Philology Matters / ISSN: 1994 – 4233
Uzbekistan State World Languages University

2022 Vol. 41 No. 2
THEORY AND PRACTICE OF TRANSLATION
DOI: 10.36078/987654707

Received: April 29, 2022
Accepted: June 15, 2022
Available online: June 25, 2022

Zahriddin Xaitqulov

O'zbekiston davlat jahon tillari universiteti
tayanch doktoranti
Toshkent, O'zbekiston
E-mail: haitkulovzakhriddin@gmail.com

TARJIMA QILISHGA O'RGATISHDA ELEKTRON KORPUSLARDAN FOYDALANISH USULLARI

ANNOTATSIYA

Bugungi kunda inson faoliyatining deyarli barcha sohalariga kirib ulgurgan, global axborot makonini tashkil etuvchi kompyuter texnologiyalari muhim o'rin tutadi. Har bir soha rivoji uchun kompyuter texnologiyalarining qo'llanilishi bugun juda dolzarb va muhimdir. Xuddi boshqa sohalarning kompyuter tizimida dasturlar, platformalar, dasturiy ta'minotlar, korxona va tashkilotlarning elektron boshqarish bazalari singari tilshunoslik va tarjimashunoslik sohasida ham kompyuterning o'rni sezilarli darajada oshdi. Tarjimani o'rgatishda korpusga asoslangan yondashuv so'nggi o'n yil ichida jadal rivojlandi, hozirda parallel korpus, ko'p tilli korpus (multilingual corpus) va taqqoslanadigan korpuslardan (comparable corpus) foydalangan holda tadqiqotlar natijasida ko'plab ma'lumotlar paydo bo'ldi. Korpora, ayniqsa, parallel korpus, tarjima tadqiqotlari va ta'lim sohasida qimmatli ma'lumot manbalari ekanligini isbotlaydi. Shuni ham ta'kidlash mumkinki, oxirgi yigirma yil ichida korpus va korpus tahlilining dasturiy ta'minoti tarjima bo'yicha treningda ma'lumotlarga boy o'quv muhitini ta'minlash uchun ko'p adabiyotlarda, turli xil kasblarda tarjima sifatini yaxshilash uchun samarali vosita va metodologiya sifatida taklif qilingan. Tarjima vazifalarini amalga oshirish uchun korpus dasturining talablari akademik soha ehtiyojlaridan mutlaqo farq qiladi va shuning uchun dasturiy ta'minotni loyihalashda alohida e'tiborga olinishi kerak. Tarjimonning korpus vositasidan foydalanishi qulay bo'lishi, korpus ma'lumotlarini yaratishni osonlashtirishi va tarjimonlar o'zlarining hujjatlashtirish faoliyati uchun odatlangan vositalarga o'xshash bo'lishi muhim hisoblanadi.

Zakhriddin Khaitkulov

PhD Student, Uzbekistan State World
Languages University
Tashkent, Uzbekistan

METHODS OF USING ELECTRONIC CORPORA IN TEACHING TRANSLATION

ANNOTATION

In modern society, an important place is occupied by computer technologies that have penetrated into the sphere of human activity and constitute the global information space. Nowadays the use of computer technologies for the development of any field is considered very relevant and important. Along with programs, platforms, software, electronic databases for the management of enterprises and organizations in the computer system of other areas, the role of computers in the field of linguistics and translation studies has increased significantly. This corpora-based approach to teaching translation has evolved rapidly over the past decade, and now a wealth of evidence has emerged from research using parallel corpora, multilingual corpora, and comparable corpora. Corpora, especially parallel corpora, have proven to be valuable sources of information in translation research and education. We can also note that over the past two decades, corpus and corpus analysis software has been proposed as an effective tool and methodology for improving the quality of translation in various professions in many literatures to provide an information-rich learning environment in teaching translation. The requirements for corpus software for translation tasks are very different from the needs of the academic field and therefore must be taken into account when developing software. It is important that the corpus tool be easy for a translator to use, to facilitate the creation of corpus data, and be similar to the tools translators use in their documentation work.

This author of the article discusses the use of electronic corpus methods in teaching translation and creating translation programs. The role of

E-ISSN: 2181 – 1237

Ushbu maqolada tarjimani o‘rgatishda va tarjima dasturlarini yaratishda elektron korpus usullaridan foydalanish turlari o‘rganilgan. Kompyuter lingvistikasi va korpus tilshunosligida tarjimaning o‘rni va uning boshqa sohalar bilan bog‘liqlik jihatlari yoritilgan.

Kalit so‘zlar: kompyuter lingvistikasi, elektron korpus, korpus tilshunosligi, mashina tarjimasi, tarjima, korpus turlari, matnlar korpusi, elektron tarjimon dasturlari, tarjimani avtomatlashtirsh, tarjima jarayonini qo‘llab turuvchi platformalar, tarjima xotirasi.

translation in computational linguistics and corpus linguistics, as well as its connection with other areas, is emphasized.

Key words: computational linguistics, electronic corpus, corpus linguistics, machine translation, corpus types, text corpus, electronic translation software, translation automation, translation support platforms, translation memory.

KIRISH

1980-yillardan boshlangan korpus lingvistikasi bugun yanada ommalashib bormoqda. Korpus lingvistikasi nima? Nega keng miqyosda dolzarblashib bormoqda? U qanday qo‘llaniladi va tarjima uchun qanday imkoniyatlar yaratadi? Bu savollarga Maykl Stubbsning nazariy qarashlaridan javob topish mumkin. Maykl Stubbsning fikricha, “korpus tilshunosligi chastota va tipiklik, misol va norma o‘rtasidagi munosabatlarni o‘rganadi, ushbu nazariya amaliyotda tasdiqlangan, ammo g‘ayrioddiy narsani talqin qilish uchun asos bo‘lishi kerakligi bois u tipik nazariyaga qaratilgan [Stubbs M., 2001; 151]. Sodda qilib aytganda, korpus lingvistikasi lingvistik hodisalarni mashinada o‘qiladigan katta hajmdagi matnlar to‘plami orqali o‘rganis demakdir. Korpus yozma yoki og‘zaki shakldagi har qanday matnni o‘zida jamlaydi. U tilning haqiqatda qanday qo‘llanilishining empirik dalillariga asoslanadi. Shunday qilib, korpus lingvistikasi leksikografiya va lug‘at dizayni, grammatika hamda ingliz va boshqa tillarni turli metodlar orqali o‘rgatish, shuningdek, ikkinchi tilni o‘zlashtirishda katta ahamiyat kasb etadi” [Stubbs M., 2001; 152–153].

Shu nuqtai nazardan hozirgi kunda korpus lingvistikasi juda muhim va dolzarb masala bo‘lib kelmoqda. Qolaversa, korpus lingvistikasi tarjimashunoslik sohalari o‘rtasida ham ko‘plab yangi imkoniyat va tadqiqotlar olib borishga yo‘l ochib bermoqda. Bugun zamonaviy kompyuterlar, texnik tarjimon, yuridik tarjimon yoki boshqa sohalardagi tarjimonlar mavjudligiga qaramay, tarjimon uchun zamonaviy kompyuter eng muhim vosita hisoblanadi. Tarjima jarayonida elektron formatdagi ensiklopediyalar bilan birga bir tilli va ikki tilli lug‘atlar asosiy yordamchi hisoblanadi. Mashinada o‘qiladigan korpuslar boshqa turdagi lug‘atlarga nisbatan bir qancha afzalliklarga ega. Birinchidan, mashinada o‘qiladigan korpuslarni oddiy tarzda qidirish mumkin. Ikkinchidan, korpusdagi ma’lumotlarni kattalashtirish va qo‘shimcha ma’lumotlar bilan boyitish mumkin.

Axborot asrida odamlar katta hajmdagi elektron ma’lumotlarga tez-tez duch kela boshladilar. Ma’lumotlar esa asosan, elektron matnga asoslanadi. Ayrim sohalarda hanuzgacha strukturaviy ma’lumotlar bazalari va yarim tizimli veb-sahifalar, shuningdek, ko‘plab tuzilmaga ega elektron matnlar mavjud. Katta hajmdagi elektron ma’lumotlarga qanday munosabatda bo‘lish insonning raqamli mavjudligining eng real muammolaridan biriga aylandi. Bunday qiyinchilik va imkoniyatlarga duch kelgan

turli fanlar muammolarni o‘z nuqtai nazaridan hal qiladi va shu orqali nazariy tizimini boyitadi va takomillashtiradi. Shunga qaramay, odamlar yangi intizom yo‘nalishlarini ishlab chiqish uchun yangi rivojlanish bosqichida taqdim etilgan shartlardan ham foydalanishlari mumkin. Shuni ham aytish kerakki, korpus lingvistikasi ana shunday davrda paydo bo‘lgan, rivojlanayotgan tilshunoslik fanidir [Di Wu, 2022; 1].

