

Adriana Mezeg
Filozofska fakulteta, Univerza v Ljubljani

Ali sploh še potrebujemo prevjalce? Strojno prevajanje iz francoščine v slovenščino

1 Uvod

Globalizacija in tehnološki razvoj sta v preteklem stoletju močno vplivala na skoraj vsa področja družbe, prava revolucija pa se je v 90. letih prejšnjega stoletja začela z nastankom interneta ter z nenehnim razvojem in napredkom na področju informacijskih in komunikacijskih tehnologij ter umetne inteligence, s čimer se je življenje, kot smo ga poznali, korenito spremenilo. Velike spremembe so se zgodile tudi na področju prevajanja, zlasti z razvojem nevronskega strojnega prevajalnika, ki zadnjih nekaj let zaznamuje delovanje prevajalskega sektorja ter vplivajo na poklic prevajalca in samo izobraževanje bodočih prevajalcev.

Če je pred vzponom digitalnih tehnologij in razvojem strojnih prevajalnikov prevajanje praviloma potekalo s pomočjo slovarjev v tiskani obliki, je takšna praksa danes prava redkost. Papirnate slovarje so med drugim nadomestili elektronski slovarji in v še večji meri brezplačni spletni slovarji, elektronski korpusi (gl. Mezeg, 2020) in brezplačna spletna orodja za strojno prevajanje, kot sta Google Prevajalnik in DeepL. Zaradi vse večje izbire strojnih prevajalnikov in njihovega precejšnjega izboljšanja v zadnjih letih se pogosto sprašujemo, ali bi ta orodja nekoč lahko nadomestila človeškega prevajalca. Še nedolgo nazaj bi večina prevodoslovcev to odločno zanimala, z razvojem prevajalnikov na osnovi nevronske mreže, ki predstavljajo velik napredek v kakovosti strojnih prevodov, zaradi česar se je njihova uporaba v zadnjih letih močno povečala (gl. Ginovart Cid idr., 2020), pa odgovor vendarle ni več tako preprost in enoznačen.

Strojno prevajanje, ki se mu želimo posvetiti v članku, je aktualna tema na Oddelku za prevajalstvo ljubljanske Filozofske fakultete, saj učitelji že nekaj časa ugotavljamo, da ga študenti uporabljajo za potrebe študija. Osredotočili se bomo na prevajanje iz francoščine v slovenščino, saj razen Kolšek (2022), ki je v magistrski nalogi obravnavala strojno prevajanje francoskih tehničnih besedil v slovenščino, za to jezikovno kombinacijo še ni bilo opravljenih raziskav. Podobno raziskavo je leta 2020 sicer izvedla Vaupot, vendar v obratni smeri (iz slovenščine v francoščino).



Po kratkem pregledu zgodovine in področja strojnega prevajanja bomo najprej predstavili rezultate ankete o uporabi strojnih prevajalnikov, ki smo jo januarja 2023 izvedli med študenti francoščine ljubljanskega Oddelka za prevajalstvo. Z anketo smo po eni strani želeli ugotoviti, katere prevajalnike uporabljajo študenti, v kolikšni meri in za katere namene, po drugi strani pa sta nas zanimala njihovo stališče do strojnih prevodov in ocena njihovih zmožnosti popravljanja strojnih prevodov. Nato bomo na osnovi štirih kratkih odlomkov francoskih besedil analizirali prevode v slovenščino, ki smo jih pridobili z Google Prevajalnikom in DeepLom, in sicer bomo primerjali napake v googlovih prevodih iz leta 2019 in 2023 ter v prevodih z DeepLom iz leta 2023. S tem želimo preveriti napredek googlovega prevajalnika v zadnjih štirih letih, oceniti trenutne zmogljivosti obeh uporabljenih prevajalnikov ter ponuditi razmislek o težavnosti popravljanja strojnih prevodov in izzivih, ki jih to prinaša za izobraževanje bodočih prevajalcev.

2 Strojni prevajalniki in popravljanje strojnih prevodov

Začetki strojnega prevajanja segajo v 50. leta prejšnjega stoletja (Hutchins, 1995). Prvi sistemi za strojno prevajanje so temeljili na pravilih (slovarji in slovnice) (Vintar, 2013, 222), po začetnem optimizmu pa se je z zavedanjem omejenih zmogljivosti računalniških sistemov in kompleksnosti jezika zamisel o izgradnji popolnoma samodejnih prevajalskih sistemov, ki bi dosegli kakovost na ravni človeškega prevajanja, izkazala za nemogočo (Hutchins, 1995). Na pravilih temelječi prevajalniki so sicer prevladovali do konca 80. let (Vintar, 2013, 222), vendar so bili »z današnjega vidika še daleč od uporabnosti« (Vintar, 2021). V 90. letih so jih nadomestili statistični prevajalniki, ki so na osnovi podatkov iz vzporednih in enojezičnih korpusov za vsako besedo oz. besedno zvezo s pomočjo prevodnega in jezikovnega modela predvideli najverjetnejšo prevodno ustreznico, vendar brez upoštevanja konteksta (prav tam). Sledili so hibridni sistemi, ki so delovali na osnovi statistike in pravil, okrog leta 2014 pa so se pojavili t. i. nevronske prevajalniki, ki se prav tako učijo na velikih vzporednih korpusih, njihova glavna značilnost pa je nevronska omrežje, »ki posnema delovanje živčnega sistema bioloških organizmov« (Donaj idr., 2018, 63) in upošteva širši kontekst (prav tam, 62).

Po prvem valu objav o nevronskih prevajalnikih so se ti hitro uveljavili v akademskem svetu in industriji, saj so se izkazali za boljše od statističnih sistemov (Moorkens, 2018, 375). Nedavne raziskave kažejo, da so prevodi, izdelani z nevronskimi prevajalniki, v splošnem bolj tekoči (Castilho idr., 2017, 107) in idiomatski (gl. Nurminen idr., 2020, 160), izboljššan je tudi besedni red, opaziti pa je nepredvidljive težave na ravni izpustov in napačnih prevodov (Castilho idr., 2017; Moorkens, 2018, 375). Po drugi strani to, da je prevod bolj tekoč, še ne pomeni, da

je tudi ustrežnejši (Castilho idr., 2017, 118), obenem pa je v takšnem prevodu še težje odkriti pomenske in druge napake (Castilho idr., 2017; Nurminen idr., 2020, 161). Martindale idr. so v raziskavi iz 2018 (v Nurminen idr., 2020, 161) namreč ugotovili, da so se udeleženci veliko močneje odzvali na napake, ko prevod ni bil dovolj tekoč, kakor na pomenske napake, ki se jih morda sploh niso zavedali in jih je zlasti težko odkriti, če nimamo na voljo izvornika ali pa ga ne razumemo (Koponen idr., 2015, v Nurminen idr., 2020, 161).

