

Kontrastivna analiza prijevoda na engleski i hrvatski jezik uporabom Google prevoditelja

Matijević, Maja

Undergraduate thesis / Završni rad

2019

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Electrical Engineering, Computer Science and Information Technology Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:200:720941>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-01-03**

Repository / Repozitorij:

[Faculty of Electrical Engineering, Computer Science and Information Technology Osijek](#)



**SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU
FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA I
INFORMACIJSKIH TEHNOLOGIJA**

Sveučilišni studij

**KONTRASTIVNA ANALIZA PRIJEVODA NA
ENGLJSKI I HRVATSKI JEZIK UPORABOM GOOGLE
PREVODITELJA**

Završni rad

Maja Matijević

Osijek, 2019.

SADRŽAJ

| | |
|---|----|
| 1. UVOD | 1 |
| 1.1. Zadatak završnog rada..... | 1 |
| 2. STROJNO PREVOĐENJE | 2 |
| 2.1. Razvoj i proces strojnog prevođenja..... | 2 |
| 2.2. Vrste strojnog prevođenja | 4 |
| 2.3. Primjena i problemi strojnog prevođenja..... | 5 |
| 3. GOOGLE PREVODITELJ | 7 |
| 3.1. Funkcije i značajke..... | 7 |
| 3.2. Prednosti i ograničenja | 8 |
| 4. ANALIZA PRIJEVODA GOOGLE PREVODITELJA | 9 |
| 4.1. Analiza prijevoda znanstvenog rada s engleskog na hrvatski jezik..... | 9 |
| 4.2. Analiza prijevoda znanstvenog rada s hrvatskog na engleski jezik..... | 14 |
| 4.3. Analiza prijevoda popularno-znanstvenog rada s engleskog na hrvatski jezik | 20 |
| 4.4. Analiza prijevoda popularno-znanstvenog rada s hrvatskog na engleski jezik | 24 |
| 5. ZAKLJUČAK | 29 |
| LITERATURA | 30 |
| SAŽETAK | 32 |
| ABSTRACT | 32 |
| ŽIVOTOPIS | 33 |

1. UVOD

Za problem nastalih grešaka u prijevodu koristeći strojni prevoditelj i dalje je potrebno pronaći sva rješenja iako se tehnologija strojnoga prevođenja znatno razvila u posljednjih nekoliko godina. Bitan čimbenik koji pomaže u razvoju i poboljšanju prijevoda strojnih prevoditelja jest analiza i dodavanje novih mogućnosti, riječi i izraza prevoditelju. U ovom će se završnom radu prikazati usporedna analiza prijevoda koristeći alat za strojno prevođenje. Bit će razmotreni razlozi nastajanja najčešćih grešaka te potencijalni načini za popravak istih. Strojno prevođenje ima vrlo široku uporabu u svijetu, a prijevod velikih količina teksta često dovodi do smanjenja točnosti prijevoda i pogreške u komunikaciji ako se prijevod koristi u te svrhe. U današnje vrijeme postoji velik broj alata za strojno prevođenje, ali je Google prevoditelj bez sumnje ispred konkurencije u području strojnog prevođenja. Uzevši to u obzir, Google prevoditelj najbolji je izbor za analizu trenutnih mogućnosti modernih alata za strojno prevođenje.

Kroz ovaj će se rad objasniti razvoj i proces strojnog prevođenja te pojedine vrste uz njihove prednosti i nedostatke. Razmotrit će se primjena strojnog prevođenja i mane koje ono sadrži. Bit će opisane osnovne karakteristike *Google Translator* alata i njegova široka primjena.

U glavnom će se dijelu rada analizirati četiri različite skupine teksta od kojih su dvije na engleskom, a dvije na hrvatskom jeziku. Nakon unošenja teksta u prevoditelj, analizirat će se dobiveni prijevodi uz naglasak na nastale greške i propuste strojnog prevoditelja. S obzirom na kompleksnost hrvatskog jezika i pojedine izraze u oba jezika koje nije moguće doslovno prevesti, nego je potrebno prepoznati kontekst rečenice, pretpostavka je da će doći do grešaka koje ni jedan strojni prevoditelj još nije uspješno riješio.

1.1. Zadatak završnog rada

Zadatak je završnog rada analizirati prijevode s engleskog na hrvatski jezik i s hrvatskog na engleski jezik koristeći Google prevoditelj. Za potrebe analize koristit će se dva znanstvena i dva popularno-znanstvena rada, pri čemu je jedan od svakog na engleskom, a drugi na hrvatskom jeziku.

2. STROJNO PREVOĐENJE

Strojno je prevođenje proces prevođenja određenog polaznog teksta s jednog jezika na drugi korištenjem računalnog sustava, s ljudskom pomoći ili bez nje [1]. Na najjednostavnijoj razini, strojno prevođenje uzima riječi iz jednog jezika i traži im ekvivalente u drugom, no takav način rijetko dovodi do točnog prijevoda. Sustavi za strojno prevođenje mogu biti dizajnirani za jedan par jezika ili više od dvaju jezika te mogu prevoditi u jednom ili u oba smjera. U svrhu rješavanja problema prijevoda razvijaju se sve napredniji strojni prevoditelji koji koriste različite tehnike, koje će u ovom radu biti dodatno pojašnjene, za dobivanje boljeg prijevoda.

2.1. Razvoj i proces strojnog prevođenja

Pojam strojnog prevođenja u prvotnom smislu postojao je kao ideja još u 17. stoljeću, kada su Descartes i Leibniz stvorili ideju univerzalnog jezika pomoću kojega bi ekvivalentne ideje različitih jezika označili istim simbolom. No, tek u 20. stoljeću došlo je do razvoja praktičnih rješenja strojnog prevođenja. 1949. godine u memorandumu američkog znanstvenika i matematičara Warrena Weavera pojavio se pojam “strojni prijevod” [2]. On je bio prvi koji je spomenuo mogućnost prijevoda dokumenata između različitih jezika koristeći računalo. Taj je memorandum bio iznimno utjecajan jer je formulirao ciljeve i metode strojnog prevođenja prije nego što je opća populacija bila svjesna širokih mogućnosti računala. Time je započeo znatno brži razvoj strojnog prevođenja pa je već u sljedećih nekoliko godina došlo do brojnih istraživanja u tom području. Prvi simpozij o strojnom prevođenju održan je 1952. godine na MIT-u gdje je postalo jasnije na koji će način napredovati buduća istraživanja. Raspravljalo se o temama kao što su rješavanje problema sintakse, potreba za ljudskim faktorom pri prevođenju prije ostvarenja potpuno točnog automatskog prijevoda te rješavanje ostalih problema koji mogu nastati u prijevodu, primjerice dvosmislenosti riječi. Nakon toga, bilo je potrebno demonstrirati tehničku mogućnost strojnog prijevoda, što je u siječnju 1954. godine rezultiralo prvom javnom demonstracijom sustava strojnog prevođenja. Odabran je određen broj rečenica na ruskom jeziku koje su prevedene na engleski koristeći vrlo ograničen rječnik od samo 250 riječi i 6 gramatičkih pravila. Iako nije značajno doprinijela razvoju strojnog prevođenja, ta je demonstracija bila dovoljno zanimljiva i pokrenula je druge značajnije projekte u svijetu. U drugom razdoblju razvoja strojnog prevođenja (60-ih godina 20. stoljeća), u izvješću ALPAC-a (Automatic Language Processing Advisory Committee) donesen je 1966. godine zaključak o strojnom prevođenju. Ono je utvrđeno kao neisplativo, sporo i dvostruko skuplje od ljudskog prijevoda te je takav zaključak imao negativne posljedice na daljnji razvoj strojnog prevođenja [3].

Sedamdesetih godina 20. stoljeća, iako je razvoj strojnog prevođenja i dalje bio slab, nada je ipak rasla sa stvaranjem nekoliko sustava i projekata fokusiranih na strojno prevođenje, prvenstveno u Japanu, Rusiji i SAD-u. Do 1980. godine, još se dodatno povećao broj i raznolikost instaliranih sustava za strojno prevođenje. Tih su godina japanske tehnološke tvrtke znatno napredovale, a nastojeći prestići konkurenciju uvodile su programsku podršku za prijevod s engleskog i na engleski jezik. Do kraja spomenutog desetljeća, došlo je do razvoja brojnih novih metoda strojnog prevođenja. Makoto Nagao, japanski informatičar, i skupina znanstvenika s kojima je radio na IBM-u razvili su sustav za strojno prevođenje koji je baziran na statističkim metodama. Prijevod se dobiva na temelju velikog broja primjera prevedenih rečenica. Mana je takvog načina prijevoda zanemarivanje sintaktičkih i semantičkih pravila jezika [4]. Potom, osnovane su brojne tvrtke koje su se bavile strojnim prevođenjem, uključujući tvrtku Trados, koja je prva osnovala tehnologiju prevođenja pomoću prijevodne memorije.

Strojno se prevođenje na Internetu pojavilo 1996. godine kada je SYSTRAN ponudio besplatan prijevod kratkih tekstova, a 1997. pod njihovim vodstvom pojavio se i AltaVista Babelfish koji je dosegao do 500 000 zahtjeva za prijevod dnevno. Taj je alat predviđen za prijevod mrežnih stranica u željeni jezik. Rezultati su, naime, znatno varirali u ovisnosti o kompleksnosti teksta. Upravo zbog toga mnogi su mediji na primjeru alata Babelfish i sličnih alata više stavili naglasak na nedostatke, nego na prednosti strojnog prevođenja [5]. Bez obzira na to, napredak strojnog prevođenja raste. 2005. godine pojavila se prva mrežna stranica za automatski prijevod. 2012. godine Google je objavio informaciju kako Google Translate dnevno prevodi količinu teksta ekvivalentnu milijunu knjiga.

Kako bi razumjeli čimbenike koje mora sadržavati relativno uspješan strojni prevoditelj, potrebno je proučiti ljudsko prevođenje. Ono se sastoji od dvaju koraka, a to su dekodiranje značenja izvornog teksta te ponovno kodiranje u željeni jezik. Iza tih dvaju jednostavnih koraka kriju se kompleksne operacije. Kako bi se izvorni tekst pravilno razumio, ljudski prevoditelj mora analizirati sve njegove karakteristike i poznavati zakonitosti tog jezika kao što su gramatika, semantika, sintaksa ili idiomi. Osim poznavanja svega spomenutog na izvornom jeziku, za točan je prijevod potrebno poznavati i jezik na koji želi prevesti tekst. Svi bi se ti čimbenici trebali na neki način uspješno prenijeti na računalo kako bi ono moglo prevesti tekst i dobiti rezultat koji je sličan ljudskom prijevodu. To i danas nije u potpunosti moguće iako sve više napreduje. Prijevodi postaju sve točniji, no bez ljudske pomoći strojni prevoditelji i dalje nailaze na neizbježne greške zbog kompleksnosti jezika.