Sun’iy intellekt, internet va boshqa zamonaviy axborot texnologiyalarining rivojlanishi yoki konsepsiyasi bilan hayot, ish va o‘qish zamonaviy axborot texnologiyalari bilan tobora ko‘proq integratsiyalashmoqda va odamlar kompyuter texnologiyalari va sun’iy intellekt texnologiyasiga qaram bo‘lib qolmoqda [Zhijie Z., 2017; 241-246]. Kompyuter yordamidagi tarjima texnologiyasi (CAT), bu – tarjimonga asoslangan mashina va kompyuter tarjimasi bilan yordam beradigan tilni tarjima qilishning yangi turidir. Dunyoda kompyuter yordamida tarjima texnologiyasining rivojlanishi qisqa vaqtni o‘z ichiga oladi [Gao T., Wang X., 2020; 32-59]. Mashina tarjimasi ilk bor ishlab chiqilganda, algoritmlarning cheklanishi va boshqa shart-sharoitlar tufayli tarjimonlarga faqat ma’lum darajada tushuncha va ma’lumot berish imkoniyatiga ega bo‘lgan, biroq mashina tarjimasi asosli, ravon va ideal tarjimalar yarata olmadi.

Mashinali tarjima tizimining (to‘liq avtomatik) tarjima aniqligi uzoq vaqt davomida 70% atrofida qoldi. Tarjimaning o‘qilishi, tizimning til hodisalarini qamrab olishi, tizimning mustahkamligi, ayniqsa, ochiqligi deyarli qoniqarli emas. Insoniyat real matnlarni (ayniqsa, ommaviy onlayn matnlarni) keng miqyosda qayta ishlashga muhtoj va bugungi jamiyatning mashinali tarjima tizimlariga, keng ko‘lamli real matnni qayta ishlashga bo‘lgan ehtiyoji ortib bormoqda. Kompyuter yordamida tarjima qilish (CAT) g‘oyasi aynan shu maqsadda ixtiro qilingan. To‘liq avtomatik mashina tarjimasi tizimi bilan taqqoslaganda, mashina yordamida tarjima tizimi inson–mashina interaktiv tizimdir [Frerot C., 36-61]. Ushbu tarjima jarayonida kompyuter tarjimonga yordam berish vazifasini o‘taydi. U tarjimonga nafaqat lug‘at, terminologiya va iboralar tarjimasi bo‘yicha muayyan bilimlarni beradi, balki tarjimonning muammoni hal qilishi uchun tarjima qilingan matndan bir xil yoki o‘xshash jumlalar tarjimasini izlaydi. Shu asnoda tarjimaning samarali chiqish ehtimoli ortadi. Kompyuter yordamida tarjima qilishning ilg‘or g‘oyasi (jumladan, tarjima xotirasi texnologiyasiga asoslangan tarjima xotirasi texnologiyasi va taxminiy misol rejimiga asoslangan tarjima texnologiyasi) tarjima xotirasida bir xil yoki o‘xshash jumla yoki iborani (ikki tilli moslashtirish kutubxonasi) va misol rejimida qidirish kabi jarayonlarni amalga oshiradi. Shu tariqa tarjimonlar va tarjimadan foydalanuvchilar mavjud tarjima ma’lumotlaridan to‘liq va samarali foydalanadilar. Ushbu yordamchi tarjima mexanizmi, ayniqsa, uzoq va tez-tez takrorlanadigan til hodisalari bo‘lgan ilmiy va texnologik monografiyalar, ilmiy va texnologik adabiyotlar, mahsulot qo‘llanmalari, o‘quv dasturiy qo‘llanmalar, Birlashgan Millatlar Tashkiloti hujjatlari va boshqalar kabi matnlarni tarjima qilish uchun juda mos keladi [Bingbing P., 2022; 72].

Tarjimani kompyuterlashtirish ya’ni inson tomonidan bajariladigan tarjimani avtomatlashtirish, mashina tarjimasi, axborot texnologiyalari fanda qo‘llanilishining boshidanoq eng muhim muammolardan biri sifatida belgilangan. Avtomatik mashina

tarjimasini yaratish orzusi olimlarning ulkan g‘oyalaridan biri edi va ushbu g‘oyani amalga oshirishlari esa olimlarning eng katta yutug‘i hisoblanadi. Bugungi kunga kelib mashina tarjimasining ixtiro qilinganligi barcha sohalarda qo‘llanib kelinmoqda. Buning natijasida esa tarjima sohasiga bo‘lgan talablar keskin ortib bormoqda. Qolaversa, ushbu muvaffaqiyatlar, inson tafakkurini va ushbu jarayonlarni muhandislik-lingvistik modellashtirishning og‘zaki muloqot qobiliyatlarini o‘rganishdagi yutuqlarga sabab bo‘lmoqda. Hozirgi kunda tarjimonning ishini shaxsiy kompyutersiz tasavvur qilib bo‘lmaydi. Undan haqiqiy tarjima va tegishli muammolarni hal qilish uchun, masalan, asosiy ma’lumotlarni qidirish va terminologiyani o‘rganish uchun ham foydalaniladi. Ma’lumki, tarjima inson intellektual faoliyatining murakkab shakli, manba tilidan ko‘zlangan tilga o‘tishdir. Tarjima ko‘p qirrali murakkab hodisa va uning ayrim jihatlari turli fanlarning o‘rganish predmeti bilan bog‘liq. Ya’ni tarjima amaliyoti madaniyatlararo muloqot jarayoni bo‘lib, unda manba matni tarjimasini tahlil qilish asosida yangi lingvistik va madaniy muhitda asl nusxani almashtirish uchun ikkinchi darajali tarjima matn yaratiladi. Ushbu jarayon bevosita tarjimaning boshqa fan va sohalar bilan uzviy bog‘liq jihatlarini namoyon etadi.

Tarjima jarayonida tarjimon lingvistik tildan tashqari boshqa bilimlardan ham foydalanadi va bu jarayon bir-biridan tubdan farq qiluvchi ikki xil bosqichni o‘z ichiga oladi. Mazkur bosqichlar quyidagilardir: a) *asl tildagi matnni tushunish*, b) *ko‘zlangan tildagi matnni sintez qilish*. Tarjima jarayonining murakkabligi tufayli u haqidagi fan (Tarjimashunoslik) fanlararo xususiyatga ega va u tilshunoslik, adabiyotshunoslik, kognitiv fan va madaniyat antropologiyasi bilan bog‘liq.

Tarjima masalasi doimo tilshunoslarning diqqatini tortgan va hozir ham eng muhim masalalardan biri hisoblanadi. Tarjimaning ayrim qiyinchiliklarini yengish uchun kompyuter texnikasining yutuqlaridan foydalaniladi. Hozirda kompyuter texnologiyalarining rivojlanishi bilan tarjimonlar tomonidan qo‘llaniladigan vositalar soni bir qancha oshdi.

Tarjimani o‘rgatishda korpusning imkoniyatlari. Hozirgi davrda elektron lug‘atlar, ensiklopediyalar, ma’lumotnomalar kabi manbalar qog‘ozga o‘xshash manbalar o‘rnini egalladi. Raqamlashtirilgan resurslarning asosiy afzalligi ular bilan ishlashni soddalashtirish va tezlashtirishdir, bu, ayniqsa, tilshunoslar va tarjimonlar uchun resurslarni jadal rivojlantirishga yordam beradi. Ma’lumki, leksik noaniqlikni hal qilish so‘zlarni qayta ishlashning markaziy vazifalaridan biridir. Maqsad – so‘zlar yoki qo‘shma atamalarning ma’nosini ular ishlatilgan kontekstga muvofiq ochishdir. Leksik noaniqlikni hal qilish matnlarni tasniflash va klasterlash usullarining aniqligini oshirish, ma’lumotlarni qidirish va boshqa ilovalarning tarjima sifatini oshirish uchun ishlatiladi. Muammoni hal qilish uchun so‘zlarning taxminiy varianti bo‘lgan ma’nolarini va bu ma’nolar bilan so‘zlar qo‘llanilgan kontekst o‘rtasidagi munosabatni aniqlash kerak. Tilshunoslar tarjimani taxminiy va aniq ma’nolar o‘rtasidagi aloqani o‘rnatish uchun tezaurslar, semantik tarmoqlar va boshqa maxsus tuzilmalarni tuzadilar. Biroq, bunday resurslarni yaratish juda katta kuch talab qiladi.

