

3-SHO‘BA. KOMPYUTER LINGVISTIKASI VA KORPUS LINGVISTIKASI MUAMMOLARI MATNNI AVTOMATIK QAYTA ISHLASH XUSUSIDA

Toirova G.I.

BuxDU dotsenti, filologiya fanlari doktori

Oxirgi o‘n yilliklarda matnning avtomatik tahlil instrumentlarini ishlab chiqishda va matnni tabiiy tilida razmetka (belgilash) sohasiga oid yangiliklarda misli ko‘rilmagan darajada sifatli o‘zgarishlar yuz berdi [1,2,8,9]. Bu hodisani soha mutaxassisleri sifatli sakrash deb baholashdi. Bu sakrash zamonaviy kompyuter lingvistikasi sohasida erishilgan yutuqlar, usullarni faol o‘zlashtirish va kompyuter fani hamda statistikasi, mashina o‘qiydigan matn ma’lumotlari hajmining eksponensial o‘sishi, shuningdek elektron lingvistik resurslarning rivojlanishi sababli yuz berdi. Lingvistik tahlilning zamonaviy instrumentlari matnning nafaqat grammatik (morfologik, sintaktik) xususiyatlarini, balki matnning ba’zi semantik va pragmatik xususiyatlarini ham avtomatik tarzda razmetka (belgilash)ga imkon berdi. Tabiiy tilni mashina to‘laonli tushinishi, avvalgidek futurologiya sohasi zimmasiga qolyapti, ammo kompyuter lingvistikasi texnologiyasi matn materiallarini tadqiq qilishning yangi imkoniyatlarini ochib berdi.

Kompyuter lingvistikasining zamonaviy usullari qo‘llanilgan razmetka (belgilash)larning avtomatlashtirilgan instrumentariysini ishlab chiqishdan va uni miqdoriy tadqiqotlar uchun qo‘llashning namoyishidan iborat[7,8,9]. Mazkur yondashuv, matn tilini avtomatik vositalar bilan va badiiy asarni filologik tadqiqot vazifalari bilan tadqiq qilish imkoniyatlari orasida yuzaga keladigan yuqorida qayd etilgan qiyinchiliklarni barataraf qilishga imkon yaratadi. Personajlar tizimini miqdoriy tahlili uchun mo‘ljallangan badiiy matning mashina o‘qiydigan semantik razmetka (belgilash)lari, A.Navoiyning asarlari misolida semantik razmetka (belgilash)larni amalga oshirish mumkin. Shunday qilib, badiiy asarning mazmuniy elementlarini tadqiq qilish uchun zamonaviy kompyuter lingvistikasi instrumentlaridan va ma’lumotlarni statistik usulda tahlil qilishning imkoniyatlari qo‘rsatib berilgan. Shu yo‘sinda, yirik shakldagi nasriy yoki she’riy asarning personajlar tizimining matndagi reprezentatsiyasi tadqiqot obyekti bo‘lib xizmat qiladi. Personajlarning parametrik tavsiflari tadqiqotning predmeti bo‘lib xizmat qilgan.

Asar personaji – bu, bir tomondan, soddalashtirilgan (sxematik, stereotipik tarzda), va yoki o‘zining badiiy timsolida juda ham murakkab tuzilishga ega bo‘lgan badiiy obraz, shaxsning ko‘lda yaratilgan modelidir. Shu bilan bir qatorda eng murakkab badiiy obraz ham matnda aniq ketma-ketlikda tilga olinganda vujudga keladi – Y.V. Nedoshivina ning bu kabi funksional ta’rifini taqqoslaymiz: “Adabiy personaj – bu, mohiyatan, bir matn chegarasida birta shaxsning paydo bo‘lishining ketma-ketlikdagi seriyasidir” [6]. Personajning matnga kirishlari turli xil bo‘lishi mumkin: ismlarning variantlari (Farhod, Shirin, Majnun, Layli, Sulton Husayn Boyqaro, Badiuzzamon), titullar va gonorotivlar (janobi oliylari), anaforik eslanishlar (u), boshqa personajlarning nutqidagi murojaatlar (muhtaram zot, siz),