Proces strojnog prevođenja unutar nekog sustava strojnog prevođenja odvija se u više koraka. Za početak je potreban unos teksta koji se želi prevesti. Taj se tekst može podijeliti u više kategorija, ovisno o težini prijevoda. Pojedine rečenice mogu biti nejasne, izražavati pretpostavke ili biti dvosmislene te tako mogu znatno otežati proces prevođenja. Drugi je korak svojevrsno formatiranje teksta. Izvorni jezik može sadržavati pojedine znakove, dijagrame i simbole koje nije potrebno prevoditi pa je bitno da se takvi dijelovi izdvoje od onih koji će se prevoditi. Zatim se dodatno uređuje tekst, a neki sustavi dijele duže rečenice u kraće radi jednostavnosti prevođenja. Na kraju napredniji sustavi obavljaju morfološku i sintaktičku analizu teksta, koliko je to moguće. Pomoću tih analiza pokušava se doći do optimalnog rezultata kako bi se kontekst rečenice pravilno preveo. U sustavima za strojno prevođenje koriste se formalne gramatike koje odgovaraju korištenim jezicima [6]. Uzevši u obzir sve karakteristike nekog jezika, može se dobiti mnogo točniji prijevod. Bez obzira na sve korištene metode, rezultati su po efikasnosti i točnosti prilično daleko od ljudskog prijevoda.

2.2. Vrste strojnog prevođenja

S razvojem strojnog prevođenja rastao je broj vrsta strojnog prevođenja u postojanju. Bit će spomenute neke osnovne vrste strojnog prevođenja [7].

Prva je vrsta statističko strojno prevođenje (eng. *Statistical Machine Translation* - SMT), koje je već prethodno spomenuto u ovom radu. U takvoj vrsti prevođenja radi se s velikom broju višejezičnih (izvorni tekst i prijevodi) i jednojezičnih tekstova. Sustav traži statističke poveznice između izvornih tekstova i prijevoda, kako za cijele ulomke, tako i za kraće izraze i time gradi prijevodni model. Takav način prijevoda ne uzima u obzir pravila danog jezika niti gramatiku, no najčešće je korišten način strojnog prevođenja danas. Koristi se i kao jezgra sustava koje upotrebljavaju Google Translate i Bing Translator. Dobre su strane takvog sustava to što nije predviđen ni za jednu specifičnu skupinu jezika, postoje velike količine višejezičnih tekstova koje se mogu usporediti i nije potrebno definirati lingvistička pravila za pojedine jezike. S druge strane, takav je prijevod dobar samo za osnovne potrebe i ne donosi uvijek željene rezultate, posebice ako se tekst koji prevodimo znatno razlikuje po sadržaju od onoga na kojem je sustav učio. Primjerice, sustav koji je za učenje prijevoda koristio tehničke tekstove imat će problema prilikom prevođenja tekstova u razgovornom stilu.

Sljedeća je vrsta strojnog prevođenja zasnovana na gramatičkim pravilima (eng. *rule-based machine translation* - RBMT). Takvi sustavi koriste gramatiku i rječnike nekog jezika kako bi pokrili semantičko, morfološko i sintaktičko područje jezika. Koriste se skupinom određenih

pravila pojedinog jezika kako bi došli do točnog prijevoda. Takvi sustavi trebaju duže vrijeme kako bi postigli određenu točnost s obzirom da je potrebno uložiti puno posla i truda u definiranje mogućih pravila. Također, takvi prevoditelji ne mogu se uspješno suočiti s tekstovima koji sadrže žargon ili metafore. Zbog toga su ovakvi sustavi uglavnom zamijenjeni statističkim ili hibridnim prevoditeljima koji su mnogo efikasniji, ali se i dalje koriste pri prijevodu tekstova kod kojih ne postoje velike količine tekstova za razvoj statističkih sustava.

Strojno prevođenje temeljeno na primjerima (eng. *example-based machine translation* - EBMT) još je jedna vrsta strojnog prevođenja. To je metoda u kojoj prevoditelj “uči” na parovima paralelnih tekstova napisanih na dvama jezicima i traži analogiju između istih. Kada je potrebno prevesti neki tekst, prevoditelj traži primjere slične onome u tekstu i prema njima donosi prijevod. S obzirom da je teško pronaći podudarne cijele rečenice, takav se prevoditelj uglavnom koristi dijelovima rečenice i izrazima iz nekih primjera kako bi došao do prijevoda nove rečenice. Što je rečenica sličnija onoj koju prevoditelj ima kao primjer, to će točnost prijevoda biti veća.

Postoji i hibridno strojno prevođenje koje je mješavina više metoda prevođenja, najčešće statističkog prevođenja (SMT) i prevođenja zasnovanog na gramatičkim pravilima (RMBT). Pri tome, koristi prijevodnu memoriju, što poboljšava kvalitetu prijevoda. No, takav način ima puno nedostataka, od kojih se najviše ističe stalna potreba za uređivanjem teksta, što znači da je ljudska uključenost u proces neophodna.

Zadnja je vrsta strojnog prevođenja koja će biti spomenuta strojno prevođenje temeljeno na neuronskim mrežama (eng. *Neural Machine Translation* - NMT). Pri takvom načinu prevođenja koriste se trenirane neuronske mreže kako bi se dobili željene rezultate. Neuronske mreže rade na principu ljudskog mozga koristeći uzorke i strukture kako bi zadani tekst bio što točnije preveden. Iako je takav način prijevoda relativno nov, izazvao je veliku zanimaciju među stručnjacima jer vidljivo poboljšava rezultate, stvarajući fluidnije rečenice koje sve više izgledaju kao ljudski prijevod. Do 2016. godine neuronske su mreže počeli koristiti svi veći sustavi strojnog prevođenja. Danas model prijevoda pomoću neuronskih mreža koriste Google, Microsoft, Yandex i PROMT.

2.3. Primjena i problemi strojnog prevođenja

Bez obzira na određene nedostatke strojnog prevođenja, ono je i dalje široko korišteno diljem svijeta. Sve se više ulaže u razvoj sustava za strojno prevođenje.

Strojnim se prevoditeljima pojedinci služe u privatne svrhe, pri tome najčešće koristeći besplatne alate dostupne široj populaciji kao što je Google prevoditelj. Sustave strojnog prevođenja koriste i brojne tvrtke za korisničku podršku, rudarenje podataka ili objavu sadržaja na više jezika, pri čemu bi zapošljavanje ljudskog prevoditelja bilo neisplativo zbog količine teksta. Zatim, strojno prevođenje koristi i vlada za prikupljanje podataka te poboljšanje kvalitete ljudskih prijevoda pri prevođenju dokumenata na više jezika. Isto tako, strojni prevoditelji mogu se pokazati korisnima i u vojne svrhe, prvenstveno za jednostavniju komunikaciju vojske s civilima. Dakako, postoje brojne druge primjene strojnog prevođenja pomoću kojih i ono samo sve više napreduje.

Unatoč prednostima koje sustavi strojnog prevođenja donose, postoje i vidljivi problemi. Glavni je problem loša kvaliteta prijevoda, a može nastati iz više razloga. Najveće razlike između izvornog teksta i prijevoda nastaju pri prijevodu jezika složenih gramatika. Prevoditelju je problem shvatiti kontekst neke rečenice ili fraze, osobito ako je specifična za izvorni jezik, a ne postoji u sličnom obliku u drugom jeziku. Osim toga, višeznačnice mogu predstavljati još jedan problem. S obzirom da računalo nema ljudsku kreativnost i osjećaj za kulturu, ne može prepoznati neke vrlo bitne značajke jezika.

3. GOOGLE PREVODITELJ

Google prevoditelj (eng. *Google Translate*) internetska je usluga tvrtke Google. Nudi besplatno višejezično strojno prevođenje unutar internetskog preglednika. Neke od usluga Google prevoditelja uključuju i mobilne aplikacije za Android i iOS, online sučelje te API koji pomaže programerima u razvoju proširenja za preglednike i kreiranju aplikacija.

3.1. Funkcije i značajke

Google prevoditelj usluga pokrenuta je u travnju 2006. godine kao usluga za statističko strojno prevođenje. U početku je koristila zapise Ujedinjenih naroda i Europskog parlamenta za prikupljanje lingvističkih podataka. Najprije se prevodio tekst na engleski jezik, a zatim na ciljani jezik. Prilikom prijevoda, Google prevoditelj traži uzorke među milijunima dokumenata kako bi došao do boljeg prijevoda.

Google prevoditelj podržava preko 100 jezika svijeta. Osim vrlo poznatog prevoditelja unutar internetskog preglednika, koji je korišten za primjere koji će se prevoditi u nastavku ovog rada, Google prevoditelj sadrži i druge načine prijevoda, primjerice putem mobilne aplikacije, prijevod slike, govora, ručno pisanog teksta, cijelih dokumenata ili pak mrežnih stranica.

Google prevoditelj unutar internetskog preglednika sadrži, osim mogućnosti prijevoda, mogućnost slušanja izgovora originalnog i prevedenog teksta. Također, moguće je odabrati opciju "Otkrij jezik", ako korisnik nije siguran na kojem je jeziku izvorni tekst. Osim običnog unosa teksta preko tipkovnice, moguće je uključiti i glasovni unos. Ako se prevodi samo jedna riječ, moguće je dobiti uvid u definiciju te riječi na izvornom jeziku te pored prijevoda učestalost pojavljivanja svakog od mogućih prijevoda te riječi.

Google prevoditelj od 2010. je godine integriran u Google Chrome preglednik kao prevoditelj za automatski prijevod mrežne stranice. U studenom 2016. godine, Google je objavio prelazak Google prevoditelja na sustav za prevođenje pomoću neuronskih mreža (GNMT) koji prevodi pune rečenice, a ne dio po dio. Takav sustav koristi širi kontekst rečenice kako bi došao do najbližeg značenja, a zatim rečenice prilagođava i sortira prema gramatičkim pravilima ciljanog jezika. U početku je bilo podržano samo nekoliko jezika za takav sustav, ali GNMT se danas koristi za sve više i više jezika. Za sada je ovaj sustav moguć samo za prijevod nekog jezika s engleskog i na engleski jezik, a ostali parovi jezika još nisu podržani.

3.2. Prednosti i ograničenja

Jedna je od prednosti Google prevoditelja to što je usluga besplatna. Prijevod teksta od strane stručnjaka obično je skup, a u slučajevima kada prijevod ne mora biti savršen, Google prevoditelj može poslužiti kao dobar izvor. Nadalje, ovaj je alat brz. U nekoliko sekundi, na ekranu se pojavi prijevod dužeg teksta, što je čovjeku prevoditelju nemoguće ostvariti. Također, Google prevoditelj podržava velik broj jezika i odličan je za osnovne potrebe komunikacije.

