

Ljudska evaluacija sustava za neuralno strojno prevođenje

Mikulić, Andrea

Undergraduate thesis / Završni rad

2020

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, University of Zagreb, Faculty of Humanities and Social Sciences / Sveučilište u Zagrebu, Filozofski fakultet**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:131:516896>

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

Download date / Datum preuzimanja: **2024-07-14**



Sveučilište u Zagrebu
Filozofski fakultet
University of Zagreb
Faculty of Humanities
and Social Sciences

Repository / Repozitorij:

[ODRAZ - open repository of the University of Zagreb
Faculty of Humanities and Social Sciences](#)



SVEUČILIŠTE U ZAGREBU
FILOZOFSKI FAKULTET
ODSJEK ZA INFORMACIJSKE I KOMUNIKACIJSKE ZNANOSTI
Ak. god. 2019./2020.

Andrea Mikulić

Ljudska evaluacija sustava za neuralno strojno prevođenje

Završni rad

Mentor: prof.dr.sc. Sanja Seljan

Zagreb, kolovoz 2020.

Izjava o akademskoj čestitosti

Izjavljujem i svojim potpisom potvrđujem da je ovaj rad rezultat mog vlastitog rada koji se temelji na istraživanjima te objavljenoj i citiranoj literaturi. Izjavljujem da nijedan dio rada nije napisan na nedozvoljen način, odnosno da je prepisan iz necitiranog rada, te da nijedan dio rada ne krši bilo čija autorska prava. Također izjavljujem da nijedan dio rada nije korišten za bilo koji drugi rad u bilo kojoj drugoj visokoškolskoj, znanstvenoj ili obrazovnoj ustanovi.

(potpis)

Sadržaj

1. Uvod.....	5
2. Povijesni razvoj strojnog prevođenja.....	6
2.1. Začetci strojnog prevođenja	6
2.2. Razdoblje sveopćeg optimizma (1956 – 1966).....	8
2.3. Tihi period (1966 – 1975)	9
2.4. Operativni i komercijalni sustavi (1976 – 1989).....	10
2.5. Nova era (1989 -)	10
2.6. Problemi strojnog prevođenja	11
3. Evaluacija strojnog prevođenja.....	14
3.1. Ljudska evaluacija strojnog prijevoda.....	15
3.2. Automatska evaluacija strojnog prevođenja	16
4. Pristupi strojnom prevođenju.....	18
4.1. Strojno prevođenje temeljeno na pravilima	18
4.2. Statističko strojno prevođenje	21
4.3. Neuralno strojno prevođenje	25
5. Istraživanje.....	32
5.1. Systran.....	35
5.2. Google Translate	36
5.3. Yandex.Translate.....	37
5.4. Bing Microsoft Translator.....	38
6. Zaključak.....	41
7. Literatura.....	42
8. Popis slika.....	46
9. Popis tablica.....	47

10.	Sažetak	48
11.	Summary	49

1. Uvod

Povezivanjem svijeta u tzv. „globalno selo“ povećava se potreba za prevođenjem pa rad prevoditelja nije dovoljan. Nastaje sve više dokumenta i informacija koje je potrebno prevesti u što kraćem roku te sam prevoditelj više nema dovoljno vremena niti znanja za kvalitetnu obradu svih nastalih tekstova. Kučiš (2010, p. 20)¹ naglašava kako jedan profesionalni prevoditelj „ne može kvalitetno obraditi više od 4 – 6 kartica teksta, odnosno oko 2500 riječi dnevno“. Sve veća očekivanja potaknula su i razvoj računalnih sustava za online prevođenje. Iako strojni prijevodi još nisu dosegнули kvalitetu ljudskog prijevoda sve su kvalitetniji te omogućuju svakodnevnim korisnicima besplatan prijevod teksta. Također, ponekad nije niti potrebno imati visokokvalitetni prijevod već je bitnije dobiti brzi prijevod, stoga se u takvim slučajevima koristi strojno prevođenje.

U radu su analizirani osnovni pristupi strojnom prevođenju, kao što su statističko strojno prevođenje, neuralno prevođenje i prevođenje temeljeno na leksičkim i gramatičkim pravilima jezika. Rad je podijeljen u dva dijela; teorijski i praktični. U teorijskom djelu bit će prikazan kratak pregled povijesti strojnog prevođenja, postupk evaluacije strojnog prijevoda te općenite pogreške i problemi s kojima se strojno prevođenje susreće. Zatim slijedi praktični dio u kojem je napravljeno istraživanje na 300 rečenica za englesko-hrvatski jezični par za četiri online alata za strojno prevođenje temeljenim na neuralnom prevođenju ili hibridnom prevođenju. Na kraju slijedi zaključak, literatura, popis slika i tablica te sažeci na hrvatskom i engleskom.

¹ Kučiš, V. (2010). Prevodilački alati u funkciji kvalitete prijevoda. *Informatologia*, 43(1), str. 19-33.

2. Povijesni razvoj strojnog prevođenja

Prema Dovedan i sur. (2002)², automatsko stajno prevođenje postavilo je svoje korijene već u 17. stoljeću kada se prvi puta pojavilo računalo. Šimić i Uglarik (2010, p. 84)³ tvrde da „strojno prevođenje ima tradiciju dugu približno kao i digitalna računala na čijoj potpori je zasnovano”. Intenzivniji razvoj računalnog prevođenja započinje u vojne svrhe 50-ih godina 20. stoljeća za vrijeme tako zvanog „hladnog rata“ između SAD-a i Sovjetskog saveza kada su obje strane htjele što prije saznati informacije kojima protivnik raspolaže.

Razvoj strojnog prevođenja imao je veliki utjecaj na svakodnevni život, ali i na prevodilačku profesiju. Kao što Kučiš (2010, p. 19)⁴ navodi: „brzi razvoj jezičnih tehnologija bitno je utjecao na prevodilačku industriju i prevodilačku profesiju omogućavajući efikasnije i brže obavljanje profesije”. Alati za strojno prevođenje mogu se koristiti u postupcima prevođenja, ali i zbog pristupa do svakodnevnih informacija.

2.1. Začetci strojnog prevođenja

Prema Hutchinsu (2001, p. 2)⁵, ideje o strojnom prevođenju počinju se formirati već u 17. stoljeću s idejama o univerzalnom ili filozofskom jeziku i mehaničkim rječnicima. Tada nije postojala potrebna tehnologija za strojeve za prevođenje te se stoga može reći kako koncentriraniji razvoj počinje 1930-ih godina kada Georges Artsrouni i Petr Troyanskii prijavljuju prve patente za strojeve koji prevode (engl. *translation machines*), prema Hutchins (2020)⁶. Prema Dovedan i sur. (2002, p. 2)⁷, Artsrouni je 1937. godine patentirao pohranu informacija na papirnatoj vrpici koja se zatim koristila za pronalaženje jednakih riječi na ciljnom jeziku. Neovisno o ovom istraživanju, Troyanskii je „imao viziju dvojezičnog ili višejezičnog mehaničkog prevođenja koje bi se provodilo u tri koraka: logička i sintaktička analiza izvornog jezika, transformacija temeljnog

² Dovedan, Z., Seljan, S. i Vučković, K., 2002. Strojno prevođenje kao pomoć u procesu komunikacije. *Informatologia*, 35(4), pp.283-291

³ Šimić, J., i Uglarik, D. (2010). 'Mogućnosti i ograničenja strojnog prevođenja', *Praktični menadžment*, 1(1), str. 81-85.

⁴ Kučiš, V. (2010). Prevodilački alati u funkciji kvalitete prijevoda. *Informatologia*, 43(1), str. 19-33.

⁵ Hutchins, W., 2001. Machine Translation over fifty years. *Histoire Épistémologie Langage*, 23(1), pp.7-31

⁶ Hutchins, J., 2020. *The History Of Machine Translation In A Nutshell*. (online) Translationdirectory.com. Dostupno na: <<https://www.translationdirectory.com/article411.htm>>

⁷ Dovedan, Z., Seljan, S. i Vučković, K., 2002. Strojno prevođenje kao pomoć u procesu komunikacije. *Informatologia*, 35(4), pp.283-291

oblika i funkcija u ekvivalentan slijed ciljnog jezika i prijevod u govorni ciljni jezik“ (Dovedan i sur., 2002, p. 2)⁸.

Za rani razvoj strojnog prevođenja bitan je i Warrena Weaver “koji je predložio da se prevođenju pristupa računalno” (Dunđer, 2015, p. 10)⁹. Kako Dovedan i sur. (2002, p. 2)¹⁰ navode, Weaver je jedan od začetnika teorije komunikacije koji je 1949. godine izložio teoriju da bi se strojno prevođenje moglo ostvariti korištenjem tehnika vojnog dešifriranja. Dakle, kako bi računalo moglo simulirati rad prevoditelja trebalo bi svaku riječ potražiti u dvojezičnom rječniku, zatim usporediti riječi iz ulaznog teksta s onima iz rječnika, odabrati točan prijevod te nakon što tako obradi cijelu rečenicu, sve prevedene riječi treba pravilno složiti u izlaznoj rečenici u skladu s pravilima ciljnog jezika.

Strojno prevođenje se u javnosti prvi puta spomenulo 1951. godine. Tri godine kasnije je izvršena prva javna i uspješna demonstracija izvodljivosti strojnog prevođenja u SAD-u koja je bila rezultat suradnje IBM-a i Sveučilišta u Georgetownu (Hutchins, 2020)¹¹. Za projekt se koristio „uzorak od 49 rečenica na ruskom jeziku preveden (...) na engleski, uz uporabu vrlo ograničenog rječnika od 250 riječi i 6 gramatičkih pravila,“ (Dovedan i sur., 2002, p. 3)¹². Kako navode Dovedan i sur. (2002, p. 3)¹³ te iste, 1954. godine, osnovan je i *Machine Translation*, prvi časopis koji se bavio strojnim prevođenjem.

Iduće, 1955. godine, izlazi i prva knjiga radova na temu strojnog prevođenja (Finka i László, 1962, p. 119)¹⁴. Nadalje, u razvoj strojnog prevođenja uključuju se SSSR i Velika Britanija 1956. godine, a Jugoslavija se kao deseta zemlja ovim naporima priključuje 1959. godine (Finka i László, 1962, p. 119)¹⁵.

⁸ Dovedan, Z., Seljan, S. i Vučković, K., 2002. Strojno prevođenje kao pomoć u procesu komunikacije. *Informatologia*, 35(4), pp.283-291

⁹ Dunđer, I., 2015. Sustav za statističko strojno prevođenje i računalna adaptacija domene. Doktorska disertacija. Sveučilište u Zagrebu. Filozofski fakultet.

¹⁰ Dovedan, Z., Seljan, S. i Vučković, K., 2002. Strojno prevođenje kao pomoć u procesu komunikacije. *Informatologia*, 35(4), pp.283-291

¹¹ Hutchins, J., 2020. *The History Of Machine Translation In A Nutshell*. (online) Translationdirectory.com. Dostupno na: <<https://www.translationdirectory.com/article411.htm>>

¹² Ibid.

¹³ Ibid.

¹⁴ Finka, B., i Laszlo, B. (1962). 'Strojno prevođenje i naši neposredni zadaci', *Jezik*, 10(4), str. 117-121

¹⁵ Ibid.

Kako navodi Dunder (2015, p. 11)¹⁶: „Yehoshua Bar-Hillel, znanstvenik s MIT-a (engl. *Massachusetts Institute of Technology*), prvi je istraživač koji je u punom radnom vremenu istraživao strojno prevođenje potpomognuto rječnicima.“ U lipnju 1952. godine, Bar-Hillel je organizirao i prvu međunarodnu konferenciju o strojnom prevođenju kojoj su mnogi znanstvenici koji su se u tom trenutku bavili strojnom prevođenjem (Hutchins, 2001, p. 2)¹⁷.

2.2. Razdoblje sveopćeg optimizma (1956 – 1966)

Prema Dovedan i sur. (2002)¹⁸, nakon začeca strojnog prevođenja nastupa period generalnog optimizma oko svih mogućnosti strojnog prevođenja koje je trajalo do sredine 1960-ih. Tada su rani sistemi za prijevod koristili uglavnom dvojezične rječnike te nekoliko pravila za izradu pravilnog poretka riječi u rečenici. Ovo razdoblje obilježilo je mnogo pozitivnih predviđanja za strojno prevođenje te se uspostavljaju grupe za istraživanje i izradu strojnog prevođenja diljem svijeta.

Međutim, početni optimizam je ubrzo splasnio kada su istraživači susreli problem semantike te nisu uspjeli napredovati kako se javnost nadala. Zbog velikih financijskih izdataka, Sjedinjene Američke Države su 1964. godine zahtijevale ispitivanje stanja i brzinu napretka strojnog prevođenja (Dunder, 2015, p. 12)¹⁹. Tako je osnovano savjetodavnog tijela za automatsku obradu jezika (engl. *Automatic Language Processing Advisory Committee (ALPAC)*). ALPAC je 1966. godine objavio izvještaj u kojem je naveo da je strojno prevođenje sporo, manje precizno i duplo skuplje od ljudskog prevođenja, te da je „bez perspektive u bližoj budućnosti, istaknuvši i kritiku post-uredništva“ (Dovedan i sur., p. 4)²⁰. Nije bila prepoznata nikakva potreba za strojnim prevođenjem niti njegov potencijal, već se sav napor preusmjerio na razvijanje sustava koji bi bili

¹⁶ Dunder, I., 2015. Sustav za statističko strojno prevođenje i računalna adaptacija domene. Doktorska disertacija. Sveučilište u Zagrebu. Filozofski fakultet.

¹⁷ Hutchins, W., 2001. Machine Translation over fifty years. *Histoire Épistémologie Langage*, 23(1), pp.7-31

¹⁸ Dovedan, Z., Seljan, S. i Vučković, K., 2002. Strojno prevođenje kao pomoć u procesu komunikacije. *Informatologia*, 35(4), pp.283-291

¹⁹ Dunder, I., 2015. Sustav za statističko strojno prevođenje i računalna adaptacija domene. Doktorska disertacija. Sveučilište u Zagrebu. Filozofski fakultet.

²⁰ Dovedan, Z., Seljan, S. i Vučković, K., 2002. Strojno prevođenje kao pomoć u procesu komunikacije. *Informatologia*, 35(4), pp.283-291

od pomoći ljudskom prevoditelju. Time je započeo period zatišja (Hutchins, 2020)²¹. Međutim, istraživanja na području strojnog prevođenja su nastavljena u Kanadi, Francuskoj i Njemačkoj.

