

PREVODOSLOVJE
IN UPORABNO
JEZIKOSLOVJE



Uredila Špela Vintar

SLOVENSKI PREVODI SKOZI KORPUSNO PRIZMO

SLO
PRE
SKO
KOR
PRI

Uredila Špela Vintar

SLOVENSKI PREVODI SKOZI KORPUSNO PRIZMO

Zbirka Prevodoslovje
in uporabno jezikoslovje

Ljubljana 2017

SLOVENSKI PREVODI SKOZI KORPUSNO PRIZMO

ZBIRKA PREVODOSLOVJE IN UPORABNO JEZIKOSLOVJE

ISSN 2335-335X

Urednica: Špela Vintar

Recenzenta: Erich Prunč, Vojko Gorjanc

Uredniški odbor zbirke: Špela Vintar, Vojko Gorjanc, Nike Kocijančič Pokorn

Tehnični urednik: Martin Grad

Založila: Znanstvena založba Filozofske fakultete Univerze v Ljubljani

Izdal: Oddelek za prevajalstvo

Za založbo: Roman Kuhar, dekan Filozofske fakultete

Ljubljana, 2017

Prva izdaja/First edition

Elektronska izdaja/Digital edition

DOI: 10.4312/9789612375652

Oblikovna zasnova: Kofein, d. o. o.

Publikacija je brezplačna./ Publication is free of charge.

Publikacija je dostopna na:/Publication is available on: <https://e-knjige.ff.uni-lj.si>

Delo je ponujeno pod licenco Creative Commons Attribution-ShareAlike 4.0

International License (priznanje avtorstva, deljenje pod istimi pogoji).

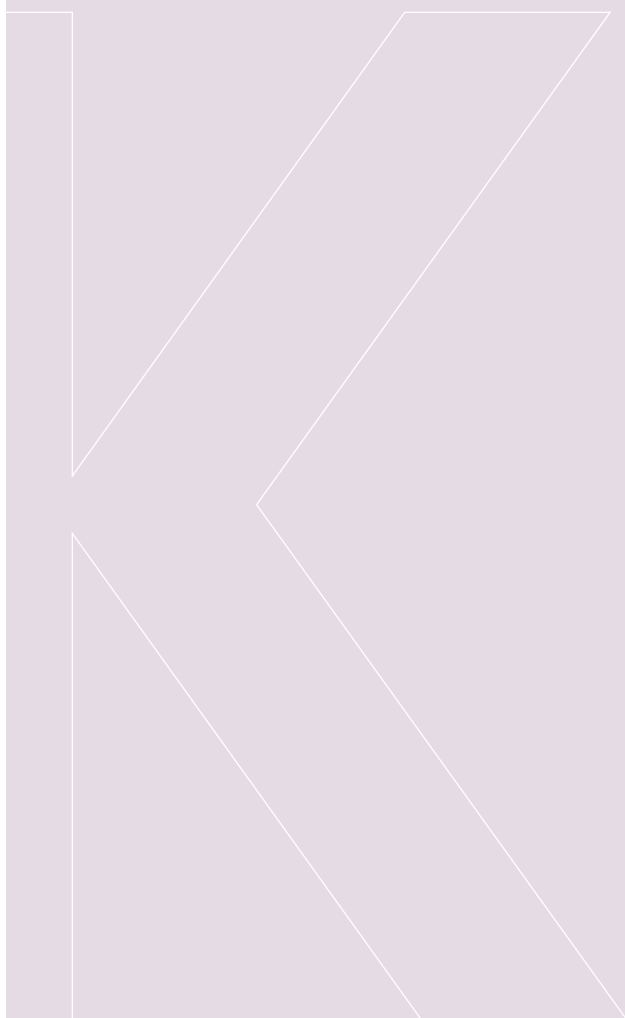


Kataložni zapis o publikaciji (CIP) pripravili
v Narodni in univerzitetni knjižnici v Ljubljani

COBISS.SI-ID=274604032

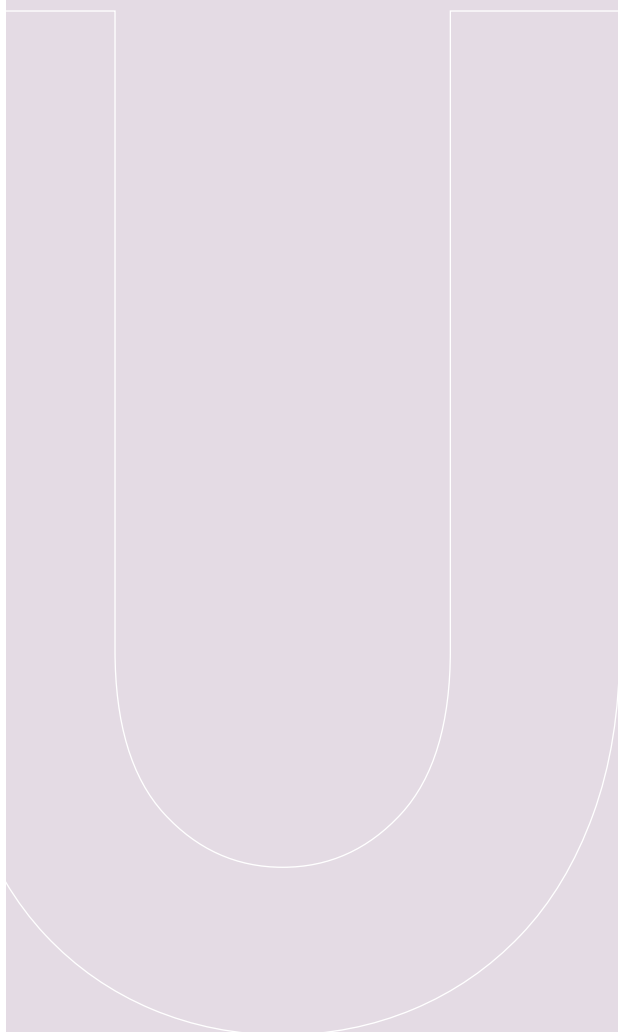
ISBN 978-961-237-565-2 (pdf)

