

Govoriš nevronske?

Kako ljudje razumemo jezik sodobnih strojnih prevajalnikov

David Bordon

Oddelek za prevajalstvo, Filozofska fakulteta, Univerza v Ljubljani
Aškerčeva 2, 1000 Ljubljana
david.bordon@ff.uni-lj.si

Povzetek

Namen prispevka je predstaviti raziskavo preverjanja razumljivosti nerevidiranih strojno prevedenih spletnih besedil. Primarni udeleženci v raziskavi so bili splošni bralci in ne izurjeni prevajalci ali popravljavci strojnih prevodov. Gre za prvo tovrstno raziskavo, ki je bila izvedena za slovenski jezik. Cilj raziskave je bil preveriti, v kolikšni meri so nerevidirani strojni prevodi razumljivi splošnemu bralstvu, pri čemer sem se posvetil tudi vplivu besedilnega in slikovnega konteksta. Preverjal sem prevode prevajalnikov Google Translate in eTranslation. Raziskava je bila izvedena z anketo, v kateri so udeleženci odgovarjali na vprašanja, ki so preverjala razumevanje spremljajočega besedilnega segmenta, v katerem je bila napaka. Rezultati nudijo vpogled v trenutno stopnjo razvoja strojnih prevajalnikov, ne z vidika storilnosti pri njihovem popravljanju, ampak z vidika, koliko jih razume ciljno bralstvo.

Do you Speak Neuralese?

The aim of this paper is to present a study on the comprehensibility of unedited machine-translated web texts. The primary participants in the study were general readers, not trained translators or post-editors, and it is the first study of its kind to be conducted for the Slovene language. The aim of the study was to examine the extent to which unedited machine translations are comprehensible to general readers, while giving focus to the influence of textual and pictorial context. The translations were obtained from Google Translate and eTranslation. The survey was conducted by means of a questionnaire, in which participants answered questions that tested their understanding of a text segment that included an error. The results provide an insight into the current state of development of machine translation engines, not from the point of view of PEMT, but from the point of view of how well machine translations are understood by the target readership.

1. Uvod

Članek obravnava raziskavo razumljivosti strojno prevedenih spletnih besedil pri bralcih, ki ne vedo, da prebirajo strojne prevode. Uporabil sem naključno izbrana angleška spletna besedila, slovenske prevode pa sem pridobil z nevronskega strojnima prevajalnikoma Google Translate in eTranslation. Prevodi niso bili revidirani, saj sem želel replicirati okoliščine, v katerih bi jih dejansko lahko našli – na spletu, kjer so zaradi (za nekatere) dovolj visoke kakovosti in cenovne nepremagljivosti (namreč so brezplačni) vedno bolj pogosta, kar velja tudi za prevajalske vtičnike, ki so vgrajeni v sodobne brskalnike in aplikacije.

Vprašanje razumljivosti v taki obliki je postalo aktualno samo v zadnjem času, saj so starejši, statistični modeli prevajalnikov slovnično nekonsistentni in jezikovno okorni, sodobni nevronske prevajalniki pa proizvajajo tekoča besedila, ki so težje ločljiva od človeških, hkrati pa je že profesionalnim pregledovalcem prevodov težje ugotoviti, kje so storili napako (Donaj in Sepesy Maučec, 2018).

Te napake nastanejo predvsem zaradi težav pri razdvajanju večpomenskih besed in prevajanju besed, ki jih ni v podatkovni zbirki, s katero smo prevajalnik urili (Thi-Vinh et al. 2019, 207; Koehn in Knowles 2017, 28, 31–33; Sennrich et al. 2016, 3). Kljub morebitnim posamičnim napačno prevedenim besedam pa lahko ljudje pomen razberemo iz sobesedila. Pri preverjanju razumljivosti sem v vseh primerih vključil še kontekst, saj se v stvarnosti bralci nikoli ne srečujejo z izoliranimi besedami, ampak z zaključenimi besedili, ker pa se osredotočam na spletno okolje, sem besedilnemu kontekstu dodal še slikovnega, ki je inherentna lastnost sodobnega spleta.

2. Namen članka

Namen članka je predstaviti grobo oceno razumljivosti prevodov NMT-sistemov (ang. *Neural machine translation*) v času, ko so taka besedila na spletu vedno bolj pogosta, pri čemer me zanima predvsem, kako slikovno gradivo v besedilnem kontekstu vpliva na rezultate. Tovrstna raziskava za slovenščino še ni bila izvedena.