2.3. Tihi period (1966 – 1975)

Nakon ALPAC-ovog izvješća, razvoj strojnog prevođenja je prekinut u Sjedinjenim Američkim Državama te se fokus prebacio na razvijanje sustava kojima je potrebna ljudska pomoć za izradu prijevoda (Hutchins, 2001, p. 1).²²

Međutim, razvoj strojnog prevođenja nije u potpunosti bio zapostavljen u svijetu jer su potrebe u Kanadi i Europi bile drugačije od onih u SAD-u. Kanadska dvokulturna politika stvorila je veliku potrebu za englesko-francuskim jezičnim parom koja nije mogla biti zadovoljena samo ljudskim prijevodom. Problemi prijevoda su bili jednako bitni i u Europi zbog sve veće potražnje za prijevodima znanstvenih, tehničkih, administrativnih i pravnih dokumenata sa i na sve jezike tadašnje Europske zajednice pa se u Europi i Kanadi strojno prevođenje i dalje razvijalo (Hutchins, 2001, p. 7).²³

U tom periodu nastaju Systran i Météo koji se, uz male preinake tijekom vremena, i danas koriste. Systran je sustav za strojno prevođenje koji je razvilo ratno zrakoplovstvo SAD-a 1970. godine, a Météo je sustav za prijevod vremenskih prognoza s engleskog na francuski i obrnuto koji je razvijen u Kanadi koji za svoj rad koristi podjezik (Hutchins, 2020)²⁴. Uz Météo, u ovom razdoblju javljaju se još dva sustava temeljena na podjeziku, što je bio jedan uvjeta za razvoj tadašnjih sustava (Seljan, 2000)²⁵. Riječ je o TITUS-u, multijezičnom sustavu koji prijevodi sažetke pisane kontroliranim jezikom, a razvijen je 1970. godine na Institut Textile de France, i sustavu Kineskog Sveučilišta u Hong Kongu CULT, koji je predstavljen 1972. godine, a dizajniran je za prevođenje matematičkih tekstova s kineskog na engleski (Hutchins, 2001, p. 8)²⁶.

²¹ Hutchins, J., 2020. *The History Of Machine Translation In A Nutshell*. (online) Translationdirectory.com. Dostupno na: <<https://www.translationdirectory.com/article411.htm>>

²² Hutchins, W., 2001. Machine Translation over fifty years. *Histoire Épistémologie Langage*, 23(1), pp.7-31

²³ Ibid.

²⁴ Hutchins, J., 2020. *The History Of Machine Translation In A Nutshell*. (online) Translationdirectory.com. Dostupno na: <<https://www.translationdirectory.com/article411.htm>>

²⁵ Seljan, S. (2000). Sublanguage in Machine Translation. Mipro 2000: Computers in Intelligent Systems.

²⁶ Hutchins, W., 2001. Machine Translation over fifty years. *Histoire Épistémologie Langage*, 23(1), pp.7-31

2.4. Operativni i komercijalni sustavi (1976 – 1989)

U početku razvoja strojnog prevođenja računala su bila veoma skupa te su zauzimala nekoliko prostoriya te su se onda koristila samo u vojne i vojno-industrijske svrhe. Pojavom mikroračunala 1980-ih godina stvorilo se tržište za jeftinije sustave za strojno prevođenje (Hutchins, 2020)²⁷. Kučiš (2010, p. 20)²⁸ navodi da je „ekonomska potreba za pronalaženjem bržeg, ujednačenog i isplativijeg prijevoda u međunarodnoj komunikaciji, rezultirala (...) intenzivnim razvojem prevodilačke tehnologije“. Stoga se naporno radilo na razvitku stroja koji bi mogao obavljati prevoditeljski posao brže i jeftinije od ljudskog prevoditelja, ali da kvaliteta bude ista. Kako bi se taj cilj ostvario potrebno je osigurati da stroj razumije ulazni tekst. U ovom periodu razvoja strojnog prevođenja postalo je jasno koliko je znanja o svijetu potrebno za prijevod čak i najjednostavnijih rečenica. Iz toga se rađa ideja da strojevi za prevođenje ne samo da trebaju imati ugrađene rječnike, već im je potrebna i univerzalna enciklopedija (Dovedan i sur., 2002, p. 4)²⁹.

Ovo razdoblje je karakterizirano realističnim očekivanjima. Za razliku od početne faze razvoja strojnog prevođenja, u ovom periodu njegovog razvoja nedostaci i pogreške su očekivane, javno obznanjene i prihvaćene te se strojno prevođenje općenito smatralo kao samo pomoć u prevođenju (Dovedan i sur., 2002, p. 4)³⁰.

70-e i 80-e godine prošlog stoljeća bile su početak razvoja specijaliziranih sustava za strojno prevođenje koji su bili namijenjeni određenoj domeni (Hutchins, 2001, p. 8)³¹.

2.5. Nova era (1989 -)

Završetkom hladnog rata i povezivanjem svijeta u međunarodno tržište povećava se potreba za bržim i efikasnijim prevođenjem u cilju prijenosa informacija na više jezika. Ova potreba je „istovremeno ukazivala na neučinkovitost i sporost ljudskog procesa prevođenja, posebno na području stručnog prevođenja, intenzivirajući interes i potrebu za stvaranjem prevoditeljskog

²⁷ Hutchins, J., 2020. *The History Of Machine Translation In A Nutshell*. (online) Translationdirectory.com. Dostupno na: <<https://www.translationdirectory.com/article411.htm>>

²⁸ Kučiš, V. (2010). Prevodilački alati u funkciji kvalitete prijevoda. *Informatologia*, 43(1), str. 19-33.

²⁹ Dovedan, Z., Seljan, S. i Vučković, K., 2002. Strojno prevođenje kao pomoć u procesu komunikacije. *Informatologia*, 35(4), pp.283-291

³⁰ Ibid.

³¹ Hutchins, W., 2001. Machine Translation over fifty years. *Histoire Épistémologie Langage*, 23(1), pp.7-31

stroja, koji bi prevođenje učinio bržim i ekonomičnijim“ (Kučiš, 2010, p. 20)³². Sve veći broj dokumenata koje je trebalo svakodnevno prevoditi te se još više potaknuo razvoj strojnog prevođenja. Nadalje, razvojem i širenjem osobnih računala raste potreba i potražnja za strojnim prevođenjem. Prema Hutchinsu (2001, p. 12)³³ dominacija strojnog prevođenja temeljenog na pravilima je prekinuta izlaskom na tržište novih metoda temeljenih na korpusima kao što je slučaj sa statističkim prevođenjem i prevođenjem temeljenim na primjerima (Brkić i sur., 2009)³⁴.

Ranih 1990-ih godina IBM je objavio rezultate eksperimentalnog programa koji je koristio statističko strojno prevođenje – Candide. Uz to, Japan je počeo koristiti metode temeljene na primjerima. Oba pristupa su se razlikovala od prvotnih programa jer nisu koristila sintaktička i semantička pravila za izradu prijevoda već su koristili korpuse. Ovo je također bilo doba mijenjanja fokusa strojnog prevođenja sa samo eksperimentalnog i istraživačkog pristupa na praktične primjene (Hutchins, 2020)³⁵.

Hutchins (2020)³⁶ tvrdi da se u kasnijim 90-im godinama strojno prevođenje još više širi zbog osobnih računala te se počinje sve više koristiti u tvrtkama.

2000-e godine su, prema Dunderu (2015, p. 14)³⁷, „razdoblje sustava za statističko strojno prevođenje koje se temelji na empirijskim opažanjima“. Takvi sustavi su tipično namijenjeni za jednu domenu s ograničenim vokabularom jer se tako postižu bolji rezultati. Strojno prevođenje se i danas razvija, a najnoviji pristup strojnom prevođenju jest neuralno prevođenje koje je na tržište izašlo 2015. godine.

2.6. Problemi strojnog prevođenja

Iako je strojno prevođenje znatno napredovalo od začetnih ideja još uvijek postoje mnogi problemi koji i dalje nisu riješeni. Neki problemi su bili očiti već i na početku razvoja strojnog prevođenja

³² Kučiš, V. (2010). Prevodilački alati u funkciji kvalitete prijevoda. *Informatologia*, 43(1), str. 19-33.

³³ Hutchins, W., 2001. Machine Translation over fifty years. *Histoire Épistémologie Langage*, 23(1), pp.7-31

³⁴ Brkić, M., Vičić, T., Seljan, S. (2009). Evaluation of the statistical machine translation service for Croatian-English. *INFuture 2009 : Digital resources and knowledge sharing*, 319-332

³⁵ Hutchins, J., 2020. *The History Of Machine Translation In A Nutshell*. (online) Translationdirectory.com. Dostupno na: <<https://www.translationdirectory.com/article411.htm>>

³⁶ Ibid.

³⁷ Dunder, I., 2015. Sustav za statističko strojno prevođenje i računalna adaptacija domene. Doktorska disertacija. Sveučilište u Zagrebu. Filozofski fakultet.

te, iako se današnji programi puno bolje nose s takvim problemima, još uvijek nisu u potpunosti riješeni. Već u ranim fazama strojnog prevođenja učestalo su se događale određene vrste pogrešaka te ih je Bar-Hillel (1953, p. 217)³⁸ podijelio u četiri kategorije, a to su: sintaksa prirodnih jezika, međuprevodivost (engl. *intertranslatability*) prirodnih jezika, idiomi i univerzalne sintaktičke kategorije.

Problemi sa strojnim prevođenjem su razni, a prema Charoenpornasawat i sur. (2002, p. 1)³⁹, glavna podjela grešaka strojnog prevođenja koristi dvije kategorije, a to su; krivo značenje i krivi redoslijed riječi u rečenici.

Iako postoje brojne kategorije pogrešaka, u radu je korištena postojeća kategorizacija pogrešaka kako bi se mogle usporediti vrste i količina pogrešaka s prethodnim istraživanjima. U radu Kučić i sur. (2009)⁴⁰ analizirani su strojni prijevodi za njemačko-slovenski jezični par prema kategorijama leksičkih pogrešaka, pravopisnih pogrešaka i interpunkcije te sintaktičkih i stilističkih pogrešaka. Brkić i sur. (2013)⁴¹ analizirali su pogreške na morfološkoj razini, leksičkoj, sintaktičkoj i semantičkoj razini. Seljan i sur. (2015)⁴² sistematiziraju kategorije pogrešaka za hrvatski jezik na sljedeće kategorije: neprevedene riječi, višak prevedenih riječi, morfološke pogreške i krivi završeci, leksičke pogreške i krivi prijevodi, sintaktičke pogreške i interpunkcije kako bi utvrditi koliko se često pojavljuju pogreške za dva online sustava i njihov utjecaj na kriterije fluentnosti i adekvatnosti. Rezultati pokazuju da na online statističkim sustavima ima više pogrešaka u smjeru rusko-hrvatski za kriterij fluentnosti na sustav Yandex.Translate, te za englesko-hrvatski za kriterij adekvatnosti na sustavu Google Translate. Sustavi su imali najveći broj morfoloških pogrešaka, iza kojih slijede leksičke i sintaktičke. Dunder (2015)⁴³ je proveo opsežno istraživanje za englesko-hrvatski jezični par s obzirom na računalu adaptaciju domene sustava za strojno prevođenje treniranog na statističkim modelima.

³⁸ Bar-Hillel, Y., 1953. Some Linguistic Problems Connected with Machine Translation. *Philosophy of Science*, 20(3), pp.217-225

³⁹ Charoenpornasawat, P., Sornlertlamvanich, V. i Charoenporn, T., 2002. Improving translation quality of rule-based machine translation. *COLING-02 on Machine translation in Asia*

⁴⁰ Kučić, V., Seljan, S., Klasnić, K. (2009). Evaluation of electronic translation tools through quality parameters. *INFuture 2009: Digital Resources and Knowledge Sharing*. 341-351.

⁴¹ Brkić, M., Seljan, S., Vičić, T. (2013). [Automatic and human evaluation on english-croatian legislative test set](#). *Intelligent Text Processing and Computational Linguistics*, Springer, 311-317.

⁴² Seljan, S., Tucaković, M. and Dunder, I., 2015. Human Evaluation of Online Machine Translation Services for English/Russian-Croatian. *New Contributions in Information Systems and Technologies*, pp.1089-1098.

⁴³ Dunder, I., 2015. Sustav za statističko strojno prevođenje i računalna adaptacija domene. Doktorska disertacija. Sveučilište u Zagrebu. Filozofski fakultet.

Ljubas (2017, p. 28)⁴⁴ također koristi klasifikaciju pogrešaka koja raspodjeljuje pogreške strojnog prevođenja u osam kategorija, a to su „neprevedene, izostavljene i dodane riječi, leksičke, stilističke, pravopisne i morfosintaktičke pogreške, pogreške u redu riječi te u dvije dodatne kategorije: pogreške u dosljednosti i nejasne rečenice“.

Također Šimić i Uglarik (2010, p. 83)⁴⁵ ističu kako je „kod prevođenja pojedinih riječi odmah (...) vidljiv najčešći, i najteži problem prevođenja - višeznačnost (ambiguity)“. Statističko strojno prevođenje pokušalo je riješiti problem višeznačnosti uz uporabu statistike, tako da se gleda koje je značenje najvjerojatnije, ali ni taj pristup nije uvijek uspješan. Najbolji način na koji bi se taj problem mogao riješiti jest da čovjek pregleda izvorni tekst prije prevođenja te ga uredi tako da ukloni sve višeznačnosti i nejasne riječi Šimić i Uglarik (2010, p. 84)⁴⁶.

⁴⁴ Ljubas, S., 2017. Analiza pogrešaka u strojnim prijevodima sa švedskog na hrvatski. *Hieronymus-časopis za istraživanja prevođenja i terminologije*, 4, pp.28-64

⁴⁵ Šimić, J., i Uglarik, D. (2010). 'Mogućnosti i ograničenja strojnog prevođenja', *Praktični menadžment*, 1(1), str. 81-85.

⁴⁶ Ibid.

3. Evaluacija strojnog prevođenja

Kako bi se pratio razvoj kvalitete strojnog prevođenja bilo je potrebno razviti i metode za ocjenjivanje strojnog prevođenja. Postoje različite metode evaluacije strojnog prevođenja koje ovise o svrsi prijevoda, svrsi ocjenjivanja, vremenu potrebnom za prijevod, cijeni prevođenja i drugim faktorima. Prema Brkić i sur. (2009, p. 323)⁴⁷ evaluacija strojnog prevođenja nije jednostavan zadatak jer različiti prevoditelji različito prevode isti izvorni tekst. Evaluacija strojnog prevođenja može biti ručna ili automatska (Brkić i sur., 2009, p. 323)⁴⁸.

