Прегледни рад

Марија Д. ВУКОВИЋ*

Факултет за саобраћај, комуникације и логистику Будва, Црна Гора

Јелена Д. РАЈОВИЋ**

Висока техничка школа струковних студија из Урошевца, Лепосавић

УТИЦАЈ МАШИНСКОГ ПРЕВОЂЕЊА НА КВАЛИТЕТ ПРЕВОДА СТРУЧНИХ ТЕКСТОВА СА ЕНГЛЕСКОГ ЈЕЗИКА НА СРПСКИ

Айсџиракџи – Употреба интернет машинских преводиоца има велику и честу праксу у свакодневном животу, као један најдоступнији, а уједно и најлакши начин за превођење како стручне литературе, тако и разних других текстова. Такође, све више можемо видети објављене преводе у којима има доста грешака. У овом раду се налази пилот истраживање, засновано на анализи десет студентских радова анализе машинских превода кратких, истих, стручних текстова у временском распону од три школске године. Циљ овог рада је да покаже колико су поуздани машински преводиоци били некада, а колико су сада као и да ли има промена и којих. Врсте грешака, као и проценти у којој мери су се јављале су анализиране, извршен је контролни превод у обрнутом смеру и на основу тога су донети закључци наведени у раду. Временски распон од три године је показао да су разлике у преводу значајне што упућује на то да је управљање машинским превођењем могуће и то у делу развоја информацио-них технологија и сарадње лингвиста са стручњацима за информационе технологије. Та сарадња у погледу машинског превођења је неопходна и мора бити стална.

Кључне ријечи: енглески језик, управљање утицајем, машински преводилац, употреба интернет преводиоца, анализа грешака, квалитет превода.

УВОД

Потребе увођења информационих технологија на свим пољима доводи нас до све веће употребе компјутера и интернета у свакодневном животу. Стручна литература која је највећим дијелом на страном језику приморава нас да је преводимо и прилагођавамо својим потребама. Током деценијског рада са студентима, кроз преглед њихових писаних радова примећена је

* Предавач, mariavukov@gmail.com

** Предавач, gvozdena2001@yahoo.com

велика употреба машинског преводиоца, веома често на начин који је био неприхватљив. Текстови који су на тај начин кориштени нису били адекватни за употребу и преводи су били веома лоши, такорећи сводили су се на просту употребу машинског преводиоца у смислу самог копирања датог превода без испављања и анализирања истог. Отуда се јавила идеја за спровођење једног пилот истраживања које се сводило на заједничку анализу текстова на часовима енглеског језика. Кроз такву анализу би и студентима постало јасно које се све грешке јављају приликом машинског превода што би допринело подизању свести о начину паметне употребе машинских преводиоца. Зарад лакшег препознавања грешака, одлучено је да се те анализе раде на кратким стручним текстовима који су били на енглеском језику, а тражени превод на матерњем – српском језику. Гугл преводилац (*Google Translate*), машински преводилац, као један од најдоступнијих, кориштен је за спроведено истраживање којем би се показало и да ли је могуће управљање утицајем машинског преводиоца на квалитет превода кратких стручних текстова са енглеског језика на српски кроз анализу грешака.

ТЕОРИЈСКИ ПРИСТУП: ПРЕВОЂЕЊЕ – МАШИНСКО ПРЕВОЂЕЊЕ

Језик је једно од најважнијих средстава комуникације међу људима. Како би комуникација била успешна, учесници се морају међусобно разумети, то јест треба да на исти начин интерпретирају оно што желе рећи. „Када би се превођење сводило само на однос између два језика, као два семиотичка система, онда би главни, ненадмашни и једини пример задовољавајућег превођења био двојезични речник. Али могло би се рећи да се то, у најбољем случају, коси са здравим разумом који речник сматра средством за превођење, а не преводом“. (Еко 2011: 39) Стога се превођење „не одвија између система већ између текстова“. (Еко 2011: 39) У том смислу можемо рећи да је превођење облик комуникације међу људима, јер је главна сврха превођења разумевање поруке. Термин превођење, као процес, подразумева сложене мисаоне радње које укључују тумачење изворног текста и преношење значења на циљном језику. Са развојем преводилаштва јавила се потреба за анализирањем начина и самих резултата превођења. (Сибиновић 1990) Готово је немогуће замислити процес превођења без размишљања, јер захтева неисприно знање граматике, синтаксе, семантике, вишезначности, идиома, сленга, жаргона и познавање многих других лингвистичких правила изворног језика као и познавање културе његових говорника. Исто такво знање је неопходно и за преношење значења на циљном језику. Треба напоменути још и то да свака врста превођења преводиоцу као захтев поставља поседовање специфичних знања и вештина. Ипак, постоји скуп знања и