Kazalo vsebine



Uvodnik: O rojstvu korpusa SPOOK in njegovih prvih sadovih <i>Špela Vintar</i>	6
Seznam del v korpusu SPOOK	10
Vzporedni korpus SPOOK: označevanje, zapis in iskanje <i>Tomaž Erjavec</i>	14
Pogostost vzročno-pojasnjevalnega <i>namreč</i> v izvirniku in prevodu: korpusni pristop <i>Nataša Hirci</i>	32
Samostalniki <i>človek</i> med generično in specifično rabo v korpusu prevedenih in izvirnih besedil v slovenščini <i>Agnes Pisanski Peterlin</i>	54
Neosebne glagolske oblike v prevodni in izvirni slovenščini: primer <i>deležja</i> <i>Tamara Mikolič Južnič</i>	70
Pridevniški polstavki v francoščini in prevedeni slovenščini <i>Adriana Mezeg</i>	94
Primerjava slovenskih prevodnih variant povezovalca <i>mais</i> v francosko-slovenskem vzporednem korpusu <i>Mojca Schlamberger Brezar</i>	120
Pogled na pojme med jeziki in jezikovnimi viri <i>Darja Fišer in Kristina Bizjak</i>	144
Ptičji pogled na leksikalno ustvarjalnost v slovenskih prevodih <i>Špela Vintar</i>	166
Pragmatični pomen in semantična prozodija v medjezični perspektivi: primer slovenščine in angleščine <i>Mojca Šorli</i>	180
Imensko kazalo	206

Uvodnik: O rojstvu korpusa SPOOK in njegovih prvih sadovih



It is important to avoid perfectionism in corpus building.
(Sinclair, 2005)

V pričujoči knjigi vam predstavljamo devet izvirnih in svežih raziskav, ki z različnih zornih kotov obravnavajo slovenske prevode in se posvečajo različnim jezikovnim in prevodnim pojavom v skupno štirih jezikovnih parih. Veže jih pogled skozi korpusno prizmo; prav vse namreč uporabljajo korpusno metodologijo, in razen ene vse temeljijo na prvem slovenskem prevodoslovnem korpusu SPOOK. Čeprav večina avtorjev v uvodnem delu govori o uporabljenem gradivu, tehnične vidike izdelave korpusa pa natančno opisuje prispevek Tomaža Erjavca, se vendarle spodobi, da takšni monografiji za popotnico povemo zgodbo o nastanku korpusa, predstavimo njegovo sestavo in pojasnimo razloge zanjo.

Slovenski PrevOdOslovni Korpus je nastal v okviru projekta Slovensko prevodoslovje – viri in raziskave, ki je potekal na Filozofski fakulteti v partnerstvu z Institutom Jožefa Stefana od maja 2009 do aprila 2012. Projekt je pod šifro J6-2009-0581 financirala Javna agencija za raziskovalno dejavnost Republike Slovenije. Kot je razvidno iz knjige, ki jo držite v rokah, sta glavna cilja projekta dosežena:

- Prvi je bil zgraditi prevodoslovni korpus, ki bi zajemal med seboj primerljive vzporedne korpuse prevodov v slovenščino iz štirih tujih jezikov, angleščine, francoščine, italijanščine in nemščine, ter njihove izvirnike, tretji nevzporedni del korpusa pa bi bil zgrajen iz del, izvirno napisanih v slovenščini.
- Drugi je bil uporabiti korpus za temeljne empirične raziskave o značilnostih slovenskih prevodov; prvo bero teh raziskav prinaša ta knjiga, upamo pa, da jih bo korpus SPOOK obrodil še veliko.

Korpus v sedANJI obliki zajema 95 leposlovnih del; od tega je 23 izvirnih slovenskih romanov, ostalo pa so prevodi in njihovi izvirniki. Prvotni načrt za sestavo korpusa je bil bistveno bolj ambiciozen, saj smo želeli zagotoviti reprezentativnost tudi po žanrski sestavi in smo v neleposlovni del nameravali vključiti še uporabniške priročnike, publicistična in strokovna besedila. Kmalu po pričetku zbiranja besedil smo morali prvotni načrt revidirati, saj je postalo jasno, da s človeškimi in finančnimi viri, ki smo jih imeli na voljo, tak ne bo izvedljiv. Tako smo v končni korpus vključili le leposlovna dela, natančneje prozna leposlovna dela, prevedena v slovenščino iz izbranih štirih jezikov v zadnjih desetih letih, ob tem pa smo si zastavili še željo po čim večji žanrski raznolikosti, različnih prevajalcih in različnih založbah, prioriteto pa naj bi imela odmevna, brana, priljubljena dela.

Takoj na začetku smo trčili ob dve pomembni vprašanji: kako priti do teh del v izvirniku in prevodu, po možnosti v elektronski obliki, s tem pa je neločljivo povezano tudi vprašanje avtorskih pravic.

Poti do knjig so bile zelo različne. Za francosko-slovenski in italijansko-slovenski del korpusa se je kmalu pokazalo, da zbirki pravzaprav že obstajata, saj sta sodelavki projekta Adriana Mezeg in Tamara Mikolič Južnič v okviru svojih doktorskih disertacij nedavno tega zgradili vsaka svoj korpus, ki je že zajemal potrebno količino leposlovnih del v izvirniku in prevodu. Obe sta bili pripravljeni prispevati svoja korpusa, tako da smo v okviru projekta izvedli le še pretvorbo v enotni SPOOK format in dela priključili ostalim. Za angleško-slovenski in nemško-slovenski jezikovni par te sreče nismo imeli in je bilo treba začeti z ničle. Najprej smo za vsak podkorpus sestavili »seznam želja«, se pravi prioritetnih del, ki bi jih želeli vključiti. Elektronske izvirnike teh del smo iskali na različnih javnih in manj javnih delih spleta, nekatere smo tudi kupili kot e-knjigo, nekaterih ni bilo mogoče pridobiti nikakor in smo jih digitalizirali sami. Slovenske prevode smo večinoma pridobili neposredno od prevajalcev, nekaj pa smo jih pridobili tudi od založb. Pri več delih se je zgodilo, da smo imeli na razpolago prevod, izvirnika pa ni bilo moč najti nikjer – v enem primeru niti tiskanega izvoda v slovenskih knjižnicah nismo našli; taka dela smo morali črtati s seznama in nadomestiti z drugimi. Končni izbor je tako nujen kompromis med željami in možnostmi, a smo kljub temu uspeli za vse jezikovne pare zagotoviti minimalno raven heterogenosti. Slovenske izvirnike smo brez izjeme vzeli iz korpusa Fidaplus, spet smo izbirali po načelu čim večje žanrske in avtorske raznolikosti, pri datumu nastanka pa smo se omejili na leto 1998 in mlajša.

Pri vprašanju avtorskih pravic smo se jasno zavedeli, da bi pravice za slovenske prevode in izvirnike morda celo uspeli zbrati, čeprav bi tudi to gotovo močno upočasnilo gradnjo korpusa in korenito vplivalo na izbor del. Med tujejezičnimi izvirniki so tudi zelo odmevna in tržno uspešna dela, ki so avtorskopravno zaščiteni do zob in čez, zato nam pravic za njihovo vključitev v korpus skoraj zagotovo ne bi uspelo pridobiti. Če bi morali taka dela izključiti in se omejiti denimo na tista, katerih avtorji so že dlje časa pokojni, bi bil tak korpus bistveno manj zanimiv za raziskovanje prevodov in sodobnih prevajalskih strategij. Tako smo s težkim srcem sprejeli odločitev, da korpus ne bo javno dostopen, do njega pa imajo odprto pot vsi raziskovalci, ki ga želijo uporabiti za kakršne koli nekomercialne namene.