2.1. Sorodne raziskave

Raziskav na področju razumevanja nerevidiranih strojnih prevodov pri naključnem splošnem bralstvu je razmeroma malo, saj je z vidika omejenosti na stroko in gospodarstvu bolj zanimive analize storilnosti pri popravljanju prevodov veliko več raziskav osredotočenih zgolj na prevodno prebivalstvo.

Razširjenost prakse popravljanja strojnih prevodov lahko opazimo že v zapisih o najboljših praksah pri popravljanju prevodov, ki so zapisani v blogih večjih ponudnikov jezikovnih rešitev, kot so denimo MemoQ (Lelner, 2022), Crowdin (Voronjak, 2022) in Memsources (Zdarek, 2020).

Na Univerzi v Gentu je bila v sklopu projekta ArisToCAT izvedena raziskava o razumevanju izmišljenih besed in samostalniških besednih zvez (Macken et al. 2019). Primeri, ki so bili iz angleščine v nizozemščino prevedeni s strojnima prevajalnikoma Google Translate in DeepL, so bili predstavljeni samostojno ali v kontekstu povedi, pri tem pa udeleženci niso imeli dostopa do izvirnega besedila. V povprečju je bilo 60 % odgovorov napačnih; rezultati so bili boljši, če je bil primer predstavljen v kontekstu povedi.

V sklopu istega projekta je bila izvedena še analiza bralnega razumevanja človeškega prevoda na eni in nepopravljenega strojnega prevoda na drugi strani.

Človeški prevodi so bili ocenjeni bolje z vidika jasnosti podajanja informacij, z vidika končnega razumevanja pa je bila razlika manjša (Macken in Ghyselen, 2018).

Castilho in Guerberof Arenas (2018) sta izvedli primerjalno analizo bralnega razumevanja za statistični in nevronske model strojnega prevajalnika v primerjavi s človeškim izvornikom. Glede na omejen vzorec (6 udeležencev) in nedoslednost rezultatov je končna ugotovitev, da sistemi-NMT izkazujejo najboljše rezultate, občasno še boljše kot angleški izvornik, nedokončna.

Martindale in Carpuat (2018) sta v raziskavi obravnavali odziv bralcev na tekočnost in natančnost nevronske strojne prevode, ob tem pa sta preverjali stopnjo zaupanja informacijam v besedilu. Ugotovili sta, da bralce zelo zmotijo prevodi, ki niso tekoči, medtem ko se ob samo natančnost informacij obregne veliko manjši delež bralstva.

Izsledke potrjuje tudi Popović (2020). V njenem eksperimentu so bralci v 30 % primerov zaradi zavajajoče tekočnosti sprejeli popolnoma napačno informacijo, še 25 % dodatnih primerov pa je bilo skoraj popolnoma (narobe) razumljivih.

Na tem mestu velja omeniti, da so se nedavno začele pojavljati bolj eksperimentalne metode prevajanja, katerih značilnost je upoštevanje multimedijskega konteksta, denimo zvočnega ali slikovnega. Lala in Specia (2018) sta razvila model multimedijskega leksikalnega prevajanja, katerega namen je prevajanje dvoumnih večpomenskih besed s pomočjo slikovnega konteksta. Sulubacak et al. (2020) so predstavili sorodne raziskave, uporabne podatkovne zbirke in metode raziskovanja na področju multimedijskega strojnega prevajanja, ki so vezane na prevajanje z zvokom, sliko in videom. Med novjšimi raziskavami Liu (2021) ponuja nevronske model vizualno-tekstovnega enkodiranja in dekodiranja.

Pričakujemo lahko, da se bo to področje v bodoče še hitreje razvijalo, predvsem zaradi tehnološkega napredka v drugih panogah (prepoznavanje slik, sinteza govora, avtomatsko podnaslavljanje ipd.).

3. Metoda

Raziskava je bila zasnovana okrog vprašalnika, ki je vseboval primere štirih vrst napak v slovenskih strojnih prevodih splošnih angleških spletnih besedil. Preverjal sem prevajalnika Google Translate in eTranslation, pri čemer je bil vsak zastopan z 12 vprašanji. Poseben pomen sem posvetil slikovnemu gradivu v sobesedilu.

3.1. Izbor besedil

Besedila sem zbiral glede na verjetnost, da bi se bralci z njimi lahko dejansko srečali na spletu. Analiza prevajalskega trga je pokazala, da večje prevajalske agencije popolnoma obvladujejo sektorje, ki nudijo največ dobička in hkrati zahtevajo človeško revizijo (tehnika, zdravstvo, pravo, finance ipd.) (Evropska komisija, 2020). V manj dobičkonosnih sektorjih, kjer človeška revizija ni tako bitna, obstaja večja verjetnost objave nerevidiranih strojnih prevodov.