No, ovaj sustav i dalje sadrži mnogo grešaka. Često se u prijevodu gubi značenje i smisao rečenice jer ne postoji način za shvaćanje konteksta rečenice. Što je tekst kompleksiji, veće su šanse da će nastati greška i besmislen prijevod. Osim toga, prevoditelj često stvara gramatičke greške, osobito kada se gramatike dvaju jezika znatno razlikuju [8]. Kvaliteta prijevoda ovisi i o paru jezika koje se koristi za prijevod s obzirom da su u bazi podataka koju prevoditelj koristi neki jezici mnogo više zastupljeni nego drugi. Jezici zapadnih država, kao što su engleski i španjolski, daju najtočnije rezultate, dok je točnost prijevoda afričkih jezika najlošija, a potom azijskih te europskih.

4. ANALIZA PRIJEVODA GOOGLE PREVODITELJA

4.1. Analiza prijevoda znanstvenog rada s engleskog na hrvatski jezik

Za potrebe analize prijevoda znanstvenog rada s engleskog na hrvatski jezik, odabran je izvorni znanstveni članak na engleskom jeziku [9].

Primjer 1

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|--|--|
| The first suggested classification is into static (traditional carpooling as the first type of problem) and dynamic carpooling, and then the classification of the dynamic into casual (the second type of problem) and technology facilitated dynamic carpooling (the third type of problem). | Prva predložena klasifikacija je statična (tradicionalni carpooling kao prva vrsta problema) i dinamičko carpooling, a zatim klasifikacija dinamičkog u povremeni (drugi tip problema) i tehnologija olakšava dinamično carpooling (treći tip problema). |

U prvom primjeru, već na samom početku prevedene rečenice, nastao je svojevrsan propust. Riječ *into* strojni je prevoditelj u potpunosti zanemario u prijevodu te pridjev *statična* izravno dodijelio imenici *klasifikacija* time ju opisujući, što je utjecalo na značenje rečenice. Iako je izostavljanje validna prijevodna strategija [10], treba odmjeriti kada ju koristiti, što strojni prevoditelj, očito, nije mogao učiniti. Drugim riječima, za pravilno razumijevanje bilo je potrebno iskazati podjelu klasifikacije na statički i dinamički suprijevoz. Zatim, strojni je prevoditelj prethodno spomenutu imenicu *carpooling* na više mjesta u rečenici ostavio u njenom izvornom obliku. Iako se u hrvatskome jeziku često koriste anglizmi, u ovome primjeru ipak je poželjno taj izraz zamijeniti hrvatskim prijevodom *suprijevoz* jer nema razloga zašto bi se ostavilo u izvornom obliku ako važeći prijevod postoji. Ista se tehnika strojnoga prevođenja (korištenje imenice u izvornome obliku i na izvornome jeziku) spominje i u radu "*Održavanje baza alata za strojno prevođenje*" [11] gdje je prikazano, da uz *Google*, i drugi alati za strojno prevođenje, točnije *Bing* i *Yandex*, koriste istu tehniku posebno kada se radi o prijevodu novijih tehničkih riječi koje još nisu ušle u hrvatski rječnik. Štoviše, u spomenutom radu prikazano je da alat za strojno prevođenje *Yandex*, prilikom prijevoda riječi koje nema u svom unutarnjem rječniku, ostavlja te riječi u pismu ćirilici (*Yandex* je ruski alat za strojno prevođenje) čak i u slučaju kada postavljeni ciljni jezik ne koristi ćirilčno pismo, čime posve onemogućava

razumijevanje riječi čitateljima koji ne koriste ćirilčno pismo. Sljedeći se prijevodni problem pojavljuje kod umetnutoga dijela teksta (u ovome slučaju zagrada). Strojni prevoditelj pri svakome korištenju zagrada u rečenici nailazi na problem nakon točnog prevođenja skupine riječi prije zagrada te riječi unutar zagrada. Nakon prevođenja teksta unutar zagrada, ponašanje je prevoditelja kao na početku rečenice, točnije strojni prevoditelj nije uspio povezati dio rečenice prije i poslije zagrade. Tako se u konačnici dobiva nepovezan i netočno preveden posljednji dio rečenice i to se posebno ističe u zadnjemu dijelu rečenice gdje je umjesto glagolskog pridjeva trpnog *tehnologijom olakšani*, korišten glagol *olakšava*. Također, umjesto korištenja pridjeva *dinamički suprijevoz*, s obzirom da se radi o kategorizaciji i opisivanju vrste suprijevoza, strojni je prevoditelj ponudio prilog *dinamično carpooling*, čime se izgubilo pravo značenje rečenice.

Primjer 2

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|--|---|
| The technology facilitated dynamic carpooling lies in the focus of this study. | Tehnologija omogućila je dinamično zajedničko korištenje automobila u fokusu ove studije. |

U drugome primjeru, slično kao u prvome, strojni prevoditelj ima problema s prepoznavanjem vrsta riječi u rečenici ovisno o kontekstu rečenice, što je greška koja se neprestano ponavlja. U engleskome se jeziku riječ *facilitate* najčešće koristi kao glagol, no u određenim slučajevima, kao što je vidljivo u ovome primjeru, ona stoji u ulozi pridjeva. Prevoditelj takvu riječ u svakoj vrsti rečenice interpretira kao glagol i upravo na sličnim mjestima nastaju najveće greške. Uzevši u obzir spomenutu grešku, strojni prevoditelj zanemaruje jedini glagol analizirane rečenice *lies* bez kojeg se ne može shvatiti pravo značenje rečenice. Nije jasno zašto prevoditelj povremeno izostavi riječi iz prevedene rečenice. Zasižno takav način selektivnog odabira riječi za prijevod ne može donijeti ispravne rezultate osim ako prijevod sam po sebi ne zahtijeva gubljenje riječi sa svrhom boljeg razumijevanja prevedenog teksta, što u ovoj situaciji nije slučaj. U primjeru je strojni prevoditelj trebao prepoznati cijeli prvi dio rečenice (do predikata) kao subjekt rečenice. Umjesto toga, došlo je do raznih grešaka u koju je uključena i ona netočnog poretka riječi u rečenici, a uz sve to, govorniku hrvatskog jezika prevedena rečenica bila bi poprilično neshvatljiva i neprirodna. Budući da se slične sintagme mogu koristiti kao dvije riječi ili sa spojnicom, pretpostavili smo da bi dodavanjem spojnice

olakšali posao strojnome prevoditelju naznačujući da riječi *technology* i *facilitated* pripadaju jedno drugome i nedvojbeno ukazujući da riječ *facilitated* ne može biti glagol. Stoga smo, iz znanstvenih razloga, izvornoj rečenici promijenili oblik zadajući strojnome prevoditelju istu izvornu rečenicu, samo s *technology-facilitated* oblikom. Međutim, prijevod je ostao isti, odnosno čak ni nakon dodavanja spojnice i vrlo jasne naznake o zajedništvu dviju riječi, strojni prevoditelj to nije mogao prepoznati i točno prevesti. Ono što je zanimljivo jest to da prevoditelj u ovom primjeru riječ *carpooling* prevodi u *zajedničko korištenje automobila*, a u prethodnome ju je primjeru ostavio u izvornome obliku, što ukazuje na nedosljedno prevođenje istih riječi.

Primjer 3

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|--|---|
| These two groups of respondents represent the active population, considered to be commuting every day in the Zagreb transport network. | Ove dvije skupine ispitanika predstavljaju aktivno stanovništvo koje se svakodnevno svakodnevno vozi na prometnu mrežu u Zagrebu. |

U primjeru 3, pola je rečenice prevedeno na korektan način. Razlog tomu vjerojatno leži u jednostavnosti i običnom poretku riječi u engleskom jeziku koji, kada se prevede na hrvatski, ostaje nepromijenjen. Neobjašnjiva pogreška nastala je u drugome dijelu rečenice kod prijevoda priložne oznake vremena *every day*. Iz nepoznatoga je razloga, strojni prevoditelj dva puta preveo danu priložnu oznaku *svakodnevno* iako se time nije utjecalo na razumijevanje rečenice. Nadalje, strojni je prevoditelj na kraju rečenice ubacio prijedlog *na* koji u prijevodu nije potreban. Ako se izbaci taj prijedlog, mogući bi prijevod bio *...svakodnevno vozi zagrebačkom prometnom mrežom*, no onda bismo od prevoditelja očekivali uspješnu pretvorbu vlastite imenice *Zagreb* u pridjev *zagrebački*. Predloženi bi prijevod bio najbolji mogući i vrlo vjerojatno odabir ljudskoga prevoditelja, a zbog kompleksnosti pretvorbe dviju vrsta riječi, teško je bilo očekivati da strojni prevoditelj ponudi tako napredan prijevod. Iako, bilo bi vrlo zanimljivo vidjeti je li strojni prevoditelj u stanju ispravno koristiti veliko i malo početno slovo, odnosno prebaciti se s korištenja velikog početnog slova u imenu *Zagreb* na malo početno slovo u pridjevu *zagrebački* budući da se pridjevi izvedeni od sličnih imena, različito koriste u engleskom i hrvatskom jeziku (English i engleski). Ako ne očekujemo toliko napredan prijevod, dostatna zamjena za nj mogla bi glasiti *...vozi prometnom mrežom u Zagrebu*. Uz takav bi se prijevod s lakoćom moglo razumjeti o čemu se radi u rečenici. Međutim, i u tome prijevodu

nailazimo na problem neadekvatnoga prijevoda i iskrivljavanja semantike rečenice. Naime, glagol *commute* odnosi se na vožnju/putovanje s/na posao. Drugim riječima, glagol *drive/voziti* svojevrsni je hiperonim u odnosu na glagol *commute/putovati na posao*, stoga uporabom glagola *voziti* iskrivljuje se značenje rečenice jer vožnja ne znači nužno putovanje s/na posao, što je svrha rečenice.

Svaki je jezik vrlo fleksibilan i većinom za kompleksnije rečenice postoji više mogućih načina prijevoda, no strojni prevoditelj, unatoč velikom napretku posljednjih godina, još nije dovoljno razvijen kako bi u svakoj situaciji odabrao prijevod koji će biti prilagođen pravilima hrvatskog jezika. Potrebno je dodatno “poučiti” strojni prevoditelj kako bi mogao lakše prevesti rečenicu u ispravnom kontekstu i pravilno prepoznati redosljed rečeničnih dijelova.