Morda je zanimivo, da so se že besedilodajalci »veselili« prevodoslovnega korpusa in potencialnih rezultatov raziskav. Ko smo se denimo na Mladinski knjigi pogovarjali z urednikoma Miho Kovačem in Tatjano Cestnik o ciljih projekta, je v pogovoru na površje splavalo več raziskovalnih hipotez, domnev o značilnostih prevodov v slovenščino, o različnih praksah založniških hiš, o lektorskih

politikah, o recepciji prevodov in še marsikaj. Prevajalci, ki so prispevali svoje prevode, so bili v nekaterih primerih negotovi (»A zdaj boste pa skritizirali moj prevod ...«), v nekaterih pa radovedni in pozitivni (»Končno, skrajni čas, da to nekdo razišče ...«).

In res je bil skrajni čas, čeprav – kot ugotavlja večina avtorjev v sklepnih mislih svojih prispevkov – je korpus SPOOK bistveno premajhen, da bi dopuščal poplajevanje pridobljenih rezultatov. Vseeno pa se iz posameznih kamenčkov v mozaiku zarisuje podoba sodobnega slovenskega leposlovnega prevoda, ki v marsičem odstopa od še vedno globoko zakoreninjenih predsodkov, da so prevodi po nekih težko oprijemljivih kriterijih inferiorni od izvirne jezikovne produkcije, da so slovenski prevajalci ujeti v primež okostenele jezikovne norme, da iz prevodov komaj zastrte v bralca boljšijo prevzete strukture izvirnega jezika, da si prevajalci tuje kulturne svetove v maternem jeziku le redko upajo slikati s tako pogumnimi potezami kot avtorji.

Za pomoč pri gradnji korpusa SPOOK se zahvaljujemo vsem prevajalcem, ki so prispevali svoja besedila za naš projekt: Sandri Baumgartner Naylor, Vasji Cerar, Branku Gradišniku, Nini Grahek Križnar, Jakobu J. Kendi, Tini Mahkota, Štefanu Vevarju in Urši Vogrinc Javoršek.¹ Zahvaljujemo se tudi založbi Mladinska knjiga za pozitiven odnos in prispevek besedil, čeprav ta na koncu niso bila vključena v korpus. Pri pridobivanju besedil iz različnih manj tradicionalnih virov se je odlično izkazal Miha Mele, ki je igral odločilno vlogo tudi pri tehnični obdelavi in poravnavi, ob pomoči Špele Gašparič. Za končno pretvorbo besedil sta skrbela Senja Pollak in Tomaž Erjavec. Zahvaljujemo se vsem avtorjem besedil in recenzentom, prav posebne zasluge pri nastanku te knjige pa gredo tehničnemu uredniku Martinu Gradu.

Preostane le še želja, da bi iz rezultatov naših raziskav prevajalci in prevodoslovci črpali navdih za še boljše prevode in še več raziskovalnih spoznanj.

Špela Vintar

Ljubljana, november 2012

¹ V zahvali navajamo le tiste prevajalce, ki so svoja dela odstopili konkretno našemu projektu. Korpus SPOOK sicer vsebuje še številne druge prevode, ki so jih prevajalci bodisi odstopili že prej Adriani Mezeg ali Tamari Mikolič, bodisi so bila vključena v korpus Fidaplus, iz katerega smo prav tako črpali prevode.

SEZNAM DEL V KORPUSU SPOOK

Angleščina – slovenščina

- Dan Brown: *The Da Vinci Code*. New York: Doubleday, 2003.
- Dan Brown: *Da Vincijeva šifra*. Prevajalka Nataša Müller, Ljubljana: Mladinska knjiga, 2004.
- Eoin Colfer: *The Supernaturalist*. New York: Hyperion, 2004.
- Eoin Colfer: *Rešitelji*. Prevajalka Urša Vogrinc, Ljubljana: Mladinska knjiga, 2005.
- Colin Dexter: *The way through the woods*. London: Macmillan, 1992.
- Colin Dexter: *Pot skozi gozdove: inšpektor Morse in Zlovešči verzi*. Prevajalka Nina Grahek Križnar, Ljubljana: Mladinska knjiga, 2002.
- Mark Haddon: *The Curious Incident of the Dog in the Night-Time*. Oxford: David Fickling, 2003.
- Mark Haddon: *Skrivnostni primer ali kdo je umoril psa*. Prevajalec Vasja Cerar, Ljubljana: Mladinska knjiga, 2006.
- Doris Lessing: *The Fifth Child*. London: Harper Perennial, 2007.
- Doris Lessing: *Peti otrok*. Prevajalec Dušan Ogrizek, Radovljica: Didakta, 2008.
- J. K. Rowling: *Harry Potter and the Half-Blood Prince*. London: Bloomsbury, 2005.
- J. K. Rowling: *Harry Potter. Polkrvni princ*. Prevajalec Branko Gradišnik, Ljubljana: Epta, 2006.
- J. K. Rowling: *Harry Potter and the Deathly Hallows*. London: Bloomsbury, 2007.
- J. K. Rowling: *Harry Potter. Svetinje smrti*. Prevajalec Jakob J. Kenda, Ljubljana: Mladinska knjiga, 2008.
- Zadie Smith: *White Teeth*. London: Penguin Books, 2001.
- Zadie Smith: *Beli zobje*. Prevajalka Tina Mahkota, Ljubljana: Študentska založba, 2004.
- J.R.R. Tolkien: *Lord of the Rings: The Two Towers*. London: Harper Collins, 2001.
- J.R.R. Tolkien: *Gospodar prstanov: Stolpa*. Prevajalec Branko Gradišnik, Ljubljana: Mladinska knjiga, 2002.

Nemščina – slovenščina

- Günter Grass: *Katz und Maus*. Göttingen: Steidl, 1993.
- Günter Grass: *Mačka in miš*. Prevajalec Štefan Vevar, Radovljica: Didakta, 2000.
- Elfriede Jelinek: *Der Tod und das Mädchen III (Rosamunde)*. Berlin: Berliner Taschenbuch Verlag, 2004.
- Elfriede Jelinek: *Smrt in deklica III (Rozamunda)*. Prevajalci Urška P. Černe, Anja Uršič in Sandra Baumgartner, Celovec: Mohorjeva družba, 2004.
- Sten Nadolny: *Die Entdeckung der Langsamkeit*. München: Piper Verlag, 1983.
- Sten Nadolny: *Odkritje počasnosti*. Prevajalec Štefan Vevar, Ljubljana: Cankarjeva založba, 2004.
- Bernhard Schlink: *Der Vorleser*. Zürich: Diogenes, 1995.