Matematik, tilshunos va tarjimon olimlar tarjimaning aniq va sifatli kontekstdagi ma’nosini foydalanuvchilar tomonidan qabul qilishlari uchun ko‘proq qulaylik yaratish

maqsadida maxsus korpuslar yaratish muhimligini olg‘a surganlar. Buning natijasida tarjima bilan bog‘liq muammolarni sezilarli darajada yengish ko‘zda tutilgan. Shu bilan birga korpuslar nafaqat tarjima sohasida, balki tilshunoslik va boshqa sohalarda ham butkul o‘zgartirilishi olimlarning diqqat markazida bo‘lgan.

Hozirgi kunda tarjima jarayonini o‘rganish maqsadida bo‘lgan korpus lingvistikasiga muvofiq ko‘plab ilmiy tajribalar bir qancah olimlar tomonidan olib borilmoqda. Jumladan, M.Uilkinson, V.N. Shevchuk, N.V. Vladimova va R.K. Koshkin elektron matnlar korpusi tarjima qilingan nutqning “noto‘g‘riligi” bilan bog‘liq “salbiy ta’sir” to‘plamiga olib kelgan omillarni aniqlash va hal qilish vositasi sifatida nazariy fikrlarni bildirganlar [Shevchuk V.N., 2010; 40-70].

O‘zbek tilshunoslaridan tadqiqotchi A.Ahmedova o‘z ilmiy izlanishida kompyuter lingvistikasi va korpus lingvistikasi bir-biridan ajralmas soha ekanligini boshqa olimlar qatori ta’kidlaydi [Ahmedova A., 2022; 17]. Kompyuter tilshunosligi va korpus lingvistikasi o‘zaro quyidagi farqli jihatlarga ega: “Kompyuter tilshunosligi korpus lingvistikasi (KL) uchun vositalarni (ya’ni dasturlarni) yaratadi. Shu ma’noda, ular bir-birini to‘ldiradi. Masalan, KL korpusda so‘zlarni avtomatik tarzda belgilash uchun vositalarga muhtoj. Agar 100 yoki 200 million so‘zli matn mavjud bo‘lsa, har bir so‘zning teglash jarayonini to‘laligicha qo‘lda amalga oshirib bo‘lmaydi, shu bois, maxsus dasturiy ta’minot talab qilinadi. Aynan mazkur dasturiy ta’minotni kompyuter lingvistikasi taqdim etadi, lekin bu – KL uning bir bo‘lagidir, degan tushunchani keltirib chiqarmaydi” [Ataboyev N., 2020; 19].

N.Ataboyev o‘z tadqiqotlarida ingliz tilidagi korpuslarning funksional xususiyatlarini COCA korpusi misolida o‘rganiish, korpus tahlil natijalariga asoslangan holda korpus lingvistikasining chastotali terminologik lug‘atini yaratishga doir fikrlarini taqdim etgan [Ataboyev N., 2020; 20].

N.Ataboyev korpuslar tasnifini maxsus printsiplarga asoslangan holda keltirish mumkinligi va ularni ma’lum ma’noda qo’llash sohalariga ko’ra turkumlash shartli ravishda yakuniy shaklga keltirilganligi va mavjud terminlardan foydalanilganligini ko‘rib ciqqan. Unga ko‘ra, lisoniy korpuslarni qo’llash mumkin bo’lgan sohalar qatori quyidagilardan tashkil topgan: A) tarjima amaliyoti va chog’ishtirma lingvistika: *maxsus/umumiy korpus; qiyosiy korpus; bir/ko’p tilli korpus; og’zaki/ yozma korpus; parallel korpus va h.k.*; B) leksikografiya: *maxsus/umumiy korpus; havolali korpus; og’zaki/yozma korpus; qo’sh korpus; bir/ko’p tilli korpus va h.k.*; C) adabiyot nazariyasida: *parallel korpus; mualliflik korpusi; adabiy korpus va h.k.* D) lingvodidaktika va tillarni o’qitish: *maxsus/ umumiy korpus; o’rganuvchi/ pedagog korpusi; kuzatuv korpusi; og’zaki/yozma korpus va h.k.*; E) etimologiya va til tarixi tadqiqi: *maxsus/ umumiy korpus; tarixiy/zamonaviy; kuzatuv korpusi; hududiy korpus va h.k.*. Bu ro’yxatni korpus tahlillarga tayanuvchi barcha til xususidagi nazariy va amaliy tadqiqot sohalari bilan davom ettirish mumkinligini olim o‘z nazariy qarashlarida e’tirof etgan [Ataboyev N., 2020; 14].

Mualliflik korpusini yaratishning genderga oid metaforalarning kontseptual tadqiqini G.Ergasheva Britaniya milliy korpusi asosida amalga oshirgan [Ergasheva G., 2018; 61]. J.Jumabaeva tildagi graduonimiyalarni ingliz tili korpusi yordamida

o’rganib, semantik xususiyatlarini aniqlagan. Mualliflik korpusini yaratishning lingvistik asoslari Sh. Xamroeva tomonidan tadqiq etilgan [Xamroeva Sh., 2018; 52]. Shuningdek, korpus va korpus tahlillari amaliyoti, korpus lingvistikasining nazariy kontseptologik asoslari bo’yicha G.Lich, K.M. Aulif, M.Malberg, P.Milin, A.Murakami, P.Tompson, B.Vinter, V.Zaxarov, V.Martin, S.Kubler, A.Laurens, E.Etuell, Boizou, M.Keneri, J.Grafmiller, J.Grieve, N.Grum, S.Hansson, R.Peych, A.Schembri kabi olimlar tadqiqotlarida kuzatiladi [Hunston S., 2002; 234].

TADQIQOT USULLARI

Kompyuter qidiruv tizimi yordamida boshqariladigan katta hajmdagi to‘plam korpus hisoblanadi. Dastlab korpus namunasi XIII asrda Injilda amalga oshirilgan. Uning har bir sahifasidagi so‘zlar satr bo‘yicha qo‘l mehnati bilan indekslangan. Injilda keltirilgan so‘zlarning parchalarda kelish o‘rni, kalit so‘zlarning alifbo tartibida berilishini tadqiqotchilar uchun konkordans yaratishning dastlabki urinishlardan biri deb ko‘rtsatish mumkin. Konkordans so‘zining etimalogiyasi lotinchadan olingan va *sit* “bilan” va *kor* “yurak” ma’nosini anglatadi [Carthy M., 2010; 3].

Korpusning inglizcha “corpus” so‘zi lotincha “body – tana” so‘zidan kelib chiqqan va har qanday yozma yoki og‘zaki matnga murojaat qilish uchun ishlatiladi. Biroq, zamonaviy korpus tilshunosligida korpus muayyan matnni mashinada o‘qiladigan shaklda saqlashni anglatadi. Til materiallarini saqlash uchun muayyan maqsadlar yoki tildan foydalanish uchun matnlar ishlatiladi. Korpusga asoslangan tadqiqot korpus lingvistikasidir. Olimlar korpus lingvistikasining ta’rifini turli jihatlardan izohlab berdilar. Alfurayhning fikriga ko‘ra, “korpus lingvistikasi, bu – tilni matn materiallari asosida o‘rganuvchi fan hisoblanadi”. Korpus tili, asosan, korpusni qurish va kompilyatsiya qilishni, korpusni qayta ishlash va boshqarish texnologiyasini, til tadqiqotida korpusdan foydalanishni va kompyuter tilshunosligida korpus lingvistikasini qo‘llashni o‘rganadi. Shuningdek, ilmiy tadqiqot o‘tkazish uchun uni qo‘llab-quvvatlaydigan kuchli korpusga ehtiyoj tug‘iladi [Bingbing P., 2022; 72–73].