Uporaba strojnih prevajalnikov je v zadnjih letih močno narasla tako v mednarodnih organizacijah in institucijah kakor tudi v prevajalskih in drugih podjetjih. Raziskava ELIS¹ za leto 2022 kaže, da je na evropski ravni 58 % vprašanih podjetij že uvedlo tehnologijo strojnega prevajanja, 20 % jih to še namerava, 70 % vprašanih, ki samostojno opravljajo jezikovni poklic, pa strojno prevajanje do določene mere že uporablja (ELIS, 2022, 25). Kljub morebitnemu občutku ogroženosti oz. strahu, da bodo strojni prevajalniki nekoč nadomestili človeške prevajalce (gl. Morokens, 2018, 376), so ti danes realnost, ki ji ni mogoče ubežati, vendar zaenkrat še nobeden ni sposoben izdelati prevoda, ki bi bil popolnoma enak človeškemu (ISO 18587, 2017), kar potrjujejo nekatere nedavne raziskave (npr. Guerberof-Arenas idr., 2020; Kenny idr., 2020). Strojni prevodi namreč kljub očitnemu izboljšanju v zadnjih letih (Canfora idr., 2020) niso brezhibni, končni izdelek pa je odvisen od človeških prevajalcev in njihove sposobnosti popravljanja strojnih prevodov (prav tam). Ločimo dve vrsti popravljanja, in sicer »full post-editing« oz. polno popravo (Vintar, 2013, 225), kjer je končni izdelek primerljiv s človeškim prevodom (ISO 18587, 2017, 3.1.5), in »light post-editing« oz. površno popravo (Vintar, 2013, 225), s katero pridobimo zgolj razumljivo besedilo, ne da bi poskušali doseči kakovost človeškega prevoda (ISO 18587, 2017, 3.1.6). Popravljanje strojnih prevodov je lahko celo zahtevnejša naloga od prevajanja samega (Krings 2001, 319, v do Carmo, 2020, 45), saj zahteva kompleksno sprejemanje odločitev, da bi bilo učinkovito (do Carmo, 2020, 52), v primerjavi s prevajanjem od začetka pa je pri tem vprašljivo tudi izboljšanje produktivnosti (Hu idr., 2016, 351), saj so kognitivne obremenitve pri popravljanju strojnih prevodov lahko večje² kot pri običajnem prevajanju. Poleg tega raziskave kažejo, da imajo popravljalci prevodov težave z odkrivanjem napak, ki jih povzročijo nevronske prevajalniki, zato bi morali strojne prevode popravljati izključno šolani popravljalci, ki so usposobljeni za odkrivanje napak (Canfora idr., 2020, 61–62).

1 European Language Industry Survey.

2 To je npr. odvisno od kakovosti surovega strojnega prevoda in usposobljenosti popravjalcev strojnih prevodov.

3 Uporaba strojnih prevajalnikov med študenti francoščine

Ko smo leta 2018 med magistrskimi študenti francoščine na ljubljanskem Oddelku za prevajalstvo izvedli anketo o uporabi različnih jezikovnih virov in orodij (Mezeg, 2020), so se strojni prevajalniki, kot je Google Prevajalnik,³ uvrstili na predzadnje mesto na seznamu,⁴ kar je ustrezalo poročanju takratnih študentov, da jih uporabljajo zelo redko (prav tam, 112). Nevronske tehnologije so prinesle »skokovit napredek« (Vintar, 2021) na področju strojnega prevajanja, s čimer se je povečalo tudi zanimanje za rabo strojnih prevajalnikov. Google Prevajalnik je avgusta 2017 dobil tekmeča, imenovanega DeepL, ki naj bi dosegal boljše rezultate od drugih prevajalskih sistemov in zato veljal za najboljšega na svetu (DeepL, b. d.). DeepL je sprva omogočal prevajanje med nekaterimi bolj razširjenimi jeziki, npr. angleščino, nemščino, francoščino in španščino, pozneje pa so bili dodani nekateri drugi jeziki, med njimi slovenščina marca 2021 (prav tam). Da bi ugotovili, kakšna je danes uporaba strojnih prevajalnikov med našimi študenti, smo januarja 2023 izvedli spletno anketo, ki jo je izpolnilo 23 študentov⁵ francoščine na 1. (13 študentov) in 2. (10 študentov) stopnji Oddelka za prevajalstvo Filozofske fakultete Univerze v Ljubljani, rezultate pa bomo predstavili v nadaljevanju. Anketna vprašanja so se večinoma nanašala na strojno prevajanje iz francoščine v slovenščino, pri čemer smo za primerjavo vključili nekaj vprašanj o prevajanju iz slovenščine v francoščino, posamezna vprašanja pa so bila povezana s strojnim prevajanjem na splošno.⁶

Približno enako število študentov je prevajalnike začelo uporabljati na fakulteti (39 %) oz. že v osnovni šoli (35 %), nekaj manj pa v srednji šoli (26 %). Izstopa (Graf 1) raba DeepLa, ki ga 35 % študentov uporablja občasno in 35 % pogosto, 9 % (tj. dva študenta) ves čas, 4 % redko in 17 % nikoli, medtem ko je Google Prevajalnik med študenti veliko manj priljubljen: 44 % ga ne uporablja, 30 % le redko, 13 % občasno, 9 % pogosto in 4 % (tj. en študent) ves čas. Še manj razširjen je eTranslation Evropske komisije,⁷ ki ga redko uporabi le 13 % študentov, preostalih 87 % pa nikoli. Od drugih brezplačnih orodij za strojno prevajanje nekaj študentov, kot so dodali v komentarju, včasih uporabi še Reverso⁸ in Glosbe,⁹ en študent pa tudi Matecat.¹⁰

3 Google Translate (<https://translate.google.com/>), ki je prišel na trg 2006, je sprva temeljil na statističnem modelu, od 2016 pa deluje na osnovi nevronske mreže.