- Bernhard Schlink: *Bralec*. Prevajalka Sandra Baumgartner, Ljubljana: Cankarjeva založba, 2001.
- W. G. Sebald: *Austerlitz*. München: C. Hanser, 2001.
- W. G. Sebald: *Austerlitz*. Prevajalec Štefan Vevar, Ljubljana: Študentska založba, 2005.
- Martin Walser: *Ein liebender Mann*. Reinbek bei Hamburg: Rowohlt, 2008.
- Martin Walser: *Ljubeči moški*. Prevajalec Slavo Šerc, Celje: Celjska Mohorjeva družba: Društvo Mohorjeva družba, 2010.
- Feridun Zaimoglu: *Leyla*. Köln: Kiepeheuer & Witsch, 2006.
- Feridun Zaimoglu: *Leyla*. Prevajalec Slavo Šerc, Ljubljana: Cankarjeva založba, 2007.

Francoščina – slovenščina

- Pascal Bruckner: *L'Amour du prochain*. Éditions Grasset & Fasquelle, 2004.
- Pascal Bruckner: *Ljubezen do bližnjega*. Prevajalec Jaroslav Skrušný, Ljubljana: Študentska založba, 2007.
- Marie Darrieussecq: *Truismes*. P.O.L. Éditeur, 1997.
- Marie Darrieussecq: *Svinjarije*. Prevajalec Marko Crnkovič, Ljubljana: DZS, 1997.
- Fatou Diome: *Le ventre de l'Atlantique*. Éditions Anne Carrière, 2003.
- Fatou Diome: *Trebuha Atlantika*. Prevajalka Ana Barič, Ljubljana: Sanje, 2007.
- Jean Echenoz: *Je m'en vais*. Les Éditions de Minuit, 1999.
- Jean Echenoz: *Grem*. Prevajalka Suzana Koncut, Ljubljana: Modrijan, 2008.
- Laurent Gaudé: *Eldorado*. Actes Sud, 2006.
- Laurent Gaudé: *Eldorado*. Prevajalka Tjaša Mohar, Celje: Društvo Mohorjeva družba, 2007.
- Philippe Grimbert: *Un secret*. Éditions Grasset & Fasquelle, 2004.
- Philippe Grimbert: *Skrivnost*. Prevajalka Jožica Grum, Radovljica: Didakta, 2008.
- Hervé Guibert: *Fou de Vincent*. Les Éditions de Minuit, 1989.
- Hervé Guibert: *Nor na Vincenta*. Prevajalec Brane Mozetič, Ljubljana: ŠKUC, 2007.
- Michel Houellebecq: *Plateforme*. Éditions Flammarion, 2001.
- Michel Houellebecq: *Platforma*. Prevajalka Mojca Medvedšek, Ljubljana: Cankarjeva založba, 2002.
- Andrei Makine: *Le testament français*. Éditions Mercure de France, 1995.
- Andrei Makine: *Francoski testament*. Prevajalka Nadja Dobnik, Študentska založba Litera, 2005.
- Pierre Mérot: *Mammifères*. Éditions Flammarion, 2003.
- Pierre Mérot: *Sesalci*. Prevajalec Jan Jona Javoršek, Ljubljana: Mladinska knjiga, 2007.
- Dai Sijie: *Balzac et la petite tailleuse chinoise*. Éditions Gallimard, 2000.
- Dai Sijie: *Balzac in kitajska šiviljica*. Prevajalka Alenka Moder Saje, Tržič: Učila, 2002.

Shan Sa: *Impératrice*. Éditions Albin Michel S.A., 2003.
 Shan Sa: *Cesarica*. Prevajalka Ana Barič, Tržič: Učila, 2007.

Italijanščina – slovenščina

Niccolò Ammaniti: *Io non ho paura*. Einaudi, 2001.
 Niccolò Ammaniti: *Ni me strah*. Prevajalka Nataša Kos, Ljubljana: Študentska založba, 2004.
 Alessandro Baricco: *Seta*. RCS Libri, 1996.
 Alessandro Baricco: *Svila*. Prevajalec Maja Novak, Vale-Novak, 1999.
 Alessandro Baricco: *Novecento*. Feltrinelli, 1994.
 Alessandro Baricco: *Devetsto*. Prevajalec Maja Novak, Vale-Novak, 2000.
 Sveva Casati Modignani: *Vaniglia e cioccolato*. Sperling & Kupfer Editori, 2000.
 Sveva Casati Modignani: *Vanilija in čokolada*. Prevajalec Lucija Marija Mal, Prešernova družba, 2003.
 Umberto Eco: *Baudolino*. Bompiani, 2000.
 Umberto Eco: *Baudolino*. Prevajalec Vasja Bratina, Mladinska knjiga, 2003.
 Oriana Fallaci: *La rabbia e l'orgoglio*. Rizzoli, 2001.
 Oriana Fallaci: *Bes in ponos*. Prevajalec Vasja Bratina, Učila International, 2004.
 Anotnio Tabucchi: *Sogni di sogni*. Sellerio, 1992.
 Anotnio Tabucchi: *Sanje o sanjah*. Prevajalec Teo Šinkovec, Ark, 2003.

Slovenščina

Renata Ažman: *[Japajade*. Tuma, 2005.
 Boštjan Cvetič: *Umri*. Karantanija, 2003.
 Aleš Čar: *Pasji tango*. Študentska založba, 1999.
 Dušan Čater: *Patosi*. Karantanija, 1999.
 Zoran Hočevár: *Šolen z Brega*. Založba / *cf., 1998.
 Alojz Ihan: *Romanje za dva --- in psa*. Študentska založba, 1998.
 Igor Karlovšek: *V objemu lože*. DZS, 1997.
 Nina Kokelj: *Milovanje*. Študentska založba, 1998.
 Aaron Kronski: *Lutke*. Karantanija, 2004.
 Dušan Merc: *Slepi potnik*. ,1999.
 Vinko Möderndorfer: *Omejen rok trajanja*. Cankarjeva založba, 2003.
 Vinko Möderndorfer: *Ljubezni Sinjebradca*. DZS, 2005.
 Andrej Morovič: *Seks, ljubezen in to*. Študentska založba, 2006.
 Edo Rodošek: *Skoraj enaki*. Karantanija, 2003.
 Franček Rudolf: *Pošlji deklico*. Karantanija, 2006.
 Peter Semolič: *Krogi na vodi*. Študentska založba, 2000.
 Ivan Sivec: *Julija iz Sonetnega venca*. Karantanija, 2006.
 Andrej Skubic: *Grenki med*. DZS, 1999.
 Tomo Snoj: *Koledar Evinih razpoloženj*. Tuma, 2005.
 David Šalamun: *New York v štirih letnih časih*. Založba / *cf., 2005.