Maxsus matnlardan tashkil topgan korpora terminologiya va mazmun ma’lumotlarining foydali manbai bo‘lishi mumkin. Sinfda tarjima gipotezalarini tasdiqlash va muayyan matn bilan bog‘liq dolzarb tarjima muammolariga mumkin bo‘lgan yechimlarni taklif qilish uchun taqqoslanadigan korpuslardan foydalanish mumkin. Shuningdek, ular tillar bo‘ylab o‘xshash domenlar yoki subdomenlarni tekshirish uchun vositani taqdim etishi mumkin. Maxsus taqqoslanadigan korpus (comparable corpus) terminologiya va tushunchalar, ma’lum bir kontekstda iboralarning tasdiqlanganligi haqida ma’lumot berishi mumkin [Pearson J., 1996; 203-216].

Matnlar korpusi to‘rtta asosiy parametr bilan tavsiflanadi:

1) matnlar korpusi yetarli darajada katta hajmda bo‘lishi kerak;

2) matnlar korpusi tuzilgan yoki belgilangan bo‘lishi kerak;

3) matnlar korpusi aniq bir korpusni tashkil etuvchi matnlar elektron shaklda bo‘lishi kerak;

4) “elektron korpus” tushunchasi qoida tariqasida ushbu korpus bilan ishlash uchun maxsus dasturiy ta’minotdir.

Korpusning qiymati quyidagicha aniqlanadi:

– korpus tilning leksik va grammatik tuzilishini, shuningdek, ma’lum vaqt oralig‘ida til o‘zgarishining uzluksiz jarayonlarini o‘rganishga imkon beruvchi til ma’lumotlarini haqiqiy muhitida ko‘rsatadi;

– korpus reprezentativ yoki matnlarning muvozanatli tarkibi bilan tavsiflanadi;

– korpus qidiruv mashinasini, morfologiyasini, tarjima tizimlarini sinab ko‘rish va turli lingvistik tadqiqotlarda foydalanish uchun ishlatilishi mumkin;

– korpus tarjimani o‘rgatish uchun juda muhim, chunki u notanish so‘z yoki grammatik shakldan foydalanish xususiyatlarini tez va samarali tekshirishi mumkin.

Tarjimaning intellektual muammo ekanligi hanuz kun tartibida. Shu bois avtomatlashtirilgan tarjima uchun kompyuterdan foydalanish imkoniyatiga shubha bilan qaralishi tabiiy. Biroq olimlarning MT tizimlari ustidan tinmay tadqiqot olib borilayotganliklari natijasida mashina tarjimasi sun’iy intellekt sinfiga kiritildi. Zotan, *mashina tarjimasi* birikmasidagi *mashina* so‘ziga *sun’iy intellekt* ma’nosida insonga xos xususiyatlar: o‘qish, tekshirish, tarjima va tahrir qilish jarayonlarining bajaruvchisi sifatida qaraladi. Kompyuter termini inglizcha *compute – hisoblamoq* so‘zidan olingan. *Kompyuter tarjimasi* zamonaviylik kasb etsa-da, mohiyatan bu to‘g‘ri emas. Ba’zi o‘rinlarda *avtomatik tarjima* birikmasi ham ishlatiladi. Lekin bu birikma mashina so‘ziga nisbatan tor doirada, ya’ni faqat dastur orqali bajarish ma’nosida qo‘llanadi.

Amaliy jihatdan, mashina tarjimasi muammosi matnlarni bir tildan boshqa tilga tez va sifatli tarjima qilish muhimligi bilan ajralib tursa, nazariy jihatdan, mashina tarjimasi nazariyasining rivojlanishini o‘rganish alohida qiziqish uyg‘otadi, chunki mashinali tarjima tizimlari lingvistik prosessorlarning bir qismi bo‘lib, lingvistik avtomatlarning rivojlanishida muhim rol o‘ynaydi [Voronovich V.V., 2013; 19].

Tarjimaning muhim jihatlaridan yana biri, bu – tarjimonlar tarjima muammosini hal qilish zarurati bilan duch kelgan har bir holatda o‘tkazishi kerak bo‘lgan mikrotadqiqotlar to‘plamidir. Tarjimonlar o‘z tarjimalarida zaif tomonlarni qanchalik ko‘p ko‘rsalar, leksik va grammatik birliklardan foydalanish to‘g‘risidagi ma’lumotlarni izlashda turli manbalarga, jumladan, matnlar korpusiga qanchalik tez-tez murojaat qilsalar, tarjima sifati shunchalik yuqori bo‘ladi. Bundan tashqari, korpusdan tarjima amalga oshirilayotgan domenga oid qo‘shimcha ma’lumot manbai sifatida muayyan muayyan foydalanish kerak. Turli fanlarga oid ixtisoslashtirilgan elektron ilova matnlarini yanada rivojlantirish, ularni tarjima amaliyotiga joriy etish tarjimon ishini ratsionallashtirish imkonini bermoqda.

Abstrakt yoki belgilash korpusning asosiy xususiyati bo‘lib, uni zamonaviy internetda keng tarqalgan raqamli to‘plamlar, kutubxonalar, ensiklopediyalardan ajratib turadi. Matnni belgilash, bu – matnni osonroq tahlil qilish uchun ma’lum ma’lumotlar atributidir [Chepik E., 2015; 69].

Belgilashning bir qancha turlari mavjud: 1) matnni bir butun deb tavsiflovchi metamatik belgisi (muallif, sarlavha, sana, hajm, matn mavzusi va boshqalar); 2)

strukturaviy belgi – bir so‘zni boshqasidan ajratish, chegaraviy iboralarni, jumlalarni, matnni ajratib ko‘rsatish imkonini beruvchi matnning tuzilishi haqidagi ma’lumot; 3) lingvistik belgi – bu matn ma’lumotlarining muayyan lingvistik birliklariga (inkor yoki so‘roq gaplar, boshqaruv yoki qo‘shnilik va boshqalar) tegishlidir. Tarkib qanchalik boy va xilma-xil bo‘lsa, korpusning ilmiy va ma’rifiy ahamiyati shunchalik yuqori bo‘ladi [Chepik E., 2015; 70].

NATIJA VA MULOHAZALAR

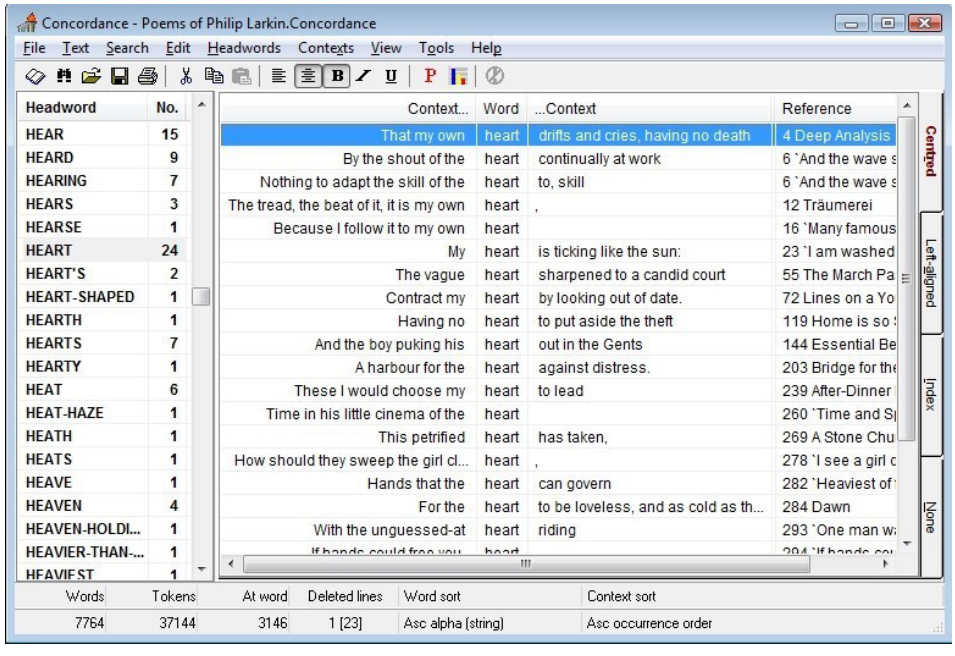
Korpusda ma’lumotni qidirish samaradorligi maxsus dasturiy ta’minotga bog‘liq korpus-menejerlar yoki konkordanser-dasturlar deb ataladi. Elektron matn korpusining maydoni olamning lisoniy manzarasini modellashtirish istiqbollarini taklif qiluvchi elektron muvofiqliklardan samarali foydalanish imkonini beradi. Konkordans, bu – ixtisoslashtirilgan tilni qo‘llash dasturi bo‘lib, u orqali namunaga avtomatik ravishda elektron matnlarning til birliklari beriladi [Bovtenko M.A., 2005; 216]. Shunday qilib, korpusga konkordans, bu – barcha kontekstlarga havola qilingan tokenlar va korpus elementlarining so‘z belgilaridagi ro‘yxati hisoblanadi. Lug‘atlar va konkordanslar o‘rtasidagi farqlar reprezentativlik, invariant, semantik yoki grammatik tahlilga yo‘naltirilganligidadir.