Ljudsko ocjenjivanje je prvo i najstarije ocjenjivanje strojnog prevođenja. Ova metoda je najkvalitetnija, ali je također skupa i spora te je sama po sebi i subjektivna (Sepesy Maučec i Donaj, 2020, p. 8)⁴⁹. Automatsko ocjenjivanje prijevoda je, kako Sepesy Maučec i Donaj (2020, p. 8)⁵⁰ tvrde, besplatna alternativa ljudskom ocjenjivanju koja se može koristiti u procesu treniranja sustava kako bi se vidio napredak. U radovima (Seljan i Dunder, 2015⁵¹; Brkić i sur., 2011⁵²) navode se prednosti, ali i nedostaci automatskih metrika, kojima je cilj ostvariti procjenu evaluacije što bližu ljudskoj procjeni.

Također, koristi se i metoda ocjenjivanja prijevoda po kojoj se dobiveni prijevod zadanog teksta ponovno prevodi na originalni jezik istim alatom za prevođenje (engl. *round-trip translation*). Iako je ovo možda najlakša metoda, idealna za laike, ona nije pouzdana jer se neki točni prevodi prilikom ponovnog prevođenja mogu krivo prevesti dok neki originalno netočni privodi prilikom ponovnog prijevoda na izvorni jezik mogu doći do originalnog teksta.

⁴⁷ Brkić, M., Seljan, S. i Vičić, T., 2009. Evaluation of the statistical machine translation service for Croatian-English. pp.319-332

⁴⁸ Ibid.

⁴⁹ Sepesy Maučec, M. and Donaj, G., 2020. Machine Translation and the Evaluation of Its Quality. *Recent Trends in Computational Intelligence*

⁵⁰ Ibid.

⁵¹ Seljan, S., Dunder, I. (2015). Machine Translation and Automatic Evaluation of English/Russian-Croatian. International Conference "Corpus Linguistics", 72.79.

⁵² Brkić, M., Seljan, S., Matetić, M. (2011). Machine translation evaluation for croatian-english and english-croatian language pairs. NLPCS Workshop: Human-Machine Interaction in Translation. Copenhagen: Copenhagen Business School, 93-104.

3.1. Ljudska evaluacija strojnog prijevoda

Papineni i sur. (2002, p. 311)⁵³ tvrde kako je ljudsko ocjenjivanje strojnog prevođenja najdetaljnije, ali tako i najskuplje te može potrajati mjesecima i taj napor se onda kasnije ne može ponovno iskoristiti. Ručno ocjenjivanje najbolje vrše bilingvalni stručnjaci tako što uspoređuju strojni prijevod sa ulaznim tekstom te provjeravaju točnost prenesenih informacija, stil pisanja, gramatiku itd. Naravno takvi stručnjaci nisu uvijek na raspolaganju za ocjenjivanje pa stoga taj zadatak mogu izvršiti i izvorni govornici ciljnog jezika. Strojne prijevode ocjenjuju isključivo prema izlaznom tekstu, bez ikakvog znanja o ulaznom tekstu (Brkić i sur., 2009, p. 324)⁵⁴.

Ljudska evaluacija strojnog prijevoda ocjenjuje mnoge aspekte prijevoda uključujući i tečnost prijevoda (engl. *fluency*), adekvatnost (engl. *adequacy*) i vjernost prenesenih informacija (engl. *fidelity*) (Papineni i sur., 2002, p. 311)⁵⁵. Brkić i sur. (2009, p. 324)⁵⁶ tvrde da se tečnost teksta odnosi na gramatiku i odabir riječi te da postoje tri aspekta tečnosti teksta, a to su jasnoća, prirodnost teksta i stil. Nadalje, adekvatnost, prema Brkić i sur. (2009, p. 324)⁵⁷, preispituje da li je neki dio izvornog teksta izgubljen ili promijenjen ili je nešto dodano u izlaznom tekstu. Brkić i sur. (2009, p. 324)⁵⁸ naglašavaju kako su presude o tečnosti i adekvatnosti obično su povezane, što ukazuje ili na poteškoće u razlikovanju dvaju kriterija ili samo na činjenicu da negramatične rečenice i pogrešan odabir riječi imaju manje značenja. Ocjene fluentnosti i adekvatnosti se stavljaju na skalu od 1 do 5, gdje 1 označuje potpuno nerazumljiv tekst, a 5 označuje savršen prijevod. Brkić i sur. (2009, p. 324)⁵⁹ preporučaju da ocjenjivači prvo pročitaju izlazni tekst, a tek onda referentni prijevod jer bi u protivnom ljudski um upotpunio informacije koje nedostaju. Fluentnost se također može mjeriti i kroz vrijeme potrebno za čitanje izlaznog teksta i cijenu naknadnog ljudskog uređivanja strojnog prijevoda (Brkić i sur., 2009, p. 324)⁶⁰.

⁵³ Papineni, K., Roukos, S., Ward, T. and Zhu, W., 2001. BLEU: a method for automatic evaluation of machine translation. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02*

⁵⁴ Brkić, M., Seljan, S. i Vičić, T., 2009. Evaluation of the statistical machine translation service for Croatian-English. pp.319-332

⁵⁵ Papineni, K., Roukos, S., Ward, T. and Zhu, W., 2001. BLEU: a method for automatic evaluation of machine translation. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02*

⁵⁶ Brkić, M., Seljan, S. i Vičić, T., 2009. Evaluation of the statistical machine translation service for Croatian-English. pp.319-332

⁵⁷ Ibid.

⁵⁸ Ibid.

⁵⁹ Ibid.

⁶⁰ Ibid.

Ovakva evaluacija se, kako Brkić i sur. tvrde (2012, p. 2144)⁶¹, smatra „zlatnim standardom“ iako je inherentno subjektivna te zahtjeva puno vremena, resursa i ljudskih napora. Stoga prema Papineni i sur. (2002, p. 311)⁶², nije primjerena za testiranje razvoja alata za strojno prevođenje gdje je potrebno nadzirati dnevne promjene kako bi se mogle razaznati dobre ideje od onih loših.

3.2. Automatska evaluacija strojnog prevođenja

Automatska evaluacija prijevoda je proces u kojem posebni programi ocjenjuju strojni prijevod bez značajne pomoći ljudi. Za rad automatskog ocjenjivanja strojnog prevođenja potrebni su mnogi referentni prijevodi zbog velike varijabilnosti čak i u ljudskim prijevodima (Sepesy Maučec i Donaj, 2020, p. 9⁶³; Brkić i sur., 2012⁶⁴). Ti prijevodi se koriste kako bi se usporedio prijevod strojnog prevođenja s profesionalnim ljudskim prijevodom te pronašle korelacije i odstupanja na temelju kojih se tekst i ocjenjuje. Primarni cilj automatske evaluacije jest visoka korelacija s ljudskom evaluacijom koja se smatra „zlatnim standardom“. Automatska evaluacija je generalno brza, jeftina i zahtjeva minimalni ljudski trud, ali zahtijeva specifična znanja programiranja i poznavanje područja računalne obrade jezika (Brkić i sur., 2012, p. 2144⁶⁵; Seljan i Dunder, 2015⁶⁶).

Jedan od prvih sustava za automatsko ocjenjivanje strojnog prijevoda jest BLEU (engl. *Bilingual Evaluation Understudy*). Nakon njega nastali su i drugi sustavi poput NIST (engl. *National Institute of Standards and Technology*), METEOR (engl. *Metric for Evaluation of Translation with Explicit Ordering*) i TER-a koji svi daju bolje ocjene onom prijevodu koji je najbliži ljudskom referentnom prijevodu. Sepesy Maučec i Donaj (2020, p. 9)⁶⁷ zato naglašavaju da svi ti sustavi trebaju referentne tekstove s kojima bi usporedili dobivene strojne prijevode.

⁶¹ Brkić, M., Seljan, S. i Vičić, T., 2012. BLEU Evaluation of Machine-Translated English-Croatian Legislation. LREC, pp.2143-2148

⁶² Papineni, K., Roukos, S., Ward, T. and Zhu, W., 2001. BLEU: a method for automatic evaluation of machine translation. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02*.

⁶³ Sepesy Maučec, M. i Donaj, G., 2020. Machine Translation and the Evaluation of Its Quality. *Recent Trends in Computational Intelligence*

⁶⁴ Brkić, M., Seljan, S. i Vičić, T., 2012. BLEU Evaluation of Machine-Translated English-Croatian Legislation. LREC, pp.2143-2148

⁶⁵ Ibid.

⁶⁶ Seljan, S., Dunder, I. (2015). Automatic quality evaluation of machine-translated output in sociological-philosophical-spiritual domain. *Information Systems and Technologies (CISTI 2015)*, 1-4.

⁶⁷ Sepesy Maučec, M. i Donaj, G., 2020. Machine Translation and the Evaluation of Its Quality. *Recent Trends in Computational Intelligence*

BLEU se uglavnom koristi za cijele korpuse prijevoda te ih ocjenjuje na skali od 0 do 1 gdje 1 dobiju samo prijevodi koji su identični referentnom te stoga niti većina ljudskih prijevoda neće dobiti ocjenu 1 (Papineni i sur., 2002, p. 316)⁶⁸. Wu i sur. (2016, p. 8)⁶⁹ naglašavaju da zbog svog dizajna za ocjenjivanje korpusa, BLEU ima neka nepoželjna svojstva kada je u pitanju ocjenjivanje pojedinačnih rečenica.

Brkić i sur. (2012, p. 2144)⁷⁰ tvrde kako se BLEU ocjenjivanje provodi tako što se uparaju nizovi riječi (tzv. N-grame) iz izlaznog teksta s onima u referentnom prijevodu te se broje postignute parove na razini rečenice i tako za cijeli tekst. Nadalje, Papineni i sur. (2002, p. 314)⁷¹ ističu da ovaj sustav penalizira pretjerano korištenje riječi koja se pojavlja u referentnom tekstu. U ovom sustavu, adekvatnost se ocjenjuje na osnovi preciznosti riječi, dok se tečnost teksta ocjenjuje ovisno o preciznosti duljih n-grama (Brkić i sur., 2012, p. 2144)⁷².

S obzirom na to da je BLEU automatska metrika koja ne ovisi o jeziku, Brkić i sur. (2012, p. 2144)⁷³ ističu da ne uzima u obzir morfološke varijante riječi. Ova metoda zahtjeva identične parove riječi u referentnom prijevodu i onome koji ocjenjuje što predstavlja problem kod ocjenjivanja flektivnih jezika kao što je hrvatski.

⁶⁸ Papineni, K., Roukos, S., Ward, T. and Zhu, W., 2001. BLEU: a method for automatic evaluation of machine translation. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02*

⁶⁹ Wu, Y., Schuster, M., Chen, Z., Le, Q., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, Ł., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., Stevens, K., Kurian, G., Patil, N., Wang, W., Young, C., Smith, J., Riesa, J., Rudnick, A., Vinyals, O., Corrado, G., Hughes, M. i Dean, J., 2016. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation

⁷⁰ Brkić, M., Seljan, S. i Vičić, T., 2012. BLEU Evaluation of Machine-Translated English-Croatian Legislation. LREC, pp.2143-2148

⁷¹ Papineni, K., Roukos, S., Ward, T. and Zhu, W., 2001. BLEU: a method for automatic evaluation of machine translation. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02*

⁷² Brkić, M., Seljan, S. i Vičić, T., 2012. BLEU Evaluation of Machine-Translated English-Croatian Legislation. LREC, pp.2143-2148

⁷³ Ibid.

4. Pristupi strojnom prevođenju

Razvojem interdisciplinarnog područja računalne obrade jezika, strojnog učenja, podatkovne znanosti i računalne lingvistike razvija se i računalno prevođenje. Što su računala moćnija, to se prevođenje može sve više oslanjati na njih te se s toga sve moderniji modeli strojnog prevođenja sve više oslanjaju na računala i zahtijevaju sve manje intervencija od strane ljudi. Tako su se i razvili različiti pristupi računalnom prevođenju od strojnog prevođenja temeljenog na pravilima, preko statističkog strojnog prevođenja do neuralnog prevođenja. U ovome poglavlju bit će prikazan svaki od navedenih pristupa.

4.1. Strojno prevođenje temeljeno na pravilima

Konvencionalni sustavi za strojno prevođenje koriste jezična pravila o izvornom i ciljnom jeziku kao izvoru znanja (Sumita i Iida, 1991, p. 185)⁷⁴. Ovakav pristup prijevodu naziva se strojno prevođenje temeljeno na pravilima (engl. *Rule-based machine translation - RBMT*). S obzirom na to da su jezična pravila proučavali i pisali ljudi, strojno prevođenje temeljeno na pravilima je prva strategija koju su istraživači u 70-im godinama prošlog stoljeća upotrijebili za razvoj strojnog prevođenja (Charoenpornasawat i sur., 2002 p. 1)⁷⁵ te je ovaj pristup bio najpopularniji do ranih 2000-ih godina (Zhang i Zong, 2020, p. 1)⁷⁶. Prednost ovakvog pristupa, prema Charoenpornasawat i sur. (2002, p. 1)⁷⁷, jest ta da omogućuje dubinsku analizu na sintaktičkoj i semantičkoj razini. Strojno prevođenje temeljeno na pravilima se, prema Sepesy Maučec i Donaj (2020, p. 2)⁷⁸, dalje dijeli na tri podskupine, a to su interlingua strojno prevođenje, strojno prevođenje temeljeno na transferu i strojno prevođenje temeljeno na rječnicima.