4 Tj. na 12. mesto po uporabnosti od 13 ponujenih jezikovnih virov in orodij.

5 Tj. 54 % vseh študentov, ki so v študijskem letu 2022/2023 vpisani na smeri slovenščina-angleščina-francoščina.

6 Npr. kdaj so študenti začeli uporabljati strojne prevajalnike in ali menijo, da strojni prevajalniki ogrožajo poklic prevajalca.

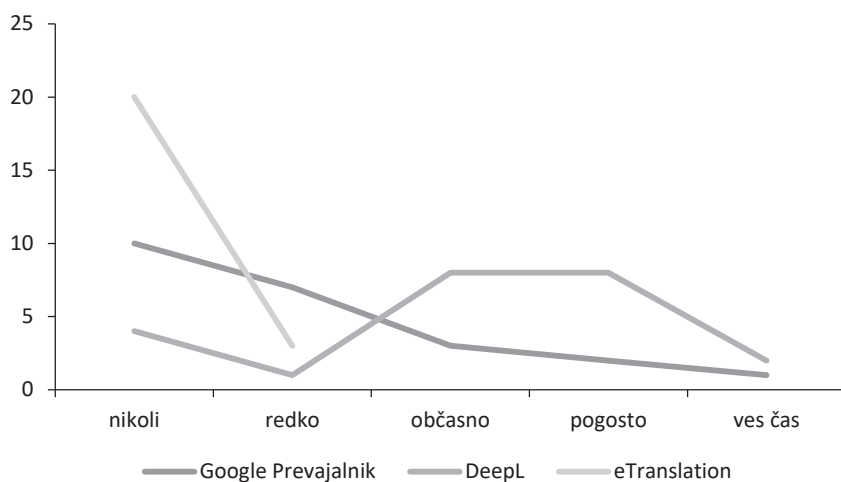
7 Dostopno na https://commission.europa.eu/resources-partners/etranslation_en.

8 Dostopno na <https://www.reverso.net/text-translation>.

9 Dostopno na <https://sl.glosbe.com/>.

10 Dostopno na <https://www.matecat.com/>.

Za potrebe prevajanja prevajalnike uporabljajo vsi študenti, od tega 44 % občasno, 39 % pogosto, 4 % (tj. en študent) ves čas in 13 % redko. Kot kaže Graf 2, pa prevajalnike uporabljajo tudi za druge namene: približno polovica jih občasno ali pogosto uporablja za pisanje spisov (48 %) in drugih pisnih sestavkov (52 %) ter za učenje študijske literature (52 %, od tega dva ves čas), medtem ko jih velika večina ne uporabi nikoli ali le redko za pisanje izpitov (92 %), e-sporočil (87 %) in brskanje po spletu (83 %). Dobra tretjina (35 %) respondentov si s prevajalniki pomaga tudi pri pisanju seminarskih nalog, med drugimi razlogi pa jih je nekaj v prostem odgovoru zapisalo, da jih uporabljajo za razumevanje nepoznanih besed, iskanje sopomenk in sporazumevanje v tujini.

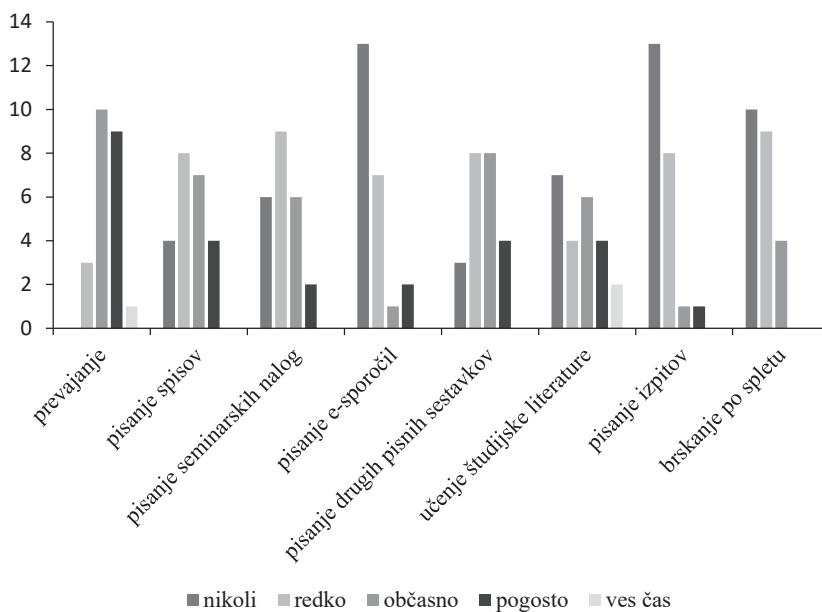


Graf 1: Pogostnost uporabe Google Prevajalnika, DeepLa in eTranslationa med študenti francoščine Oddelka za prevajalstvo.

Za veliko večino študentov so prevajalniki nepogrešljiv pripomoček pri prevajanju in pisanju, dobra tretjina (35 %) pa se s tem ne strinja. Zanimalo nas je, ali so študenti v stiski, kadar morajo besedilo prevesti ali napisati brez uporabe prevajalnikov, pri čemer so odgovori deljeni. Pri prevajanju iz francoščine v slovenščino slabi polovici (44 %) to ne povzroča nobene stiske, preostali pa so zaradi tega v manjši ali večji stiski. Kadar morajo besedilo v francoščini napisati sami, pa je zaradi tega v stiski precej več študentov, medtem ko na slabo tretjino to nima nobenega vpliva.