Pregleda tržnega deleža spletnih iskalnikov, ki jih uporabljamo v Sloveniji je pokazal, da 96 % vseh uporabnikov spleta uporablja iskalnik Google.¹ Na osnovi

najbolj iskanih pojmov v brskalniku² sem izločil spletišča, ki nimajo prevodnega potenciala (družbena omrežja, spletni portali v slovenščini, slovenski mediji). S tem sem prišel do končnega izbora besedilnih področij: spletno nakupovanje, turizem, elektronika, multimedija in videoigre, luksuzne storitve, moda, osebno zdravje (telesna vadba in prehrana).

3.2. Prevodi besedil

Pri preizkušanju strojnih prevajalnikov se je izkazalo, da Googlov prevajalnik nudi drugačne prevodne rešitve glede na to, kako besedilo naložimo v obdelavo. Če besedilo prevajamo v pogovornem oknu vmesnika ali v brskalniku prevedemo spletno stran kot celoto, so rezultati boljši kot tisti, ki jih dobimo s funkcijo prevajanja dokumenta. Od štirih različnih specializiranih domen, ki jih nudi eTranslation, je najboljše rezultate nudil prevajalnik za splošna besedila (General Text). Uporabil sem najboljše možne prevode – omenjeno domeno v eTranslation, v Googlu pa sem prevajal v pogovornem oknu.

Prevod iz vnosnega polja oz. samodejni prevod strani	Prevod, pridobljen s funkcijo »prevedi dokument«	Izvornik
Naj bo topla - mikrovalovna pečica ohranja hrano, kot so zelenjava, juhe, jedi, gravija, omake in sladice, topla in okusna v pečici, dokler niso pripravljene za postrežbo.	Naj bo topla funkcija - Mikrovalovna ohranja živila, kot so zelenjava, juhe, nerazporejenega d'oeuvres, gravies, omake in sladice toplo in okusno v pečici, dokler oni pripravljene, da služijo.	Keep Warm Feature Maintains Food Temperature Keeps foods like vegetables, soups, hors d'oeuvres, gravies, sauces and desserts warm and delicious in the oven until they're ready to serve.

Tabela 1: Razlike v prevodih glede na način obdelave; Google Translate.

Prevod modela »General Text« prevajalnika eTranslation
Ohraniti toplo funkcijo - Microwave ohranja hrano, kot so zelenjava, juhe, predjed d'oeuvres, omake, omake in sladice tople in okusne v pečici, dokler niso pripravljene za postrežbo

Tabela 2: Prevod enakega segmenta; eTranslation.

¹ <https://gs.statcounter.com/search-engine-market-share/all/slovenia>

² <https://ahrefs.com/keyword-generator>

3.3. Kategorizacija napak

Previde sem analiziral in določil štiri kategorije najpogostejših napak, ki niso vezane na jezikovni sistem oz. predpis.

- **Neprevedena beseda;** v prevodu se pojavlja beseda v enaki obliki kot v izvorniku. Dopustil sem možnost spremembe začetnih ali končnih morfemov, če je prevajalnik besedo samo preoblikoval³.
- **Napaka pri razdvoumljanju večpomenske besede;** denotativni pomen večpomenske besede ali besedne zveze ne ustreza pomenu v izvorniku.
- **Hujša pomenska napaka;** napaka, ki otežuje razumevanje celotnega besedila.
- **Izmišljena beseda;** prevajalnik si izmislil novo besedo, ki je na prvi pogled videti slovenska, a ne spada v slovensko besedišče – t. i. »nevronščina«.

3.4. Kontekst

Izbranim besedilom sem glede na inherentne lastnosti spletne pojavitve dodal kontekst. Kontekst je lahko bil več vrst:

- izključno besedilni,
- besedilni in slikovni; slika ne vpliva na razumevanje,
- besedilni in slikovni; slika vpliva na razumevanje,
- izbor ene izmed več predlaganih slik glede na to, kaj piše v besedilu.

Slikovni kontekst sem vključil pri besedilih, ob katerih so se na spletu pojavljale fotografije, ki so pri nekaterih primerih bile zgolj vizualni dodatek, pri drugih pa je bilo pravilno razumevanje besedila vezano na prepoznavanje pravega vizualnega elementa.