rolli va vaziyatli tavsiflar (yosh yigit, kelayotgan kishi, yaralangan) va boshqalar. Ammo o'ta muhim bo'lgan bu kabi kirishlar va shakllantiriladigan lingvistik belgilar aniq chegaralarga ega (atoqli otlar, kishilik olmoshlari) va shuning uchun ham ular kompyuter lingvistikasi vositalarida mavjud bo'lgan razmetkalarga mos tushadi. Shunday qilib, personaj badiiy asarning shunday elementi bo'lib kiradiki, u o'zida badiiy murakkablik va o'ta teranlik bilan matnda avtomatlashtirilgan razmetka imkonini o'zida mujassamlashtiradi.

"Personaj bir nechta sahifadan so'ng yana qayta paydo bo'lish va o'sib borayotgan birlikka yana birta bo'g'in qo'shish uchun yo'qoladi, o'zini o'rnini boshqalarga bo'shatib beradi. Takrorlanuvchi, u yoki bu darajada barqaror belgilar personajning xususiyatini tashkil etadi" [6]. Matnda lingvistik nuqtai nazardan har bir yangi kirish, uni o'rab turgan kontekstning o'ziga xosligiga va muvofiq leksemalarning grammatik (so'z o'zgartiruvchi) xususiyatlariga ishora qiluvchi bir qator belgilari bilan tavsiflanishi mumkin. Bu personajlarning miqdoriy tadqiqotlari uchun keng imkoniyatlar ochib beradi.

Badiiy asardagi personajlari nutq matnini semantik razmetka (belgilash)ga quyidagi usulini ishlab chiqish va uning aprobatsiyasini taklif etamiz: badiiy asar personajlari tizimini modellashtirish; matnda personajlar qo'llanilishini avtomatik razmetka (belgilash); mavjud birta aniq personajning yagona koreferent zanjiriga mukammal personaj identifikatori yordamida bog'langan.

Matnni avtomatik qayta ishlashning zamonaviy vositalari orqali ahamiyatli ma'lumotlarni va badiiy asarlar personajlari tizimini semantik belgi shaklida strukturalash uchun qo'llanilishi mumkin. Tayyorlangan razmetkalar personajlar tizimi tahlilini miqdoriy metodlar bilan amalga oshirishga imkon beradi (takrorlanishlar tahlili, ko'p omilli statistik tahlil, korrelyatsion tahlil, tarmoqli tahlil). Razmetkalardan olingan ma'lumotlarni tahlil qilishning aniq usulini (metodini) tanlash, personajlarning aynan qaysi tizimi olingan modelda qo'llanilishga bog'liq bo'ladi.

Qodiriyning "O'tkan kunlar" romanidagi quyidagi parchaga e'tiborimizni qaratsak:

Saroy ahli kunduzgi ish kunlaridan bo'shab hujralariga qaytqanlar, ko'b hujralar kechlik osh pishirish ila mashg'ul, shuning uchun kunduzgiga qarag'anda saroy jonliq: kishilarning shaqillashib so'zlashishlari, xoxolab kulishishlari saroyni ko'kka ko'targudek.

Abdulla Qodiriy "O'tkan kunlar" asari matnini avtomatik-semantik razmetkalashda quyidagi modeldan foydalanadi:

t:nut.metonim: - nutqiy metonimiya

O'zbek xalqining chin ma'noda o'ziga xos milliyligini aks ettiradigan mazkur romanning ushbu parchasida metafora leksik-stilistik usuli kechlik osh pishirish ila mashg'ul, saroy jonliq, saroyi ko'kka ko'targudek kabi birliklarda metafora qo'llangan. Bu yerda xujralar, saroy so'zlari orqali odamlar, savdogarlar ifodalanmoqda.