Za izradu ovog sustava za strojno prevođenje potrebni su stručnjaci iz domene računalne obrade jezika (engl. *natural language processing – NLP*) koje ujedinjuju znanja iz područja informatike, računarstva, računalne lingvistike i drugih područja, kako bi prikupili formalna morfološka,

⁷⁴ Sumita, E. i Iida, H., 1991. Experiments and prospects of Example-Based Machine Translation. Proceedings of the 29th annual meeting on Association for Computational Linguistics

⁷⁵ Charoenpornasawat, P., Sornlertlamvanich, V. i Charoenporn, T., 2002. Improving translation quality of rule-based machine translation. *COLING-02 on Machine translation in Asia*

⁷⁶ Zhang, J. i Zong, C., 2020. Neural Machine Translation: Challenges, Progress and Future

⁷⁷ Charoenpornasawat, P., Sornlertlamvanich, V. i Charoenporn, T., 2002. Improving translation quality of rule-based machine translation. *COLING-02 on Machine translation in Asia*

⁷⁸ Sepesy Maučec, M. and Donaj, G., 2020. Machine Translation and the Evaluation of Its Quality. *Recent Trends in Computational Intelligence*

sintaktička i semantička pravila jezika koja bi se onda koristila kao baza za program (Sepesy i Donaj, 2020, p. 3)⁷⁹. Prema Zhang i Zong (2020, p. 1)⁸⁰ za izradu ovog sustava potrebni su dvojezični jezični stručnjaci koji su odgovorni za dizajniranje specifičnih pravila za analizu izvornog jezika, transformaciju iz izvornog jezika u ciljni i generiranje ciljnog jezika. Jedan od mogućih modela korištenja formalne gramatike je i leksičko-funkcionalna gramatika koja se koristi u izradi pravila transfera (Seljan, 2003⁸¹; Borra i sur., 2007⁸²; Way, 1999⁸³; Seljan, 2003b⁸⁴). Potreban je veliki ljudski napor kako bi se pripremila sva pravila i jezični resursi kao što su morfološki analizatori, sustavi za označavanje vrsta riječi (engl. *part-of-speech taggers*) i sustavi sintaktičkog raščlanjivanja (engl. *syntactic parsers*), zatim dvojezični rječnici, pravila transfera, morfološki generator, pravila za promjenu redoslijeda itd. (Sreelekha, 2017, p. 1)⁸⁵. Za razliku od ostalih sustava za prevođenje, strojno prevođenje temeljeno na pravilima ne zahtijeva paralelne korpuse niti bilo kakve druge tekstove na izvornim i ciljnim jezicima jer funkcionira po pravilima te nije potrebno dodavati primjere. Ovo je velika prednost kada su u pitanju jezici koji nemaju velike korpuse ili uopće nemaju dvojezične korpuse sa željenim ciljnim jezikom.

Softver strojnog prevođenja temeljenog na pravilima raščlanjuje tekst (A pioneer and global leader in translation solutions, 2020)⁸⁶. Ovaj postupak zahtijeva opsežne leksikone s morfološkim, sintaktičkim i semantičkim informacijama, kao i velike skupove pravila. Softver koristi ove složene skupove pravila, a zatim prenosi gramatičku strukturu izvornog jezika na ciljni jezik (What is Machine Translation? Rule Based vs. Statistical | Systran, 2020)⁸⁷.

Najjednostavnije objašnjenje rada sustava za strojno prevođenje temeljeno na pravilima jest da se ulazni tekst na nekom jeziku analizira na više razina te se zatim uz pomoć sintaktičke, semantičke

⁷⁹ Sepesy Maučec, M. and Donaj, G., 2020. Machine Translation and the Evaluation of Its Quality. *Recent Trends in Computational Intelligence*

⁸⁰ Zhang, J. i Zong, C., 2020. Neural Machine Translation: Challenges, Progress and Future

⁸¹ Seljan, S. (2003). Leksičko-funkcionalna gramatika hrvatskoga jezika: teorijski praktični modeli. Doktorska disertacija. Sveučilište u Zagrebu.

⁸² Borra, A., Chan, E.A., Lim, C.I., Tan, R.B., & Tong, M.C. (2007). LFG-Based Machine Translation Engine for English and Filipino.

⁸³ Way, Andy. (1999). A Hybrid Translation Model using LFG-DOP.

⁸⁴ Seljan, S. (2003). The Role of the Lexicon in Lexical-Functional Grammar-Example on Croatian. International Language Technologies Conference IS-LTC 2006, 198-203.

⁸⁵ Sreelekha, S., 2017. Statistical vs rule based machine translation; a case study on Indian language perspective

⁸⁶ Systransoft.com. 2020. *A Pioneer And Global Leader In Translation Solutions*. (online) Dostupno na: <<https://www.systransoft.com/systran/>>

⁸⁷ Systransoft.com. 2020. *What Is Machine Translation? Rule Based Vs. Statistical | SYSTRAN*. (online) Dostupno na: <<https://www.systransoft.com/systran/translation-technology/what-is-machine-translation/>>

i morfološke analize ciljnog jezika slaže izlazni tekst. Sreelekha (2017, p. 2)⁸⁸ ovaj proces opisuje u fazama. Prvo nastupa analiza ulaznog teksta gdje se ekstrahiraju informacije vezane uz morfologiju, dijelove govora (engl. *Part of speech*), fraze i uklanjanje višeznačnosti riječi tj. određivanje na što se određena riječ točno odnosi (Sreelekha, 2017, p. 2)⁸⁹. Zatim slijedi faza leksičkog prijenosa za koju Sreelekha (2017, p. 2)⁹⁰ nadovi da sadrži dva koraka, jedan je prijevod riječi, a drugi je analiza gramatike. Kod prijevoda riječi, dana je korijenska riječ (dakle onakva kakva se nalazi u rječniku) tj. lema na izvornom jeziku koja je zamijenjena lemom na ciljnom jeziku uz pomoć dvojezičnog rječnika. Prijevod gramatike označava dodjeljivanje sufiksa. U fazi proizvodnje ispravljaju se rodovi prevedenih riječi te se provodi slaganje s ostalim riječima u rečenici što osigurava da se sve riječi u frazama slažu u rodu, broju i padežu te da se i fraze međusobno slažu (Sreelekha, 2017, p. 2)⁹¹.

Strojno prevođenje temeljeno na pravilima radi na principu točnog uparivanja, pa ako ne postoji pravilo koje u potpunosti pokriva dani ulazni tekst onda ga strojno prevođenje temeljeno na pravilima jednostavno ne može prevesti (Sumita i Iida, 1991, p. 192)⁹². Za razliku od, na primjer, prevođenja temeljenog na primjerima koje će u slučaju da ne nađe potpuni par, odabrati drugi najbolji par (Sumita i Iida, 1991, p. 192)⁹³.

Pristup temeljen na pravilima zahtjeva veliko znanje interdisciplinarnog područja računarstva, jezičnog inženjeringa i računalne lingvistike. Zhang i Zong (2020, p. 1)⁹⁴ naglašavaju da je zbog svoje subjektivnosti i potrebe za intenzivnim radom, strojno prevođenje temeljeno na pravilima teško nadograditi. Nemoguće je predvidjeti sve moguće scenarije i greške te napisati pravila za njihovo rješavanje, s obzirom na nebrojene kombinacije i izuzetke (Charoenpornasawat i sur., 2002, p. 1)⁹⁵. Srodno tome, potrebni su i dobri jezični izvori, kako gramatike, tako i rječnici. Sumita i Iida (1991, p. 191)⁹⁶ naglašavaju da je samo stvaranje potrebne baze pravila koja su uzajamno

⁸⁸ Sreelekha, S., 2017. Statistical vs rule based machine translation; a case study on Indian language perspective

⁸⁹ Ibid.

⁹⁰ Ibid.

⁹¹ Ibid.

⁹² Sumita, E. i Iida, H., 1991. Experiments and prospects of Example-Based Machine Translation. Proceedings of the 29th annual meeting on Association for Computational Linguistics

⁹³ Ibid.

⁹⁴ Zhang, J. i Zong, C., 2020. Neural Machine Translation: Challenges, Progress and Future

⁹⁵ Charoenpornasawat, P., Sornlertlamvanich, V. i Charoenporn, T., 2002. Improving translation quality of rule-based machine translation. *COLING-02 on Machine translation in Asia*

⁹⁶ Sumita, E. i Iida, H., 1991. Experiments and prospects of Example-Based Machine Translation. Proceedings of the 29th annual meeting on Association for Computational Linguistics

ovisna i vrlo kompleksna, a naknadno dodavanje dodatnih pravila još više otežava stvar. Ovako velika i kompleksna baza pravila teško se i sporo ugrađuje u program što navodi programera da odustane od pokušaja poboljšanja sustava.

Kako bi se poboljšao rezultat strojnog prevođenja temeljenog na pravilima, pravila o generiranju ili analizi prijevoda bi se morala dodati ili modificirati što zahtjeva puno vremena i znanja, a ne garantira bolji prijevod. Štoviše, mogao bi čak i uzrokovati pogoršanje trenutnog stanja (Charoenpornasawat i sur., 2002, p. 1)⁹⁷.

Nadalje, različiti fenomeni u jeziku predstavljaju problem za ovakvo prevođenje jer se ne mogu pokriti generalnim pravilima (Sumita i Iida, 1991, p. 186)⁹⁸. Najpoznatiji takav fenomen su idiomi.

Iako su kompleksna pravila vezana uz sintaksu, semantiku, kontekst i stvarni svijet potrebna za strojno prevođenje, što se više razrađuje baza pravila za strojno prevođenje temeljno na pravilima, Sumita i Iida (1991, p. 191)⁹⁹ zaključuju da ono postaje sve sporije i teže za poboljšati. Stoga se može zaključiti da održavanje i poboljšavanje ovog sustava nije nimalo jednostavno ni lako, a čini se da nije ni isplativo.

4.2. Statističko strojno prevođenje

Statističko strojno prevođenje se, kao što i samo ime kaže, temelji na statistici te ono izračunava „vjerojatnost da će se neka riječ prevesti nekom drugom ili da će se prijevoditi dviju riječi koje se nalaze jedna pokraj druge također nalaziti jedni pored drugih“ (Ljubas, 2018, p. 72)¹⁰⁰. Postoje podvrste statističkog prevođenja kao što su prevođenje temeljno na sintaksi (engl. *syntax-based*), prevođenje temeljeno na frazama (engl. *phrase-based*) koje prevodi nizove riječi ili fraze različitih duljina (Wu i sur., 2016, p. 2)¹⁰¹ i prevođenje temeljeno na riječima (engl. *word-based translation*)

⁹⁷ Charoenpornasawat, P., Sornlertlamvanich, V. i Charoenporn, T., 2002. Improving translation quality of rule-based machine translation. *COLING-02 on Machine translation in Asia*

⁹⁸ Sumita, E. i Iida, H., 1991. Experiments and prospects of Example-Based Machine Translation. Proceedings of the 29th annual meeting on Association for Computational Linguistics

⁹⁹ Ibid.

¹⁰⁰ Ljubas, S., 2018. Prijelaz sa statističkog na neuronski model: usporedba strojnih prijevoda sa švedskog na hrvatski jezik. *Hieronymus-časopis za istraživanja prevođenja i terminologije*, 4, pp.72-91

¹⁰¹ Wu, Y., Schuster, M., Chen, Z., Le, Q., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, Ł., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., Stevens, K., Kurian, G., Patil, N., Wang, W., Young, C., Smith, J., Riesa, J., Rudnick, A., Vinyals, O., Corrado, G., Hughes, M. i Dean, J., 2016. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation

koje prijevodi riječ po riječ. Prema Wu i sur. (2016, p. 2)¹⁰², statističko strojno prevođenje je desetljećima bilo dominantna prevoditeljska paradigma, sve do razvoja neuralnog prevođenja. Toliko je bilo bitno za strojno prevođenje da neki autori poput Shofner (2020)¹⁰³ tvrde kako je baš statistički pristup oblikovao strojno prevođenje u ovakvo kakvo ga danas poznajemo.

Statistički modeli strojnog prevođenja, prema Dunđeru (2015, p. 17)¹⁰⁴, “proces generiranja podataka dijele u manje korake koji se zatim zasebno modeliraju i statistički opisuju, a konačan rezultat nastaje kombiniranjem svih pojedinačnih koraka”.

Dunđer (2015, p. 17)¹⁰⁵ ističe kako je za treniranje statističkog strojnog prevođenja koristite paralelni i jednojezični korpusi koji su osnova ovog sustava, kako bi moglo izračunati vjerojatnost pojave određene riječi kod prijevoda, te se mogu koristiti i jezično neovisni alati. Prema Shofner (2020)¹⁰⁶, dobar korpus za statističko prevođenje sadrži oko 100 milijuna riječi ili 1 milijun uparenih rečenica kako bi program u određenoj mjeri bio uspješan. Domene tekstova koji se koriste za treniranje sustava imaju veliki utjecaj na njegov konačan prijevod jer će najbolji prijevod biti na iz iste domene na kojoj je sustav treniran. Jedna velika prednost ovog sustava je što ne ovisi o jeziku već se program može iskoristiti za bilo koji jezični par. Statističko strojno prevođenje se dalje dijeli na podskupove: temeljeno na riječima, na frazama, na sintaksi i hijerarhijskoj frazi (Shofner, 2020)¹⁰⁷.

Model prevođenja temeljeno na frazama se često koristi. Osnovna ideja ovakvog prevođenja jest „segmentirati skup rečenica izvornog jezika (engl. *test set*) u fraze, odnosno nizove riječi koje se zatim prevode (tj. zamjenjuju) u ciljni jezik“ (Dunđer, 2015, p. 17)¹⁰⁸. Dakle rečenica se dijeli na manje dijelove koje je zatim lakše obrađivati. Ovakve fraze se onda prevode te se konačan prijevod

¹⁰² Wu, Y., Schuster, M., Chen, Z., Le, Q., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, Ł., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., Stevens, K., Kurian, G., Patil, N., Wang, W., Young, C., Smith, J., Riesa, J., Rudnick, A., Vinyals, O., Corrado, G., Hughes, M. i Dean, J., 2016. Google’s Neural Machine Translation System: Bridging the Gap between Human and Machine Translation

¹⁰³ Shofner, K., 2020. *The Pros And Cons Of Statistical Machine Translation*. (online) Unitedlanguagegroup.com. Dostupno na: <<https://www.unitedlanguagegroup.com/blog/pros-and-cons-statistical-machine-translation>>

¹⁰⁴ Dunđer, I., 2015. Sustav za statističko strojno prevođenje i računalna adaptacija domene. Doktorska disertacija. Sveučilište u Zagrebu. Filozofski fakultet.

¹⁰⁵ Ibid.

¹⁰⁶ Shofner, K., 2020. *The Pros And Cons Of Statistical Machine Translation*. (online) Unitedlanguagegroup.com. Dostupno na: <<https://www.unitedlanguagegroup.com/blog/pros-and-cons-statistical-machine-translation>>

¹⁰⁷ Ibid.

¹⁰⁸ Dunđer, I., 2015. Sustav za statističko strojno prevođenje i računalna adaptacija domene. Doktorska disertacija. Sveučilište u Zagrebu. Filozofski fakultet.

na ciljni jezik dobiva sastavljanjem prevedenih fraza koje se zovu prijevodni kandidati (engl. *candidate translations*), prema Dunder (2015, p. 17)¹⁰⁹. Kako bi sustav odredio koji je prijevodni kandidat najbolji pridružuje mu se, kako Dunder (2015, p. 18)¹¹⁰ navodi, „skup pripadajućih vrijednosti težina značajki modela“.