V svoji raziskavi besed nisem nikoli predstavil v izolaciji, kot so to denimo storili v raziskavi Macken in drugi (2019), saj to niso realne okoliščine – napake v objavljenih strojnih prevodih bodo vedno del nekega besedila. Besedil nisem popravljala, anketirancem so bila predstavljena vključujoč vse slovnične in pomenske napake, take, kot bi jih našli v divjini.

3.5. Oblikovanje vprašalnika, format odgovorov na vprašanja in udeleženci

Anketo sem ustvaril na platformi Google Forms, ki nudi podporo za prikaz slik in dober vmesnik za pregled in izvoz rezultatov. Pomembno je poudariti, da anketirancem nisem razkril, da bodo brali strojno prevedena besedila. Omenil sem, da bodo »prebrali več kratkih besedil, ki so napisana v nekoliko okorni slovenščini«.

Vrste odgovorov so bile omejene s funkcionalnostjo platforme Google Forms in niso sledile nobeni logični metodi; določil sem jih subjektivno glede na vsebino primera in vrsto napake. Gre za najbolj nezanesljivo spremenljivko v metodi, saj bi s formulacijo vprašanja lahko sugeriral pravi odgovor, zanimalo pa me je predvsem to, če prihaja do večjega odstopanja glede na tip odgovora, denimo, če bi bili odgovori odprtega tipa, kjer anketiranci vnesejo svoj odgovor v prazno vnosno polje, bistveno slabši kot tisti, kjer izbirajo med štirimi

predlaganimi odgovori. S tem bi lahko preveril konsistenco pravilnosti oz. odstopanja glede na vrsto odgovora.

Vprašalnik sem delil na družbenih omrežjih Facebook in Instagram in znance pozval, naj ga posredujejo naprej svojcem in svojim znancem, če je le mogoče starejšim. Demografskih podatkov nisem zbiral, kar je mogoče ena izmed pomanjkljivosti raziskave. Glede na razmeroma majhen vzorec sodelujočih in morebiten efekt odmevne komore bi bilo vsekakor raziskavo potrebno nadgraditi in ponoviti na bolj naključnem in predvsem večjem vzorcu, toda glede na čas zbiranja odzivov, ki je sovpadal s prvo omejitvijo gibanja vezano na epidemijo Covid-19, nisem imel druge izbire.

Na vprašalnik sem prejel 120 odgovorov.

Pearl P-3000D Demon enojna pedala

- Pedal za bas bobne
- Eno bas bas pedala
- Neposredna vožnja od pedala do udarca
- Ninja kroglični ležaj
- Duo Deck Longboard
- Zeleni latenci U-sklepov
- Control Core bat
- Preklopite na kontrabas s pedalom
- Primer vključen

Ali je torba vključena poleg pedala? *

Da

Ne

Zakaj? *

Vaš odgovor

Slika 1: Primer vprašanja. Izbor z razlago.

4. Rezultati

Rezultate predstavljam po naslednjih parametrih:

- splošno razumevanje,
- razumevanje glede na prevajalnik,
- razumevanje glede na tip napake,
- razumevanje glede na tip konteksta,
- razumevanje glede na tip odgovora.

4.1. Splošno razumevanje

Vprašalnik je obsegal 24 vprašanj, s 120 odzivi je bilo vseh možnih odgovorov 2880. Vseh pravih odgovorov je bilo 1697 oz. 58,96 %. Daljša razčlemba je na voljo v celotni raziskavi (Bordon, 2021).