- Dušan Šarotar: *Potapljanje na dah*. Študentska založba, 1999.
Jani Virk: *Pogled na Tycho Brahe*. Študentska založba, 1998 .
Lenart Zajc: *Hevimetal*. Tuma, 2004.
Pavle Zidar: Čebele iz Bachov. Društvo 2000, 1996.
Darko Žlebnik: *Prešustvo*. Študentska založba, 1999.

Vzporedni korpus SPOOK: označevanje, zapis in iskanje terminoloških virov

Tomaž Erjavec

Odsek za tehnologije znanja, Institut »Jožef Stefan«
tomaz.erjavec@ijs.si

Abstract

The paper discusses the technical aspects of the compilation and use of the SPOOK parallel corpus, in particular the morphosyntactic tagging and lemmatisation, the encoding of the corpus, and the concordancer developed for searching the corpus. The morphosyntactic tagging of the Slovene texts was performed using previously developed methods and the tagset elaborated in the JOS and MULTEXT-East projects. For tagging the English, French, Italian and German parts of the corpus we used the program TreeTagger, which, however, uses incompatible tagsets for the languages. For SPOOK we developed mappings between the TreeTagger tagsets and the multilingual MULTEXT(-East) recommendations, thus harmonising the tagsets across all the included languages, which enables easier use of the corpus. The computer encoding of the corpus is based on the XML standard. In the process of the compilation of the corpus we used a simple set of Slovene language elements, while the final version follows the Text Encoding Initiative Guidelines, TEI P5. The Web-based concordancer CUWI was also developed for SPOOK, which supports searching the text and annotations through regular expressions, displaying of concordances and frequency lexicons, multilingual searching and display, and downloading results in several formats.

Keywords: morphosyntactic tagging, corpus encoding, concordancer

Ključne besede: oblikoskladenjsko označevanje, kodiranje korpusov, konkordančnik

1 UVOD

Izdelava vzporednega korpusa ima več korakov, začeni z digitalizacijo besedil, poravnavo izvirnega besedila in prevoda po stavkih do zapisa metapodatkov za posamezna besedila – v prispevku teh korakov ne obravnavamo, temveč se osredotočimo na jezikoslovno označevanje ter na računalniški zapis korpusa. Te informacije niso zanimive samo s stališča priprave korpusa, temveč tudi s stališča njegove uporabe, saj so jezikoslovne lastnosti, pripisane posameznim besedam v korpusu, lahko koristne pri iskanju po korpusu ter pri luščenju raznih kvantitativnih lastnosti besedil. Korpus, označen z metapodatki o posameznih besedilih in jezikoslovnimi lastnostmi posameznih besed, je dostopen preko konkordančnika, ki je tudi predstavljen v prispevku. Konkordančnik je bil razvit za namene prevodoslovnih raziskav in omogoča raznovrstna iskanja po jezikoslovno označenih vzporednih korpusih.

2 OBLIKOSKLADENJSKO OZNAČEVANJE IN LEMATIZACIJA

2.1 Postopek označevanja

Oblikoskladenjsko označevanje je postopek, pri kateri vsaki besedni pojavnici v besedilu glede na sobesedilo pripišemo ustrezno oblikoskladenjsko oznako. Avtomatsko označevanje tipično poteka s statistično zasnovanimi in jezikovno neodvisnimi programi, ki se modela jezika naučijo na ročno označenem korpusu. Modeli so večinoma sestavljeni iz dveh delov:

- Leksikon besednim oblikam pripiše množice oblikoskladenjskih oznak, npr. »čebule → *Sozer:8 + Sozmt:1*, torej 8x samostalnik, občno_ime, ženski spol, ednina, rodilnik in 1x v tožilniku množine.
- Kontekstne značilke, s pomočjo katerih iz množice leksikalnih oznak program izbere pravo glede na kontekst besede. Tipične značilke so kar n-terčki oznak s pripadajočo frekvenco iz učnega korpusa, npr. *Zk-mer Ppnmer Somer:46*, primer niza je »tega mestnega orkestra« ali »onega narodnega veljaka«

Za označevanje slovenskega dela korpusa smo uporabili orodje ToTaLe (Erjavec et al., 2005), ki besedilo najprej tokenizira, torej razdeli na pojavnice (besede in ločila), nato tagira (oblikoskladenjsko označi) in na koncu lematizira (besednim pojavnici pripiše osnovno obliko). Program se je modela jezika (leksikona in kontekstnih značilk označevalnika, kot tudi modela lematizacije) naučil na korpusu jos1M (Erjavec in Krek 2008), ki vsebuje milijon besednih pojavnici, označenih z oblikoskladenjsko oznako in lemo. Program je uporabil tudi pomožni

leksikon, izluščen iz korpusa FidaPLUS (Arhar in Gorjanc, 2007) – ta je sicer zelo velik, vendar vsebuje napake avtomatskega označevanja in je zato manj zanesljiv kot leksikon iz jos1M. Nad korpusom SPOOK nismo naredili evalvacije točnosti označevanja, je pa bila ta narejena v okviru projekta JOS (Erjavec in Krek 2008), kjer je oblikoskladenjsko označevanje doseglo točnost 86,6 %, pri čemer je bila evalvacija narejena z učenjem in desetkratnim križnim preverjanjem neposredno na korpusu jos100k. Ker je po eni strani učni korpus in pomožni leksikon modela za označevanje SPOOK bistveno večji, kot je bil pri evalvaciji točnosti z jos100k, po drugi pa besedilni tip besedil drugačen kot v jos100k, ni jasno, koliko se točnost označevanja slovenskih besedil v SPOOK razlikuje od te vrednosti.