Konkordans funksiyasini matn muharriridagi qidiruv funksiyasi bilan solishtirish mumkin, lekin konkordansiyaning imkoniyatlari ancha katta: u bir vaqtning o‘zida bir emas, bir nechta matn yoki elektron matnlar korpusini tahlil qiladi. Texnik imkoniyatlarga qarab, konkordans muayyan til birligidan foydalanish chastotasi va muvofiqligi haqida ma’lumot berishi mumkin, shuningdek, misol topilgan muayyan matnga kirish imkonini beradi [Bovtenko M.A., 2005; 219].

Konkordans dasturlarning keng qamrovli xususiyatlari matnlar korpusidagi qidiruv ma’lumotlarining lingvistik murakkabligi bilan cheklanadi. Boshqa tahlillarda, konkordans, bu – tildagi so‘z va iboralarning qo‘llanilishini aniqlash uchun katta hajmdagi matnlarni tahlil qilish imkonini beruvchi dastur. Konkordans (Concordance) – dastur korpusda so‘ralgan so‘zni qidiradi va so‘z yoki ibora ishlatiladigan turli matnlardan bir nechta jumlalar bo‘laklari bilan yangi oyna hosil qiladi. Korpus natijalaridan so‘zning ma’nosini kontekstdan tashqarida qidirish va uning tilda qo‘llanilishi tahlilini olish mumkin. Qidiruv natijalari tildagi ayrim so‘z va iboralarni qo‘llash qoidalarini aniqlashtirish va bartaraf etish, shuningdek, tilning grammatik

tuzilishini o‘rganish uchun ishlatilishi mumkin [Sysoev P.V., 2010; 12-13].

1-rasm.



1-rasmda “Heart” (yurak) so‘zi uchun Konkordans (Concordance) 3.3 dasturidan qidiruv natijalari oynasi ko‘rsatilgan.

Ba’zi tadqiqotchilar konkordansni dastur sifatida emas, balki uning ishi natijasi deb hisoblashadi. V.N. Shevchukning talqinida “konkordans – elektron korpusda alifbo tartibida so‘zlar paydo bo‘lishining vertikal ro‘yxati hisoblanadi. So‘z o‘zining chap va o‘ng muhiti bilan birga ta’minlanadi” deb ta’kidlaydi olim [Shevchuk V.N., 2010; 45-57]. Korpusga kirish imkoniga ega bo‘lgan tarjimon bir necha soniya ichida matnning millionlab so‘zlaridan so‘z va iboralarning barcha namunalarini ko‘rishi mumkin. Cheklanmagan (doimiy rivojlanib boruvchi) monitor korpuslari lug‘at tarkibida katta rol o‘ynaydi, chunki ular yangi so‘zlarni kuzatib borish, tilni tekshirish yoki oldindan mavjud bo‘lgan so‘zlar ma’nosini o‘zgartirish yoki uslubga muvofiq bo‘lgan foydalanish imkonini beradi. Bundan ko‘rinadiki, korpus matnlarini taqdim etish va saqlash usuli asosan zamonaviy kompyuter texnologiyasiga tayanadi [Shevchuk V.N., 2010; 45].

Matnlar korpusini quyidagi vositalardan biri deb hisoblash mumkin, ular foydalanish uchun muayyan hollarda til hodisalarini tahlil qilishning zaruriy sharti sifatida qaralishi kerak. Masalan, zamonaviy elektron korpusda yuz millionlab so‘z belgilari mavjud bo‘lib, bu ularning til kompetensiyasi nuqtai nazaridan hayotiyligi haqida gapirishga imkon beradi.

N.V. Vladimovning ta’kidlashicha, “tadqiqotchi korpusni tahlil qilishida mavjud bo‘lgan asosiy protseduralarga quyidagilar kiradi [Vladimov N.V., 2005; 25]:

– korpusda ko‘rsatilgan so‘zlar va iboralarni qidirish;

– muayyan sohaning xususiyatlarini hisobga olgan holda qidiruv natijalarini ko‘rsatish;

– korpusda so‘zning ishlatilishiga oid misollar sonini hisoblash;

– kerakli parametrlar bo‘yicha qidiruv natijalarini saralash.

Bu jarayonlarning barchasi kompyuter konkordans kompilyatsiyasi (ekvivalentlarni izlash) dasturi yordamida tez va aniq amalga oshiriladi, bu esa tadqiqotchilarga kerakli matn va tarjimani tez va aniq topish imkonini beradi”. Hozirgi avtomatlashayotgan kompyuter davrida matnlar korpusi sifatida Internetdan foydalanish imkoniyati keng ommalashmoqda. Bu “Web as well Corpus” deb nomlangan onlayn veb-konkordanserlar paydo bo‘lganidan keyin hosil bo‘ldi. Ushbu resurs bir qancha kamchiliklarga ega va uning avtonom rejimda ishlaydigan tematik korpuslarga qaraganda samaraligi nisbatan kamroq. Biroq, vaqt tanqisligi va ixtisoslashgan korporatsiyaga tayyor onlayn-konkordanserlarning yetishmasligini hisobga olsak, ushbu resurs ishonchli lingvistik ma’lumotlar manbai bo‘lishi mumkin [Vladimov N.V., 2005; 26].

Virtual (ixtisoslashtirilgan) korpus tarjimonlar uchun muayyan lingvistik ma’lumotlar matnini tanlash uchun hajm jihatidan ancha katta bo‘lgan maxsus mavzular bo‘yicha matnlardan tuziladi. Matnlar turli manbalardan (davriy nashrlar, ensiklopediyalar, internet va boshqa) qat’iy belgilangan kategoriyada olinadi va har doim elektron shaklda taqdim etiladi. Tarjimon tomonidan muayyan mavzular bo‘yicha ishlab chiqarilgan virtual korpus unga quyidagi hollarda yordam berishi mumkin [Shevchuk V.N., 2010; 50-65]:

– so‘zlarning leksik va grammatik mosligini aniqlash;

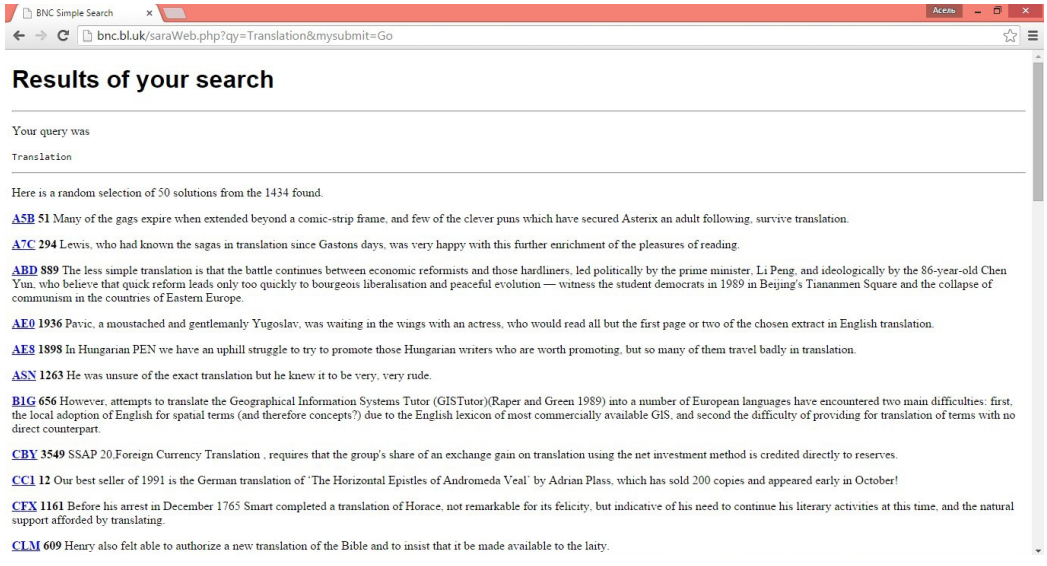
– turli lug‘atlarda yoki Internetda taklif qilingan asl so‘zning leksik ekvivalentini bir nechta variantlardan tanlash;

– tarjimon tomonidan intuitiv ravishda tanlangan qarorlarni tasdiqlash;

– mavzu bo‘yicha qo‘shimcha ensiklopedik ma’lumotlarni topish;

– terminologik dubletlar, antonimlar, holonimlar, meronimlar, atamalarning aniqlovchi nomlari va ta’riflarini topish [Shevchuk V.N., 2010; 45-57]:

2-rasm.