4.2. Razumevanje glede na prevajalnik

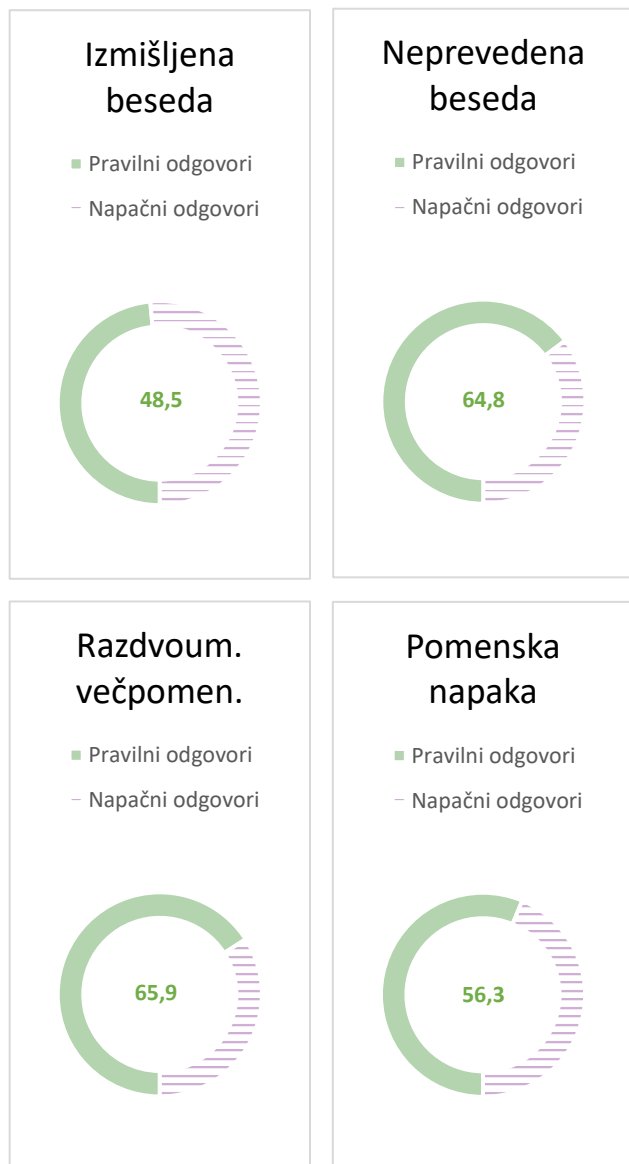
Odgovori na vprašanja, vezana na prevajalnik Google Translate so bili pravilni v 51,3 % primerov oz. 739 od 1440 odgovorov. Prevajalnik eTranslation je pokazal boljše rezultate, delež pravih odgovorov je znašal 66,6 %.

³ Denimo, prevod za rob zaslona (ang. *bezel*, je prevajalnik prevedel kot »bezela«).

4.3. Razumevanje glede na tip napake

V vprašalniku so bili vključeni štiri tipi različnih napak. V alinejah nizam tip napake in odstotek pravih odgovorov:

- izmišljena beseda: 48,5 %,
- neprevedena beseda: 64,8 %,
- napačno razdvoumljene večpomenske besede: 65,9 %,
- hujša pomenska napaka: 56,3 %.



Slika 2: Diagrami 1–4. Rezultati glede na tip napake v %.

4.4. Razumevanje glede na kontekst

V naslednjem segmentu predstavljam delež pravih odgovorov vezanih na kontekst.

- Izključno besedilni: 60,4 %,
- besedilni in slikovni; slika ne vpliva na razumevanje: 44 %,
- besedilni in slikovni; slika vpliva na razumevanje: 69,8 %,
- izbor ene izmed več predlaganih slik glede na to, kaj piše v besedilu: 64,2 %.

4.5. Razumevanje glede na tip odgovora

V tem segmentu predstavljam rezultate glede na način izbora odgovora. Primarna funkcija te analize je preveriti konsistenco oz. morebitna odstopanja npr.; če so odgovori odprtega tipa, kjer anketiranci v prazno vnosno polje vnesejo poljuben odgovor, bistveno slabši kot tisti, kjer imajo na voljo denimo štiri predlagane odgovore, izberejo pa enega.

- Odgovor odprtega tipa (vnosno polje): 36,3 %,
- odgovor zaprtega tipa (A, B, C ali D): 60,8 %,
- izbor z razlago (A ali B, zakaj?): 68,3 %.

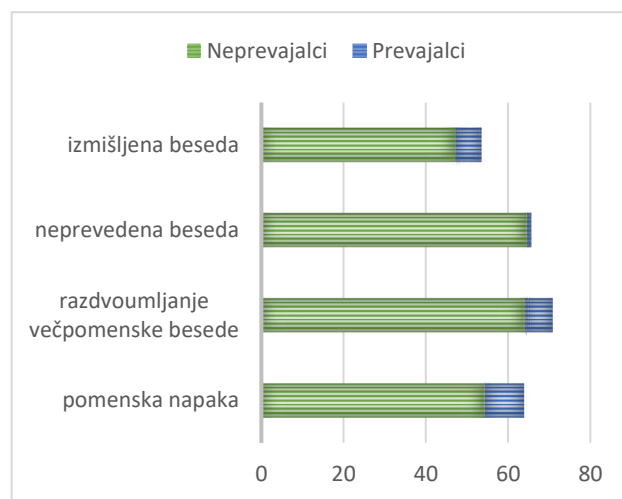
Slabši rezultat pri odgovorih zaprtega tipa je potrebno jemati z rezervo, saj so bili primeri s tako vrsto odgovora zgolj štiri. Samo določanje pravilnosti odgovora je pri takih primerih težje, osebno pa sem bil strog ocenjevalec, saj sem vse odgovore, ki niso bili popolnoma pravilni, označil za napačne.