Sepesy Maučec i Donaj (2020, p. 4)¹¹¹ navode tri komponente sustava statističkog strojnog prevođenja temeljenog na frazama koje utječu na kvalitetu prijevoda, a to su: model prevođenja fraza, model preslagivanja redoslijeda riječi (engl. *reordering model*) i jezični model (engl. *language model*). Ove značajke za vrijeme treniranja modela dobivaju određenu vrijednost težine (engl. *weights*) koje potom „utječu na logiku modela statističkog strojnog prevođenja temeljenog na frazama“ (Dunder, 2015, p. 18)¹¹². Dakle prijevod pomoću modela statističkog strojnog prevođenja se dobiva kroz niz ulančanih i uvjetovanih odluka koje su osnovane na vrijednosti težine prijevodnih kandidata. Kako bi se prijevod u potpunosti generirao u modelu statističkog strojnog prevođenja Dunder (2015, p. 19)¹¹³ tvrdi da su potrebne „tri ključne komponente: jezični model (engl. *language model*), prijevodni model (engl. *translation model*) i dekodera“. Jezični model se trenira na korpusu ciljnog jezika kako bi se koristio kao pomoć u odabiru točnih riječi i redoslijeda riječi te tako osigurao fluentnost i prikladnost strojnog prijevoda (Dunder, 2015, p. 23)¹¹⁴. Prijevodni model se razvija pomoću paralelnog korpusa kako bi mogao upariti nizove riječi iz izvornog jezika u ciljni jezik (Dunder, 2015, p. 39)¹¹⁵. Zadaća dekodera bi stoga bila naći prijevod s maksimalnom vjerojatnošću (Dunder, 2015, p. 58)¹¹⁶.

Statističko strojno prevođenje dakle funkcionira u tri temeljne faze, koje Dunder (2015, pp. 17-18)¹¹⁷ navodi kao: uparivanje ili sravnavanje segmenata, ekstrakcija fraznih parova te na kraju se radi izračun vjerojatnosti za svaki frazni par.

¹⁰⁹ Dunder, I., 2015. Sustav za statističko strojno prevođenje i računalna adaptacija domene. Doktorska disertacija. Sveučilište u Zagrebu. Filozofski fakultet.

¹¹⁰ Ibid.

¹¹¹ Sepesy Maučec, M. i Donaj, G., 2020. Machine Translation and the Evaluation of Its Quality. Recent Trends in Computational Intelligence,.

¹¹² Dunder, I., 2015. Sustav za statističko strojno prevođenje i računalna adaptacija domene. Doktorska disertacija. Sveučilište u Zagrebu. Filozofski fakultet.

¹¹³ Ibid.

¹¹⁴ Ibid,

¹¹⁵ Ibid.

¹¹⁶ Ibid.

¹¹⁷ Ibid.

U modelu statističkog strojnog prevođenja segmenti izvornog jezika se uparuju sa segmentima iz ciljnog jezika, taj proces se još naziva i sravnavanje. Takvi sravnjeni korpusi predstavljaju osnovu za strojno prevođenje i računalno potpomognuto prevođenje (Seljan i Pavuna, 2006¹¹⁸). Nadalje, same fraze se zatim dobivaju iz sravnjenih riječi u izvornom i ciljnom jeziku te valja naglasiti da se ne samo fraze moraju sravnjavati već se i riječi unutar fraza moraju međusobno upariti (Dunđer, 2015, p. 17)¹¹⁹.

Prednost ovakvog sustava strojnog prevođenja je ta što, kako Dunđer (2015, p. 19)¹²⁰ navodi, obazire se na morfološke i leksičke varijacije riječi što je veoma važno za jezike poput hrvatskog koji je morfološki bogat.

Statističko strojno prevođenje također ima i mnoštvo neriješenih problema i nedostataka. Jedan od problema je taj što mu je prevođenje tekstova koji nisu slični onima na kojima je treniran otežano. Shofner (2020)¹²¹ tvrdi da statistički modeli mogu dati izvrsne prijevode za tekstove koji su u sličnim domenama korištenima za trening, koji su tehnički i pisani jednostavnim stilom, dok će se mučiti s tekstovima koji sadrže žargone, idiome ili ležerniji stil pisanja. Jedan od problema je nemogućnost pregledavanja pravopisnih pogrešaka tj. ako je neka riječ krivo napisana onda je sustav uvodi pod potpuno novu riječ i ne može ju povezati s njenim pravim značenjem jer za vrijeme treniranja takva riječ se nije pojavila pa se tako nije niti statistički opisala (Dunđer, 2015, p. 21)¹²².

Nadalje, postoji i problem s redosljedom riječi u rečenici. S obzirom na to da sintaksa, koja definira načela i pravila slaganja rečenica prirodnog jezika, nije integrirana u klasične modele statističkog strojnog prevođenja takvi modeli onda ne mogu znati kako posložiti riječi u rečenicama. Ovakvi problemi su pogotovo zamjetni kada se prevodi s jezika s relativno slobodnim poretkom riječi u rečenici, kao što je hrvatski, na jezik s fiksnim poretkom riječi, kao što je engleski, i obrnuto (Dunđer, 2015, p. 22)¹²³. Nadalje, ako željeni jezični parovi koriste drugačiju

¹¹⁸ Seljan, S., Pavuna, D. (2006). Translation Memory Database in the Translation Process. Information and Intelligent Systems IIS 2006, 327-332.

¹¹⁹ Dunđer, I., 2015. Sustav za statističko strojno prevođenje i računalna adaptacija domene. Doktorska disertacija. Sveučilište u Zagrebu. Filozofski fakultet.

¹²⁰ Ibid.

¹²¹ Shofner, K., 2020. *The Pros And Cons Of Statistical Machine Translation*. (online) Unitedlanguagegroup.com. Dostupno na: <<https://www.unitedlanguagegroup.com/blog/pros-and-cons-statistical-machine-translation>>

¹²² Dunđer, I., 2015. Sustav za statističko strojno prevođenje i računalna adaptacija domene. Doktorska disertacija. Sveučilište u Zagrebu. Filozofski fakultet.

¹²³ Ibid.

jezičnu konstrukciju (na primjer: subjekt-objekt-predikat i subjekt-predikat-objekt) onda su problemi s redosljedom riječi gotovo pa i neizbježni.

Jedan od najvećih problema statističkog prevođenja je i višeznačnost riječi. Najbolji primjer ovog problema višeznačnosti mogao bi biti prijevod engleske riječ „can“ na hrvatski gdje ona može značiti „limenka“, „konzervirati“, „toalet“ ili može biti pomoćni glagol „moći“. Ovisno o kontekstu, statističko prevođenje trebalo bi predvidjeti koje je značenje najvjerojatnije, ali i tada se događaju greške. Povezano s ovim problemom je i problem izbacivanja ili neprevođenja nekih riječi iz izvornika u ciljni jezik.

4.3. Neuralno strojno prevođenje

Neuralno strojno prevođenje predstavlja najnoviji pristup računalnom prevođenju koje se razvilo uz razvoj umjetne inteligencije (engl. *artificial intelligence* - AI). Ova vrsta strojnog prevođenja je, prema Zhang i Zong (2020, p. 1)¹²⁴, znatno napredovala zadnjih godina te se sada koristi u gotovo svim online programima za strojno prevođenje. O njegovom brzom napretku govori i činjenica koju Koehn (2017, p. 6)¹²⁵ ističe, a to je da je 2015. godine na natjecanju organiziranom od strane Konferencije za strojno prevođenje (engl. *Conference on Machine Translation* - WMT) bio samo jedan potpuno neuralni program, te iako je bio konkurent, nije pobijedio jer su tada statistički programi bili kvalitetniji. Međutim, već iduće godine neuralni stroj za prevođenje je pobijedio na natjecanju, a godine 2017. gotovo svi prijavljeni programi radili su na principu neuralnog strojnog prevođenja (Koehn, 2017, p. 6)¹²⁶. Zhang i Zong (2020, p. 17) ističu kako je 2018. godine sustav za neuralno strojno prevođenje za prevođenja vijesti s kineskog na engleski dosegnuo ljudsku kvalitetu te nije bilo značajne razlike između ocjena strojnog prijevoda i profesionalnog ljudskom prijevoda. Prema Cho i sur. (2014, p. 1)¹²⁷ i Ljubas (2018, p. 74)¹²⁸ ovaj pristup prevođenju inspiriran je neuronskim mrežama i trendom dubokog učenja (engl. *deep*

¹²⁴ Zhang, J. i Zong, C., 2020. Neural Machine Translation: Challenges, Progress and Future.

¹²⁵ Koehn, P., 2017. *Statistical Machine Translation*. Johns Hopkins University, pp.1–113

¹²⁶ Ibid.

¹²⁷ Cho, K., van Merriënboer, B., Bahdanau, D. i Bengio, Y., 2014. On the Properties of Neural Machine Translation: Encoder–Decoder Approaches. *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*

¹²⁸ Ljubas, S., 2018. Prijelaz sa statističkog na neuronski model: usporedba strojnih prijevoda sa švedskog na hrvatski jezik. *Hieronymus-časopis za istraživanja prevođenja i terminologije*, 4, pp.72-91

learning) koje koristi nekoliko razina za progresivno izdvajanje značajki više razine iz sirovog unosa. Najveća prednost koju ovaj pristup ima nad ostalima je njegova sposobnost direktnog učenja preslikavanja od ulaznog teksta u pridruženi izlazni tekst (Wu i sur., 2016, p. 1)¹²⁹. Prema Zhang i Zong (2020, p. 4)¹³⁰ neuralno strojno prevođenje se formalno definira kao sekvencionalno predviđanje u kojem su skrivene četiri pretpostavke, a to su: ulazni tekst je rečenica, a ne odlomak ili cijeli dokument; izlazni tekst se generira autoregresivno slijeva na desno; neuralno strojno prevođenje je kroz treniranje optimizirano za dvojezične podatke koji bi trebali uključivati puno paralelnih rečenica; neuralno strojno prevođenje obrađuje čiste tekstove (tokene, riječi i rečenice), a ne govor i videozapis.

Neuralno prevođenje u pravilu koristi dvije ponavljajuće neuronske mreže, tako da jedna obrađuje ulazni tekst, a druga onda generira izlazni tekst, ili sustav za kodiranje i dekodiranje gdje koder ekstrahira prikaz određene duljine iz izvorne rečenice bilo koje dužine, a dekodeer onda iz te reprezentacije generira ispravan prijevod (Cho i sur., 2014 p. 1)¹³¹.

Dakle, neuralno strojno prevođenje može raditi na principu ponavljajuće neuronske mreže (engl. *Recurrent neural network - RNN*) ili na principu kodiranja i dekodiranja (engl. *encoder – decoder network*).

Kao što je slučaj s ostalim principima za prevođenje, i neuralno strojno prevođenje zahtjeva mnogo resursa, uglavnom paralelne korpuse (Zhang i Zong, 2020, p. 1)¹³². Alat Google Translate je tako na primjer iskoristio dokumente iz Ujedinjenih Naroda i Europskog Parlamenta pisane standardnim jezikom bez žargona i nejasnoća koje su već profesionalni prevoditelji preveli na strane jezike pa su tako idealni paralelni korpusi za treniranje strojnog prevođenja (Google Translate, 2020)¹³³. Iako je Google Translate ranije koristio statistički pristup, krajem 2016. godine

¹²⁹ Wu, Y., Schuster, M., Chen, Z., Le, Q., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, Ł., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., Stevens, K., Kurian, G., Patil, N., Wang, W., Young, C., Smith, J., Riesa, J., Rudnick, A., Vinyals, O., Corrado, G., Hughes, M. i Dean, J., 2016. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation.

¹³⁰ Zhang, J. i Zong, C., 2020. Neural Machine Translation: Challenges, Progress and Future.

¹³¹ Cho, K., van Merriënboer, B., Bahdanau, D. i Bengio, Y., 2014. On the Properties of Neural Machine Translation: Encoder–Decoder Approaches. *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*

¹³² Zhang, J. i Zong, C., 2020. Neural Machine Translation: Challenges, Progress and Future.

¹³³En.wikipedia.org. 2020. Google Translate. (online) Dostupno na: <https://en.wikipedia.org/wiki/Google_Translate> (Pristupljeno: 28 kolovoz 2020).

počinje koristiti kao osnovu neuralno prevođenje te danas koristi kombinaciju neuralnog i statističkog strojnog prevođenja kako bi postigao najbolje rezultate (Google Translate, 2020)¹³⁴.

Neuralno strojno prevođenje koristi ponavljajuću neuronsku mrežu koja unutar sebe koristi jedinice dugog kratkotrajnog pamćenja (engl. *Long short-term memory - LSTM*) za svoj rad prepoznavanja uzoraka u nizu informacija. Dugo kratkotrajno pamćenje sadrži informacije izvan normalnog toka ponavljajuće neuronske mreže u posebnoj ograđenoj ćeliji (engl. *gated cell*). Ovdje se informacije mogu spremati, čitati i pisati. Prema Cho i sur. (2014, p. 2)¹³⁵ ponavljajuća neuronska mreža radi na nizovima različite dužine $x = (x_1, x_2, \dots, x_T)$ tako što održava skriveno stanje (engl. *hidden state*) h gdje se skriveno stanje $h(t)$ ažurira svaki t puta. Formula za ovakav model RNN-a je:

$$h^{(t)} = f(h^{(t-1)}, x_t),$$

Figure 1: Formula za RNN¹³⁶

Ovdje je f aktivacijska funkcija. Skriveno stanje za neko vrijeme t jest $h(t)$, a $h(t-1)$ je prijašnje skriveno stanje.

Ponavljajuća neuronska mreža u obzir dakle uzima ne samo trenutačnu ulaznu informaciju već i prijašnje ulaze te oni utječu na način na koji će se trenutačna ulazna informacija obraditi (Cho i sur., 2014, p. 2)¹³⁷. RNN stoga ima dva izvora ulaznih informacija; trenutačnu ulaznu informaciju i onu koja je direktno prethodila te se obje onda spajaju kako bi odredile kako će sustav reagirati na novu informaciju.

Cho i sur. (2014, p. 2)¹³⁸ također predlažu i rekurzivnu konvolucijsku neuronsku mrežu (engl. *recursive convolutional neural network*). To je još jedan pristup obradi nizova različite duljine

¹³⁴En.wikipedia.org. 2020. Google Translate. (online) Dostupno na: <https://en.wikipedia.org/wiki/Google_Translate> (Pristupljeno: 28 kolovoz 2020).

¹³⁵ Cho, K., van Merriënboer, B., Bahdanau, D. i Bengio, Y., 2014. On the Properties of Neural Machine Translation: Encoder–Decoder Approaches. *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*

¹³⁶ Ibid.

¹³⁷ Ibid.

¹³⁸ Ibid.

gdje se parametri svake razine dijele s cijelom mrežom. Ovaj pristup također koristi ogradačene ćelije koje mu omogućuju da odjednom sazna strukturu cijele ulazne rečenice.