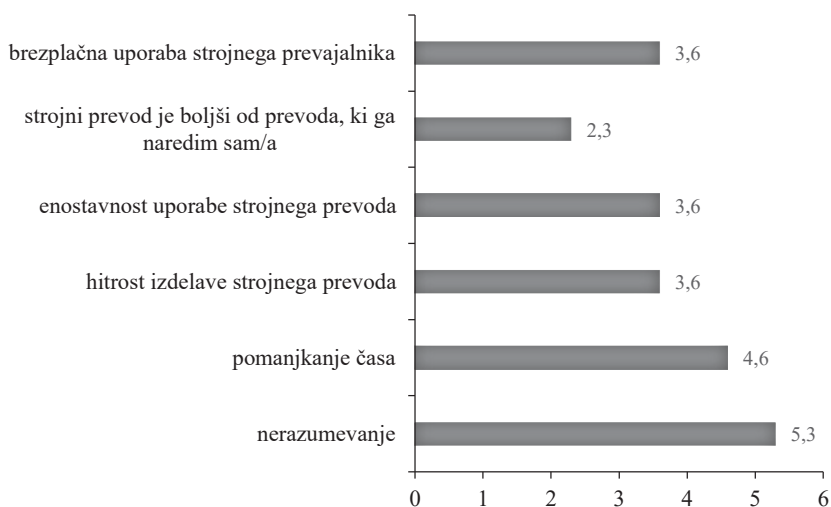
Kot najpogostejši razlog, zakaj pri prevajanju iz francoščine v slovenščino uporabljajo prevajalnike (Graf 3), so študenti na lestvici od 1 do 6 označili nerazumevanje (besed, besednih zvez ali povedi) v francoščini (povprečje 5,3) pred pomanjkanjem časa (4,6), srednjo pomembnost (3,6) so pripisali brezplačni uporabi strojnega

prevajalnika, hitrosti izdelave strojnega prevoda in enostavnosti uporabe strojnega prevajalnika, najmanj pogosto (2,3) pa po prevajalnikih posežejo, ker bi mislili, da je strojni prevod iz francoščine v slovenščino boljši od prevoda, ki ga opravijo sami. Pri tem se mnenje pričakovano razlikuje med študenti 1. in 2. stopnje: 52 % študentov, pretežno z 2. stopnje, se namreč nikakor ne strinja s tem, medtem ko večina študentov na 1. stopnji zaradi premalo izkušenj morda še ne zaupa povsem v svoje prevajalske kompetence, saj meni, da prevajalniki izdelajo boljši prevod iz francoščine v slovenščino. Da so njihovi prevodi iz francoščine boljši zaradi strojnih prevajalnikov, pa se strinja 39 % in delno strinja 52 % študentov, torej pretežna večina.



Graf 2: Namen uporabe strojnih prevajalnikov med študenti francoščine.

Brez uporabe prevajalnika francosko besedilo v slovenščino prevedeta samo dva študenta, nekaj več kot polovica (61 %) jih besedilo prevede sama, pri čemer včasih upošteva predloge strojnega prevoda, slaba tretjina (30 %) pa besedilo prevede s prevajalnikom in strojni prevod nato temeljito popravi. Ob tem se zastavlja vprašanje, ali imajo študenti dovolj jezikovnega in drugega potrebnega znanja za popravljanje strojnih prevodov: nihče od študentov ne misli, da nima dovolj znanja, 44 % (štirje na 1. in šest na 2. stopnji) jih ocenjuje, da imajo dovolj znanja, 56 % (večinoma študenti 1. stopnje) pa tega ne ve, pri čemer so nekateri dodali, da potrebujejo več vaje. Nemalo jih je v komentarju dopisalo, da je to sicer odvisno tudi od ciljnega jezika: v slovenščini s tem nimajo težav, v francoščini se ne počutijo kompetentne, v angleščini pa jim gre bolje.



Graf 3: Razlogi za uporabo strojnih prevajalnikov med študenti francoščine.

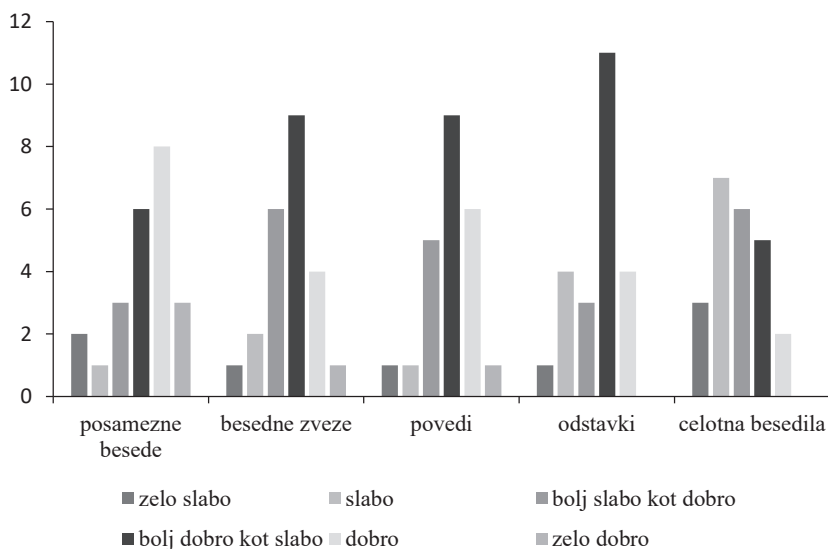
Kar zadeva prevajanje iz francoščine v slovenščino, približno polovica respondentov posamezne strojno prevedene besede in besedne zveze preveri v drugih virih, pri prevajanju iz slovenščine v francoščino pa je preverjanja manj, saj to pogosto ali vedno stori le dobra tretjina vprašanih. To bi lahko pomenilo, da študenti pri prevajanju v tuji jezik bolj zaupajo strojnemu prevajalniku in manj svojemu znanju, kar se glede na našo pedagoško prakso spremeni na magistrski ravni študija, ko imajo študenti boljše jezikovno znanje in prevajalske kompetence.

V povezavi z oceno kakovosti strojnih prevodov iz francoščine v slovenščino (Graf 4) je med 61 % in 74 % vprašanih na pozitivni strani lestvice označilo, da se jim zdijo strojni prevodi posameznih besed, besednih zvez, povedi in odstavkov dobri, dobra tretjina ali manj pa se s tem ne strinja. Več kot dve tretjini jih slabo ocenjuje tudi strojne prevode celotnih besedil. Študenti so nato v prostih odgovorih dodali, da je kakovost strojnih prevodov med drugim odvisna od jezikovne kombinacije (po njihovem mnenju so boljši prevodi v tuji jezik kakor v slovenščino), področja (strojni prevodi humanističnih, umetnostnih in strokovnih besedil vsebujejo veliko napak) ter vrste in zahtevnosti besedila: v pomoč so npr. pri preprostih besedah in povedih, pri terminologiji so lahko zavajajoči, pri prevodih kulturnospecifičnih izrazov, frazemov, stilističnih prvin in zahtevnejše skladnje pa pogosto odpovejo, zato se nanje ne moremo zanesti, saj je človeški prevod v teh pogledih po mnenju številnih boljši.