4.6. Skupina prevajalcev

Edini demografski podatek, ki sem ga zbiral, je, ali se oseba, ki odgovarja na vprašalnik, ukvarja s prevajanjem. Pritrdilno je odgovorilo 24 udeležencev od 120. Pri teh osebah sem analiziral odgovore glede na vrsto napake in jih primerjal z neprevajalci. Nasploh so bili njihovi rezultati za 6 % boljši (63,7 %), po kategorijah pa:

- izmišljena beseda 53,5 % (+ 6,3 %),
- neprevedena beseda 65,6 % (+ 1 %),
- razdvoumljanje večpomenske besede 70,8 % (+ 6,7 %),
- pomenska napaka 63,9 % (+ 9,6 %).

Ostalih demografskih podatkov nisem zbiral, kar je ena od slabosti raziskave. V primeru da bi podatki sovpadali z mojo predpostavko, da niso relevantni, jih ne bi vključil, sedaj pa preprosto nimam podatkov, na katerih bi lahko utemeljil svojo odločitev.



Graf 1: Rezultati skupine prevajalcev proti ostalim.

5. Razprava

Pri pregledu rezultatov sem ugotovil, da povprečna stopnja razumevanja znaša 59 %. Od vseh 2880 odgovorov je bilo 1697 pravih.

Na tej točki je potrebno izpostaviti primer št. 6, ki je bil nasploh najslabše razumljen in je znižal povprečje rezultatov v vseh kategorijah, v katerih se je nahajal. Daljša razlaga z razčlenbo je na voljo v celotni raziskavi (Bordon, 2021).

Izvirnik	Prevod
En zmagovalec bo prejel grafično kartico GeForce RTX 2080 Ti Cyberpunk 2077 Edition.	One winner will receive the GeForce RTX 2080 Ti Cyberpunk 2077 Edition graphics card.
Vstop v predavanje je enostaven: 1. Prijavite se na forume ali ustvarite forumski račun . 2. Komentirajte to temo (BREZ CITIRANJA TE POSTAJE) in nam povejte, kaj želite narediti najbolj v Cyberpunku 2077. 3. Za potrditev vpisa vpišite svoje uporabniško ime v naš pripomoček za oddajo.	Entering the giveaway is easy: Sign in to the forums or create a forum account. Comment on this thread (WITHOUT QUOTING THIS POST) and tell us what you want to do most in Cyberpunk 2077. Sign your username in our giveaway widget to confirm your entry.
KAKO VSTOPITI: Če želite vstopiti, vnesite mednopni vložek in sledite navodilom za vstop v nagradne igrače.	HOW TO ENTER: To enter, submit your entry during the Sweepstakes Period and follow the directions to enter the Sweepstakes.

Tabela 3: Primer št. 6; »Mednopni vložek.«

eTranslation je bil v povprečju za 15 % boljši od prevajalnika Google Translate, v katerem je bil omenjen primer. Nasploh pa je eTranslation kazal boljše rezultate. Najboljši rezultati glede na tip napake so bili vezani na razdvoumljanje besednega pomena (65,9 %), kar kaže, da znamo ljudje nasploh dobro razbrati pomen iz sobesedila, na drugem mestu pa so bile neprevedene besede (64,8 %), kar lahko pripišemo dobremu znanju angleščine med udeleženci v anketi.

Rezultati so bili slabši, ko je prevajalnik napravil hujšo pomensko napako, ki je oteževala razumevanje celotnega segmenta (56,3 %), daleč najslabše rezultate pa je bilo moč opaziti v kategoriji izmišljena beseda (48,5 %), v kateri je sicer bil prej omenjeni primer št. 6.

Glede na tip konteksta so bili najboljši rezultati pri primerih, kjer je slika vplivala na razumevanje (69,8 %) in kjer so udeleženci morali izbrati sliko, na katero se je nanašalo besedilo (64,2 %). Rezultati so bili nekoliko slabši v izključno tekstovnem kontekstu (60,4 %), najslabši rezultati pa so bili v kategoriji, kjer je bila besedilo priložena slika, ki ne vpliva na razumevanje oz. potencialno zmede udeleženca (44 %) – v tej kategoriji je bil tudi primer št. 6. Izkazalo se je, da slikovni kontekst, ki lahko potencialno vpliva na razumevanje besedilnega segmenta,

pri strojnih prevodih v realnih okoliščinah, torej na spletu, z vsem pomožnim gradivom, igra pomembno vlogo.