2-rasmda “British National Corpus” ning veb-versiyasida so‘z tarjimasi qidiruv natijalari misoli ko‘rsatilgan [http://www.natcorp.ox.ac.uk/].

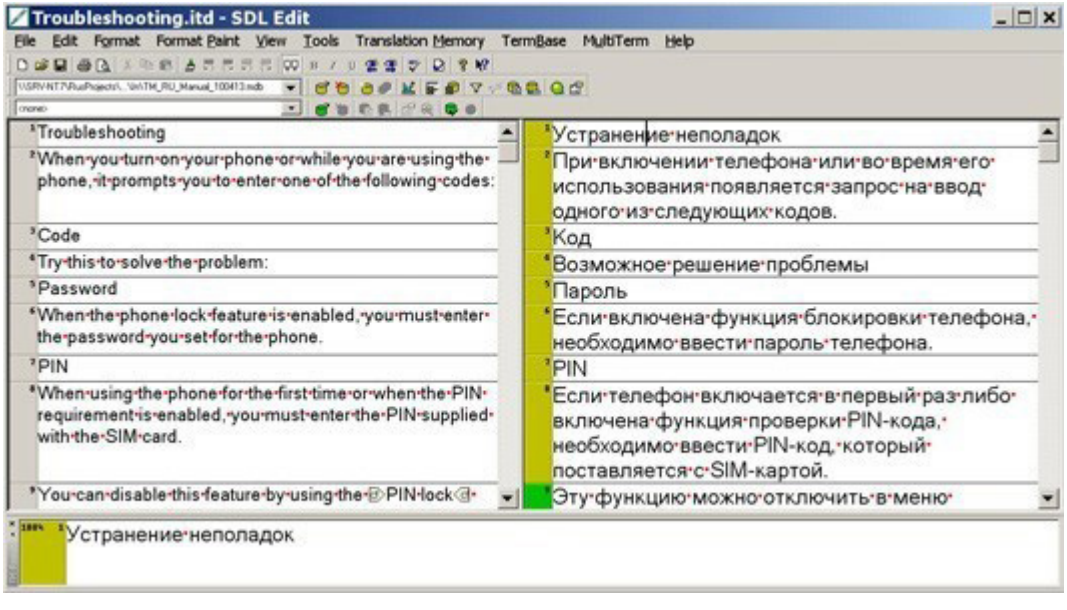
Korpus lingvistikasi sohasidagi ko‘plab tadqiqotchilar korpus tarjimonga tilni to‘g‘ri yo‘naltirish va tarjima jarayonida ma’lum miqdordagi lingvistik va til bilan bog‘liq muammolarni hal qilish imkoniyatini berishini ta’kidlaydilar [Shevchuk V.N., 2010; 45-57]. V.N. Shevchuk ta’kidlaganidek, “bu kuchli va ishonchli elektron resurs bo‘lib, amalda ona tilida so‘zlashuvchining o‘rnini bosadi va “virtual native speaker – virtual ona tilida so‘zlovchi” deb nomlanadi. Undan foydalanish shubhasiz, tarjima sifatiga hissa qo‘shadi. Korpus zamonaviy tilda amal qiladigan leksik, grammatik, stilistik, imlo va tinish belgilari haqida aniq tasavvur beradi” [Shevchuk V.N., 2010; 50-57].

Tarjima uchun qo‘llab-quvvatlovchi dasturlar yoki CAT (Computer assisted translation) kompyuter yordamida tarjima – bu tarjimonning muntazam operatsiyalarini avtomatlashtiradigan, uning vaqtini intellektual vazifalarga yo‘naltirgan kompyuter dasturlari sirasiga kiradi, desak adashmaymiz. Mashina tarjimasidan farqli o‘laroq, inson aralashuvi faqat lug‘atni yuklab olish va matnning natijasini tahrirlash uchun talab qilinadi, CAT dasturlari yordamida esa tarjima jarayoni doimo va to‘liq odamlar nazorati ostida bo‘ladi. Avtomatlashtirish avval tarjima qilingan matn qismlarining bir nechta tarjimalaridan qochish va tarjima variantlarini Tarjima xotirasi yoki (TM – Translation Memery) deb ataladigan tizimda saqlashni anglatadi. Misol uchun, agar tarjimon *aktuator korpusining qopqog‘i vintlari va olti burchakli gaykalar* iborasini allaqachon tarjima qilgan bo‘lsa, keyingi safar u matnda ushbu iboraga duch kelganida, dastur avtomatik ravishda o‘sha ilgari qilingan tarjimani taklif etadi. CAT dasturlari tarjimonga (TM – Translation Memory) Tarjima Xotirasi ikki tilli so‘z juftliklarida “original – tarjima” to‘plash imkonini beradi. Janr va mavzu jihatidan o‘xshash matnlar ustida ishlash jarayonida TX (tarjima xotirasi) to‘ldirilishi hisobiga tarjima jarayoni

tezlashadi [Grabovskij V.N., 2004; 57–62]. Tarjima birligini (TU-Translation Unit) asl nusxaning segmenti va uning tarjima segmenti bilan taqqoslanishini ifodalaydi. Segmentning roli odatda gap orqali amalga oshiriladi. Shu sababli ham TX dasturlari bir nechta mezonlar bilan ajralib turadi. Shunday qilib, texnik amalga oshirish orqali mahalliy TX dasturlari va onlayn rejimda mavjud bo‘lgan dasturlar bir-biridan ushbu jihatlari orqali ajratiladi. Onlayn versiyalardan foydalanganda masalan, *Google Translator Toolkit* va *Word fast Anywhere* ish brauzer oynasida amalga oshiriladi [Grabovskij V.N., 2004; 57–62].

Ikkinchi mezonga ko‘ra, qo‘shimcha funktsional imkoniyatlar: terminologiya ma’lumotlar bazasiga kirishni ta‘minlash va tarjima xotirasini (TM) to‘ldirish uchun o‘rnatilgan dasturlar. Bu yagona korxona yoki domen doirasida til siyosatiga muvofiqligini ta‘minlash uchun tarjimada terminologiyaning izchil bo‘lishiga xizmat qiladi. Tarjima xotirasini yaratish uchun asl matn va tarjima segmentlarini keyinchalik ma‘lumotlar bazasi dasturiga kiritish uchun solishtirish muhim. Shunday qilib, TM dasturlari matn segmentlarini baholash funktsiyasi bilan ta’minlangan (odatda jumlalar). Bundan tashqari, to‘g‘ridan-to‘g‘ri tarjima jarayonida ma’noni ajratuvchi leksik birliklarni hal qilish uchun so‘zni kontekstda ko‘rish talab qilinishi mumkin. TM dasturlari, qoida tariqasida, so‘zning muvofiqligini yaratish imkoniyatini beradi: kontekstdagi so‘zning barcha hodisalari ro‘yxatini topish, shu bilan birga har bir holatda leksik birlikni tarjima qilish usullarini taqdim etish ham muhim hisoblanadi.

3-rasm.



3-rasmda **Troubleshooting TX** dasturning interfeysi ikki tilda ingliz va rus tillarida ko‘rsatilgan.