Primjer 4

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|--|--|
| The younger the carpooling user, the greater the benevolence of using the new and advanced technologies. | Što je mlađi korisnik carpooling, to je veća naklonost korištenja novih i naprednih tehnologija. |

U primjeru 4, imenica *carpolling* ponovno se koristila u izvornome obliku. Razlog ovakvoga postupka ostaje nejasan posebice zato što je ta riječ opetovano korištena u izvorniku i opisno prevedena u prijevodu, stoga je nejasno kako je programiran program za strojno prevođenje i zašto se pri svakome korištenju iste riječi ne isporučuje njen prijevod. Dakle, postavlja se pitanje na kojemu principu radi program za strojno prevođenje i je li programirani način ili je taj broj nasumičan. Iako je jedan od prijevoda imenice *benevolence naklonost*, u nastavku se rečenice više trebao uzeti u obzir kontekst rečenice, što bi ljudski prevoditelj zasigurno učinio, i taj dio rečenice prevesti koristeći se tehnikom skraćivanja na način *...to je veće korištenje/to se povećava upotreba...*, odnosno bez upotrebe imenice *naklonost*. Nakon čitanja cjelokupne rečenice, uviđa se potreba za drugačijim rasporedom riječi u rečenici na njenome početku gdje je trebalo reći *Što je korisnik zajedničkoga automobila/vožnje/suprijevoza mlađi...* kako bi se postigao učinak istovjetan izvorniku. Drugim riječima, red riječi u rečenici izvornika i prijevoda ne treba (i ne može) uvijek biti jednak, a prednost uvijek treba dati prenošenju učinka i poruke pred strukturom rečenice i teksta.

Primjer 5

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|--|---|
| The purpose of the performed research is to improve the connections of the users of joint rides by passenger cars by using the advanced technologies and services based on them (ICTS and LNTS). | Svrha provedenog istraživanja je poboljšati povezanost korisnika zajedničkih vožnji osobnim automobilima korištenjem naprednih tehnologija i usluga na njima (ICTS i LNTS). |

Peti primjer pokazuje napredak alata za strojno prevođenje u usporedbi s prethodnim primjerima. S obzirom da je poredak riječi u izvornoj rečenici prikladan hrvatskom jeziku, nema mnogo mjesta na kojima bi moglo doći do greške. Pogreška koju bi tek prvostupnik/magistar hrvatskoga jezika i/ili lektor za hrvatski jezik primijetio jest u obaveznom poretku riječi u rečenici na hrvatskom jeziku, odnosno postavljanju zanaglasnica nakon prve naglašene riječi. Stoga, rečenica bi trebala glasiti *Svrha je provedenog istraživanja poboljšati...* Međutim, ovo je nešto na čemu se još treba raditi, a iako ne utječe na značenje rečenice, ne treba zanemariti na standardnost jezika i potrebu suradnje prvostupnika/magistara hrvatskoga jezika i/ili lektora za hrvatski jezik i programera.

Nadalje, sljedeći je prijevodni problem opet u domeni ciljnoga jezika. Iako je dio *...povezanost korisnika zajedničkih vožnji osobnim automobilima...* točan prijevod izvorne rečenice, na hrvatskom jeziku on može biti dvosmislen. Može se shvatiti da korisnici zajedničkih vožnji već koriste osobne automobile za tu svrhu, a može i da će se povezanost poboljšati tako što će korisnici zajedničkih vožnji početi koristiti osobne automobile u svrhu zajedničkih vožnji. Dakle, red bi se riječi u rečenici trebao reorganizirati kako bi se izbjegla mogućnost dvosmislenosti. Primjerice, moglo bi se reći *...poboljšati povezanost korisnika zajedničkih vožnji upotrebom osobnih automobila...*; međutim, još nismo naišli na slučaj da alat za strojno prevođenje primjenjuje tehniku dodavanja jezičnih jedinica u svrhu jasnijega prijevoda, što je nešto na čemu se treba raditi. Na samom kraju rečenice, *...services based on them...* prevedeno je kao *...usluga na njima...*, umjesto *...usluga temeljenih na njima....* Zbog toga je u prijevodu kraj rečenice pomalo nejasan jer se ne zna na što se odnosi sintagma *na njima*. Značenje predloženoga i potrebnoga prijevoda nije isto jer se *usluge na njima* izravno pružaju na nečemu, dok se *usluge temeljene na njima* samo temelje na nečemu, a pružatelj može biti netko drugi. No, čak i sa spomenutom greškom, prijevod je mnogo razumljiviji govorniku hrvatskog jezika nego dosadašnji primjeri.

4.2. Analiza prijevoda znanstvenog rada s hrvatskog na engleski jezik

Za analizu prijevoda s hrvatskog na engleski jezik odabran je još jedan znanstveni rad, čiji je izvorni jezik hrvatski [12].

Primjer 6

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|-----------------------------|-------------------------------|
| Postoji više pristupa radu. | There is more access to work. |

U šestome primjeru, strojni prevoditelj dostavlja netočan prijevod vrlo kratke i jasne rečenice. Predikat rečenice *Postoji* prevodi u izraz *There is*, a kod prijevoda ostatka rečenice u potpunosti zanemaruje sve prethodne i sljedeće riječi. Rečenica je prevedena ne obraćajući pažnju na smisao i povezanost riječi u rečenici. Kada bi program za strojno prevođenje sagledao rečenicu u cijelosti, vjerojatno bi se učestalost sličnih grešaka znatno smanjila. U ovom se primjeru lako uočava prijevod riječ-po-riječ, a takva vrsta prijevoda u većini je slučajeva loša praksa. Strojni prevoditelj ne vodi računa o broju imenica; uz govor o postojanju “više nečega”, to nešto ostavlja u jednini. Drugim riječima, prekršena je jednostavna gramatička sročnost subjekta i predikata, što je vrlo čest slučaj prilikom prijevoda grupa jezika koji imaju slaganje riječi u rečenici po rodu, broju, padežu i drugim gramatičkim kategorijama (hrvatski) i onih koji to nemaju (engleski). Iako ne utječe na samo razumijevanje rečenice, gramatički je netočan prijevod, stoga bi programeri, radeći zajedno s jezičarima, trebali dodati sve sročne oblike u unutarnje rječnike strojnih prevoditelja.

Nadalje, imenica se *access* najčešće odnosi na pristup u fizičkome smislu, stoga bi praktičnije bilo koristiti imenicu *approach* jer na bolji način opisuje pristup određenom zadatku ili poslu i kolokacijski je češća s imenicom *zadatak/posao*. Izvađena iz konteksta, *rad* je imenica koju je vrlo teško točno prevesti čak i ljudskome prevoditelju jer se rad može odnositi na *posao*, *djelatnost*, *tvorevinu*, a može biti i znanstveni/stručni rukopis. Ključan je za točan prijevod kontekst. U ovome je primjeru strojni prevoditelj dobro prepoznao o kojoj se vrsti rada radi i točno je preveo imenicu.

Primjer 7

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|---|---|
| Tradicionalan pristup oblikovanju uveza | A traditional approach to shaping a tattoo, |

| | |
|--|--|
| umara, što izravno uzrokuje obavljanje posla bez razmišljanja. | which directly causes the job to be done without thinking. |
|--|--|

U sedmome primjeru nastaje neobjašnjiva greška u prijevodu. Imenica *tattoo* pojavljuje se u prijevodu iako u izvornoj rečenici ne postoji ni približno slična riječ. Ovdje se postavlja pitanje kako je strojni prevoditelj došao do te riječi i zašto je upotrijebio nasumičnu riječ umjesto onih koje se nalaze u rečenici. U ovome bi slučaju čak i prijevod riječ-po-riječ bio mnogo bolja opcija od umetanja nasumičnih riječi, najvjerojatnije korištenih u drugim primjerima s kojima se prevoditelj susreo. Drugog objašnjenja nemamo jer se *tetoviranje* eventualno može povezati s *urezivanjem*, no ovdje nije riječ o *urezivanju*, nego *uvezivanju* koje se automatski vezuje uz knjige. Nadalje, kolokacija imenice *tattoo* i glagola jest *napraviti (do)*, *imati (have)*, *ukloniti (remove)* i sl., odnosno imenici *tattoo* ne pripisuje se glagol *shape (oblikovati)*, što ukazuje na to da se neke riječi u rečenice doslovno prevode (u ovome slučaju glagolska imenica *oblikovanje*), dok se umjesto nekih koriste, naizgled, nasumično odabrane nimalo sličnoga značenja imenice (*tattoo*). Nadalje, temeljni se glagol iz rečenice (*umarati*) posve izostavlja u prijevodu, što utječe na značenje rečenice koje je nepotpuno budući da u hrvatskome jeziku rečenica ne može funkcionirati bez radnje. Drugi je dio rečenice relativno razumljiv u odnosu na prvi jer je poredak riječi u izvornoj rečenici prikladniji engleskome jeziku.

Primjer 8

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|---|--|
| Unošenjem u posao svaki pojedinac nastoji oblikovati uvez dajući mu odgovarajuću vanjsku formu i vodeći računa o samom sadržaju djela koje uvezuje. | By entering into a job each individual seeks to formulate a binding by giving it an appropriate external form and taking into account the content of the tying itself. |

U navedenome primjeru riječ *entering* ne pripada kontekstu ove rečenice jer se takav prijevod ponajviše odnosi na ulazak u fizičkome smislu. Stoga bi u prijevodu bolje bilo upotrijebiti prikladniju riječ pomoću koje bi značenje postalo značajno jasnije. Takve riječi mogle bi biti glagoli *to get into* ili *to engage in*. Slično tome, *formulate* je riječ koja čitatelja obavještava o oblikovanju misli i ideja, a ne fizičkih stvari. Poželjno je pri stvaranju prijevoda iskoristiti izdašan engleski rječnik i prethodno spomenutu riječ zamijeniti drugom pravog značenja za rečenicu iz primjera. Samo su nekoliko mogućih riječi *form*, *shape* i *style*. Sredina

rečenice zadovoljava uvjete dostatno razumljivog prijevoda, ali na kraju rečenice iznova dolazi do vidljivih nedostataka u prijevodu. Vidljiv je još jedan primjer izgubljene riječi u prijevodu, a ovoga je puta to imenica *djela*. Strojni je prevoditelj shvatio rečenicu na netočan način i umjesto da se spomene djelo koje uvezuje, on je spomenuo vezivanje samo po sebi. Izostavljen je bitan dio rečenice. Prema tome, dobiva se sveukupno neadekvatan prijevod.

Nadalje, u ovome primjeru imamo i slučaj nesistematičnog korištenja istoga vokabulara. Naime, u izvornoj se rečenici sistematično koristi terminologija *uvezivanja* knjiga u svim morfološkim oblicima riječi, dok se u prijevodu koriste *binding* i *tying* koji označavaju neku vrstu vezanja. Osim potrebe za sistematičnim korištenjem, ostaje pitanje koji će se od dvaju predloženih oblika koristiti. Ovdje se ponovno mora uzeti u obzir kontekst teksta jer će se tek tada moći odlučiti za riječ *binding* budući da je ta riječ u kolokaciji s imenicom *knjiga*, dok se *tying* povezuje uz drugačiju vrstu vezanja koje ne mora nužno biti fizičkoga oblika.