Študenti se razen dveh izjem ne strinjajo (91 %), da so strojni prevodi iz francoščine dovolj dobri in jih zato ne popravljajo, jih je pa skupno kar precejšen delež (74 %) na pozitivni strani lestvice označil, da je popraviti strojni prevod v slovenščini

lažje, kot če ga morajo prevesti sami; s tem se sicer sploh ne strinja 26 % vprašanih, vsi z magistrske stopnje, ki imajo že nekaj izkušenj s popraviljanjem strojnih prevodov. Da popraviljanje prevodov nikakor ni zamudno opravilo, meni 26 % študentov, vsi s 1. stopnje, večina pa se vendarle delno ali povsem strinja, da popraviljanje strojnih prevodov vzame veliko časa.

Nazadnje smo želeli preveriti še, ali študenti menijo, da strojni prevajalniki ogrožajo poklic prevajalca, s čimer se ne strinja 39 % vprašanih, več kot polovica pa se do določene mere zaradi tega počuti ogrožena. V prostih odgovorih so v povezavi s tem izpostavili, da se strojni prevod ne more kosati s človeškim prevodom, saj imamo ljudje več znanja o jeziku, kulturnih specifikah in stilistiki kot stroji, zato dvomijo, da bi človeške prevajalce nekoč v celoti lahko zamenjali strojni prevajalniki. Zavedajo se, da se bo poklic prevajalca spremenil, vendar bo vedno nujen človeški pregled.



Graf 4: Študentska ocena kakovosti strojnih prevodov različnih enot jezika iz francoščine v slovenščino.

4 Analiza strojnih prevodov

Čeprav je anketa pokazala, da študenti prevajalnike uporabljajo za različne namene, se bomo v nadaljevanju posvetili zgolj prevajanju iz francoščine v slovenščino. Kljub majhnemu vzorcu lahko zaključimo, da si večina vprašanih študentov pri izdelavi prevodov pomaga s strojnimi prevajalniki, pri čemer slaba tretjina, večinoma študenti 1. in 2. letnika 1. stopnje in en študent 2. stopnje, francosko besedilo v celoti prevede s prevajalnikom in ga nato temeljito popravi v slovenščini, obenem pa ti študenti ne

vedo, ali imajo dovolj jezikovnega znanja za popravljanje strojnih prevodov. Čeprav večina magistrskih študentov v anketi trdi, da ima dovolj takšnega znanja, pa pedagoško delo z njimi pri prevajanju strokovnih besedil iz francoščine v slovenščino tega ne potrjuje, saj v strojnih prevodih pogosto ne zaznajo vseh napak ali pa imajo težave z njihovim popravljanjem.

V tem poglavju želimo preveriti in primerjati kakovost strojnih prevodov iz francoščine v slovenščino, izdelanih z Google Prevajalnikom in DeepLom januarja 2023. Pri treh besedilih bomo primerjali tudi googlove prevode iz let 2019 in 2023, da bi ugotovili napredek tega nevronskega prevajalnika v zadnjih štirih letih. Izbrali smo odlomke štirih francoskih besedil, dolgih od 100 do 200 besed, pri čemer smo število napak zaradi primerljivosti normalizirali na 100 besed: 1) časopisni članek s politično vsebino iz časnika *Le Monde*, 2) poljudno besedilo o parfumu Chanel N° 5, 3) otroško besedilo o prazniku svetega Martina in 4) poslovno pogodbo o prodaji smučarskih vozovnic. Vse surove strojne prevode smo ročno analizirali in popravili, pri čemer smo uporabili osem kategorij napak:¹¹ ločila, pravopis, besedišče, slovnica, besedni red, slog, izpust in dodajanje.

V skladu s pričakovanji smo največ napak odkrili v treh (besedila 1–3) googlovih prevodih iz 2019 (gl. Graf 5), ki so se do 2023 nekoliko izboljšali: količina napak je sicer precej primerljiva (1,2 : 1 v korist verzije 2023) v poljudnem besedilu (besedilo 2) in časopisnem članku (besedilo 1) (1,4 : 1), medtem ko smo v prevodu otroškega besedila (besedilo 3) iz 2019 odkrili dvakrat več napak kot v verziji iz 2023 (razmerje 2,2 : 1). V prevodu tega besedila, ki je skladenjsko in leksikalno najmanj zahtevno, je napak sicer precej manj kot v prevodih drugih dveh besedil, kot kaže Graf 5. V prevodih vseh besedil iz 2019 in 2023 izstopa količina napak na ravni besedišča (npr. napačni, nerodni ali dobesečni prevod), v poljudnem besedilu, kjer je skladnja bolj zahtevna, je skoraj tretjina napak slovničnih (npr. čas, sklon, oseba, število), druge napake pa so redke, pri čemer smo opazili, da se sicer nizko število slogovnih napak v prevodih iz 2023 ni zmanjšalo. Čeprav je v verziji iz 2023 napak manj (4,6–7,8 manj napak na 100 besed), pa se tu pojavijo nekatere, ki jih prevodi iz 2019 niso vsebovali (npr. napačna raba ločil in izpusti). Za boljšo predstavo navajamo prevoda dveh povedi (besedilo 2; gl. primer Izvirnik 1), ki smo ju morali precej popraviti, da smo dobili razumljivo in slogovno ustrezno besedilo v slovenščini, pri čemer bi bile ponekod mogoče tudi drugačne rešitve.