Og'ir tabiatlik, ulug' g'avdalik, ko'rkam va oq yuzlik, kelishgan, qora ko'zlik, mutanosib qora qoshliq va endigina murti sabz urgan bir yigit Muallif mazkur parchada romandagi bosh qahramonlardan biri Otabekni tasvirlashga harakat qilgan. Bu yerda murti sabzi urgan tasviriy ifodasi orqali Otabekni o'spirin bir

yigit, yoshi yigirma-yigirma ikkilar atrofida ekanligini ko'rsatib, badiiy adabiyotning o'ziga xos badiiy estetik qiymatini namoyon etmoqda.

t:nut.metonim: - murti sabz urgan

Qandog'dir bir xayol ichida o'lturg'uchi bu yigit Toshkandning mashhur ayonlaridan bo'lg'an Yusufbek hojining o'g'li — Otabek.

t:nut.metonim: - xayol ichida o'lturg'uchi

Bu yigit yaxshig'ina davlatmand bo'lsa ham, lekin shuhrati nimauchundir boyligi bilan bo'lmay, «Homid xotinboz» deb shuhratlangan, kishilar Homid orqasidan so'zlashkanda uning otig'a taqilg'an laqabni qo'shib aytmasalar, yolg'iz «Homidboy» deyish ila uni tanita olmaydilar (aslida: u haqida gapirganda), otig'a taqilgan (aslida: ismiga qo'shib aytiladigan) so'z birikmalarida metafora leksik-stilistik usul qo'llagan. Romandagi salbiy qahramon bo'lmish Homidga berilgan ta'rifdan muallif badiiy tasviriy vosita, ya'ni metaforadan mohirona foydalanish orqali yoritib bergan. Bu yerda orqasidan so'zlashgan

t:nut.metonim: otig'a taqilgan

Manimcha, uylanishdek nozik bir ish dunyoda yo'qdir, — dedi Rahmat va Otabekka yuz o'girdi. — Uylangach, xotining ta'bingga muvofiq kelsa bu juda yaxshi; yo'qsa, munchalik og'ir gap dunyoda bo'lmaslanishdek nozik bir ish” “kichik bir xato butun umrga tatigulik tashvishga aylanishi mumkin bo'lgan amal” kabi ma'nolarda qo'llanilmoqda. “Og'ir gap” birikmasi esa, yuqorida ta'kidlab o'tilgan “butun umrga tatigulik tashvish” ma'nosida qo'llanilmoqda. Aslida ishning nozigi bo'lmagandek, gapning og'iri ham bo'lmaydi. Bu yerda metaforik ko'chma ma'no orqali badiiy adabiyotning serjiloligi ko'rsatib berilmoqda Ushbu matnda metafora leksik-stilistik usuli uylanishdek nozik ish va og'ir gap birikmalarida qo'llangan.

t:nut.metonim: Og'ir gap

Bir so'z bilan aytganda, metonimiya - bu inson bilimi va munosabatini toifaga ajrata oladigan ko'plab usullardan birini aks ettiradigan umumkognitiv jarayondir.

- *Manimcha, o'risning bizdan yuqoridalig'i uning ittifoqidan bo'lsa kerak, - dedi Otabek, ammo bizning kundan-kunga orqag'a ketishimizga o'zora nizoimiz sabab bo'lmoqda, deb o'ylayman, boshqa xil aytkanda Ziyo amakining fikrlari qisman to'g'ri.* Ushbu parchada qo'llanilgan o'risning bizdan yuqoridalig'i birikmasida yuqori so'zi orqali ruslarning bizga qaraganda yuqori pog'ona, yuqori daraja, yuqori mavqeda ekanligi ma'nolarida qo'llanilgan. Ushbu gapda metonimiya so'zlardan birining tushib qolishi orqali o'z ma'nosini ifodalagan.

Razmetkalar, matnlarni TEI xalqaro kodlash formati asosida bajariladi[5,7]. Razmetkalar personajlarning identifikatsiya qilingan eslashlarini o'z ichiga oladi, shu jumladan anaforik ishoralar, adresat va адресанта personajning ko'chirma gapidagi replikalarining bir ma'noli ishoralarini ham qamrab olish kerak bo'ladi.

Foydalanilgan adabiyotlar

1. Абдурахмонова Н.З. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти (содда гаплар мисолида): Филол.фан.бўйича фалсафа доктори (PhD)...дис. афтореф. – Тошкент, 2018.