Primjer 9

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|--|--|
| Bitno obilježje e-knjiga je mogućnost čitanja offline, odnosno nakon što je knjiga preuzeta na uređaj za čitanje, više nije potreban internet. | An essential feature of an e-book is the ability to read offline, or after the book is downloaded to a reader, no longer an internet connection. |

Promatrajući prijevod ovog primjera, odmah je uočljiva pogrešna uporaba broja imenice. Imenica *e-knjige* u izvornoj rečenici korištena je u genitivu množine, no strojni ju je prevoditelj interpretirao kao imenicu u nominativu jednine s obzirom da su dva navedena oblika morfološki jednaka. Prema tome, strojni prevoditelj nije dovoljno razvijen kako bi mogao prepoznati padeže hrvatskog jezika, nego se i dalje prvenstveno vodi doslovnim prijevodom svake riječi za sebe uzimajući temeljni oblik svake riječi (nominativ jednine u slučaju uporabe imenica). Riječ *odnosno* ne bi trebala biti prevedena rastavnim veznikom *or* koji se obično koristi u rečenicama u kojima je ostvariva jedna od dviju mogućnosti. U ovome primjeru, to bi značilo da prvi i drugi dio rečenice nisu istovjetni, a važan cilj pri prevođenju bio bi dobiti dojam da se drugim dijelom rečenice dodatno pojašnjava prvi. U tome bi slučaju izraz *that is (to jest, točnije)* ili latinski izraz *id est (i.e.)* kao prijevod za *odnosno* puno bolje odgovarao smislu rečenice. Zatim, *uređaj za čitanje* preveden je u *reader*, ali s obzirom da se radi o uređaju za čitanje e-knjiga, jasniji prijevod bio bi *e-book reader* ili jednostavno *e-reader*. Bez obzira na to, nije teško shvatiti o

čemu se radi i s dostavljenim prijevodom. Ipak, posljednji dio rečenice nije toliko razumljiv. Glagol *to be needed* je izostavljen i kraj nije povezan sa sredinom rečenice na koju se odnosi. Prijevod ..., *internet is no longer needed* jedna je od nekoliko prihvatljivih opcija. Štoviše, prijevod strojnog prevoditelja i korištenje sintagme pridjeva i imenice *internetska veza* umjesto samo imenice *internet* pravilnija je uporaba u hrvatskome jeziku, stoga je ovdje strojni prevoditelj odradio bolji posao nego što mu je bilo zadano i/ili se očekivalo.

Primjer 10

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|---|---|
| Iako je zastupljenost e-knjige na tržištu velika, ispitanici ipak preferiraju tiskanu verziju jer knjiga na njih djeluje opuštajuće dok je listaju i drže u rukama. | Though the presence of e-books on the market is great, respondents prefer the printed version because the book on them is relaxing while they are rolled and held in their hands. |

Postoji više načina prijevoda veznika *iako* na engleski jezik, a jedan je od njih zasigurno *though*; međutim, ne u svim kontekstima. *Though* je veznik koji se najčešće koristi u kombinaciji s *as* (*as though*) ili *even* (*even though*). S obzirom da je u izvorniku veznik *iako*, prijevod bi trebao biti *even though* ili *although*, a nikako samo *though*. Iako pogrešna uporaba veznika ne utječe značajno na semantiku rečenice, utječe na njen stil, a svakako se treba voditi računa da se prijevodom ne narušavaju ni stilistička obilježja jezika. I u ovome se primjeru nastavlja prijevod riječ-po-riječ. Ponovno je imenica *e-knjige*, koja je jednaka u nominativu množine i genitivu jednine, prevedena u pogrešnome padežu i broju, odnosno jednako kao i u prethodnome primjeru. Nadalje, imenice *zastupljenost* (*representation*) i *prisutnost* (*presence*) nemaju isto značenje, odnosno moglo bi se reći da je imenica *pristupnost* hiperonim u odnosu na imenicu *zastupljenost* koja označava udio prisutnosti nečega, što ukazuje na pogrešan prijevod navedene imenice. Pridjev *great* obično opisuje nešto dobro, a rjeđe veliko, pa je bolje iskoristiti bogatstvo rječnika i upotrijebiti, primjerice, riječi *wide* ili *large* jer kolokacijski bolje odgovaraju danome kontekstu.

Najveći je problem u drugome dijelu rečenice. Prva je velika greška u prijevodu sintagme *na njih*. U izvornoj se rečenici želi reći kako knjiga djeluje na ispitanike, no dobiveni je prijevod doslovan i u njemu knjiga više ne djeluje na njih, već se nalazi na njima. Taj bi se dio znatno mogao popraviti uz prijevod *...has a relaxing effect on them...*, iako bi dio prijevoda koji je

prevoditelj dostavio (*is relaxing*) bio dovoljno razumljiv kada bi ostatak rečenice točno preveo. Razlog ovim pogreškama koje narušavaju i semantiku i sintaksu prijevoda jest u korištenju prijevodne tehnike riječ po riječ. Posljednji je dio prijevoda također problematičan jer se mijenja perspektiva rečenice i ljudi (ispitanici) od vršitelja radnje postaju trpitelji radnje listanja i držanja knjiga, što niti ima smisla niti je sintaktički ispravno. Iz prijevoda se ne može razaznati tko su to *they*; jesu li oni ispitanici ili dvije skupine. Ako su dvije skupine od kojih su jedna ispitanici, onda ne saznajemo tko je druga skupina. Ako su pak to ispitanici, onda su oni istovremeno i vršitelji i trpitelji radnje listanja knjige, što je nemoguće. Osim toga, glagol *listati* (knjigu) na engleski bi se jezik preveo s *browse, scan, skim* i sl., no nikako s *roll*, što je pogreška na leksičkoj razini.

Primjer 11

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|---|--|
| Posljednjih desetak godina elektroničke knjige polako zamjenjuju tiskane. | The last ten years of electronic books are slowly replacing printed. |

Iako kratka, ova rečenica obiluje prijevodnim nepravilnostima. Točnije, ovako prevedena, rečenica nema smisla i nikako ne prenosi izvorno značenje. Prva je pogreška u uporabi prijedloga *of*, koji je nepotreban. Može se zaključiti da strojni prevoditelj ni ovu rečenicu nije u potpunosti sagledao, nego je prvih nekoliko riječi koje označavaju vrijeme povezo sa sintagmom *elektroničke knjige*. Dakle, korištenje prijevodne tehnike dodavanja u ovome slučaju nije opravdano niti točno. Do dodavanja prijedloga *of* u ciljnome jeziku došlo je zbog jednojezičnih sintaktičkih pravila ciljnoga jezika. Naime, tri su temeljna načina uporabe priložnih oznaka vremena u engleskome jeziku – korištenjem prijedloga *of*, kao u primjeru 11, *for* ili *since*. Korištenje prijedloga *of* i *since* u ovome kontekstu nije prikladno tako da se trebao koristiti prijedlog *for* ...*For the last ten years...*

Nadalje, u izvornome je jeziku korišteno glagolsko vrijeme prezenta i nesvršeni vid glagola *zamjenjivati*. Hrvatski se i engleski jezik razlikuju u upotrebi glagolskih vremena pa tako sadašnje glagolsko vrijeme u hrvatskome jeziku ne odgovara uvijek sadašnjemu glagolskom vremenu u engleskome jeziku. Drugim riječima, u engleskome jeziku postoji glagolsko vrijeme *present perfect* koji se, između ostaloga, koristi za radnje koje su započele u prošlosti i traju u sadašnjosti, odnosno upravo za radnje poput ove u primjeru 11. Stoga je prijevod trebao glasiti ...*electronic books have been replacing...* Upotreba načinskog priloga *slowly* doslovni je, iako

relativno točan, prijevod priloga *polako*. Međutim, ljudski bi se prevoditelj možda odlučio na upotrebu priloga *gradually* jer postoji razlika između tih dvaju načinskih priloga. Naime, prilog *slowly* označava brzinu izvođenja neke radnje, dok prilog *gradually* označava brzinu (sporo) izvođenja neke radnje tijekom nekog dužeg razdoblja, dakle upravo ono što se želi reći u rečenici. Iako se neki prijevodni problemi alata za strojno prevođenje mogu doraditi i/ili riješiti s vremenom, uočavanje sitnih značenjskih razlika među riječima i upotreba točnih leksema vjerojatno je nešto što se neće još dugo (ako ikad) moći doraditi kod strojnoga prevođenja.

Primjer 12

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|--|--|
| Učenici i studenti, osobe koje čitanje, pisanje i učenje smatraju svojim životnim pozivom, najčešći su konzumenti elektroničke knjige. | Students and students who read, write, and learn people as their livelihood are the most common e-book consoles. |

Imenice hrvatskoga jezika *učenici* i *studenti* na engleski se jezik prevode jednakom imenicom *students*. Upravo je to strojni prevoditelj i učinio, ali kako su se obje riječi našle u jednoj rečenici, poprilično je zbunjujuće jer dolazi do ponavljanja iste riječi i nije jasno na koga se koja odnosi. Potrebno je učiniti razliku između dviju riječi s jednakim prijevodom. Moguće je riječ *učenici* promijeniti u *pupils* kako bi se razumjelo o kome se govori. Druga je mogućnost dodati pridjev koji pojašnjava razinu školovanja poput *high school students*, *university* ili *college students*. S obzirom da to ovdje nije bio slučaj, nema opravdanja za dvostruku uporabu imenice *students* u ciljnome jeziku, a i čitatelj ne bi mogao shvatiti da se radi o dvjema grupama ljudi, nego bi vjerojatno pomislio da je došlo do pogreške.

Sljedeća je greška nastala kada strojni prevoditelj nije prepoznao glagolske imenice *čitanje*, *pisanje* i *učenje*, nego su one prevedene u glagole, što bi bilo u redu da je ostatak rečenice točno preveden, a red riječi u rečenici reorganiziran. Poredak je riječi u rečenici izmijenjen tako da je imenica *osobe* izmještena u sredinu rečenice, a ne odmah nakon učenika i studenata. Pomiješan redoslijed riječi i kriva interpretacija vrste riječi dovodi do besmislenog prijevoda ostatka rečenice. Ovako prevedena rečenica mogla bi imati značenje da postoje neki učenici ili studenti koji čitaju, pišu (vjerojatno bi se mislilo da pišu o nekome) i uče (izvorni bi se govornik ovo mogao samostalno protumačiti u poučavaju) neke ljude, a to im je životni poziv i izvor prihoda, što je nemoguće jer, osim honorarnoga, ni učenici ni studenti ne mogu imati izvor prihoda. Osim

netočne sintakse, prijevod je pojedinih riječi besmislen. Tako *životni poziv* ako ga pravilno prevedemo glasi *life calling*, a prijevod koji je prevoditelj ponudio jest *livelihood*, pri čemu se obično misli na izdržavanje, točnije način na koji netko preživljava i zarađuje sredstva neophodna za život. Još je jedan netočan prijevod onaj imenice *konzumenti*, koja nije najspretnije korištena ni u izvornome jeziku. Međutim, neovisno o tome, ne pronalazimo motivaciju za uporabu imenice *konzole* (consoles) koja se u tekstovima tehničke naravi odnosi na upravljački igrački uređaj, dok se u tekstovima iz područja građevine i arhitekture odnosi na potporni stup. Dakle, neovisno o području iz kojega se u unutarnju bazu alata za strojno prevođenje učitano značenje spomenute imenice, njeno korištenje u ovakvome kontekstu niti je opravdano niti je objašnjivo.

