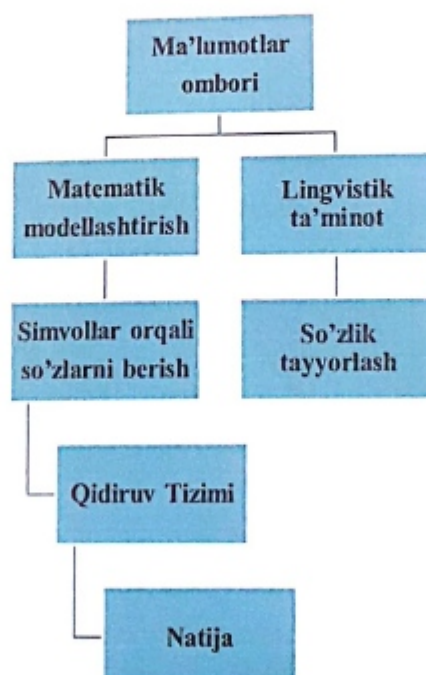
Kompyuter leksikografiyasini elektron matnlar korpusi yoki parallel matnlar korpuslarisiz tasavvur qilish mumkin emas. «Matnlar korpusi («corpus» lotincha «tana» degan ma'noni anglatadi) 1 bu elektron holda saqlanadigan ma'lum til birliklari bo'lib, ular tilshunoslar uchun turli xil muammolarni hal etish uchun tatbiq etishda va turli yo'nalishdagi tadqiqotlar uchun

zaruriyatga qarab turli shakllarda tuziladi. Bular fonema, grafema, morfemalardan tortib undan kattaroq birliklar - leksema, gap va matnlardan (badiiy yoki ilmiy asar, gazeta va jurnal matnlari) tashkil topishi mumkin» [4. 143-bet]. Ularning qay tarzda saqlanishiga qarab maxsus dasturlar yordamida har bir kerakli soʻz yoki soʻz birikmasi uchun darhol uning qoʻlanishi boʻyicha misollar topilishi, imlo boʻyicha variantlari, sinonimik qatorlari topilishi mumkin. Matnlar korpusiga oid ilmiy tadqiqotlar salmogʻining koʻpayishi natijasida tilshunoslikda korpus lingvistikasi yoʻnalishi shakllandi.

Kompyuterda yaratilgan «birinchi matnlar korpusi Braun korpusi (БК, inglizcha *Brown Corpus*, BC) hisoblanadi, u 1961-yilda Braun universitetida yaratilgan, har biri 2000 soʻzli 500 ta matn fragmentini oʻz ichiga oladi. 1970-yillarda 1 mln soʻzni oʻz ichiga olgan matnlar korpusi asosida rus tilirning chfastotali lugʻati yaratildi » [2. 75-77-bet]. 1980-yillarda Shvetsiyaning Upsala uni- versitetida ham rus tilida matnlar korpusi yaratildi. Keyinchalik kompyuter leksikografiyasining rivojlanishi natijasida katta hajmli matnlar korpusiga ehtiyoj tugʻildi. Yaʼni 1 mln ta soʻz elektron lugʻatlar bazasi uchun etarli emas. Shu asosda yirik hajmli matnlar korpusi yaratila boshlandi.

Kompyuter leksikografiyasi quyidagi vazifalar uchun kodlashtirilgan:

- Lugʻat tamoyilidan kelib chiqqan holda maʼlumotlar omborini tayyorlash – tarjima boʻlsa, qaysi til nazarda tutilsa, oʻsha tildagi muqobil variantini bir satrga joylash. Agar izohli boʻlsa, leksemaning izohini bir satrga joylash;
- Qidiruv tizimida ikki yoqlama munosabatdordlikdagi halqani yaratish: Tarjimada nechta til boʻlsa, oʻsha tildagi soʻzlar bilan tarjimada beriladigan soʻzlar orqali topa olish tizimi yoki izohlar orqali soʻzni, soʻz orqali izohni topish;
- Sabab-natija-sabab munosabat tizimini yaratish – sabab – berilayotgan soʻzlik, natija – tarjima variant



• Ushbu ma'lumotlarni hisobga olgan holda, mashina tarjimasi dasturlari ko'pincha mahalliy tilda gapiradigan kishi uchun boshqa tilda yozilgan ma'lumotlarning taxminiy ma'nosini bilib olish uchun etarlicha yaxshi ishlaydi. Qiyinchiligi muayyan usulni qo'llab - quvvatlash uchun to'g'ri turdagi ma'lumotlarga ega bo'lish kerak. Masalan, statistik metodlar uchun zarur bo'lgan katta ko'p tilli ma'lumotlar korpusining grammatikaga asoslangan usullari uchun zaruriyat yo'q. Ammo ular ishlatadigan grammatikani diqqat bilan loyihalash uchun malakali tilshunos kerak. Bir-biriga yaqin bo'lgan tillarni tarjima qilish uchun qoidaga asoslangan MT deb nomlanadigan texnik usulni qo'llash mumkin.

Xulosa o'rinda shuni aytish mumkinki, barcha tarjimonlar uchun umumiy bo'lgan tx uchun yagona ma'lumotlar bazasini tuzish lozim. Aks holda hamma alohida bir gapni qayta-qayta tarjima qilishi, bir ishni takroran bajarishi mumkin. Buni esa shu yo'l bilan hal qilish mumkin. Yana bir boshqa saytda esa tarjimonlar tomonidan hali tarjima qilinmagan (ishlanmagan) narsa (ish) lar berib borilishi kerak.

Foydalanilgan adabiyotlar:

1. Po'latov A., Muhamedova S.X. Kompyuter tilshunosligida avtomatik tahrir qiluvchi dasturning lingvistik ta'minotini yaratish asoslari // O'zbek tilshunosligining dolzarb masalalar. – Toshkent, TDPU ilmiy to'plami, 2003.
2. Rahimov A. Kompyuter lingvistikasi asoslari. – Toshkent: Akademnashr, 2011.
3. Abdurahmonova N., Inglizcha matnlarni o'zbek tiliga tarjima qilish dasturining lingvistik ta'minoti
4. (sodda gaplar misolida) Monografiya Toshkent, 2018.
5. Pulatov A., Muhamedova S. Kompyuter lingvistikasi (o'quv qo'llanma) – Toshkent, 2014; Muhamedova S. O'zbek tilidagi harakat fe'llari asosida kompyuter dasturlari uchun lingvistik taxmin yaratish. – Toshkent, 2006.