O.I. Babina va P.G. Osmininlarnng so‘zlariga ko‘ra, tarjima xotirasidan foydalangan holda tarjimaning haqiqiy jarayoni bir necha bosqichlarni o‘z ichiga

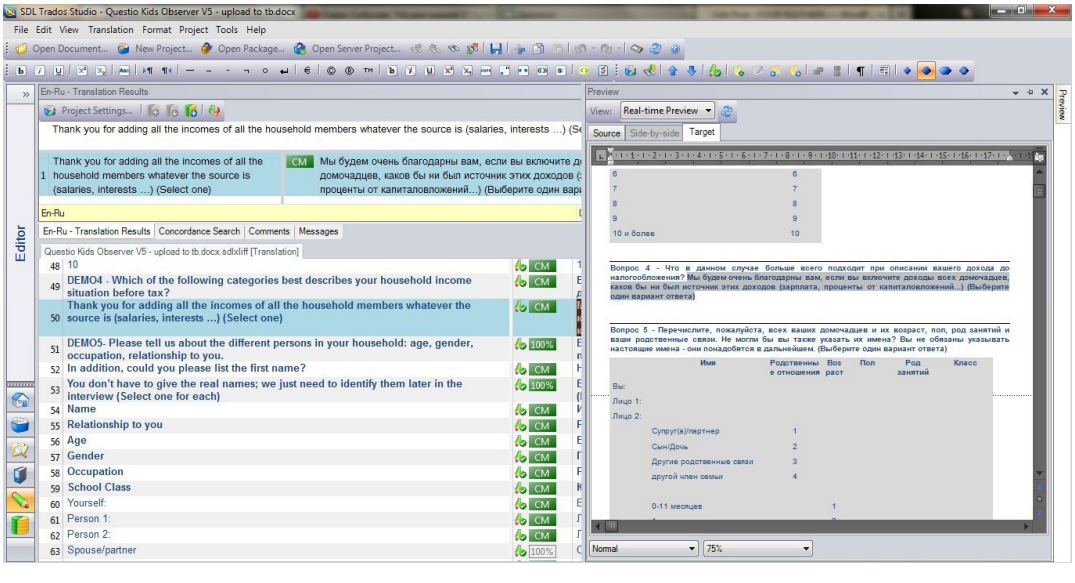
oladi [Babina O.I., Osminin P.G., 2013; 5].

1. Tayyorgarlik (Preperation): bu bosqichda yangi tarjima xotirasini yaratish muammosini hal qilish yoki matnni tarjima qilish uchun foydalaniladigan tarjima xotirasini tanlash, kerakli tuzatishlar kiritish (masalan, tarjima qilish usullariga e'tibor qaratish, oldindan yaratilgan loyihalarga mos keladigan atamalarda duch keladigan muammolarni hal qilish kerak) va tanlangan tarjima xotirasini tuzatishlar olib borish kerak bo'ladi. Masalan, ma'lum bir loyihaga muvofiq barcha tarjima birliklarida ma'lum bir atamani almashtirish yoki shunga o'xshash jarayonni amalga oshirish kerak bo'ladi.

2. Tarjima (Translation): oldindan tarjima (pre-translation), bu – manba matnning aniq segmentlarini maqsadli tildagi ekvivalentlariga avtomatik ravishda almashtirish imkoniyatini beradi. Ma'lumotlar bazasida topilmagan matn segmentlari qo'lda yoki agar mavjud bo'lsa, mashina tarjimasi tizimidan foydalangan holda tarjima qilinadi. TM – tijorat dasturlari, qoida tariqasida, atamalarni ajratib olishning o'rnatilgan tizimlariga ega (masalan, Extract, SDL MultiTerm, PROMT TerM, MonoConc Pro, Simple Concordance Programme va boshqa tarjimon dasturlar). "Xom" matn, shu jumladan avtomatik almashtirilgan terminologiya va yakuniy bosqichda matn segmentlari post-tahrirlashdan o'tkaziladi.

3. Sifat nazorati (Quality control): tarjimaning to'liqligi, grammatik aniqligi va tegishli atamalarning to'g'ri tarjimasi uchun rasmiy tekshiruvni o'z ichiga oladi, bu tarjimon tomonidan (ehtimol) mijoz tomonidan amalga oshirilishi mumkin [Babina O.I., Osminin P.G., 2013; 5].

4-rasm.



4-rasmda TRADOS TM-xotira dasturi bilan ishlash ko'rsatilgan.

USLUBIY TAVSIYALAR

Yuqoridagi tahlillardan kelib chiqib shuni ta'kidlash kerakki, tarjimaga asoslangan vazifalarning deyarli barcha kamchiliklarini bartaraf etish uchun ko'p vaqt talab qilinadi. Bunday tarjimon xotirasiga asoslangan dasturlarni amalga oshirish uchun trenerlar va o'rganuvchilar ko'p vaqt va katta kuch talab qiladigan maxsus ko'nikmalarga ega bo'lishlari kerak. Bundan tashqari, hozirda bunday turdagi tayyor materiallarga asoslangan baza (veb saytlar) deyarli yo'q. Bu shuni anglatadiki, trener maxsus topshiriqlarni ishlab chiqishi va katta hajmdagi materialni TMga joylashtirishi muhim ahamiyatga ega. Biroq tarjimani o'rgatishda korpusning o'rni katta. Darhaqiqat, oo'rganuvchilarning barcha asosiy xatolari ularning muayyan so'z yoki iboralarni ishlatishda yaxshi tajribaga ega bo'lmaganligi, turli xil matn uslublarining tuzilishini faqat nazariy jihatdan bilishi tufayli sodir bo'ladi. Talabalarga qoidaning amalda qo'llanilishini tushunish imkoniyatini berish real va zarurdir. Shuni ham ta'kidlash joizki, o'z tadqiqotlari davomida olingan bilimlar, asosan, tayyor xulosalardan farqli ravishda xotirada saqlanadi. Talabalar o'z tadqiqotlari natijalarini mamnuniyat bilan baham ko'radilar va bu ishni targ'ib qilish chet tilini o'rganish uchun eng yaxshi rag'bat bo'la oladi.

Aniq korpusga oldindan tayyorlangan vazifalar chet tili va tarjima amaliyoti darslarida qo'llanilishi mumkin. Bunda predlog – qo'shimchalarning qo'llanilishi va tarjimasini ta'minlash (masalan, about – haqida) murakkab va so'roq gaplarni tahlil qilish; *to tell/ to say aytmoq/demoq*, *listen/hear tinglash/eshitish* va shularga o'xshash turli atamalarning qo'llanishdagi farqlarini tahlil qilish; so'zlarni sanitarizatsiya qilish (masalan, turli kontekstlarda fe'llarning ishlatilishini tahlil qilish); shuningdek xalqaro so'zlarni topish va ularning ma'nosi ona tilida va tarjima qilinayotgan tilda bir xil yoki bir xil emasligini aytish va shu kabilar tarjima amaliyotida korpus vositalari orqali qo'llanilishi mumkin.

XULOSA

Korpus zamonaviy tilshunoslik va tabiiy tilni qayta ishlashdagi eng muhim manbalardan biridir. Korpusning samarali sohasida biz pozitivistik lingvistik tadqiqotlar olib borishimiz va lug'at, grammatika, pragmatika va tarjima kabi gumanitar masalalarni o'rganishimiz mumkin. Biroq korpusni boshqa tizim orqali, ya'ni korpusni turli ma'lumotlar bazasi sifatida ko'rib chiqish va kompyuterda o'rganish imkonini berish orqali ko'p sohalarda tadqiq etishimiz mumkin. Bundan tashqari, biz korpusni cheksiz, doimiy ravishda yaratilgan matn to'plamining namunaviy vakili sifatida ko'rib chiqishimiz va til ma'lumotlarini avtomatik qayta ishlash tizimlarini yaratish va baholashni o'rganishimiz va o'rgatishimiz mumkin. Matnlarning tarjima korpusini o'qitish maqsadida mavhum ma'lumot sifatida qaralishi va tarjima usullari va metodlarini o'rganishda professional tarjima namunalarini taqdim etish muhim o'rin tutadi [Di Wu, 2022; 1].