вештина који се очекује од преводиоца без обзира на врсту премођења којом се бави. „У тај скуп улази способност декодирања (то јест, разумијевања) изворне поруке и њеног поновног кодирања (то јест, изражавања) у језику примаоца, затим способност вјерног и неискривљеног преношења обавијести, те способност за успостављање комуникацијске интеракције с примаоцима преведене поруке“. (Ивир 1978:16)

Идеја о употреби рачунарских програма за нови облик премођења-машинско премођење, стара је колико и први рачунарски системи. Први машински преводи који датирају из шесдесетих година XX века, убрзо после појаве првих рачунара, представљали су постојање унапред задате базе података у којој су се налазили одговарајући парови одређених језика, такви преводи су били засновани на једноставној замени речи једног језика речима другог језика. Унапређењем рачунара, односно рачунарских програма и машинско премођење је постајало боље, тако да се каснији модели нису ограничавали само на речи, већ су користили читаве фразе, што је знатно унапредило граматичку структуру текста и превод фраза код којих дословни преводи нису били могући. Машинско премођење је потпоље рачунарске лингвистике и представља употребу рачунарских софтвера у циљу премођења текста или говора са једног језика на други. Машинско премођење, као и све остале софтверске апликације, захтева програмирање. У овом случају састављање двојезичне или вишејезичне базе термина и израза, уношење граматичких правила, граматичких структура и великог броја примера на основу којих би систем могао одредити делове циљног текста који одговарају деловима изворног текста. Рачунарски програми за машинско премођење који се данас користе дозвољавају подешавање за одређене области или професије, на тај начин побољшавају решење ограничене базе дозвољених замена. Упркос почетном ентузијазму и очекивању да ће даљи развој машинског премођења донети релативно брзе и значајне резултате, машински преводи још увек нису достигли ниво тачности који би пружио потпуно аутоматизоване преводе. Међутим његов највећи значај огледа се у помоћи коју пружа преводиоцима, при чему резултати тачности зависе од типа текста који се преводи, у појединим случајевима се тако преведени текстови користе као готов производ. Ипак, актуелни системи још увек не могу произвести превод који је истог квалитета као мануални, тј. ручни превод, посебно уколико је реч о неформалном језику.

Машинско премођење аутоматизује лакши део преводачког процеса у најбољем случају ако имамо све горе наведено у виду, остало што захтева истраживање, решавање двосмислености, препознавање грешака и њихову обраду, за сад је нажалост недостижно без укључивања људског фактора. Не сме се рећи да је машинско премођење и рад на његовом усавршавању бескористан, машинско премођење може имати корисне резултате. Неки алати за машинско

превођење користе садржај доступан на интернету који обухвата милионе преведених докумената у циљу проналажења језичких образаца и успостављања везе између више језика. Једино се проблем може препознати у томе што се такви резултати могу постићи само у случају одређених типова текстова, уз примену контролисаног језика и завршну обраду коју би вршио човек.

Ако би препустили читав преводачки процес машини, занемарили би чињеницу да се комуникација одвија у смислу који разумеју људи. Што значи да би тако језик, који је друштвени феномен, био дехуманизован, одвојен од људи, а превођење би довели до математичке и статистичке операције, где се занемарују лингвистичка и друштвена својства. Дословни превод или превођење ван контекста, је неупотребљиво у бројним случајевима и може звучати неумесно и довести до неспоразума. Превод поезије такође није одговарајући употребом машинских преводаца, „поезија, као супериорнији облик синонимије много је тежа за превођење него што је случај са уобичајеном поруком“. (Рајовић 2019) „Уколико се нешто без тешкоћа пренесе из једног језика у други, то онда јесте превод, али никако није реч о поезији.“ (Громовић 2019)

Машине још дуго неће бити способне да разумеју, што нас наводи да закључимо да је будућност машинског превођења у пружању разумљиве основе коју ће даље дотерати преводиоци и лингвисти. Машински превод код књижевних текстова још дуго ће бити неупотребљив без улоге човека као преводиоца. „Читалац (па и књижевни теоретичар, који је „ангажовани читалац“), нужно се детерминише способношћу да прати сугестије аутора и да их актуелизује. Чин перцепције дела тако постаје сродан акту стварања, јер се њиме успостављају увек нова значења.“ (Стишовић 2018)

Развој информационих технологија ће вероватно омогућити бар донекле прихватљиве преводе у будућности, поставља се само питање када у будућности. Брзи развој вештачке интелигенције, створиће велику базу знања, речника и меморију и вероватно довести до много финијег и усклађенијег превода у различитим областима.