4.3. Analiza prijevoda popularno-znanstvenog rada s engleskog na hrvatski jezik

Popularno-znanstveni rad pomoću kojeg je analiziran prijevod s engleskog na hrvatski jezik uključuje dva članka svjetski poznatih američkih dnevnih novina *The New York Times* [13], [14].

Primjer 13

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|---|---|
| She can still look sources in the eye while typing. | I dalje može tražiti izvore u oku tijekom tipkanja. |

U trinaestome je primjeru glagol *look* preveden u *tražiti* iako je njegovo pravo značenje u ovome primjeru *gledati*. Ljudski bi prevoditelj to mogao zaključiti na temelju konteksta, a posebice jer je u engleskome jeziku općepoznat idiom *to look someone in the eye* (gledati nekoga u oči). Takve i slične izraze bilo bi dobro uvrstiti u bazu alata za strojno prevođenje kako bi se mogli adekvatno prepoznati i koristiti. Doslovni prijevod fraza, frazema, poslovice i izreka većinom ne može dati njihovo pravo značenje jer su takvi izrazi najčešće ukorijenjeni u jezik iz kojeg potječu i prijevod im je na ciljni jezik značajno drugačiji [15]. Dalje, imenica *sources* odnosi se na novinarske izvore pa je dobiveni prijevod *izvore* zadovoljavajući, ali ne bi bilo loše naglasiti na koje se izvore odnosi. No, teško je bez uzimanja u obzir prethodnih rečenica znati o

čemu se točno radi pa je dobiveni prijevod te riječi sasvim dovoljan za potrebe razumijevanja zadane rečenice. Hrvatski se i engleski jezik sintaktički, između ostaloga, razlikuju i u potrebi eksplicitnoga korištenja subjekta izraženoga zamjenicom. Drugim riječima, u hrvatskome se jeziku najčešće takav subjekt može implicitno izraziti nastavcima na predikatu/glagolu, dok to nije slučaj u engleskome jeziku. Međutim, to nije uvijek slučaj ni u hrvatskome jeziku. Naime, ako je rečenica izuzeta iz konteksta, kao u primjeru 13, nemoguće je nedvojbeno utvrditi vršitelja radnje. Također, ako se u prethodnome tekstu ne spominje taj vršitelj radnje, svakako ga treba eksplicitno upotrijebiti, što se ovdje nije dogodilo, stoga vršitelj radnje u izvornome i ciljnome jeziku nije jednak, odnosno iz glagolskih nastavaka ne možemo sa sigurnošću tvrditi je li vršitelj radnje u ciljnome tekstu muškoga ili ženskoga spola. Eksplicitna uporaba vršitelja radnje u hrvatskome jeziku nije pogrešna, iako je rjeđa, stoga ne znamo zašto alat nije programiran tako da u većini slučajeva eksplicitno koristi vršitelja radnje jer tada ne bi dolazilo do dvojbi ili pogrešaka u određivanju vršitelja radnje.

Primjer 14

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|---|---|
| It's harder to get someone in a powerful position on the phone. | Teže je nekoga pozvati na moćnu poziciju. |

Drugi primjer rečenice popularno-znanstvenog rada sadrži već viđenu grešku. Loša interpretacija rečenice dovela je alat za strojno prevođenje do povezivanja pogrešnih dijelova rečenice. Nastala je rečenica u kojoj se izraz *powerful position* prevodi u akuzativ umjesto lokativ jednine. Tim prijevodom mijenja se ukupno značenje rečenice i referiranje na nekoga iz osobe koja je na moćnoj poziciji u osobu koja se poziva da bude na moćnoj poziciji. Iako morfološki i sintaktički korektna, rečenica u ciljnome jeziku krši semantička pravila i onemogućava ikakvo razumijevanje čitatelju. Dio *...on the phone* gubi se iz rečenice te, osim što se u prijevodu ne saznaje na koga se odnosi glagol, nije razumljiv ni način na koji se vrši glagolska radnja. Redosljed riječi u rečenici na engleskom i hrvatskom jeziku ne treba i ne može ostati isti pri prevođenju. Prevedena rečenica ima značenje da se teško motivira pojedinac da preuzme neku moćnu ulogu umjesto značenja da se pojedinac koji obnaša neku važnu i moćnu funkciju teško može dobiti na telefon. Izostavljanje dijelova rečenica kao i parafraziranje pokazale su se kao neprikladne prijevodne tehnike u ovome primjeru.

Primjer 15

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|-------------------------------|----------------------------------|
| Are Australians embracing it? | Jesu li ga Australci prihvaćaju? |

U primjeru jednostavne upitne rečenice, dobiveni je prijevod doslovan i bez obraćanja pažnje na vrijeme u kojemu se odvija radnja. Pomoćni glagol *to be*, odnosno *Are* izravno je preveden u *Jesu li*, što je točno ako je izvađeno iz konteksta, odnosno ostatak rečenice u kojemu se nalazi glagol potpuno je zanemaren. Upitne rečenice engleskoga jezika ne tvore se na isti način kao one hrvatskoga pa se ne može prevoditi svaku riječ zasebno. U engleskome jeziku, dvije su temeljne vrste upitnih rečenica – one koje počinju *wh-* upitnim riječima (*what, where, when, why, which, whom, whom, whose, how*) i one koje počinju pomoćnim glagolom (*do, does, did, to be, to have, itd.*). S druge strane, u hrvatskome se jeziku upitne rečenice tvore punim oblikom pomoćnih glagola *biti* ili *hijeti* i česticom *li* (jesu li) ili upotrebom glagolskoga predikata ili glagolskoga dijela imenskoga predikata i čestice *li* (prihvaćaju li). Potonja se opcija koristi kada je god to moguće, a ni u kojemu se slučaju ne kombiniraju obje opcije (jesu li + prihvaćaju), što je učinjeno u primjeru 15. Pravilan bi prijevod započinjao glagolom *i*, sukladno pravilima hrvatskog jezika, trebao bi glasiti *Prihvaćaju li ga/to Australci?*. Ako se u alat za strojno prevođenje unese rečenica *Australians are embracing it*, dobiveni je prijevod je ispravan. Zbog toga su upitne rečenice područje na kojemu bi se mnogo više poraditi za napredak strojnog prevođenja s engleskoga na hrvatski jezik.

Primjer 16

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|---|--|
| A digital revolution has transformed China into a place where nearly anything — financial services, spicy takeout, manicures and dog grooming, to name a few — can be summoned with a smartphone. | Digitalna revolucija pretvorila je Kinu u mjesto gdje se gotovo ništa - financijske usluge, pikantno uzimanje hrane, manikure i dotjerivanje pasa - mogu nazvati pametnim telefonom. |

Početak rečenice dobro je preveden, a riječju *anything* započinju greške. Upravo je ta riječ prevedena s *ništa*, a trebalo bi biti baš suprotno *sve*. Morfološki, *anything* je neodređena zamjenica koja se koristi u pitanjima i negacijskim rečenicama, a značenje joj je *nešto, bilo što* ili *sve*, no nikako *ništa*, stoga čudi ovaj odabir alata za strojno prevođenje. Kada se samo

zamjenica *anything*, izdvojena iz konteksta, zada istom alatu za strojno prevođenje, točno ju prevede neodređenom zamjenicom *nešto*, što ukazuje na to da do prijevodnih problema dolazi prilikom stavljanja zamjenice u kontekst. Možemo samo nagađati koji kontekstualni dio utječe na prijevod. Pretpostavljamo da se radi o prilogu *nearly* (*gotovo/skoro*) iako ne vidimo motivaciju za takvo što.

Spicy takeout odnosi se na *pikantnu hranu za van* ili *za ponijeti*, a u prevedenoj je rečenici ta imenica pretvorena u glagolsku imenicu (*uzimanje hrane*). Čak i da koji slučajem čitatelj uspije zaključiti da se uzimanje hrane odnosi na hranu za ponijeti, izostavljanjem pridjeva *spicy* (*pikantan/začinjen*) mijenja se značenje toga dijela rečenice. Iako je izostavljanje relevantna prijevodna tehnika, u ovakvim je slučajevima nedopuštena jer izravno utječe na promjenu značenja rečenice. Osim toga, sintagma *to name a few* (*da spomenemo nekoliko, itd.*) ignorirana je i izostavljena iz prijevoda bez ikakvoga razloga. Iako ne utječe na značenje rečenice, ne smije ju se izostavljati jer se u suprotnom postiže drugačiji učinak nego kod izvorne rečenice.

Prijevodni problemi započeti zamjenicom *anything* i crticom nastavljaju se nakon druge crtice. U izvornoj je rečenici dio rečenice unutar crtica umetnut, no alat za strojno prevođenje to ne prepoznaje i ponovno dostavlja doslovni prijevod preostalog dijela rečenice. Sveukupno, najveći je problem upravo taj kraj rečenice za koji je neophodno imati točan prijevod inače prijevod prvog dijela ne doprinosi razumijevanju cijele rečenice. Glagol *summon* na hrvatski se jezik prevodi glagolima *pozvati/dozvati* i eventualno *nazvati*, ali kada je kontekst rečenice takav da *nazvati* označava *pozivanje*, a ne *imenovanje*. Drugim riječima, potrebno je uzeti u obzir kontekst rečenice ako se glagolu *summon* želi pridati značenje *nazivanja*, što ovdje nije učinjeno jer da jest, vidjelo bi se da kontekst ne dopušta uporabu glagola *nazvati*, nego samo *pozvati*. Upravo su takve razlike još uvijek nešto što razlikuje ljudskoga prevoditelja od alata za strojno prevođenje.

Primjer 17

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|---|--|
| Chen Haokai, the co-founder of SmartAHC, said farmers don't really need facial recognition. | Chen Haokai, suosnivač SmartAHC-a, kaže da poljoprivrednici ne trebaju priznanje lica. |

Alat za strojno prevođenje ostavlja osobno ime u izvornome obliku, što je dobro. Također, pravilno prepoznaje potrebu za iskazivanjem naziva tvrtke u genitivu. No, glagol *said* u

izvorniku je u prošleme glagolskom vremenu, a alat ga za strojno prevođenje koristi u sadašnjemu glagolskom vremenu. Međutim, ako uzmemo u obzir mjesto objavljivanja ovoga popularno-znanstvenoga rada (dnevne novine) i obilježja publicističkoga stila (prednjačenje sadašnjega glagolskog vremena), onda je ova modifikacija i izbor glagolskoga vremena opravdan. Greška koja upada u oči u ovoj rečenici jest *priznanje lica* kao prijevod za *facial recognition*. Točan prijevod *prepoznavanje lica* često se koristi u današnje vrijeme pa nije sasvim jasno zašto ga alat za strojno prevođenje nije tako preveo i zašto je, uz ostale riječi koje mogu poslužiti kao pomoć, došao do takvog prijevoda.