Za označevanje angleškega, francoskega, italijanskega in nemškega dela korpusa smo uporabili označevalnik TreeTagger (Schmid 1994; TreeTagger – a language independent part-of-speech tagger), ki je prosto dostopen in ima prednost, da je na voljo skupaj z modeli za vse te jezike. To zelo olajša označevanje, saj se izognemo dolgotrajnemu in kompleksnemu postopku zagotovitve ročno označenih korpusov za vsakega od jezikov in šolanju označevalnika. Tudi za TreeTagger nismo naredili evalvacije točnosti označevanja, ta pa tudi sicer, kolikor nam je znano, ne obstaja za vse jezike, z izjemo angleščine (Schmid 1994), kjer je ocenjena na 96,3 %, in nemščine (Schmid 1995), kjer je ocenjena na 97,5 %. Te številke so sicer mnogo boljše kot za slovenski jezik, vendar so tudi nabori oblikoskladenjskih oznak bistveno manjši kot za slovenščino. Če ima slovenščina v sistemu JOS preko 1,900 različnih oznak, jih ima od drugih jezikov v SPOOK največ nemščina (52), najmanj pa francoščina (30). Razlogi za razlike so deloma jezikoslovni, saj so slovanski jeziki bolj pregibni od zahodnoevropskih, izvirajo pa tudi iz pragmatičnih ali z lokalno jezikoslovno tradicijo povezanih odločitev avtorjev posameznih naborov.

Tako ToTaLe kot TreeTagger poleg oblikoskladenjskega označevanja besedilo tudi lematizirata, vsaki besedni obliki v besedilu torej pripišeta njeno osnovno obliko. Lematizacija za slovenščino prav tako temelji na modelu, naučenem na (leksikonu) jos1M, pri čemer je program sposoben razmeroma natančno lematizirati tudi neznane besede. Za razliko od ToTaLe TreeTagger podpira samo lematizacijo znanih besed, vendar pa za lematizacijo zahodnoevropskih jezikov že to ne daje slabih rezultatov, saj so bili učeni na velikih korpusih, in gre za jezike, ki so za razliko od slovenščine pregibno precej bolj enostavni.

2.2 Oblikoskladenjska priporočila SPOOK

Modeli TreeTagger za tuje jezike korpusa SPOOK so bili naučeni iz večjih dostopnih ročno označenih korpusov za posamezne jezike, vendar pa ti korpusi uporabljajo med seboj neodvisne nabore oznak. Tako je npr. oznaka za lastno ime

za angleščino *NP*, za nemščino *NE*, za francoščino *NAM* in za italijanščino *NPR*. Takšno neskladje seveda otežuje uporabo korpusa kot celote, zato smo poskusili uskladiti nabore oznak (in oblikoskladenjskih lastnosti) za vse jezike korpusa SPOOK, tako da smo naredili preslikave izvirnih TreeTagger (in ToTaLe) oznak v analitični sistem narejen po metodologiji MULTEXT.

2.2.1 Oblikoskladenjska priporočila MULTEXT-East in JOS

Za slovenski jezik uporabljamo priporočila za oblikoslovno označevanje JOS (Erjavec in Krek 2008; Oblikoskladenjske specifikacije JOS V1.1), oz. kar ustreza slovenskemu delu večjezičnih oblikoskladenjskih priporočil MULTEXT-East (Erjavec (ur.) 2010a; Erjavec 2012a). Osnovni princip teh priporočil je, da za vsak jezik definirajo nabor oblikoskladenjskih oznak (npr. ena od oznak za slovenščino je *Ncfs̑g*), definirajo pa tudi sisteme lastnosti in enostavno preslikavo oznak v te sisteme, npr. *Ncfs̑g* $\equiv$ *Noun Type=common Gender=feminine Number=singular Case=genitive*. Priporočila podpirajo tudi prevajanje oz. lokalizacijo oznak in lastnosti v druge jezike, npr. *Ncfs̑g* $\equiv$ *Sozer* $\equiv$ *samostalnȋ vrsta=občno_ime spol=ženski število=ednina sklon=rodilnik*.

Priporočila MULTEXT-East pokrivajo 16 jezikov, med njimi veliko večino slovanskih, ne vsebujejo pa zahodnoevropskih jezikov. Oblikoskladenjski nabori oznak za te so bili sicer definirani v okviru prvotnega projekta MULTEXT (Ide in Veronis 1994), vendar bi težko dobili ročno označene korpusne za vse jezike SPOOK, ki bi uporabljali nabore MULTEXT.

Da bi uskladili nabore oznak, smo po vzoru priporočil MULTEXT-East naredili petjezična priporočila SPOOK, ki so v slovenskem delu (skoraj) enaka kot JOS, v priporočilih za tujejezično označevanja pa uporabljajo metodologijo večjezičnih definicij MULTEXT(-East), pri čemer imajo njihove oznake enostavno, 1-1 preslikavo z izvirnimi oznakami programa TreeTagger.

Priporočila imajo razmeroma enostavno strukturo. Najprej definirajo *kategorije* oz. besedne vrste, pri čemer so te »besedne vrste« ponekod tudi tehnične narave, npr. *okrajšava* ali *neuvrščeno*. Angleška in slovenska imena ter enočrkovne kode za 12 definiranih kategorij MULTEXT podaja tabela 1. V tabeli je tudi število atributov za posamezno kategorijo. Kategorijam namreč priporočila pripišejo nabor lastnosti, tj. atributov in naborov njihovih vrednosti. Te, t. i. skupne tabele, pokrivajo vse jezike in pri vsaki kombinaciji atributa in vrednosti za kategorijo tudi določijo, za katere jezike se uporablja. Primer take tabele, vzet neposredno s spleta, je podan v sliki 1.

Tabela 1: Kategorije MULTEXT-East oz. SPOOK

Kategorija (en)	Koda (en)	Kategorija (sl)	Koda (sl)	Atributov	Definirano za jezike				
Noun	N	samostalni	S	5	sl	en	de	fr	it
Verb	V	glagol	G	10	sl	en	de	fr	it
Adjective	A	pridevnik	P	7	sl	en	de	fr	it
Pronoun	P	zaimék	Z	11	sl	en	de	fr	it
Determiner	D	določilnik	I	1		en			it
Article	T	člen	C	1			de	fr	it
Adverb	R	prislov	R	2	sl	en	de	fr	it
Adposition	S	predlog	D	4	sl	en	de	fr	it
Conjunction	C	veznik	V	2	sl	en	de	fr	it
Numeral	M	števník	K	6	sl	en	de	fr	it
Particle	Q	členek	L	2	sl	en	de		it
Interjection	I	medmet	M	0	sl	en	de	fr	
Abbreviation	Y	okrajšava	O	1	sl	en	de	fr	
Residual	X	neuvrščeno	N	1	sl	en	de		it

Atributi in vrednosti - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Projekt SPOOK Open Repositories 2012 Atributi in vrednosti

Google Koledar products - elga - Picasa DOGHOUSE IMPACT@ISI Europa's Overview The Action ISO Standards Develop... IntelliText Corpus Que... Sketch Engine