Prema Cho i sur. (2014, p. 3)¹³⁹ sustav neuralnog strojnog prevođenja također koristi i uređaje za kodiranje i dekodiranje koji uz sebe istodobno koriste i tako zvanu mrežu pažnje (engl. *attention network*). Ovdje koder procesira ulaznu rečenicu bilo koje dužine i izrađuje vektor određene dužine koji predstavlja značenje te rečenice (Cho i sur., 2014, p. 3)¹⁴⁰. Dakle kako Koehn (2017, p. 48)¹⁴¹ tvrdi, zadatak enkodera je dati prikaz ulazne rečenice. Prikaz daje tako što ulaznu rečenicu pretvori u niz vektora gdje je jedan vektor jednak jednom ulaznom simbolu. Kada enkoder dođe do kraja ulazne rečenice, dakle predvidio je označivač kraja rečenice (engl. *end-of-sentence symbol* - *EOS*), onda skriveno stanje kodira značenje rečenice.

Ovo skriveno stanje se koristi za izradu prijevoda u fazi dekodiranja koja ne samo da mora imati dovoljno informacija kako bi mogla uspješno predvidjeti svaku sljedeću riječ, već mora i znati koji su dijelovi ulazne rečenice već prevedeni, a što se još treba prevesti (Koehn, 2017, pp. 47-48)¹⁴². Zatim pod utjecajem te kodirane reprezentacije, dekodeer, koji je također vrsta ponavljajuće neuronske mreže (Koehn, 2017, p. 49)¹⁴³, generira izlaznu rečenicu tako da uzima danu listu vektora te izrađuje jedan po jedan simbol dok ne naiđe do posebnog simbola koji označava kraj rečenice. Sve to povezuje mreža pažnje koja omogućuje dekodeeru da se fokusira na različite dijelove ulazne rečenice tijekom procesa dekodiranja dok ona „pamti“ ostale dijelove (Wu i sur., 2016, p. 3)¹⁴⁴.

Prema Koehnu (2017, p. 48)¹⁴⁵, u praksi ovaj model daje dosta dobre rezultate za rečenice koje se sastoje od 10 do 15 riječi, ali nije toliko uspješan s dužim rečenicama.

¹³⁹ Cho, K., van Merriënboer, B., Bahdanau, D. i Bengio, Y., 2014. On the Properties of Neural Machine Translation: Encoder–Decoder Approaches. *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*

¹⁴⁰ Ibid.

¹⁴¹ Koehn, P., 2017. *Statistical Machine Translation*. Johns Hopkins University, pp.1–113

¹⁴² Ibid.

¹⁴³ Ibid.

¹⁴⁴ Wu, Y., Schuster, M., Chen, Z., Le, Q., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, Ł., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., Stevens, K., Kurian, G., Patil, N., Wang, W., Young, C., Smith, J., Riesa, J., Rudnick, A., Vinyals, O., Corrado, G., Hughes, M. i Dean, J., 2016. Google’s Neural Machine Translation System: Bridging the Gap between Human and Machine Translation.

¹⁴⁵ Koehn, P., 2017. *Statistical Machine Translation*. Johns Hopkins University, pp.1–113

Zhang i Zong (2020, p. 17)¹⁴⁶ tvrde kako se zadnjih godina brzo razvila mreža za kodiranje i dekodiranje, od ponavljajućih neuronskih mreža, do konvulcijskih neuronskih mreža i onda do neuronske mreže zasnovane na samopozornosti (engl. *self-attention network*) zvanoj Transformer.

Prema Zhang i Zong, (2020, p. 17), u Transformeru, enkoder sadrži N identičnih slojeva i svaki sloj je sastavljen od dva podsloja: podsloj samopozornosti kojeg slijedi podsloj za prosljeđivanje unaprijed (engl. *feed-forward sub-layer*). Prvi podsloj izračunava izlazni prikaz tokena prateći sve tokene u istom sloju izračunavajući ocjenu korelacije između ovog tokena i svih susjednih te na kraju kombinira sve prikaze susjednih tokena i danog tokena (Zhang i Zong, 2020, p. 17) Završni proizvod N-tog sloja enkodera je semantički prikaz izvornog teksta (Zhang i Zong, 2020, p. 17)¹⁴⁷. Dekoder, prema Zhang i Zong (2020, p. 17)¹⁴⁸ također ima N identičnih slojeva od kojih se svaki sastoji od tri podsloja. Prema Zhang i Zong (2020, p. 17)¹⁴⁹, prvi podsloj je maskirani mehanizam samopozornosti koji sažima djelomičnu povijest predviđanja. Drugi je podsloj pažnje koderadekoderu koji određuje dinamički kontekst na izvornoj strani za trenutno predviđanje, a treći je podsloj za prosljeđivanje unaprijed (Zhang i Zong, 2020, p. 17)¹⁵⁰. Preostala veza i normalizacija sloja izvode se za svaki podsloj u koderu i dekoderu, navode (Zhang i Zong, 2020, p. 17)¹⁵¹.

Bez obzira na to što je ovo najmoderniji način strojnog prevođenja koji se uglavnom danas koristi u svim online sustavima za strojno prevođenje, ili samostalno ili u kombinaciji s drugim modelima, postoje i nedostaci. Poput statističkog strojnog prevođenja, i neuronski modeli „uvelike ovise o dostupnosti većih količina podataka i poznato je da oba pristupa daju loše rezultate kada resursa nema dovoljno” (Ramesh i Sankaranarayanan, p. 112)¹⁵². Kako bi program bio uspješan u izvršavanju svog prijevoda nužno je imati dovoljno tekstova na izvornom i ciljnom jeziku te rječnike i glosare za sve željene jezike. U protivnom sustav neće uspješno prepoznavati riječi i fraze te će davati netočne prijevode ili jednostavno neće uopće prevesti dani tekst.

¹⁴⁶ Zhang, J. i Zong, C., 2020. Neural Machine Translation: Challenges, Progress and Future.

¹⁴⁷ Ibid.

¹⁴⁸ Ibid.

¹⁴⁹ Ibid.

¹⁵⁰ Ibid.

¹⁵¹ Ibid.

¹⁵² Ramesh, S. i Sankaranarayanan, K., 2018. Neural Machine Translation for Low Resource Languages using Bilingual Lexicon Induced from Comparable Corpora. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop*

Nadalje, Cho i sur. (2014, p. 2)¹⁵³ su zaključili kako veličina vokabulara ima snažan utjecaj na prijevod što znači da je neuronskom modelu potrebno postaviti veliki rječnik kako bi mogao uparivati riječi iz izvornika u ciljni jezik. Ovo dodatno potvrđuje i tvrdnja Ramesh i Sankaranarayanan (p. 112)¹⁵⁴ koje su primijetile da ovakvi modeli prevođenja imaju lošu kvalitetu prijevoda, tj. prijevodi su im netočni, i lošu pokrivenost, tj. postoji puno riječi koje nisu u vokabularu (engl. *out-of-vocabulary rate*), zbog nedovoljno resursa za treniranje programa. No niti najveći rječnik ne sadrži sve riječi jednog jezika što uzrokuje probleme s prijevodom kada je sustav suočen s rijetkim ili zastarjelim riječima.

Dakle, neuralno prevođenje je, kako Wu i sur. (2016, p. 2)¹⁵⁵ tvrde, slabo u prevođenju rijetkih riječi. Iako je prijevod temeljno problem otvorenog tipa kada je riječ o vokabularu (imena, brojevi, datumi itd.), neuralno strojno prevođenje koristi rječnike s određenim i ograničenim riječima. Ovo uzrokuje mnogo problema s prijevodom, međutim, postoje neki prijedlozi za rješavanje ovog problema. Wu i sur. (2016, p. 7)¹⁵⁶ ističu dvije kategorije pristupa prijevodu riječi izvan vokabulara programa. Prvi pristup samo ostavlja rijetke riječi neprevedene jer su većina takvih riječi imena ili brojevi koji se ne prevode već samo kopiraju. Druga kategorija koristi podjedinice riječi kao što su simboli (Wu i sur., 2016, p. 7)¹⁵⁷. Wu i sur. (2016, p. 2)¹⁵⁸ također predlažu izradu kodirajućeg modela koji bi imitirao tradicionalni model poravnavanja (engl. *alignment model*) ili korištenje modela za pažnju. Također naglašavaju da niti jedan od ova dva pristupa nije pouzdan zato što kvaliteta poravnavanja ovisi o jezicima koji se prevode, a poravnavanja nastala kroz modele za pažnju su nestabilna ako je mreža duboka.

¹⁵³ Cho, K., van Merriënboer, B., Bahdanau, D. i Bengio, Y., 2014. On the Properties of Neural Machine Translation: Encoder–Decoder Approaches. *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*

¹⁵⁴ Ramesh, S. i Sankaranarayanan, K., 2018. Neural Machine Translation for Low Resource Languages using Bilingual Lexicon Induced from Comparable Corpora. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop*

¹⁵⁵ Ibid.

¹⁵⁶ Wu, Y., Schuster, M., Chen, Z., Le, Q., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, Ł., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., Stevens, K., Kurian, G., Patil, N., Wang, W., Young, C., Smith, J., Riesa, J., Rudnick, A., Vinyals, O., Corrado, G., Hughes, M. i Dean, J., 2016. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation.

¹⁵⁷ Ibid.

¹⁵⁸ Ibid.

Analiza koju su proveli Cho i sur. (2014, p. 2)¹⁵⁹ pokazuje da se kvaliteta prijevoda s neuronskim strojnim prevođenjem brzo smanjuje kako se duljina unesene rečenice povećava, to jest, što je rečenica dulja, to je prijevod lošije kvalitete. Dakle, kako bi se postigao najbolji rezultat prevođenja, ulazni tekst za neuronski model ne bi smio koristiti dugačke rečenice niti rijetke ili zastarjele riječi, no sustavu nije uvijek jasno što su sustavu rijetke riječi.

Neuralno strojno prevođenje je računalno skupo kako za izradu tako i za održavanje što može u potpunosti spriječiti njegovu izradu kada su u pitanju veliki modeli s puno informacija (Wu i sur., 2016, p. 1)¹⁶⁰. Proces izrade je spor i dugotrajan te zahtjeva puno računalnih resursa kako bi se sustav u potpunosti istrenirao na velikim skupovima podataka što usporava inovacije i eksperimentiranje sa sustavom (Wu i sur., 2016, p. 2)¹⁶¹.

Ljubas (2018, p. 89)¹⁶² naglašava kako program i dalje ima velik broj pogrešaka, „nerijetko i potpuno nejasne rečenice“ ali i da je „prijelaz na neuronski model korak u pravom smjeru: što se morfologije i sintakse tiče“.

¹⁵⁹ Cho, K., van Merriënboer, B., Bahdanau, D. i Bengio, Y., 2014. On the Properties of Neural Machine Translation: Encoder–Decoder Approaches. *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*

¹⁶⁰ Ibid.

¹⁶¹ Ibid.

¹⁶² Ljubas, S., 2018. Prijelaz sa statističkog na neuronski model: usporedba strojnih prijevoda sa švedskog na hrvatski jezik. *Hieronymus-časopis za istraživanja prevođenja i terminologije*, 4, pp.72-91

5. Istraživanje

Ovaj dio rada bavi se usporedbom rezultata prevođenja dobivenih korištenjem različitih besplatnih internetskih alata za strojno prevođenje, od kojih svi koriste model neuralnog strojnog prevođenja, samostalno ili u kombinaciji s drugim pristupima i resursima. Alati koji su korišteni su Systran, Google Translate, Bing Microsoft Translator i Yandex.Translate. Istraživanje je provedeno na 300 rečenica iz opće domene i medicine prevedenih s engleskog na hrvatski jezik. Dužina rečenica seže od 4 do 20 riječi za kratke rečenice i 21 do 62 riječi za duge rečenice. Sve rečenice preuzete su s interneta te koriste specifičnu terminologiju. Ukratko će se analizirati tipovi grešaka koje su se pojavljivale u prijevodima u različitim sustavima, kao i njihova učestalost.

Kategorizacija pogrešaka provedena je sukladno prethodnim istraživanjima.

U istraživanju Kučiš i Seljan (2014)¹⁶³ provedena je analiza na 100 rečenica za englesko-hrvatski jezični par na online alatima koji se temelje na statističkim modelima. Analizirane su sljedeće vrste pogrešaka: morfološke, neprevedene riječi, leksičke, višak prevedenog, sintaktičke pogreške i interpunkcija. pri čemu najveći broj pogrešaka su morfološke, zatim od navedene količine morfoloških pogrešaka slijede u 50% iznosa leksičke pogreške, u 33% slijede sintaktičke pogreške, te ostale.

U istraživanju Seljan i Dunder (2014)¹⁶⁴ provedena je evaluacija sustava englesko-hrvatski i hrvatsko-engleski jezik nakon procesa automatskog prepoznavanja govora i strojnog prevođenja. Evaluacija je provedena primjenom WER i PER metrike i uspoređena s ljudskom evaluacijom. Ljudska evaluacija je napravljena za kriterij korisnosti, nakon što je tekst računalno prepoznat kroz automatsko prepoznavanje govora za engleski i strojno preveden na hrvatski. Najviše pogrešaka ima za neprepoznate riječi, zatim datumi, brojevi i fraze. Prosječna ocjena strojnog prijevoda za englesko-hrvatski iznosi 3,24, a za hrvatsko-engleski 4,09.

Brkić i sur. (2013)¹⁶⁵ proveli su ljudsku evaluaciju za četiri online sustava za statističko strojno prevođenje za četiri različite domene (pravna, tehnička, sport, opća). Rezultati pokazuju najveći

¹⁶³ Kučiš, V., Seljan, S. (2014). The role of online translation tools in language education. *Babel* 60 (3).

¹⁶⁴ Seljan, S., Dunder, I. (2014). Combined Automatic Speech Recognition and Machine Translation in Business Correspondence Domain for English-Croatian. *Int. Journal of Computer, Information, Systems and Control Engineering*, WASET 8 (11), 1069.

¹⁶⁵ Brkić, M., Seljan, S., Vičić, T. (2013). [Automatic and human evaluation on english-croatian legislative test set](#). *Intelligent Text Processing and Computational Linguistics*, Springer, 311-317.

broj morfoloških pogrešaka, za 50% manje leksičkih pogrešaka, iza kojih slijede nepreveden riječi s obzirom na specifične domene. Seljan i sur. (2020)¹⁶⁶ proveli su istraživanje o evaluaciji strojnih prijevoda za online sustave (Google Translate i Yandex.Translate) trenirane na neuronskim mrežama i statističkom strojnom prevođenju za njemačko-hrvatski jezični par za poeziju, pri čemu su uzeti u obzir kriteriji fluentnosti i adekvatnosti, na koje utječu navedene kategorije pogrešaka. Prema izvorima (Castilho et al., 2017¹⁶⁷; Martindale i sur., 2019¹⁶⁸, Tu i sur., 2017¹⁶⁹), neuralno strojno prevođenje pokazuje bolje rezultate za fluentnost u odnosu na statističko strojno prevođenje, ali su ti rezultati nekonzistentni za kriterij točnosti i naknadno post-uređivanje.