4.4. Analiza prijevoda popularno-znanstvenog rada s hrvatskog na engleski jezik

Pri analizi prijevoda popularno-znanstvenog rada s hrvatskog na engleski jezik, korišten je popularno-znanstveni rad o kriptovalutama, izvorno na hrvatskom jeziku [16].

Primjer 18

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|--|--|
| Drugo, kako im objasniti Internet, programiranje računala, aplikacije za Android, 3D printere i slično kad oni ni struje ni telefona nikad vidjeli nisu? | Secondly, how to explain them to the Internet, computer programming, apps for Android, 3D printers and the like when they neither electricity nor phone have ever seen are not they? |

Dva objekta prvoga dijela zadane upitne rečenice međusobno su zamijenjena, što je rezultiralo greškom u dijelu rečenice koji glasi *...how to explain them to the Internet, computer programming...* umjesto *...how to explain the Internet, computer programming...to them*. Dakle, izravni se objekt (*Internet*) trebao postaviti na prvo, a neizravni (*them*) na drugo mjesto u ciljnome jeziku za razliku od redoslijeda riječi u rečenici izvornoga jezika. To je već samo po sebi značajno promijenilo značenje rečenice jer ovako prevedeno značilo bi da se netko (*them*) treba objasniti internetu, programiranju i sl. koji se antropomorfiziraju, što je besmisleno. Nakon nabiranja, alat za strojno prevođenje nastavlja prevoditi jednu po jednu riječ tehnikom

doslovnoga prijevoda zbog čega nastaje *apps for Android*, što jezično nije netočno, ali se ne koristi u ovome obliku, nego u obliku *Android apps*.

Nadalje, sastavna rečenica hrvatskoga jezika povezana niječnim veznicima *ni* najčešće odgovara izrazima *neither/nor* u engleskome jeziku, stoga se dio *...ni struje ni telefona nikad vidjeli nisu?* trebao prevesti s *...neither the electricity nor the telephone have they seen?* Kada su god ovakvi primjeri u pitanju očekuju se prijevodni problemi neovisno o tome radi li se o prijevodu s hrvatskoga na engleski ili obrnuto. Naime, u engleskome je jeziku nepravilno koristiti dvostruku negaciju. Drugim riječima, ako se koriste izrazi *neither/nor*, glagol mora biti u pozitivnome obliku jer izraz *neither/nor* daje negacijsko značenje rečenici. Za razliku od engleskoga, u hrvatskome jeziku može biti više uzastopni negacijskih oblika tako da je prevođenje ovakvih izraza izazovno i ljudskome nestručnom prevoditelju, a kamoli alatu za strojno prevođenje. O problematičnosti ovoga, za alat izazovnoga prijevoda, govori i činjenica da su na kraju rečenice redundantno dodane riječi *are they not* koje bi inače imale ulogu kratkih pitanja (question tag), no s obzirom da se to kratko pitanje ne nalazi u izvorniku, nema opravdanja za njegovo dodavanje u prijevodu čak i kada bi bilo spretno upotrijebljeno.

Primjer 19

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|---|---|
| Tu su već „dugih“ osam godina, ali još su novost. | There are already "long" eight years, but there are new ones. |

U primjeru 19, mjesni prilog *tu* trebao bi u engleskome prijevodu glasiti *there* jer se, dakako, odnosi na mjesto. Prijevod *there are* govorniku engleskog jezika ne opisuje o čemu se radi i gdje je ono na što se odnosi rečenica. Prilog *tu* pripada zamjениčkoj skupini priloga koju čine sinsemantične riječi koje su ovisne o kontekstu. Zamjениčki prilozima uspostavljaju odnos prema sudionicima govornoga čina, odnosno različiti se prilozima koriste za označavanje mjesta koje je blizu govornika (*ovdje, ovamo, ovuda*), sugovornika (*tu, tamo, tuda*) i mjesto udaljeno i od govornika i od sugovornika (*ondje, onamo, onuda*). Slično je i u engleskome jeziku gdje se prilog *here* koristi za mjesto blizu govornika, a *there* za mjesto blizu sugovornika. Uzimajući navedeno u obzir, alat za strojno prevođenje upotrijebio je mjesni prilog koji, kao pojedinačna riječ, zbilja odgovara mjesnome prilogu iz izvornoga jezika. Ono što, nažalost, nije učinio jest uzeo kontekst u obzir jer da jest, taj bi se dio preveo s eksplicitnom upotrebom subjekta *they* (*oni*) na početku rečenice budući da rečenica u ciljnome jeziku ne može sadržavati implicitan

subjekt kao što može rečenica u izvornome jeziku, a mjesni bi se prilog *there* upotrijebio nešto kasnije u rečenici, odnosno nakon predikata.

Kod predikata, odnosno glagolskoga vremena nailazimo na sljedeći problem koji je već viđen u primjeru 11. Naime, zbog razlika u jezičnim sustavima izvornoga i ciljnoga jezika, korištenje glagolskih vremena može se razlikovati. Ljudske se prevoditelje podučava da se prednost uvijek daje prenošenju značenja i učinka nauštrb strukture [18], stoga bi se ljudski prevoditelj u ovome primjeru vjerojatno odlučio za upotrebu glagolskoga vremena *present perfecta* (*They have already been there...*), dok je alat za strojno prevođenje odabrao repliciranje strukture i upotrebu jednostavnog sadašnjeg vremena (*are*) kao u izvornome jeziku. Kako bi se naglasilo trajanje postojanja kroz određeno vrijeme, može se dodati prijedlog *for* (*for "long" eight years*). Ono što treba pohvaliti jest pravilna uporaba pravopisnih znakova navodnika koji se različito koriste u hrvatskome (prvi dolje, drugi gore) i engleskome (oba gore) jeziku. Nadalje, ista je pogreška u upotrebi glagolskoga vremena nastala i u daljnjem dijelu rečenice. Naime, vremenski prilog *still* prototipičan je prilog za glagolsko vrijeme *present perfecta*, stoga je trebao poslužiti kao orijentir za upotrebu spomenutoga (*...but they have still been a novelty/news.*). Zašto alat za strojno prevođenje nije isprogramiran tako da se neki prototipični prilozi/ključne riječi povezuju s upotrebom određenih glagolskih vremena i zašto je spomenuti vremenski prilog posve izostavljen iz rečenice ostaje nedoumica. Konačno, za imenicu *novost* u engleskome jeziku postoji sasvim prikladna riječ *novelty* koja rečenici iz primjera savršeno pristaje. Sagledavanje cijele rečenice uviđamo da je neuzimanjem u obzir razlika dvaju jezika, kršenjem jezičnih pravila ciljnoga jezika i upotrebom prijevodne tehnike riječ po riječ, dobivena rečenica posve nerazumljiva govorniku engleskoga jezika.

Primjer 20

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|--|---|
| Jednom prilikom grupa je japljana isklesala rai, no putem nazad snašla ih je oluja i havarija: golemi je disk potonuo. | On one occasion, the group of Japanese were scratched, but in the back there was a storm and havoc: the huge disc sank. |

U primjeru 20 pojavljuje se nekoliko vrsta grešaka. Japljani nisu Japanci već stanovnici otoka Yap u atlantskome oceanu. Kod ovakvih riječi koje se rijetko pojavljuju u tekstovima sasvim je jasno zašto dolazi do grešaka u prijevodu. Riječi koje se često koriste zastupljenije su u bazi prevoditelja, pa je tako, koristeći rijetke riječi, velika vjerojatnost nastajanja pogreške.

Međutim, s ovom bi vlastitom imenicom prijevodnih problema imao i ljudski prevoditelj jer bi mu kontekstualna uporaba sugerirala da se radi o vlastitoj imenici, dok bi korištenje maloga početnog slova bilo u koliziji s time. Također, vrlo je vjerojatno da ni ljudski prevoditelj ne bi znao o kojim se stanovnicima radi i morao bi provjeriti u nekim dostupnim internetskim bazama, što ne bi nužno urodilo plodom osim ako se prevoditelj ne bi dosjetio pa grafem *j* zamijenio grafemom *y*.

Sljedeće su dvije riječi *isklesala rai* prevedene na još nejasniji način prijevodom *were scratched*. U prijevodu nije vidljiva sličnost s izvornim riječima i ponovno se postavlja pitanje na koji način prevoditelj odlučuje koje će riječi izabrati za prijevod u nedostatku pravih riječi. Prijevod *carved rai* bio bi razumljiv i onome tko nije čuo za kamen koji se na jeziku naroda s otoka Yap naziva rai. Dakle, u ovome dijelu rečenice bilježimo dva prijevodna problema – izostanak prijevoda imenice *rai* i, posljedično, promjena perspektive u rečenici jer se iz prijevoda iščitava da su *Japljani* (ili *Japanci* kako predlaže alat za strojno prevođenje) *isklesani/izgrebi* (oni su trpitelji radnje), dok je u izvorniku ta grupa ljudi vršitelj radnje klesanja. Nadalje, iako glagoli *carve* i *scratch* podrazumijevanju oštećivanje neke površine, razlika njihova značenja jest u vrsti površine. Kod glagola *scratch*, riječ je o površini ljudskoga organa kože, dok je kod glagola *carve* riječ o površini drveta, kamena ili nečega sličnog. Dakle, iako se radi o sličnome postupku, postoji razlika u upotrebi i značenju tih dvaju glagola.

Nadalje, dio izvorne rečenice *...putem nazad...* nije ekvivalentan prijevodu *...in the back...*, već *...on their way back...* ili *...on the way back...* (ako nije jasno na čijem putu). Prijedložni padežni izraz *...in the back...* odnosio bi se na nešto u pozadini što se događa, a time bi se potpuno promijenilo značenje izvorne rečenice.

Primjer 21

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|---|--|
| Iako je u srcu kriptovaluta kriptografija, samo šifriranje nije nikakva novost. | Although cryptographic cryptography is at the heart, only encryption is no news. |

Prijevod je riječi *kriptovaluta* u *cryptographic* prijevod imenice u pridjev sasvim neke druge riječi. No, s ovim primjerom bilo je već očekivano da ta riječ neće biti prevedena pravilno vodeći se prijašnjim primjerima. Zapravo, alat za strojno prevođenje ponudio je imenički (*cryptography*) i pridjevski (*cryptographic*) oblik jedne te iste riječi (*kriptografija*), a posve izostavio imenicu kriptovaluta (*cryptocurrency*). Redoslijed riječi u prevedenoj je rečenici

preraspodijeljen i prijevod bi možda bio bolji kada bi izvorna rečenica zvučala prirodnije, primjerice *Iako je kriptografija u srcu kriptovaluta...* Ne pronalazimo motivaciju za korištenje prijedloga *at* umjesto *in* u dijelu ...u *srcu/at the heart...* jer se radi o istome prijedlogu i u izvornome i u ciljnome jeziku.