2.3.8. Predlog

Tabela atributov in vrednosti za predlog

P	Atribut (en)	Vrednost (en)	Koda (en)	Atribut (sl)	Vrednost (sl)	Koda (sl)	slovenščina	angleščina	nemščina	francoščina	italijanščina
0	CATEGORY	Adposition	S	besedna vrsta	predlog	D	sl	en	de	fr	it
1	Case	nominative	n	sklon	imenovalnik	i	sl				
		genitive	g		rodilnik	r	sl				
		dative	d		dajalnik	d	sl				
		accusative	a		tožilnik	t	sl				
		locative	l		mestnik	m	sl				
		instrumental	i		orodnik	o	sl				
2	Type	preposition	p	vrsta	predlog	p		en	de		it
		possessive_endings	s		svojilna končnica	s		en			
		to	o		to	t		en			
		postposition	t		postpozicija	t			de		
		right_part	r		desni del	d			de		
3	Binding	article	a	vezava	člen	l			de		
4	Formation	simple	s	sestavljenost	enostaven	e				fr	
		compound	c		prিপসকি	p				fr	
		article	t		člen	c					it

Slika 1: Definicija za kategorijo »predlog« v skupni tabeli oblikoskladenjskih priporočil

Vsak jezik ima nato v priporočilih še svoj razdelek, kjer so zapisane podobne tabele, kot so skupne, vendar samo za kategorije in lastnosti, ki so za ta jezik definirane. Razdelki za posamezne jezike pa vsebujejo tudi nabore oblikoskladenjskih oznak, ki so na enostaven način sestavljene iz kode kategorije in kod za posamezne vrednosti naštetih atributov – te oznake se, ker so kratke, uporablja za označevanje korpusov, čeprav so formalno enake sistemom lastnosti. Primer seznama definiranih oznak za italijanščino je dan v sliki 2, kjer je vsaka oznaka podana dvojezično, pripisani pa so ji tudi primeri uporabe, pri čemer so ti vzeti kar iz označenega korpusa.

Priporočila SPOOK za francoski jezik - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Projekt SPOOK Open Repositories 2012 Priporočila SPOOK za francoski jezik

nljjs.si/spook/msd/html-sl/msd-fr.html#msd-fr

Google Koledar products - olga - Picas... DOGHOUSE IMPACT@ISI Europress > Overview The Action ISO Standards Develop... IntelliText Corpus Que...

subjunctive	s	VForm	Verb	subjunktiv	u	oblika	glagol
-------------	---	-------	------	------------	---	--------	--------

3.4.16. Seznam oblikoskladenjskih oznak za francoščino

V tabeli so našteje oblikoskladenjske oznake s svojimi oblikoskladenjskimi lastnostmi, preslikava iz nabora oznak, ki jih za ta jezik uporablja označevalnik *TreeTagger* in primeri najbolj pogostih uprab posameznih oznak iz korpusa SPOOK, ki je bil avtomatsko označen s *TreeTaggerjem*. *TreeTaggerjevi* oznake je v tabeli pripisan kratek komentar iz priročnika, kopiran od [tu](#).

Tabela oblikoskladenjskih oznak za francoščino

MSD (sl)	Lastnosti (sl)	MSD (en)	Lastnosti (en)	TreeTagger Opis	Primeri
So	samostalnik vrsta=občno_ime	Nc	Noun Type=common	NOM noun	m./monsieur, pays, temps, fois/fois, vie, monde, ans/an, années/année, homme
Sl	samostalnik vrsta=lastno_ime	Np	Noun Type=proper	NAM proper name	Etat, France, Etats-Unis, Europe, Charlotte, Luo, Paris, Valérie, Cour
G-i	glagol čas=imperfekt	V-i	Verb Tense=imperfect	VER.impf verb imperfect	était/être, avait/avoir, avais/avoir, étaient/être, avaient/avoir, étais/être, faisait/faire, pouvait/pouvoir, allait/aller
Gps	glagol oblika=povednik čas=sedanjik	Vip	Verb VForm=indicative Tense=present	VER.pres verb present	est/être, a/avoir, sont/être, ont/avoir, ai/avoir, peut/pouvoir, dit/dire, suis/suivre/être, fait/faire sera/être, aura/avoir, seront/être,

Slika 2: Začetek tabele oblikoskladenjskih oznak SPOOK za francoščino.

Pri preslikavi oznak *TreeTagger* v sistem *MULTTEXT-East* največ problemov povzroča razlika med analitičnimi oznakami slednjega in pogosto sintetičnimi oznakami prvega. Tako ima npr. angleščina oznako *VBZ* za *Verb, 3rd person singular present*, kar je preslikano v *G-iste* oz. *glagol vrsta=nedoločeno oblika=povednik čas=sedanjik oseba=tretja število=ednina*. To sicer ni problematično, vendar ima angleščina tudi oznako *VBP* oz. *Verb, non-3rd person singular present*, za katero je dosti manj jasno, kako jo ustrezno preslikati. Ker smo želeli ohraniti bijektivno (1:1) preslikavo med obema sistemoma oznak, smo se v tem primeru odločili, da osebi in številu pustimo nedoločeno vrednost, torej *glagol oblika=povednik čas=sedanjik* torej *G-is--* oz. kar *G-is*.

Problematične so tudi oznake *TreeTagger*, ki so lahko pripisane posamezni besedi, npr. oznaka *CHE* za italijansko besedico *che*. Take oznake se uporabljajo za zelo

dvoumne funkcijske besede, ki jih oblikoskladenjski označevalniki ne zmorejo razdvoumiti, z besedi lastno oznako pa se temu problemu enostavno izognejo. Za zapis takšnih primerov v priporočilih SPOOK besedi pripišemo (idealno zanjo najbolj prototipično) besedno vrsto, nato pa jo kvalificiramo s posebno lastnostjo. Tako se *CHE* preslika v *členek vrsta=che* oz. *Lh*.

2.2.2 Oblikoskladenjska priporočila IMP

Kot rečeno, so slovenske oblikoskladenjske oznake oz. lastnosti že v sistemu JOS oz. MULTEXT-East. Priporočila SPOOK vsebujejo več ali manj kopijo le-teh, vendar z manjšo razliko. Ker imajo ostali jeziki dva sistema oznak, pri čemer jih ločijo 30–50, smo tudi v slovenska priporočila (in korpus) dodali preslikavo v manjši nabor oznak.

Ta nabor smo izdelali za označevanje korpusov starejšega slovenskega jezika IMP (Erjavec 2012b; Jezikovni viri starejšega slovenskega jezika IMP), kjer je bil poudarek na označevanju besed s posodobljenimi oblikami, vseeno pa smo želeli v korpusu imeti tudi ročno preverjene oznake vsaj za besedno vrsto in manjši nabor leksikalnih lastnosti. Zato smo izdelali priporočila IMP, ki so enaka kot JOS, samo da izpustijo vse pregibne lastnosti, s čimer dobimo nabor 32 oznak, kar se tudi lepo sklada s številom oznak ostalih jezikov.