Udeleženci, ki se sicer ukvarjajo s prevajanjem, so na splošno odgovarjali boljše od povprečja. Njihov delež uspešnosti je bil največji v kategoriji hujša pomenska napaka (+ 9,6 %), kar kaže na to, da zaradi »poklicne deformacije« bolj učinkovito razumejo kontekst.

6. Zaključek

V članku sem predstavil raziskavo o razumljivosti nerevidiranih strojno prevedenih spletnih besedil pri končnih uporabnikih, ki niso bili posebej obveščeni, da prebirajo strojne prevode. Razumevanje besedilnih segmentov, ki so vključevali štiri različne tipe napak, ki nastanejo pri strojnem prevajanju NMT-sistemov, sem preverjal z anketo. Ta je vsebovala strojne prevode splošnih besedil, ki sem jih prevedel s prevajalnikoma Google Translate in eTranslation. Besedila so bila nerevidirana, vsebovala so napake, ki so bile predstavljene v več različnih kontekstih, bodisi s slikovnim gradivom bodisi brez.

Rezultati so pokazali, da je splošna stopnja razumevanja 59 %, pri čemer se je izkazalo, da so prevodi eTranslationa nasploh razumljivejši od prevodov Googlovega prevajalnika. Število pravih odgovorov je bilo najvišje v kategoriji razdvoumljanja večpomenskih besed, kar nakazuje na to, da ljudje lažje razumemo pomen strojnih prevodov, če nam je dan kontekst. Pri tem je bilo najbolj učinkovito slikovno gradivo, s katerim so si lahko udeleženci v raziskavi pomagali razjasniti pomen določenega besedilnega segmenta. Druga najuspešnejša kategorija je bila razumevanje neprevedenih besed, kar pomeni, da je bilo znanje angleškega jezika med udeleženci na visokem nivoju.

Po analizi se je izkazalo, da je bil nekoliko problematičen način izbire odgovorov, saj sem anketirancem naključno vnaprej določil, na kakšen način bodo odgovarjali. Odgovori odprtega tipa so kazali slabše rezultate kot izbirni odgovori in odgovori zaprtega tipa, toda zaradi majhnega števila vprašanj je težko izpeljati kakšen razumen zaključek. Podobno velja za samo metodo odgovarjanja na anketo, ki je bila pogojena pandemičnemu času. Za bolj relevantne rezultate bi bilo potrebno izvajati test razumljivosti v živo, na razpravljalen način. Enako velja tudi za vzorec sodelujočih – večji in bolj raznolik vzorec bi dal jasnejše rezultate.

V bodoče bi bilo zanimivo raziskati, če se razumevanje nerevidiranih strojno prevedenih besedil izboljšuje skupaj z nadgradnjami strojnih prevajalnikov, hkrati pa bi se lahko osredotočil še na avtomatsko generirana besedila in jezik spletnih robotov.

Menim, da bo v prihodnje nekoliko manj raziskav storilnosti pri popravljanju strojnih prevodov in veliko več raziskav, ki bodo vezane na razumljivost strojno prevedenih ali avtomatsko generiranih besedil v praktičnih situacijah. Končni bralec se vedno bolj pogosto srečuje s takimi besedili, lahko pa pričakujemo, da bo zaradi še dodatnih izboljšav strojnih prevajalnikov, novih metod in razširjenosti prakse tovrstnih potencialnih stikov med stroji in bralci brez vmesnega posega človeškega popraviljalca vedno več.