Rezultati u Tablici 1. prikazuju kategorije i broj pogrešaka za četiri različita online sustava za strojno prevođenje. Najmanji broj pogrešaka identificiran je sustavu Google Translate te u sustavu Yandex.Translate. Najlošiji rezultat pokazuje sustav Systran, koji je i najkasnije razvijen za hrvatski jezik, dok se sustav Bing Microsoft Translator nalazi u sredini s obzirom na broj pogrešaka.

Tablica 1: Popis grešaka

Tip pogreške	Yandex.Translate	Bing Translator	Systran	Google Translate
Kriva riječ	34	23	37	21
Krivo značenje	36	43	56	37
Neprevedeno	4	9	16	4
Neslaganje	7	32	53	10
Nedostaje riječ	6	11	16	5
Nepotrebna riječ	2	8	9	5

¹⁶⁶ Seljan, S., Dunder, I., Pavlovski, M. (2020). Human Quality Evaluation of Machine-Translated Poetry. MIPRO

¹⁶⁷ Castilho, S., Moorkens, J., Gaspari, F., Calixto, I., Tinsley, J., Way, A. (2017). Is Neural Machine Translation the New State of the Art?, The Prague Bulletin of Mathematical Linguistics 108, 109–120.

¹⁶⁸ Martindale, M., Carpuat, M., Duh, K., McNamee, P. (2019). Identifying Fluently Inadequate Output in Neural and Statistical Machine Translation, Proc. of Machine Translation Summit XVII volume 1: Research Track, 233–243.

¹⁶⁹ Tu, Z., Liu, Y., Shang, L., Liu, X., Li, H. (2017). Neural Machine Translation with Reconstruction, Proc. of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI'17), 7.

Krivi raspored riječi	0	2	6	1
Neprirodno za hrvatski	1	4	10	3
Nejasna rečenica	2	6	10	3
Ukupno	92	138	213	89

Kriva riječ odnosi se na odabir krivog sinonima, ali točnog značenja riječi. Krivo značenje je pogreška koja se dogodi kada je višeznačna riječ točno prevedena, ali u danom kontekstu odabrano značenje nije točno, kao što je slučaj s engleskom riječi „medicine“ koja može značiti „lijek“ i „medicina“. Neslaganje označava pogrešku u odabiru roda, broja, padeža ili vremena određene riječi. Ako alat nije preveo neku riječ iz ulaznog teksta onda to spada pod tip pogreške „neprevedeno“, a ako je riječ izbacio iz izlaznog teksta onda se radi o tipu pogreške „nedostaje riječ“. U suprotnom, ako je alat za prevođenje u izlazni tekst ubacio riječ koja se nije nalazila u izvornom tekstu, bilo da je riječ o nepotrebnom dodavanju zamjenice ili dodavanjem informacija onda se radi o tipu pogreške „nepotrebna riječ“. Kategorija pogrešaka „neprirodno za hrvatski“ označava prijevod koji tehnički nije netočan, tj. sve riječi iz izvornog jezika su točno prevedene u ciljni jezik, međutim, one se kao takve, u tom značenju ili rasporedu ne koriste u ciljnom jeziku. Neprirodan prijevod ne rezultira nerazumijevanjem značenja rečenice. Ovakva pogreška se događa kod idioma ili direktnog prijevoda npr. fraza „say-so“ je direktno prevedena kao „reći-tako“ kada je trebala biti prevedena s riječi „odobrenje“. Obje riječi „say“ i „so“ su točno prevedene, ali u danom kontekstu one mijenjaju svoje značenje što utječe na pravilnost prijevoda rečenice. Nejasna rečenica označava one rečenice u kojima su pogreške bile takve da znatno utječu na razumijevanje te je teško odrediti što je bilo originalno značenje ulazne rečenice.

5.1. Systran

Systran je online sustav osnovan 1968. godine koristeći statističko strojno prevođenje te je prvi koristio hibridno strojno prevođenje (A pioneer and global leader in translation solutions, 2020)¹⁷⁰. Danas on koristi neuralno strojno prevođenje, u kombinaciji s ostalim pristupima (A pioneer and global leader in translation solutions, 2020)¹⁷¹. U ovome dijelu rada analizirat će se količina i vrsta pogrešaka u strojnom prijevodu. Tablica 2 prikazuje izoliranje greške koje je Systran radio i njihove postotke zastupljenosti.

Tablica 2: Pogreške za sustav Systran

Tip pogreške	Kriv a riječ	Kriv o znač.	Neprev	Neslag	Nedost . riječ	Nepotr . riječ	Kriv i ras. riječi	Neprir . Za hrv	Nejasn o
Broj pogrešaka	37	56	16	53	16	9	6	10	10
Zastupljeno st pogrešaka	17%	26%	8%	25%	8%	4%	3%	5%	5%

Testirao se prijevod s engleskog na hrvatski te je program u usporedbi s ostalim testiranim alatima dao najlošije rezultate (vidi Tablicu 1.) Sustav je sveukupno imao 213 greške i najviše nejasnih rečenica; 10 od 300 rečenica je bilo nejasno. Kao što se vidi iz Tablice 2., imao je najviše problema s krivim značenjem riječi i neslaganjem u rodu, broju, padežu ili vremenu što je zajedno ispunilo čak 51% svih pogrešaka koje je alat napravio. Ovo je dakako imalo snažan utjecaj na razumijevanje rečenica i njihovu tečnost ili fluentnost. Poput sustava Yandex.Translate, imao je nekih problema s osobnim imenima te tako nije mogao prepoznati njihov rod. Nadalje, Systran je imao daleko najviše neprevedenih riječi što pokazuje da ima najmanji vokabular. To ne samo da mu je predstavljao problem kod prevođenja stručnih termina već i kod prevođenja rjeđih općenitih riječi

¹⁷⁰ Systransoft.com. 2020. *A Pioneer And Global Leader In Translation Solutions*. (online) Dostupno na: <<https://www.systransoft.com/systran/>>

¹⁷¹ Ibid.

kao što je „coax“. Zbog toga je imao i problema s prevođenjem dužih rečenica te je tada najviše ubacivao nepotrebne riječi ili izbacivao glagole.

Također je imao i problema s idiomima te je jedini krivo preveo „fall ill“ s doslovnim prijevodom „pasti bolestan“, te je birao krive riječi čak 37 puta tj. davao je točan prijevod, ali u određenom kontekstu neki drugi sinonim bi bio bolji izbor.

5.2. Google Translate

Online sustav Google Translate je najpopularniji besplatni alat za prevođenje na internetu, a dnevno prevede oko 100 milijardi riječi (Google Translate, 2020)¹⁷². Ovaj alat je također u početku koristio statistički model strojnog prevođenja, ali se 2016. godine prebacio na neuralni model (Google Translate, 2020)¹⁷³. Tablica 3 prikazuje broj pogrešaka koje je Google Translate napravio u prijevodu i postotak zastupljenosti tipa pogreške.

Tablica 3: Sustav Google Translate

Tip pogreške	Kriv a riječ	Kriv o znač.	Neprev .	Neslag .	Nedost . riječ	Nepotr . riječ	Kriv i ras. riječi	Neprir . Za hrv	Nejasn o
Broj pogrešaka	21	37	4	10	5	5	1	3	3
Zastupljenost pogrešaka	24%	42%	4%	11%	6%	6%	1%	3%	3%

Alat Google Translate je dao najbolje rezultate sa samo 89 grešaka. Rečenice su rijetko bile nejasne i neprirodne (samo 3 puta), a krivi raspored riječi se dogodio samo jednom. Najvećih problema je imao s krivim značenjem (37) i odabirom krivih riječi (21), što mi je od ukupnog broja grešaka

¹⁷²En.wikipedia.org. 2020. Google Translate. (online) Dostupno na: <https://en.wikipedia.org/wiki/Google_Translate> (Pristupljeno: 28 kolovoz 2020).

¹⁷³ Ibid.

činilo 42% i 24% grešaka. Međutim, u usporedbi s ostalim alatima te brojke nisu prevelike te je, uz alat Yandex.Translate, imao najmanji broj neprevedenih riječi.

Online sustav Google Translate je najvećih problema imao s idiomima i riječima i frazama koje spadaju u općeniti vokabular, a ne onaj specifičan za domenu. Tako je na primjer slučaj sa „get it down“ kada je značio „progutati“, a Google Translate je dao prijevod „spustiti“ što je u direktnom značenju točno, ali u zadanom kontekstu riječ je o idiomu te ovaj prijevod nije adekvatan. Još jedan primjer krivog značenja jest prijevod rečenice „The nurse will dress this cut for you“ gdje je riječ „dress“ koja ima više značenja potrebno prevesti kao „zaviti“, a alat Google Translate ju je preveo kao „odjenuti“. Kriva riječ je odabrana u prijevodu riječi „exact“ koja je u danom kontekstu trebala biti prevedena kao „precizna“, a alat ju je preveo s riječi „točna“. Sustav je jedini točno preveo riječi „anti-cold vaccine“ sa „cjepivo protiv prehlade“ dok su svi ostali davali prijevode poput „anti.konvulzivno cjepivo“ ili „antihladno cjepivo“.

Google Translate nije imao nikakvih značajnijih problema s prijevodom dužih rečenica. Štoviše većina dužih rečenica je prevedena potpuno točno.

5.3. Yandex.Translate

Online sustav Yandex.Translate se originalno koristio statističkim strojnim prevođenjem, no 2017. godine prebacio se na hibridni model te sada koristi kombinaciju statističkog i neuralnog strojnog prevođenja (Translate API — About machine translation — Yandex Technologies, 2020)¹⁷⁴. Sustav Yandex.Translate je davao druge najbolje rezultate od svih ispitanih alata za prevođenje. Imao je samo 3 pogreške više od online sustava Google Translate koji je na prvom mjestu te je mnogo pogrešaka slično. Tablica 4 daje prikaz broja grešaka i postotak zastupljenosti tipa pogreške.

Tablica 4: Sustav Yandex.Translate

Tip pogreške	Kriv a riječ	Kriv o znač.	Neprev .	Neslag .	Nedost . riječ	Nepotr . riječ	Kriv i ras. riječi	Neprir . Za hrv	Nejasn o

¹⁷⁴ Tech.yandex.com. 2020. *Translate API — About Machine Translation — Yandex Technologies*. (online) Dostupno na: <<https://tech.yandex.com/translate/doc/dg/concepts/how-works-machine-translation-docpage/>>

Broj pogrešaka	34	36	4	7	6	2	0	1	2
Zastupljenost pogrešaka	37%	39%	4%	8%	7%	2%	0%	1%	2%

Najviše pogrešaka je bilo zbog krive riječi ili krivog značenja, sveukupno 76% grešaka koje je alat počinio su krive riječi ili značenje. Valja naglasiti kako je Yandex.Translate jedini sustav koji nije imao niti jednu grešku kada je u pitanju bio raspored riječi u rečenici. Rijetko kada je ubacivao (samo dva puta) riječi, izbacivao ih (6 puta) ili ih ostavio neprevedenima (4). Također je imao najmanje nejasnih i neprirodnih rečenica.

Sustav Yandex.Translate je imao problema s raspoznavanjem osobnih imenima. Tako na primjer, Destiny nije uopće prepoznao kao ime, Mabel nije znao da je žensko ime, a Alexis je mislio da je množina. Međutim, on je povremeno bio jedini alat koji je uspješno preveo neke specifične termine kao što je bio slučaj s riječi „noncancerous“. Dok su je drugi alati ostavili neprevedenom, ili u slučaju sustava Google Translate preveli kao „beznačajno“, Yandex.Translate je dao točan prijevod s riječi „ne-kancerogen“. Poput alata Google Translate, Yandex.Translate je također odabrao krivu riječ za prijevod engleske riječi „exact“ u rečenici „Medicine isn't an exact science“. Također je umjesto riječi „precizna“ odabrao riječ „točna“. Također je odabrao krivi sinonim za prijevod riječi „anxious“. U danom kontekstu trebala se upotrijebiti riječ „zabrinut“, a alat Yandex.Translate je upotrijebio riječ „uzbuđen“. Nadalje, kod prijevoda rečenice „Her hip has been troublesome for quite a while, and she'll probably need surgery on it“, Yandex.Translator je krivo preveo značenje riječi „troublesome“ sa „uznemiren“, kada je trebao iskoristiti riječ „problematičan“.

5.4. Bing Microsoft Translator

Online alat Bing Microsoft Translator je prevoditelj kompanije Microsoft koji se služi neuralnim strojnim prevodenjem u kombinaciji sa statističkim modelom strojnog prevodenja (Microsoft

Translator, 2020)¹⁷⁵. Prijevodi dobiveni ovim alatom su osrednji jer je program napravio 138 grešaka. Tablica 5 je prikaz broja pogrešaka i njihove zastupljenosti.

Tablica 5: Sustav Bing Translator

Tip pogreške	Kriv a riječ	Kriv o znač.	Neprev .	Neslag .	Nedost . riječ	Nepotr . riječ	Kriv i ras. riječi	Neprir . Za hrv	Nejasn o
Broj pogrešaka	23	43	9	32	11	8	2	4	6
Zastupljenost pogrešaka	17%	31%	7%	23%	8%	6%	1%	3%	4%

Ovaj alat je imao najviše problema sa semantičkim i sintaktičkim pogreškama, tj. značenjem (43 greške), neprevođenjem riječi i neslaganjem riječi u rodu, broju, padežu ili vremenu što ukazuje na njegovu slabost kod prevođenja morfološki bogatih jezika kao što je hrvatski. Nadalje, alat se nije niti leksički iskazao jer je koristio mnoštvo krivih riječi te je imao najviše nepotrebnih i izbačenih riječi, nakon Systrana. Većina riječi na kojima je alat griješio bila su iz općenitog rječnika, a ne dio medicinske terminologije, koju je relativno dobro savladao. Nadalje, nije bilo značajnijeg pogoršana kvaliteta prijevoda kod dužih rečenica.