ОПИС ИСТРАЖИВАЊА

Истраживање спроведено за потребе овог рада односи се на анализу грешака које су се јавиле приликом само просте употребе машинског превођења на кратким стручним текстовима. Период истраживања обухвата две школске године у распону од три године. Тај распон је узет како би исти студенти могли бити укључени у истраживање. Први део истраживања се реализовано 2017. године када су студенти добили задатак да у машински преводац Г-огел унесу кратак стручни текст на енглеском језику и преведу га на српски језик. Имали смо десет група, односно десет текстова сличне дужине, оквирно 400 речи. Добијени превод су само налепили у ворд документ без икаквог

исправљања и обележили све грешке које су препознали као направљене. Добијене грешке смо поделили у пет група, и то: 1) граматичке грешке, 2) техничке грешке, 3) лексичке грешке, 4) стилске грешке и 5) остављено не-преведено. Други део истраживања спроведен је на потпуно исти начин код исте групе студената, на машинском преводу истих текстова само три школске године касније, 2019. године.

Кренули смо од претпоставке да машински преводиоци праве грешке и да је потребно одређено знање циљног језика како би се те грешке препознале, такође једна од полазних претпоставки је била да развој информатичких технологија доводи и до развоја и побољшања машинских преводиоца. У истраживању и изради рада коришћена је дескриптивна метода, затим метода контрастивне анализе где се упоређује изворни текст са преводом који смо добили. У оквиру контрастивне анализе коришћена је и метода анализе и синтезе где се на основу анализе теоријских приступа и издвојених грешака долази до одређених закључака. Резултати истраживања представљени су графички, бројчано и описно.

Период од три године је сасвим довољан временски распон на основу којег би могли увидети промене које су се десиле, наравно ограничавамо се на мали корпус, кратке стручне текстове али довољне да покажу да ли и у којем смеру се промене дешавају и колико је могуће управљати њима.

Подела грешака је проста и разумљива за све учеснике истраживања, имајући у виду слаб ниво познавања енглеског језика групе која је учествовала у истраживању, као и не баш завидан ниво познавања граматике матерњег језика.

Након добијених анализа грешака за свих десет текстова, узели смо просечни број грешака који се јављао у текстовима, јер нас је у истраживању занимало само да покушамо да утврдимо на основу временског распона од три школске године да ли број грешака варира, у којој групи и где се најчешће јављају. Лингвистичка анализа грешака није рађена, јер би као таква превазишла оквире овог рада.

РЕЗУЛТАТИ ИСТРАЖИВАЊА

Током првог дела истраживања спроведеног 2017. године дошли смо до следећег: укупан број грешака, узимајући у обзир просек на основу свих десет текстова, машинског превода стручног текста са енглеског језика на српски дужине од 400 речи је – 199 што је скоро 50% од укупног броја речи.

Током другог дела истраживања спроведеног 2019. године на истим текстовима код истих испитаника дошли смо до следећег: укупан број грешака, узимајући у обзир просек на основу свих десет текстова, машинског превода стручног текста са енглеског језика на српски дужине од 400 речи је – 54 што је скоро 13,5% од укупног броја речи. Евидентно је да је укупан број грешака

драстично смањен у току временског распона од три године, што наводи на закључак да је рачунарски програм за машинско преводјење унапређен.

Грешке које су се јавиле груписане су на следећи начин:

а) ГРАМАТИЧКЕ ГРЕШКЕ

Пре приказа резултата потребно је објаснити шта се све подразумева под граматичким грешкама.

Сви они глаголи који су преведени у погрешном глаголском времену (ограничено само на проста глаголска времена: прошлост, садашњост или будућност), затим именице које нису преведене у тачном роду или броју (мушки, женски или средњи род, односно једнина или множина), придеви који нису преведени у тачном степену поређења (позитив, компаратив или суперлатив). Као што је већ речено у опису, због слабог познавања како енглеског језика тако и граматике српског језика код испитаника, базирали смо се само на основе појмове који су свима били јасни како би могли препознати грешке у тексту.