7. Literatura

- David Bordon. 2021. Razumevati nevronščino: Kako si ljudje razlagamo jezik strojnih prevajalnikov. Magistrsko delo. Univerza v Ljubljani. Dostop 30. 5. 2022. <https://repozitorij.uni-lj.si/IzpisGradiva.php?id=125328>.
- Sheila Castilho in Ana Guerberof Arenas. 2018. Reading Comprehension of Machine Translation Output: What Makes for a Better Read?. V: Juan Antonio Perez-Ortiz, Felipe Sanchez-Martinez, Miquel Espla-Gomis, Maja Popovič, Celia Rico, Andre Martins, Joachim Van den Bogaert in Mikel L. Forcada, ur., *Proceedings of the 21st Annual Conference of the European Association for Machine Translation*, str. 79–88, Alacant, Španija. Dostop 30. 5. 2022. <http://doras.dcu.ie/23071/>.
- Gregor Donaj in Mirjam Sepesy Maučec. 2018. Prehod iz statističnega strojnega prevajanja na prevajanje z nevronskimi omrežji za jezikovni par slovenščina-angleščina. V: *Zbornik konference Jezikovne tehnologije in digitalna humanistika 2018*, str. 62–68, Ljubljana. Dostop 30. 5. 2022. http://www.sdjt.si/wp/wp-content/uploads/2018/09/JTDH-2018_Donaj-et-al_Prehod-iz-statisticnega-strojnega-prevajanja-na-prevajanje-z-nevronskimi-omrezji-za-jezikovni-par-slovenscina-anglescina.pdf.
- Evropska komisija, 2020 European Language Industry Survey 2020 Before & After Covid-19. Dostop 30. 5. 2022. https://ec.europa.eu/info/sites/default/files/2019_language_industry_survey_report.pdf.
- Philipp Koehn in Rebecca Knowles. 2017. Six challenges for neural machine translation. V: *Proceedings of the First Workshop on Neural Machine Translation*, str. 28–39. Association for Computational Linguistics, Vancouver, Canada. . Dostop 30. 5. 2022. <https://arxiv.org/pdf/1706.03872.pdf>.
- Chiraag Lala in Lucia Specia. 2018. Multimodal Lexical Translation. V: *Proceedings of the 11th international conference on language resources and evaluation (LREC)*, str. 3810–3817. Miyazaki, Japonska: European Language Resources Association (ELRA). Dostop 30. 5. 2022. <https://www.aclweb.org/anthology/L18-1602/>.
- Zsófia Lelner. 2022. »Machine Translation vs. Machine Translation Post-editing: Which One to Use and When?«. Dostop 30. 5. 2022. <https://blog.memoq.com/machine-translation-vs.-machine-translation-post-editing-which-one-to-use-and-when>.
- Jiatong Liu. Multimodal Machine Translation. Dostop 30. 5. 2022. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9547270>.
- Lieve Macken in Iris Ghysele. 2018. Measuring Comprehension and User Perception of Neural Machine Translated Texts: A Pilot Study. V: *Translating and the Computer 40 (TC40): Proceedings*, str. 120–126. Geneva: Editions Tradulex. Dostop 30. 5. 2022. <https://biblio.ugent.be/publication/8580951>.
- Lieve Macken, Laura Van Brussel in Joke Daems. 2019. NMT's wonderland where people turn into rabbits. A study on the comprehensibility of newly invented words in NMT output. V: *Computational Linguistics in the Netherlands Journal 9 (2019)*, str. 67–80. Dostop 30. 5. 2022. <https://www.clinjournal.org/clinj/article/view/93>.
- Marianna J. Martindale in Marine Carpuat. 2018. Fluency Over Adequacy: A Pilot Study in Measuring User Trust in Imperfect MT. Dostop 30. 5. 2022. <https://arxiv.org/abs/1802.06041>.
- Maja Popović. 2020. Relations between comprehensibility and adequacy errors in machine translation output. V: Raquel Fernández in Tal Linzen, ur., *Proceedings of the 24th Conference on Computational Natural Language Learning (CoNLL 2020)*, str. 256–264. Dostop 30. 5. 2022. <https://aclanthology.org/2020.conll-1.19.pdf>.
- Rico Sennrich, Barry Haddow in Alexandra Birch. 2016. Neural Machine Translation of Rare Words with Subword Units. Dostop 30. 5. 2022. <https://arxiv.org/abs/1508.07909>.
- Umut Sulubacak, Ozan Caglayan, Stig-Arne Grönroos, Aku Rouhe, Desmond Elliott, Lucia Specia in Jörg Tiedemann. 2020. Multimodal machine translation through visuals and speech. Dostop 30. 5. 2022. <https://arxiv.org/abs/1911.12798>.
- Ngo Thi-Vinh, Thanh-Le Ha, Phuong-Thai Nguyen in Le-Minh Nguyen. 2019. Overcoming the Rare Word Problem for Low-Resource Language Pairs in Neural Machine Translation. V: *Proceedings of the 6th Workshop on Asian Translation*, str. 207–214. Association for Computational Linguistics. Hong Kong, Kitajska. Dostop 30. 5. 2022. <https://arxiv.org/abs/1910.03467>.
- Diana Voroniak. Post-Editing of Machine Translation: Best Practices. Dostop 30. 5. 2022. <https://blog.crowdin.com/2022/03/30/mt-post-editing/>.
- Dan Zdarek. Machine Translation Post-editing Best Practices. Dostop 30. 5. 2022. <https://www.memsource.com/blog/post-editing-machine-translation-best-practices/>.