2. Абжалова М. Ўзбек тилидаги матнларни таҳрир ва таҳлил қилувчи дастурнинг лингвистик модуллари (расмий ва илмий услубдаги матнлар

тахрири дастури учун): Филол.фан.бўйича фалсафа доктори (PhD)...дис.– Фарғона, 2019.–Б.22.

3.Ванюшкин А. С., Гращенко Л. А. Оценка алгоритмов извлечения ключевых слов: инструментарий и ресурсы // Новые информационные технологии в автоматизированных системах. – 2017. – № 20. – С. 95–102.

4.Захаров В.П. Корпусная лингвистика: Учебно-метод. пособие. – СПб., 2005. – 48 с.

MATN TAHRIRI VA MAHALLIY TAJRIBA: “MUKAMMAL MAQOLA FORMULASI” HAMDA “OSh-MAQOLA” METODI

Egamov N.

Qarshi DU o‘qituvchisi, jurnalist

“Amaliy filologiya” yo‘nalishining “Matn tahriri” fanida talabalarga “Publitsistik matnlar tahlili va tahriri”, “Tahrir maktablari”, “Uslub va tahrir”, “Korrektura va tahrir” singari mavzularni o‘tishda mahalliy matbuot tajribalaridan unumli foydalanish yaxshi samara beradi.

Bir necha yil mahalliy gazetalarda mas‘ul kotib, muharrir sifatida faoliyat yuritib, tahrir borasida to‘plagan tajribam mazkur fan amaliy mashg‘ulotlarida asqatmoqda.

Qashqadaryo viloyati hokimligi va xalq deputatlari viloyat kengashi muassisligidagi “Qashqadaryo” gazetasi tahririyati – xalqaro “Oltin qalam” tanlovi Bosh Milliy mukofotiga birinchilardan bo‘lib sazovor bo‘lgan katta ijodiy maktab. Bu yerda o‘ziga xos tahrir maktabi yaratilgan, desak, mubolag‘a bo‘lmaydi. Ayni paytda viloyatimiz va poytaxtdagi nufuzli nashrlar hamda telekanallarda ishlab yurgan taniqli jurnalist va muharrirlarning ko‘pchiligi mana shu “qozonda qaynagan”. Mazkur tahririyatda 20 yilga yaqin vaqt davomida muxbir, bo‘lim muharriri va mas‘ul kotib sifatida faoliyat yuritib, tahrir ko‘nikmalariga bosqichma-bosqich erishib borganman.

Gazeta bosh muharririning o‘rinbosari, O‘zbekistonda xizmat ko‘rsatgan jurnalist Botir Temirov – mana shu maktabning usta tahrirchilaridan. Tahririyat qoshida bir necha yil faoliyat yuritgan “Jurnalistik mahorat maktabi” da B.Temirov bilan birga yoshlarga dars berib yurgan kezlirimizda u “MMF” degan formula kashf qilgandi. “Mukammal maqola formulasi” deb ataladigan bu soddagina metod semiz-semiz nazariy kitoblardagi uzundan uzoq qoidalaridan farqli o‘laroq sohaga havasmand yoshlarga saboq berishda bizga juda qo‘l kelgan.

B.Temirov “Mukammal maqola formulasi”ni quyidagi tarzda ifodalagan:
MMF=DSV+VEF+ETS+JT+JP...

Mazkur formuladagi shartli belgilarni birma-bir izohlashdan avval yana bir fikrni qayd etmoqchiman. Biz tajribali jurnalist va tahrirchi hamkasbimizning fikrlariga qo‘shilgan holda xuddi shu formulaga uyg‘un tarzda mukammal maqola matnini tuzishni osh tayyorlashga qiyoslagan holda tushuntirishni taklif qilamiz. Ya‘ni, palov tayyorlash retsepti deyarli hammagan yaxshi ma‘lum:
OSh=GURUCh+SABZI+GO‘ShT+YoG‘+TUZ VA ZIRAVORLAR

Xuddi shu retseptni maqola matnini tuzishga qiyoslagan holda mavzuni