Matn korpusining tahlili, korpus lingvistikasining usullari va yutuqlari chet tillarini o'qitish va tarjima sohasidagi istiqbolli yo'nalishdir. Bu sohani rivojlantirishning jahon amaliyoti ularning samaradorligini isbotlamoqda. Korpus

tomonidan taklif etilgan so‘z va iboralar bilan ishlash vositalari to‘plami tarjimonning ekvivalentlarini topishda vaqtini tejash uchun qo‘shimcha imkoniyatlar yaratadi va kommunikativ-evristik xarakterdagi bilimlarni to‘plashga hissa qo‘shadi, tarjimonga tilni murakkab usulda bilish imkonini beradi. Kontekst esa, tarjima muammolarini hal qilishning samarali strategiyasidir.

Demak, korpus bir necha sabablarga ko‘ra lug‘atlarga qaraganda qulayroq va ishonchli vositadir. Birinchidan, matnlar korpusi an’anaviy lug‘atlar kabi to‘plam tuzilmalari emas, balki doimiy ravishda yangilanib turadigan ma’lumotlar bazasidir. Uning yordamida tarjimon korpusda so‘zni qo‘llashda tilga asoslangan tahlilni rivojlantirishning so‘nggi tendentsiyalaridan xabardor bo‘lishi mumkin. Ikkinchidan, korpus hajmi jihatidan ancha katta bo‘lganligi, so‘z to‘g‘risidagi korpusdan olinadigan ma’lumotlarning xolis va to‘g‘ri bo‘lishi tufayli lug‘atdan ko‘ra keng so‘z belgilarining manbai hisoblanadi. Uchinchidan, korpus bilan ishlash lug‘atga qaraganda ancha oson va qulay. Korpus hozirda tarjimon ish joyining bir qismi bo‘lgan mashinaga (kompyuterga) joylashtirilgan. Klaviatura va “sichqoncha”ning oddiy manipulyatsiyasi yordamida tarjimon bir necha soniya ichida kerakli lingvistik ma’lumotlarni olishga qodir.

FOYDALANILGAN ADABIYOTLAR

1. boyev, N.B. (2020). Ingliz tili korpusining funksional xususiyatlari (COCA misolida). *Filol. fan. fal. dokt. diss*. Toshkent.
2. Axmedova, A.N. (2022). Parallel korpusda o‘xshatishlarning leksik-semantik munosabatlari (inglizcha-o‘zbekcha, o‘zbekcha-inglizcha parallel matnlar misolida). *Filol. fan. fal. dokt. diss*. Toshkent.
3. Babina, O.I., Osminin, P.G. (2013). *Pamyat’ perevodov v obuchenii perevodchikov*, *3*(5), (7). Chelyabinsk.
4. Bingbing, P., & Qiongfang, Q. (2022). *Construction of parallel corpus for English translation teaching based on computer aided translation software*. China.
5. Bovtenko, M.A. (2005). *Komp’yuternaya lingvodidaktika*. Moskva: Flinta.
6. Di, Wu. (2022). *Design of Chinese corpus based on semantic mining algorithm*. China.
7. Ergasheva, G. (2018). Turli tizimdagi tillarda genderga oid terminologiyaning shakllanishida lingvistik va ekstralingvistik omillar. *Filol. fan. fal. dokt. diss*. Toshkent.
8. Frerot, C. (2016). *Corpora and corpus technology for translation purposes in professional and academic environments: Major achievements and new perspectives* (pp. 36-61). https://doi.org/10.5007/21757968.2016v36nesp1p36.
9. Gao, T. Wang, X. (2022). The construction of computer-aided translation teaching platform based on borpus. *Journal of Physics*, *1648*(3), 32-59. https://doi.org/10.1088/1742-6596/1648/3/032059.
10. Grabovskij, V.N. (2004). Tekhnologiya translation memory mosty. *Zhurnal perevodchikov*, *2*, 57-62, 65.
11. Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.
12. McCarthy, M., O‘Keeffe A. (2010). *The Routledge handbook of corpus linguistics*. New York.
13. Pearson, J. (1996). Teaching terminology using electronic resources. In S. Botley, J. Glass, T. McEnery & A. Wilson (Eds.), *Proceedings of teaching and language corpora* (pp. 203-216). Lancaster.
14. Shevchuk, V.N. (2010). *Elektronnye resursy perevodchika*. Moskva: Librajt.
15. Stubbs, M. (2001). Texts, corpora, and problems of interpretation: A response to Widdow-

son. *Applied Linguistics*, *22*(2), 149-72.
16. Sysoev, P.V. (2010). *Inostrannye yazyki v shkol*e. Moskva: Metodicheskaya mozaika.
17. Vladimov, N.V. (2005). Korpusnyj podhod k resheniyu perevodcheskih problem (na materiale pis'mennyh perevodov s russkogo yazyka na anglijskij). *Diss. kand. filol. nauk*. Moskva.
18. Voronovich, V.V. (2013). *Machine translation*. Minsk.
19. Xamroeva, Sh. (2018). O'zbek tili mualliflik korpusini tuzishning lingvistik asoslari. *Filol. fan. dokt. diss. avtoref*. (52 b.). Qarshi
20. Zhijie, Z. (2017). A study on the computer aided English translation of local legal based on parallel corpus. In International Conference on *Frontier Computing* (pp. 241-256). https://doi.org/10.1007/978-981-10-7398 4_26.
21. Chepik, E.Yu. (2005). Politicheskoe slovo v strukture ehlektronnogo slovarya. *Kul'tura narodov prichernomor'ya*, 69, 205-210.
22. http://www.natcorp.ox.ac.uk/.

REFERENCES

1. Akhmedova, A.N. (2022). Lexical-semantic relations of similes in the parallel corpus (in the case of English-Uzbek, Uzbek-English parallel texts). *PhD thes. in phil.* Tashkent.
2. Ataboyev, N.B. (2020). Functional properties of the English corpus (in the case of COCA). *PhD thes. in phil*. Tashkent.
3. Babina, O.I., Osminin, P.G. (2013). *Translation memory in the training of translators*, *3*(5), (7). Chelyabinsk.
4. Bingbing, P., & Qiongfang, Q. (2022). *Construction of parallel corpus for english translation teaching based on computer aided translation software*. China.
5. Bovtenko, M.A. (2005). *Computer linguodidactics*. Moscow: Flinta.
6. Chepik, E.Yu. (2005). Political word in the structure of the electronic dictionary. *Culture of the peoples of the Black Sea region*, *69*, 205-210.
7. Di, Wu. (2022). *Design of Chinese corpus based on semantic mining algorithm*. China.
8. Ergasheva, G. (2018). Linguistic and extralinguistic factors in the formation of gender terminology in languages of different systems. *PhD thes. in phil*. Tashkent.
9. Frerot, C. (2016). *Corpora and corpus technology for translation purposes in professional and academic environments: Major achievements and new perspectives* (pp. 36-61). https://doi.org/10.5007/21757968.2016v36nesp1p36.
10. Gao, T. Wang, X. (2022). The construction of computer-aided translation teaching platform based on borpus. *Journal of Physics*, *1648*(3), 32-59. https://doi.org/10.1088/1742-6596/1648/3/032059.
11. Grabovsky, V.N. (2004). Bridges translation memory technology. *Journal of Translators*, *2*, 57-62, 65.
12. Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.
13. Khamroeva, Sh. (2018). Linguistic foundations of the composition of the author's corpus of the Uzbek language. *PhD thes. in phil*. (p. 52). Karshi.
14. McCarthy, M., O'Keeffe A. (2010). *The routledge handbook of corpus linguistics*. New York.
15. Pearson, J. (1996). Teaching terminology using electronic resources. In S. Botley, J. Glass, T. McEnery & A. Wilson (Eds.), *Proceedings of teaching and language corpora* (pp. 203-216). Lancaster.
16. Shevchuk, V.N. (2010). *Electronic resources of the translator*. Moscow: Libright.
17. Stubbs, M. (2001). Texts, corpora, and problems of interpretation: A response to Widdowson. *Applied Linguistics*, *22*(2), 149-72.
18. Sysoev, P.V. (2010). *Foreign languages at school*. Moscow: Methodical Mosaic.
19. Vladimov N.V. (2005). Corpus approach to solving translation problems (on the basis of

written translations from Russian into English). *Diss. cand. philol. scien*. Moscow.
20. Voronovich, V.V. (2013). *Machine translation*. Minsk.
21. Zhijie, Z. (2017). A study on the computer aided English translation of local legal based on parallel corpus. In *International Conference on Frontier Computing* (pp. 241-256). https://doi.org/10.1007/978-981-10-7398 4_26.
22. http://www.natcorp.ox.ac.uk/.