Просечан број граматичких грешака, на основу анализе свих десет текстова у 2017. години, је је 26 што је 6,5% од укупног броја речи. Просечан број граматичких грешака на основу анализе свих десет текстова у 2019. години је био 6 што је 1,5% од укупног броја речи, односно за 5% мање у односу на 2017. годину (слика 1).

Слика 1 – Грамаатичке грешке у 2017. и 2019. години

б) ТЕХНИЧКЕ ГРЕШКЕ

Под техничким грешкама подразумевали смо правописне грешке (мало/велико слово, нови ред), интерпункцијске грешке (тачке, зарези и остало).

Од укупног броја речи ових техничких грешака је било најмање у односу на све остале грешке. На основу анализе свих десет текстова, дошли смо до просечног броја од 14 што износи 3,5% од укупног броја речи. Просечан број техничких грешака на основу анализе свих десет текстова у 2019. години је је 7 што је 1,75% од укупног броја речи, а то је за 1,75% мање у односу на 2017. годину (слика 2).

Слика 2 – Техничке грешке у 2017. и 2019. години

с) ЛЕКСИЧКЕ ГРЕШКЕ

Лексичке грешке се односе на оне речи у тексту које могу имати два или више значења, односно, које су преведене тако да не дају адекватан смисао реченици преведеној на наш језик. На пример, енглеска реч „good“ има два значења, може бити „добро“ и може бити „роба“.

Од укупног броја речи, просечан број лексичких грешака је 30 што је 7,5% од укупног броја речи преведеног текста. Просечан број лексичких грешака на основу анализе свих десет текстова у 2019. години је је 15 што је 3,75% од укупног броја речи или за 3,75% мање у односу на 2017. годину (слика 3).

Слика 3 – Лексичке грешке у 2017. и 2019. години

д) СТИЛСКЕ ГРЕШКЕ

Стилске грешке су сличне вокабуларним грешкама, само што се оне односе на више речи или на читаву реченицу, која је у преводу са енглеског језика на српски употребом машинског преводиоца, дала резултат такве реченице да она нема смисао и разумљивост приликом читања.

За разлику од претходних грешака, овај тип грешке се не упоређује у односу на укупан број речи, већ у односу на укупан број реченица које се налазе у преведеном тексту.

Стил писања енглеског језика се генерално разликује од стила српског језика, тако да се проценат грешака тешко може извести прецизно, јер скоро свака реченица може бити другачије преформулисана, како би била у духу

српског језика. Истраживањем и анализирањем текста уочено је 43 грешке у односу на 92 реченице које се налазе у преведеном тексту, што представља 46,7% од просечног броја реченица, што је статистички релативно велики број грешака које гугл преводилац прави приликом превођења. Просечан број стилских грешака на основу анализе свих десет текстова у 2019. години је 12 што је 3% од укупног броја реченица. Што је за 43,7% мање у односу на 2017. годину (слика 4).

Слика 4 – Стилске грешке у 2017. и 2019. години

е) ОСТАВЉЕНО НЕПРЕВЕДЕНО

„Остављено непреведено“ се односи на оне речи које је машински преводилац изоставио приликом обраде текста, тј. приликом његовог превођења.

Статистички гледано укупан број грешака које је преводилац направио приликом обраде је 86 што у односу на просечан број 400 постојећих речи у тексту, представља 21,5% од укупног броја речи у тексту. Просечан број речи које су остављене непреведене на основу анализе свих десет текстова у 2019. години је 14, што је 3,5% од укупног броја речи. Што је за 18% мање у односу на 2017. годину. Овде треба нагласити да је 7 истих речи остало непреведено и у 2017. и у 2019. години што је 1,75% од укупног броја речи. Намеће се закључак да база података у којој су смештени еквиваленти за ова два језика није адекватно ажурирана (слика 5).

Слика 5 – Остављено непреведено у 2017. и 2019. години

ЗАКЉУЧНА РАЗМАТРАЊА

На основу спроведеног истраживања намеће се закључак да временски распон од три године показује изузетно побољшање превода, број грешака које је машински преводацац направио је драстично смањен, самим тим и квалитет превода је много бољи. Питање које се намеће још од шездесетих година XX века је да ли ће машинско преводјење заменити људе као преводаце и даље остаје нејасно. Тренутно је то још увек немогуће, али са брзим развојем информационих технологија, односно вештачке интелигенције, машинско преводјење ће сигурно бити знатно успешније. Машине не разумеју значења, рачунарски програм само замењују један текст другим на основу информација које су им људи унапред одредили програмом. Превод без размишљања и разумевања у најбољем случају може бити добар, ако се испуне бројни предуслови или као случајност.