Dobar prikaz problema s neslaganjem je prijevod rečenice „The nurse cleansed the wound before stitching it“ s rečenicom „Medicinska sestra je očistio ranu prije nego što ga je šivanje“. U ovoj rečenici nalazi se nekoliko primjera neslaganja. Kao prvo, glagol „očistiti“ se ne slaže sa subjektom u rodu te se zamjenica „ga“ ne slaže u rodu sa imenicom „rana“ na koju se odnosi. Nadalje, „stitching it“ je krivo prevedeno jer je odabran glagol „šivati“, koji se također nije slagao s ostatkom rečenice u vremenu, a točan prijevod je glagol „zašiti“.

Najbolji primjer semantičke pogreške je u prijevodu rečenice „You may begin experiencing backaches, headaches, muscle spasms of the neck, and muscle tension“ koju je alat Bing Microsoft Translator preveo kao „Vi svibanj početi doživljava backaches, glavobolje, grčevi mišića vrata, i

¹⁷⁵ En.wikipedia.org. 2020. Microsoft Translator. (online) Dostupno na: <https://en.wikipedia.org/wiki/Microsoft_Translator#cite_note-research.microsoft.com-4>

napetost mišića“. Riječ „may“ preveo je kao „svibanj“, a jedna riječ nije niti preveo. Nadalje i u ovoj rečenici se dogodio problem sa neslaganjem između subjekta i glagola.

Alat je također imao problema s prevođenjem pasiva što je rezultiralo krivim značenjem rečenice te je tako, na primjer, umjesto da napiše da medicinsku njegu „primaju“ pacijenti, napisao da je oni „pružaju“. Većina grešaka koje je alat radio vezano uz biranje krivih riječi i značenja dogodilo se zbog nerazumijevanja teksta. Takve greške nisu bile drastične stoga nisu bitno utjecale na razumijevanje ni tečnost teksta.

6. Zaključak

U radu je prikazan povijesni razvoj strojnog prevođenja – od začetaka do današnjeg stupnja razvoja, tj. neuralnog strojnog prevođenja, kroz povijesne faze, osnovne principe izgradnje sustava i probleme. Prikazane su smjernice za automatsku i ljudsku evaluaciju. Objasnjeni su sustavi za strojno prevođenje temeljeno na pravilima i statističko strojno prevođenje, a posebno je obrađen model neuralnog strojnog prevođenja, za koji je provedena evaluacija četiri online sustava. Analiza je provedena na 300 rečenica za četiri online sustava koja koriste neuralno strojno prevođenje, kao samostalni ili hibridni model izgradnje. Ljudska evaluacija provedena je prema kriterijima točnosti prenesenih informacija i gramatike. Sve pogreške koje su se u prijevodima dogodile su naznačene i podjeljene u 9 kategorija te se zatim svaki online sustav zasebno analizirao kako bi se prikazale najzastupljenije vrste pogrešaka. Rezultati pokazuju da svi ispitani alati imaju najviše pogrešaka koje se tiču semantike, dok su pogreške u sintaksi bile rijetke. Riječi i fraze su bile krivo prevedene, tj. značenje je bilo krivo te su se često koristili krivi sinonimi, ali točnog značenja. Glavno značenje rečenice je zadržano stoga pogreške u prijevodu nisu bitno utjecale na razumijevanje.

Iako online sustavi za strojno prevođenje još uvijek rade pogreške koje se tiču značenja i fluentnosti, njihova upotreba sve više raste te ih ima sve više i za manje govorene jezike. Primjena neuralnog strojnog prevođenja pokazuje pozitivne pomake u određenim kategorijama, iako ono ovisi o količini i kvaliteti treniranih podataka i svrsi primjene. Najčešće su i dalje semantičke pogreške, međutim, to ne utječe pretjerano na kvalitetu prijevoda jer su prijevodi i dalje pretežno razumljivi i sve informacije su sačuvane. Postojeći modeli pokazuju mjesta za daljnji napredak i razvoj te primjenu u društvu.

7. Literatura

1. Bar-Hillel, Y., 1953. Some Linguistic Problems Connected with Machine Translation. *Philosophy of Science*, 20(3).
2. Borra, A., Chan, E.A., Lim, C.I., Tan, R.B., & Tong, M.C. (2007). LFG-Based Machine Translation Engine for English and Filipino
3. Brkić, M., Seljan, S. i Vičić, T., 2009. Evaluation of the statistical machine translation service for Croatian-English.
4. Brkić, M., Seljan, S., Matetić, M. (2011). Machine translation evaluation for croatian-english and english-croatian language pairs. NLPCS Workshop: Human-Machine Interaction in Translation. Copenhagen: Copenhagen Business School.
5. Brkić, M., Seljan, S. i Vičić, T., 2012. BLEU Evaluation of Machine-Translated English-Croatian Legislation. LREC.
6. Brkić, M., Seljan, S., Vičić, T. (2013). Automatic and human evaluation on english-croatian legislative test set. *Intelligent Text Processing and Computational Linguistics*, Springer.
7. Castilho, S., Moorkens, J., Gaspari, F., Calixto, I., Tinsley, J., Way, A. (2017). Is Neural Machine Translation the New State of the Art?, *The Prague Bulletin of Mathematical Linguistics* 108.
8. Charoenpornasawat, P., Sornlertlamvanich, V. i Charoenporn, T., 2002. Improving translation quality of rule-based machine translation. *COLING-02 on Machine translation in Asia*
9. Cho, K., van Merriënboer, B., Bahdanau, D. i Bengio, Y., 2014. On the Properties of Neural Machine Translation: Encoder–Decoder Approaches. *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*,.
10. Dovedan, Z., Seljan, S. i Vučković, K., 2002. Strojno prevođenje kao pomoć u procesu komunikacije. *Informatologia*, 35(4).
11. Dunder, I., 2015. *Sustav za statističko strojno prevođenje i računalna adaptacija domene*. Doktorska disertacija. Sveučilište u Zagrebu. Filozofski fakultet.
12. En.wikipedia.org. 2020. *Google Translate*. (online) Dostupno na: <https://en.wikipedia.org/wiki/Google_Translate> (Pristupljeno: 28 kolovoz 2020).

13. En.wikipedia.org. 2020. *Microsoft Translator*. (online) Dostupno na: <https://en.wikipedia.org/wiki/Microsoft_Translator#cite_note-research.microsoft.com-4> (Pristupljeno: 28 kolovoz 2020).
14. En.wikipedia.org. 2020. *Rule-Based Machine Translation*. (online) Dostupno na: <https://en.wikipedia.org/wiki/Rule-based_machine_translation> (Pristupljeno: 28 kolovoz 2020).
15. Finka, B., i Laszlo, B. (1962). 'Strojno prevođenje i naši neposredni zadaci', *Jezik*, 10(4), str. 117-121. Preuzeto s: <https://hrcak.srce.hr/226114> (Pristupljeno: 28. kolovoz 2020.)
16. Hutchins, J., 2020. *The History Of Machine Translation In A Nutshell*. (online) Translationdirectory.com. Dostupno na: <<https://www.translationdirectory.com/article411.htm>> (Pristupljeno: 8 kolovoz 2020).
17. Hutchins, W., 2001. Machine Translation over fifty years. *Histoire Épistémologie Langage*, 23(1).
18. Koehn, P., 2017. *Statistical Machine Translation*. Johns Hopkins University.
19. Kučiš, V. (2010). Prevodilački alati u funkciji kvalitete prijevoda. *Informatologia*, 43(1), str. 19-33. Dostupno na: <https://hrcak.srce.hr/53039> (Pristupljeno: 28. kolovoz 2020.)
20. Kučiš, V., Seljan, S., Klasnić, K. (2009). Evaluation of electronic translation tools through quality parameters. *INFuture 2009: Digital Resources and Knowledge Sharing*.
21. Ljubas, S., 2017. Analiza pogrešaka u strojnim prijevodima sa švedskog na hrvatski. *Hieronymus-časopis za istraživanja prevođenja i terminologije*, 4.
22. Ljubas, S., 2018. Prijelaz sa statističkog na neuronski model: usporedba strojnih prijevoda sa švedskog na hrvatski jezik. *Hieronymus-časopis za istraživanja prevođenja i terminologije*, 4.
23. Martindale, M., Carpuat, M., Duh, K., McNamee, P. (2019). Identifying Fluently Inadequate Output in Neural and Statistical Machine Translation, *Proc. of Machine Translation Summit XVII volume 1: Research Track*
24. Papineni, K., Roukos, S., Ward, T. i Zhu, W., 2002. BLEU: a method for automatic evaluation of machine translation. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02*.
25. Ramesh, S. i Sankaranarayanan, K., 2018. Neural Machine Translation for Low Resource Languages using Bilingual Lexicon Induced from Comparable Corpora. *Proceedings of*

the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop.

26. Seljan, S. (2000). Sublanguage in Machine Translation. Mipro 2000: Computers in Intelligent Systems.
27. Seljan, S., Dunder, I. (2014). Combined Automatic Speech Recognition and Machine Translation in Business Correspondence Domain for English-Croatian. *Int. Journal of Computer, Information, Systems and Control Engineering*, WASET 8 (11), 1069.
28. Seljan, S., Dunder, I. (2015). Machine Translation and Automatic Evaluation of English/Russian-Croatian. International Conference "Corpus Linguistics"
29. Seljan, S., Dunder, I., Pavlovski, M. (2020). Human Quality Evaluation of Machine-Translated Poetry. MIPRO
30. Seljan, S., Pavuna, D. 2006. Translation Memory Database in the Translation Process. Information and Intelligent Systems IIS 2006.
31. Seljan, S., Tucaković, M. and Dunder, I., 2015. Human Evaluation of Online Machine Translation Services for English/Russian-Croatian. *New Contributions in Information Systems and Technologies*.
32. Sepesy Maučec, M. i Donaj, G., 2020. Machine Translation and the Evaluation of Its Quality. *Recent Trends in Computational Intelligence*.
33. Shofner, K., 2020. *The Pros And Cons Of Statistical Machine Translation*. (online) Unitedlanguagegroup.com. Dostupno na: <https://www.unitedlanguagegroup.com/blog/pros-and-cons-statistical-machine-translation> (Pristupljeno: 8 kolovoz 2020).
34. Šimić, J., i Uglarik, D. 2010. 'Mogućnosti i ograničenja strojnog prevođenja', *Praktični menadžment*, 1(1), str. 81-85. Dostupno na: <https://hrcak.srce.hr/67847> (Pristupljeno: 28. kolovoz 2020.)
35. Sreelekha, S., 2017. Statistical vs rule based machine translation; a case study on Indian language perspective.
36. Sumita, E. i Iida, H., 1991. Experiments and prospects of Example-Based Machine Translation. *Proceedings of the 29th annual meeting on Association for Computational Linguistics*

37. Systransoft.com. 2020. *A Pioneer And Global Leader In Translation Solutions*. (online)
Dostupno na: <<https://www.systransoft.com/systran/>> (Pristupljeno: 8 kolovoz 2020).
38. Systransoft.com. 2020. *What Is Machine Translation? Rule Based Vs. Statistical / SYSTRAN*. (online) Dostupno na: <<https://www.systransoft.com/systran/translation-technology/what-is-machine-translation/>> (Pristupljeno: 8 kolovoz 2020).
39. Tech.yandex.com. 2020. *Translate API — About Machine Translation — Yandex Technologies*. (online) Dostupno na: <<https://tech.yandex.com/translate/doc/dg/concepts/how-works-machine-translation-docpage/>> (Pristupljeno: 28 kolovoz 2020).
40. Tu, Z., Liu, Y., Shang, L., Liu, X., Li, H. (2017). Neural Machine Translation with Reconstruction, Proc. of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI'17),
41. Way, Andy. 1999. A Hybrid Translation Model using LFG-DOP
42. Wu, Y., Schuster, M., Chen, Z., Le, Q., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, Ł., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., Stevens, K., Kurian, G., Patil, N., Wang, W., Young, C., Smith, J., Riesa, J., Rudnick, A., Vinyals, O., Corrado, G., Hughes, M. i Dean, J., 2016. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation.
43. Zhang, J. i Zong, C., 2020. Neural Machine Translation: Challenges, Progress and Future.

8. **Popis slika**

Figure 1: Formula za RNN 27

9. Popis tablica

Tablica 1: Popis grešaka	33
Tablica 2: Pogreške za sustav Systran	35
Tablica 3: Sustav Google Translate	36
Tablica 4: Sustav Yandex.Translate	37
Tablica 5: Sustav Bing Translator.....	39

10. Sažetak

Strojno prevođenje u bilo kojem obliku danas pronalazi primjenu u mnogim granama ljudskog života te se zbog globalizacije i sve veće prisutnosti interneta sve više i brže razvija. Ovaj rad se bavio strojnim prevođenjem te je prošao kroz njegovu povijesti, evaluaciju i različite pristupe. U radu su analizirana tri osnovna pristupa strojnom prevođenju: prevođenje temeljeno na pravilima te statističko i neuralno strojno prevođenje, kojemu je posvećena velika pažnja, te su se naveli i njihovi nedostaci. Provedeno je istraživanje četiri online sustava za strojno prevođenje temeljeno na neuralnom prevođenju na 300 rečenica englesko-hrvatskog para te su analizirani tipovi pogrešaka. Svi navedeni pristupi su dobri u određenim segmentima strojnog prevođenja iako svaki od pristupa ima svoje nedostatke, a najčešće su bile pogreške vezane uz semantiku. Osnovnu ideju izvornog teksta sve vrste strojnog prevođenja uspijevaju vjerno prevesti što se može iskoristiti za pomoć u prijevodu ili laicima kojima nije potrebna visoka kvaliteta prijevoda, za potpuni prijevod teksta. Postoji još mjesta za napredak svih vrsta, ali zaključeno je da je za sada neuralno prevođenje najkvalitetnije te je i stručnjacima najzanimljivije. Najbolji rezultati su postignuti hibridnim prevođenjem, tj. korištenjem dva ili više sustava za strojno prevođenje u jednom alatu.

Ključne riječi: neuralno strojno prevođenje, ljudska evaluacija, pogreške

11. Summary

Nowadays machine translation in any form finds a use in many areas of human life and due to globalization and the expanding of the internet it is being developed at a faster rate. This paper dealt with machine translation, its history, evaluation and types. A description of three most prevalent types was given: rule-based machine translation, statistical MT and neural MT, which was given special attention, as well as their drawbacks. A study of four online machine translation systems based on neural translation on 300 sentences of an English-Croatian pair was conducted and the types of errors were analyzed. All listed types produce successful translations even though they all make mistakes they still produce quality translation. All types of machine translation manage to truthfully translate the gist of the original text which can be used as help for translating or layman can use it as full translations if they do not need a high quality translation. There is still room for improvement, but it was concluded that neural MT is at the moment the best and most interesting for experts. The best results were achieved through hybrid translations, i.e. using two or more types of MT in one translation tool.

Key words: neural machine translation, human evaluation, errors.