11 Pri tem smo se zgledovali po tipologijah, ki so jih mdr. uporabili Hu idr. (2016), Moorkens (2018) in Vaupot (2020).

Izvirnik 1:

Le parfum sorti le 5 mai (cinquième mois de l'année) aurait porté le numéro 5 dans une série de propositions olfactives numérotées, comme le veut la légende. La réplique la plus célèbre liée au parfum est signée Marilyn Monroe qui, en 1954, affirmait ne porter que quelques gouttes de N° 5 de Chanel pour dormir.

Prevod 1 (Google 2019):

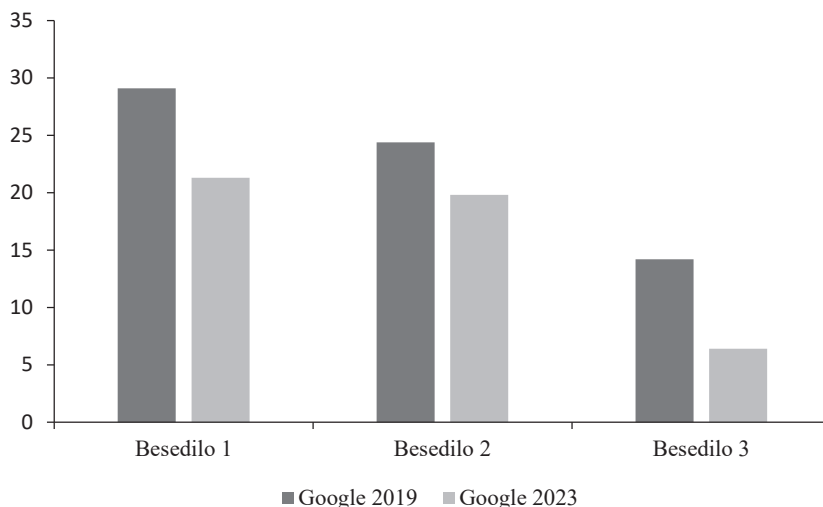
Parfum, izdan 5. maja (peti mesec v letu), bi nosil številka 5 v nizu oštevilčenih vohalnih predlogov, kot jih ima legenda. Najbolj znana replika, povezana s parfumom, je podpisana Marilyn Monroe, ki je leta 1954 trdila, da za spanje nosi le nekaj kapljic Chanel No 5.

Prevod 2 (Google 2023):

Dišava je izšla 5. maja (peti mesec v letu) bi imela številko 5 v nizu oštevilčenih vohalnih predlog, kot pravi legenda. Najbolj znana replika, povezana s parfumi, je Marilyn Monroe, ki je leta 1954 trdila, da za spanje nosi le nekaj kapljic Chanelovega No 5.

Popravljeni in izboljšani prevod:

Parfum, ki je na trg prišel 5. maja (peti mesec v letu), naj bi po legendi nosil številko 5 v seriji oštevilčenih predlogov dišave. Avtorica najbolj znane izjave o parfumu je Marilyn Monroe, ki je leta 1954 dejala, da med spanjem nima na sebi nič drugega kot nekaj kapljic chanela N° 5.



Graf 5: Število napak v posameznih besedilih glede na googlov prevod iz 2019 oz. 2023.

Nazadnje smo primerjali še prevode štirih besedilih, ki smo jih januarja 2023 pridobili s prevajalnikoma Google in DeepL. Analiza je pokazala, da so googlovi prevodi bolj okorni in vsebujejo več napak (Graf 6), vendar razlika ni velika (1,1–4,8 več napak na 100 besed kot pri DeepLu), razen v primeru prevoda besedila 2, kjer smo našli 11,6 več napak (predvsem leksikalnih in slovničnih) kot v prevodu z DeepLom, kjer smo odkrili le nekaj leksikalnih napak in izpustov. Več kot 20 napak (na 100 besed) vsebujeta oba prevoda odlomka pogodbe (besedilo 4; gl. primer Izvirnik 2), ki sta od vseh obravnavanih prevodov najslabša in bi ju bilo treba (vsaj nekatere dele) ponovno prevesti ali temeljito popraviti, kar po izkušnjah obravnave tega besedila z več generacijami magistrskih študentov prevajanja ni lahka naloga, saj sta za to potrebna dobro razumevanje izvirnika in znanje s področij prodaje in turizma.

Izvirnik 2:

Le PRENEUR s'engage à promouvoir et à commercialiser ces formules tout compris (Hébergement + Remontées Mécaniques) dans l'ensemble de ses supports de diffusion et notamment brochures, Internet... sans que cette liste ne soit exhaustive.

Prevod 1 (Google 2023):

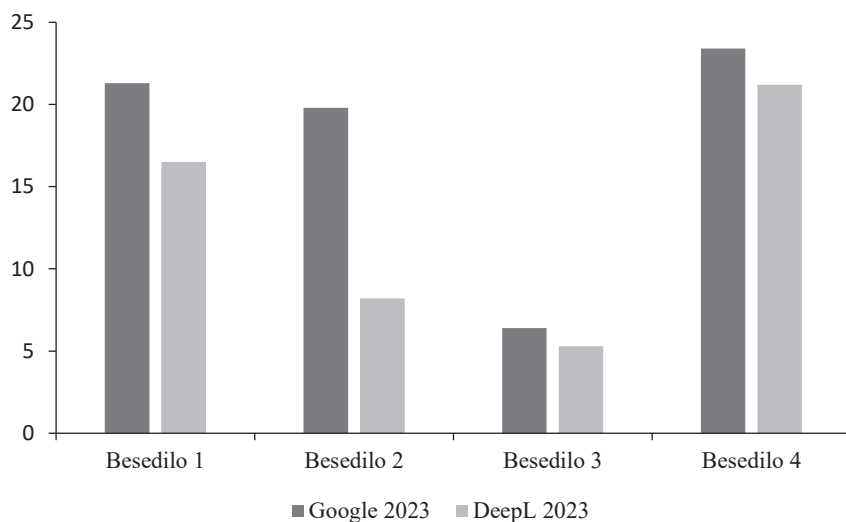
NAJEMNIK se zavezuje, da bo promoviral in tržil te vseobsegajoče formule (Nastanitev + smučarske vlečnice) v vseh svojih distribucijskih medijih in zlasti v brošurah, na internetu ... ne da bi bil ta seznam izčrpen.