Sljedeća je greška u prijevodu ...*only encryption...* umjesto čega treba biti ...*the encryption itself...* jer *samo* u smislu ove rečenice znači nešto samo po sebi. Do pogreške je došlo zbog polisemičnosti te riječi u izvornome jeziku u kojemu *samo* može biti čestica, a može biti i veznik. Odabir pravilnoga oblik u izvornome jeziku, a posljedično i prijevoda u ciljni jezik ovisi, dakako, o kontekstu uporabe, što alat za strojno prevođenje nije uzeo u obzir.

Primjer 22

| IZVORNA REČENICA | GOOGLE TRANSLATE |
|---|--|
| Protivnici kriptovaluta često ističu da su najvjerniji korisnici kriptovaluta kriminalci. | Cryptoval oppositors often point out that they are the most likely criminals of criminals. |

U posljednjemu primjeru, imenicu *kriptovaluta* treba prevesti s *cryptocurrency* jer *cryptoval* nema nikakvo značenje. Budući da se nemogućnost prevođenja imenice *cryptocurrency* ponavlja, možemo pretpostaviti da ta imenica još nije unesena u unutarnji rječnik/bazu podataka alata za strojno prevođenje. S obzirom da čestoću pojavnosti te imenice u tekstovima koji se bave relativno novim tehnologijama, trebalo bi se žurno navedenu imenicu uvrstiti u unutarnji rječnik alata za strojno prevođenje. Riječ *oppositor* točan je prijevod imenice *protivnici*; međutim, toliko je zastario da ga rijetko tko (ako itko) koristi, stoga je navedenu imenicu potrebno zamijeniti imenicom *opponents* i ažurirati interni rječnik alata za strojno prevođenje.

Besmisleno je pridjev *najvjerniji* umjesto *the most loyal* prevesti kao *the most likely*, pa je pretpostavka da je prevoditelj tu riječ shvatio kao *najvjerojatniji*. Riječ *kriptovaluta* ni u drugom dijelu rečenice nije dobro prepoznata, a ni imenica *korisnici*. Niti jedna od tih dviju riječi ne pojavljuju se u prijevodu, a *criminals* je ponovljena dva puta bez ikakvog povoda. Dakako, nepotrebno ponavljanje imenica i ne korištenje onih upotrebljenih u izvornome jeziku utječe na značenje rečenice koje je ozbiljno narušeno.

5. ZAKLJUČAK

U ovom je radu bilo govora o strojnom prevođenju te vrstama strojnog prevođenja i njihovim prednostima i nedostacima. Zatim je opisan Google prevoditelj kao alat koji je korišten za daljnju analizu tekstova. Glavni zadatak rada bio je analizirati učinkovitost prijevoda s hrvatskog na engleski i engleskog na hrvatski koristeći Google prevoditelj. Dvije različite vrste tekstova korištene su kako bi prevoditelj imao priliku prevesti različite tipove rečenica. Broj točno i netočno prevedenih rečenica je za ove dvije vrste tekstova bio podjednak, iako se u manje formalnim tekstovima može uočiti više nepravilnosti i prevoditelju je općenito teže shvatiti točno značenje. U obje vrste teksta uočene su brojne greške. Najčešće nastaju problemi zbog nedostatka razumijevanja konteksta rečenice. Prevoditelj je u velikom broju rečenica, posebice onih koje su prevedene s hrvatskog na engleski jezik, prevodio tehnikom riječ po riječ. S obzirom na različitost dvaju promatranih jezika, takav pristup nije donio dobre rezultate, osim kada je poredak riječi u rečenici za oba jezika otprilike jednak. Netočna interpretacija vrsta riječi također je dovela do grešaka, a osim toga često se ponavljala i greška izostavljanja riječi bez vidljivog razloga. To je dovelo do gubljenja značenja rečenice. Osim izostavljanja riječi, prevoditelj je u nekoliko primjera dodao sasvim nepotrebne riječi koje nisu povezane sa zadanom rečenicom. Razlike dvaju jezika u glagolskim vremenima, padežima i redosljedu riječi dovele su do dodatnih grešaka. Prijevod višeznačnica bio je još jedan od problema jer prevoditelj nije uspio pretpostaviti koje je značenje riječi potrebno uzeti kako bi rečenici ostao prvotni smisao. Osim navedenih grešaka, prevoditelj je neke riječi ostavio u izvornom obliku, a pretpostavka je da za te riječi u svojoj bazi nema adekvatan prijevod. Takve se greške mogu smanjiti neprestanim ažuriranjem baze riječi kojom prevoditelj raspolaže.

Kada bi se prijevodi dobiveni u ovom radu usporedili s drugima koji su nastali vremenski ranije, mogao bi se uočiti napredak u današnjem prijevodu. Sustav za strojno prevođenje neprestano se nadograđuje, a ono što bi znatno pomoglo jest više uključivanje ljudi u proces te nadogradnje. Potrebno je dostaviti prevoditelju mnogo veći broj primjera koji sadrže ne toliko jasan sadržaj kako bi mogao prepoznati i povezati uzorke i s većom točnošću pronaći pravi prijevod. Taj nedostatak primjera posebice je izražen u jezicima koji su manje zastupljeni na Internetu. Uz sve veći razvoj i primjenu neuronskih mreža bit će moguće ostvariti sve bolje rezultate u području strojnog prevođenja.

LITERATURA

- [1] W. J. Hutchins, H. L. Somers, *An introduction to machine translation*, Academic Press, London, 1992.
- [2] W. Weaver, *Translation*, The Rockefeller Foundation, New York, 1949
- [3] M. Amine Chérargui, *Theoretical Overview of Machine translation*, CEUR Workshop Proceedings, Vol. 867., pp. 160.-169., 2012.
- [4] Wikipedia: History of machine translation [online], dostupno na: https://en.wikipedia.org/wiki/History_of_machine_translation#The_1980s_and_early_1990s [16.8.2019.]
- [5] W.S.Bennett, Logos Corporation, *Taking the Babble Out of Babel Fish*, Language International, pp. 20-21, June 2000
- [6] World of Computing: Machine Translation Process [online], dostupno na: <http://language.worldofcomputing.net/machine-translation/machine-translation-process.html> [16.8.2019.]
- [7] Localize: Different Types Of Machine Translation [online], dostupno na: <https://localizeblog.com/types-of-machine-translation/> [16.8.2019.]
- [8] Language Connections: The Pros And Cons Of Google Translate [online], dostupno na: <https://www.languageconnections.com/blog/the-pros-cons-of-google-translate/> [16.8.2019.]
- [9] I. Grgurević, A. Stančić, M. Slavulj, *Identifying relevant factors of applying technologies in dynamic carpooling*, Tehnički glasnik, god. 9, br. 4, 2015.
- [10] V.Sharma, *The Relevance of Addition, Omission and Deletion (AOD) in Translation*, International Journal of Translation (IJT), 2015.
- [11] [Online]. D. Mitrović, D. Božić Lenard, I. Galić, *Održavanje baza alata za strojno prevođenje*, 27. Međunarodni znanstveno-stručni skup „Organizacija i tehnologija održavanja“ - OTO 2018. - Zbornik radova, Osijek : Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek, str. 23-32, 2018.
- [12] E. Dasović, G. Petković, S. Pasanec Preprotić, *Oblikovanje i budućnost knjižnog uveza u svijetu e-knjige*, Tehnički glasnik, god. 9, br. 4, 2015.
- [13] The New York Times: How an App Creates ‘Disturbingly Agile Millennial Thumbs’ [online], Feb. 2019., dostupno na: <https://www.nytimes.com/2019/02/20/technology/personaltech/agile-millennial->

thumbs.html [27.6.2019.]

- [14] Sui-Lee Wee, E. Chen, The New York Times: China's Tech Firms Are Mapping Pig Faces [online], Feb. 2019., dostupno na: <https://www.nytimes.com/2019/02/24/business/china-pig-technology-facial-recognition.html> [27.6.2019.]
- [15] G. Kovács, *An Evergreen Challenge for Translators – The Translation of Idioms*, Acta Universitatis Sapientiae, Philologica, 8(2), pp.61-77, Aug. 2019.
- [16] Tim Gutteridge: Translate meanings, not words [online], dostupno na: <http://timgutteridge.co.uk/translate-meanings-not-words/> [16.8.2019.]

SAŽETAK

Prevođenje velikih količina teksta može biti vrlo zahtjevan i dugotrajan posao. Kao posljedica toga, nastala je potreba za strojnim prevoditeljem. U ovom radu opisani su definicija i vrste strojnih prevoditelja, kao i dobre i loše strane istih. Za potrebe analize prijevoda strojnog prevoditelja, odabran je Google prevoditelj. Dva znanstvena rada i dva popularno-znanstvena rada prevođena su s hrvatskog na engleski i obrnuto. Svrha završnog rada bila je analizirati potencijalne pogreške koje nastaju prilikom prijevoda Google prevoditeljem te raspraviti razloge nastajanja grešaka. Uz pomoć rezultata ove analize, dolazi se do zaključka da, unatoč velikom napretku u zadnjih nekoliko godina, strojni prevoditelji još uvijek nailaze na poteškoće pri prevođenju pojedinih kompleksnijih i dužih rečenica. Iako su u obje vrste rada pronađene greške, u znanstvenom se radu primjećuju manje greške, s obzirom da nema žargona niti puno rečenica u kojima je nejasan kontekst.

Ključne riječi: *strojno prevođenje, Google prevoditelj, analiza prijevoda*

ABSTRACT

Translating great amounts of text can be a demanding and time consuming task. As a consequence, a need for machine translation arose. Explained in this thesis are the definition and the types of machine translators, along with their good and bad sides. For the purpose of the machine translation analysis, Google Translator was chosen. Two scientific papers and two popular-scientific papers were translated from Croatian to English and vice versa. This thesis aimed to analyze the potential mistakes which appear while translating text using the Google Translator tool and to discuss the mistakes' reasons. The research results point to a progress which has recently been made in the field of machine translation; however, machine translation tools continue to come across obstacles when translating some more complex and long sentences. Although both types of papers had mistakes in the translations, there were fewer mistakes in the scientific paper due to the fact that there is no jargon or a lot of sentences with ambiguous meaning.

Key words: *machine translation, Google Translator, translation analysis*

ŽIVOTOPIS

Maja Matijević rođena je 29. lipnja 1997. godine u Osijeku. Osnovnu školu Svete Ane u Osijeku završava 2012. godine, nakon čega upisuje III. gimnaziju Osijek (prirodoslovno-matematička). Nakon završenog srednjoškolskog obrazovanja, 2016. godine upisuje preddiplomski studij računarstva na Elektrotehničkom fakultetu u Osijeku, današnjem Fakultetu elektrotehnike, računarstva i informacijskih tehnologija Osijek.

Potpis: _____