Управљање утицајем машинског преводјења на квалитет превода је једино могуће кроз константно усавршавање рачунарских програма за машинско преводјења, као и сталним ажурирањем базе података у таквим програмима. Посебно треба обратити пажњу на оне машинске преводаце који су доступни свима и који су бесплатни, јер се ствара утисак да су они занемарени у односу на оне који се користе на вишим лествицама друштва.

Подизањем свести код корисника интернета о грешкама које се јављају приликом употребе машинских преводаца у многоме може допринети контроли таквих превода, и ограничити преузимање истих као тачних и даље ширење лоших превода.

Главна разлика између преводаца и рачунарског програма за машинско преводјење огледа се у чињеници да се човек несвесно прилагођава свим околностима. Човек има способност да инедоречен, непрецизан и двосмислен текст и једино он може адекватан начин прилагодити текст циљној сврси, препознати непостојање одговарајуће граматичке категорије или термина те грешке и одлучити како да се с њима поступи. Једино се човек може определити за одређени синоним на основу свог знања или искуства. Начин преводјења преводаца заснива се на великом броју одлука, животном искуству и разумевању значења, што се од машине не може очекивати. Машине често не доносе праву одлуку и резултат машинског превода у већини случајева није оно што се може сматрати природним језиком. „Знање“ и речника код машина је ограничено на оно што су програмери успели унети. Када причамо о недостацима машинског преводјења списак је поприличан, али машинско преводјење свакако има и позитивних страна и великих потенцијала. Првенствено текст који је машински преведен треба да послужи као полазна основа коју даље дорађују преводаци и лингвисти.

Употребу машинских преводиоца не можемо ограничити само на оне кориснике који познају добро, пре свега свој матерњи језик, па онда и онај на који преводе жељени текст, али можемо сталним истраживањима и показима резултата истих подићи свест код корисника о грешкама које се јављају приликом таквих превода. Употреба машинских преводиоца је корисна у смислу управљања временом јер у многومه скраћује време потребно за превод текста, али само уз велику опрезност у смислу исправљања, а пре свега препознавања грешака које он прави.

ЛИТЕРАТУРА

- Громовић 2019: Милан Громовић. Есејистика Момчила Настасијевића у мелодијској спирали звука и језика. *БАШТИНА*, св. 47.
- Еко 2011: Умберто Еко. *Казати тојово исиу сивар*. Београд: Paideia.
- Ивир 1978: Владимир Ивир. *Теорија и техника преводјења*. Уџбеник за I годину позивноусмереног образовања и васпитања средњег ступња преводачке струке. Карловачка гимназија, Сремски Карловци.
- Рајовић 2019: Јелена Рајовић. Елементи кохезије у оригиналу и преводу поезије на примеру песме „Стар си, оче Виљаме“ из романа алиса у земљи чуда. *БАШТИНА*, св. 47.
- Сибиновић 1990: Миодраг Сибиновић. *Нови оријинал: Увод у преводјење*. Београд: Научна књига.
- Стишовић 2018: Ана Миловановић Стишовић. Комуникација са пишчевом имажинацијом. *БАШТИНА*, св. 46.

Marija D. VUKOVIĆ

Jelena D. RAJOVIĆ

MANAGING THE IMPACT OF MACHINE TRANSLATION
ON THE QUALITY OF THE TRANSLATION OF THE TECHNICAL TEXTS
FROM ENGLISH INTO THE SERBIAN LANGUAGE

SUMMARY

The use of Internet machine translators is a common and frequent practice in everyday life, as one of the most accessible and at the same time the easiest way to translate technical literature and other variety of texts. In addition, the number of published translations containing many errors increases on a daily basis. This paper presents a pilot study which is based on ten students' papers dealing with the analysis of machine translations of short, identical, technical texts in the time frame of three school years. The aim of this paper is to investigate the reliability of machine translators in the past and present and whether there are any changes in translations and which. The types of errors as well as percentage to what extent they occurred were analysed. The reverse translation was performed and based on that the results presented in the paper are obtained. The period of three years has shown that the differences in translation are significant. This indicates that the machine translation management is possible, especially in the field of information technology development and mutual co-operation of linguists and information technology experts. The cooperation in the domain of machine translation is necessary and must be permanent.

Key words: English language, management of the impact, machine translator, use of internet translator, error analysis, quality of translation.