Prevod 2 (DeepL 2023):

NAJEMNIK se zavezuje, da bo te all-inclusive pakete (nastanitev + žičnice) promoviral in tržil v vseh svojih distribucijskih medijih, zlasti v brošurah, na internetu ..., pri čemer ta seznam ni popoln.

Popravljeni in izboljšani prevod:

KUPEC se zavezuje, da bo oglaševal in prodajal te pakete all inclusive (nastanitev in smučarska vozovnica) prek vseh svojih prodajnih kanalov, zlasti v brošurah, na spletu itn.



Graf 6: Število napak v analiziranih besedilih glede na uporabljeni strojni prevajalnik.

Precejšnje število napak smo zaznali tudi v obeh prevodih časopisnega članka (besedilo 1), pri čemer je v prevodu z DeepLom manj leksikalnih in slogovnih napak, nekatere rešitve pa so lahko tudi zavajajoče ali pa jih težje zaznamo kot neustrezne brez upoštevanja in dobrega razumevanja izvirnika. Kot kaže spodnji primer (gl. primer Izvirnik 3), sta se prevajalnika, še posebno Google, slabo odrezala na dveh mestih: »les heures supplémentaires déchargées et défiscalisées«, kjer je DeepL izpustil prevod besede »déchargées« (oprostene plačila prispevkov), Google pa jo je prevedel neustrezno (»odpuščene«), in »volontaires«, kjer nobeden ni ponudil ustrezne rešitve, saj gre za to, da zaposleni lahko delajo ob nedeljah, če se s tem strinjajo.

Izvirnik 3:

Nicolas Sarkozy compte aussi rétablir les heures supplémentaires déchargées et défiscalisées et autoriser le travail le dimanche « *sous réserve que les salariés soient volontaires et davantage rémunérés* ».

Prevod 1 (Google 2023):

Nicolas Sarkozy namerava tudi obnoviti odpuščene in davka oproščene nadure ter dovoliti delo ob nedeljah, »pod pogojem, da so zaposleni prostovoljni in boljje plačani«.

Prevod 2 (DeepL 2023):

Nicolas Sarkozy namerava tudi ponovno uvesti neobdavčene nadure in dovoliti nedeljsko delo, »če bodo zaposleni pripravljene in boljje plačani«.

Popravljeni in izboljšani prevod:

Nicolas Sarkozy namerava tudi ponovno uvesti neobdavčene nadure, oproščene plačila prispevkov, in odobriti nedeljsko delo, »če se zaposleni strinjajo z njim in so plačani več.«

Če izvzamemo tipografsko oblikovanje (npr. krepki tisk), ki se s strojnim prevodom izgubi in ga je treba naknadno urediti (gl. primer Izvirnik 3), v vseh prevodih prevladujejo napake na ravni besedišča, sledijo slovnične in slogovne napake, katerih količina je pogojena z zahtevnostjo izvirnika (največ smo jih odkrili v prevodih pogodbe), druge vrste napak pa so se v večini prevodov pojavile redko (ločila, pravopis, izpusti, besedni red) ali sploh ne (dodajanje).

5 Sklep

Kljub temu da se strojni prevajalniki izboljšujejo in se bo zaradi vedno boljših rezultatov njihova uporaba v prihodnosti povečevala (Ginovart Cid idr., 2020, 284), pa strojni prevodi še vedno vsebujejo nezanemarljivo število napak. Analiza prevodov iz francoščine v slovenščino je pokazala, da je DeepL nekoliko boljši od Google Prevajalnika, ki se je z leti sicer izboljšal, vendar kakovost strojnega prevoda iz francoščine v slovenščino še vedno ni na ravni človeškega. V veliki meri je odvisna tudi od vrste in zahtevnosti izvirnika, pri čemer je vprašljiva produktivnost, ki naj bi se povečala s strojnimi prevodi (gl. do Carmo, 2020, 45), saj je njihovo popravilje neizbežno in lahko tudi zelo zahtevno opravilo, ki zahteva številne veščine, med drugim odlično razumevanje izvirnika ter sposobnost hitrega prepoznavanja napak in njihovega ustreznega odpravljanja v ciljnem jeziku.

Rezultati ankete kažejo, da študenti uporabljajo strojne prevajalnike za različne namene, pri čemer pričakovano izstopa prevajanje in raba prevajalnika DeepL. Stališče anketiranih študentov 1. in 2. stopnje glede uporabnosti prevajalnikov se sicer razlikuje, pretežna večina pa vendarle meni, da so zaradi strojnih prevajalnikov njihovi prevodi iz francoščine v slovenščino boljši, vendar ne vedo, ali so jih sposobni ustrezno popraviti. Pri seminarju iz prevajanja francoskih strokovnih besedil v slovenščino že več let ugotavljamo, da so študenti pogosto zadovoljni z rešitvami strojnih prevajalnikov, včasih pa tudi ne, vendar ne vedo, kako bi prevod izboljšali, saj jih strojna različica zavede, le redki pa so celo ob koncu študija sposobni brezhibno popraviti strojni prevod iz francoščine v slovenščino. Zato ni dovolj, da bodoče prevajalce naučimo zgolj to, da bodo sami znali odlično prevajati, temveč se moramo pri njihovem izobraževanju osredotočiti tudi na razvoj kompetenc za popravilje strojnih prevodov, saj obstaja velika verjetnost, da bodo prevajalci nekoč, morda že kmalu, večinoma postali popraviljalci strojnih prevodov, kar je po Pymu (2013, 488) namen strojnega prevajanja.

Zahvala

Članek je nastal v okviru programa »Medkulturne literarnovedne študije« (P6-0265), ki ga iz sredstev proračuna RS financira ARRS.

Literatura

- Canfora, C. idr., Risks in Neural Machine Translation, *Translation Spaces* 9 (1), 2020, str. 58–77.
- Castilho, S. idr., Is Neural Machine Translation the New State of the Art?, *The Prague Bulletin of Mathematical Linguistics* 108, 2017, str. 109–120.
- DeepL, *Press Information*, b. d. https://www.deepl.com/press.html#press_history_article, 13. 4. 2023.
- do Carmo, F., 'Time is Money' and the Value of Translation, *Translation Spaces* 9 (1), 2020, str. 35–57.
- Donaj, G. idr., Prehod iz statističnega strojnega prevajanja na prevajanje z nevronskimi omrežji za jezikovni par slovenščina-angleščina, v: *Zbornik konference Jezikovne tehnologije in digitalna humanistika* (ur. Fišer, D. idr.), Ljubljana 2018, str. 62–68.
- ELIS, *European Language Industry Survey 2022*, 2022, str. 1–43. <https://elis-survey.org/>, 13. 4. 2023.
- Ginovart Cid, C. idr., Language Industry Views on the Profile of the Post-editor, *Translation Spaces* 9 (2), 2020, str. 293–313.
- Guerberof-Arenas, A. idr., The Impact of Post-editing and Machine Translation on Creativity and Reading Experience, *Translation Spaces* 9 (2), 2020, str. 255–282.
- Hu, K. idr., A Comparative Study of Post-editing Guidelines, *Baltic Journal of Modern Computing* 4 (2), 2016, str. 346–353.
- Hutchins, W. J., Machine Translation: A Brief History, v: *Concise History of the Language Sciences: From the Sumerians to the Cognitivists* (ur. Koerner, E. F. K. idr.), Oxford 1995, str. 431–445.
- ISO 18587:2017(en), *Translation Services — Post-editing of Machine Translation Output — Requirements*, 2017. <https://www.iso.org/obp/ui/#iso:std:iso:18587:ed-1:v1:en>, 13. 4. 2023.
- Kenny, D. idr., Machine Translation, Ethics and the Literary Translator's Voice, *Translation Spaces* 9 (1), 2020, str. 123–149.
- Kolšek, T., *Strojno prevajanje tehničnih besedil s področja avtomobilizma na primeru navodil za uporabo vozila*, magistrsko delo, Ljubljana 2022.
- Mezeg, A., Parallel Corpus vs. Bilingual Dictionary: Their Usefulness in Translator Training, v: *Translating and Comparing Languages: Corpus-based Insights* (ur. Granger, S. idr.), Louvain-la-Neuve 2020, str. 123–140.

- Moorkens, J., What to Expect from Neural Machine Translation: A Practical In-class Translation Evaluation Exercise, *The Interpreter and Translator Trainer* 12 (4), 2018, str. 375–387.
- Nurminen, M., Machine Translation and Fair Access to Information, *Translation Spaces* 9 (1), 2020, str. 150–169.
- Pym, A., Translation Skill-Sets in a Machine-Translation Age, *Meta* 58 (3), 2013, str. 487–503.
- Vaupot, S., Analyse des erreurs de traduction automatique pour la combinaison de langues slovène-français et perspectives pour une formation en post-édition, *Matices en linguas extranjerias* 14 (2), 2020, str. 83–110.
- Vintar, Š., Sodobne prevajalske tehnologije in prihodnost prevajalskega poklica, *Uporabna informatika* 21 (4), 2013, str. 221–227.
- Vintar, Š., Z nevronščino v prihodnost, *Alternator: misliti znanost*, 9. december 2021, <https://www.alternator.science/sl/daljse/z-nevrnoscino-v-prihodnost/>, 13. 4. 2023.

Ali sploh še potrebujemo prevajalce? Strojno prevajanje iz francoščine v slovenščino

Ključne besede: strojno prevajanje, popravljanje strojnih prevodov, Google, DeepL, francoščina, slovenščina

Zaradi izboljšanja strojnih prevajalnikov v zadnjih letih se pogosto sprašujemo, ali bi ta orodja nekoč lahko nadomestila človeškega prevajalca. V članku predstavimo rezultate ankete o uporabi strojnih prevajalnikov med študenti francoščine ljubljanskega Oddelka za prevajalstvo, nato pa se posvetimo analizi napak v prevodih štirih besedil iz francoščine v slovenščino, ki smo jih pridobili z Google Prevajalnikom in DeepLom. Analiza je pokazala, da je DeepL nekoliko boljši od Google Prevajalnika, ki se je z leti sicer izboljšal, vendar kakovost strojnega prevoda iz francoščine v slovenščino še vedno ni na ravni človeškega, pri čemer izstopajo leksikalne napake pred slovničnimi in slogovnimi napakami. Večina študentov si pri prevajanju iz francoščine v slovenščino pomaga s prevajalniki in meni, da so zaradi tega njihovi prevodi boljši, obenem pa ne vedo, ali imajo dovolj znanja za njihovo popravljanje. Zato ni dovolj, da študente naučimo zgolj to, da bodo sami znali odlično prevajati, temveč se moramo pri njihovem izobraževanju osredotočiti tudi na razvoj kompetenc za popravljanje strojnih prevodov, saj obstaja velika verjetnost, da bodo prevajalci v prihodnosti večinoma postali popravljalci strojnih prevodov.

Do We Even Need Translators Anymore? Machine Translation from French into Slovene

Keywords: machine translation, postediting machine translation, Google, DeepL, French, Slovene

Given the improvement of machine translation (MT) in recent years, we often wonder if these tools could one day replace a human translator. In this paper, we present the results of a survey on the use of MT among French students at the Department of Translation in Ljubljana. We then analyse the errors in MT of four texts from French into Slovene created with Google Translate and DeepL. The analysis shows that DeepL is slightly better than Google Translate, which has improved over the years, but the quality of MT from French into Slovene is still not at the level of human translations, with lexical errors outweighing grammatical and stylistic errors. Most students use MT as a translation aid because they believe it will make their translations from French into Slovene better, but they do not know if they have the skills to correct them. Therefore, it is not enough to teach students to be excellent translators themselves, as we also need to focus their education on developing skills for postediting MT, since it is very likely that translators will mainly work as posteditors in the future.

O avtorici

Adriana Mezeg je izredna profesorica za prevodoslovje na Oddelku za prevajalstvo Filozofske fakultete Univerze v Ljubljani. Raziskovalno se ukvarja s prevodoslovjem ter s francosko-slovensko kontrastivno slovnico in korpusnim jezikoslovjem.

E-naslov: adriana.mezeg@ff.uni-lj.si

About the author

Adriana Mezeg works as an Associate Professor of Translation Studies at the Department of Translation, Faculty of Arts, University of Ljubljana. Her main research interests include translation studies, French-Slovene contrastive and corpus linguistics.

Email: adriana.mezeg@ff.uni-lj.si