Preslikava oznak JOS v nabor IMP je zelo enostavna, saj oznake JOS samo okrajšamo, s čimer izgubimo pregibne lastnosti besed, ohranimo pa besedno vrsto in (nekaj) inherentnih lastnosti. Tako ima npr. nabor IMP namesto 279 oznak za pridevnike, kolikor jih ima JOS, samo 5 oznak, kjer je določena vrsta (*splošni, svojilni, deležniški*) in pa, za splošne pridevnike, stopnja (*nedoločeno, primernik, presežnik*). Slovenščina tako v SPOOK-u vsebuje preslikavo med naborom oznak JOS in naborom IMP. Kot drugi jeziki torej uporablja dva nabora oznak, čeprav je njihov status drugačen kot za druge jezike.

2.2.3 Dostopnost priporočil

Priporočila SPOOK so na voljo na enak način kakor MULTEXT-East, JOS in IMP, pri katerih smo se trudili, da bi bila čim bolj dostopna. Dostopnost ima več dimenzij (Erjavec 2009), ki so pri oblikoskladenjskih priporočilih SPOOK udeležene na sledeče načine:

- za izmenljivost, trajnost in možnost uporabe v raznovrstne namene so kodirana po mednarodnih standardih in priporočilih;

- dostopna so na spletu tako v obliki XML kot v pretvorbi v HTML, torej stavljena s kazalom in navigacijo, v obliki, primerni za branje oz. kot spletna referenca, saj imajo vsa poglavja fiksne URL-je (npr. <http://nl.ijs.si/spook/msd/html-sl/msd-de.html> msd.P-de kaže na tabelo za nemški zaimek);
- nabori oblikoskladenjskih oznak in lastnosti SPOOK so definirani tako v slovenskem kot v angleškem jeziku, v obeh je napisano tudi tako prosto besedilo priporočil kot tudi stavljenje v HTML (npr. »Kazalo« in »Table of contents«), s čimer je omogočena uporaba slovenščine za naše jezikoslovce, obenem pa razumevanje oznak za tuje raziskovalce;
- avtorskoppravno so priporočila dostopna pod licenco »Creative Commons Priznanje avtorstva-Deljenje pod enakimi pogoji« (Creative Commons licence Attribution-ShareAlike 3.0), kar pomeni, da jih je dovoljeno kopirati, distribuirati in prenašati ter tudi spreminjati in uporabljati v komercialne namene pod pogojem, da se navede izvirnega avtorja in, če se jih spreminja, da se jih deli pod enakimi pogoji.

3 RAČUNALNIŠKI ZAPIS IN OBSEG

3.1 Zapis XML

Za računalniški zapis korpusa in oblikoskladenjskih priporočil smo uporabili standard W3C XML (Extensible Markup Language (XML)), ki je prevladujoči način zapisa (tudi) digitalnih besedil. V XML lahko definiramo nabore (besedišča) oznak in dovoljena medsebojna razmerja, te nabore pa potem uporabljamo za označevanje različnih zvrsti besedil. XML ima prednost, da je po eni strani razmeroma berljiv, po drugi pa primeren za strojno obdelavo. Definicije naborov za opis tipov dokumentov so izražene v shemah XML, kar omogoča strojno preverljivost pravilnosti oznak v dokumentih, ki je zelo pomembna pri izdelavi jezikovnih virov, ker zagotavlja, da v dokumentih ni tehničnih napak.

Zapis XML je tudi dobro podprt z orodji, oz. pridruženimi standardi. Eden bolj pomembnih je XSLT (XSL Transformations (XSLT) Version 1.0), ki definira jezik za transformacije med dokumenti XML. V projektu SPOOK smo XSLT skripta uporabili za pretvorbo delovnega formata korpusa v kanoničnega (opis v naslednjem razdelku) in nato za pretvorbo kanoničnega formata v izvedene oblike, primerne za konkordančnik ali statistične analize. Tudi oblikoskladenjska priporočila so zapisana v XML, nato pa z XSLT pretvorjena v HTML za branje, ali v razne tabelarične formate, ki podajo npr. preslikavo oblikoskladenjskih oznak in lastnosti, prevedejo iz slovenščine v angleščino, podajo sortirno zaporedje oznak itd.

V procesu izdelave korpusa smo uporabili dva zapisa v XML. V delovnem formatu je bil zapisan končni rezultat prve faze obdelave korpusnih enot, ki vsebujejo korigirano besedilo, so opremljene z bibliografskimi podatki in so razdeljene na prevodne segmente, ki so poravnani med izvirnikom in prevodom. Format oz. shema XML, ki podpira ta zapis, je bila namensko napisana in uporablja slovenska imena elementov, obenem pa je tudi minimalna po številu uporabljenih elementov in torej primerna za manjše ročne korekcije.

3.2 Nabor znakov Unikod

Standard XML je osnovan na uporabi univerzalnega kodnega nabora Unikod (The Unicode Consortium), ki je tudi uporabljen v korpusu. Vendar pa je korpus v procesu izdelave moral iti preko več programov, ki lahko tako ali drugače pokvarijo kodiranje znakov. Največ problemov je povzročil ravno TreeTagger, saj ta za več jezikov predpostavlja, da bo format besedila v starem, 8-bitnem kodiranju ISO 8896-1 oz. »Latin-1«, ki pokriva zahodnoevropske jezike. Če se v takem besedilu pojavijo ne-zahodnoevropske črke ali posebna ločila, teh ni mogoče zapisati v Latin-1. Zato je bilo treba napisati programe, ki besedilo v XML najprej glede na jezik pretvorijo v format in kodiranje za TreeTagger, nato pa združijo TreeTaggerjev izhod z izvirnim besedilom, da ohranijo izvirni in s tem pravilni zapis znakov.

Tudi v delovnem zapisu XML so se še pojavljale napake kodiranja, ki smo jih poskusili popraviti z namenskim programom. Najprej smo naredili profil znakov za celoten korpus in po datotekah, nato pa pogledali, kateri znaki predstavljajo tipične napake, ki jih je mogoče popraviti, in katerih znakov ni mogoče enostavno interpretirati v naboru Unikod – prve smo s programom popravili v ustrezne znake Unikod, slednje pa nadomestili z znakom Unikod U+FFFD z imenom »REPLACEMENT CHARACTER«, ki opozarja, da je bil na tem mestu znak, ki ga v Unikodu nismo mogli predstaviti. Takšnih znakov je malo, v celotnem korpusu jih je samo 114, pojavljajo pa se v 10 dokumentih.

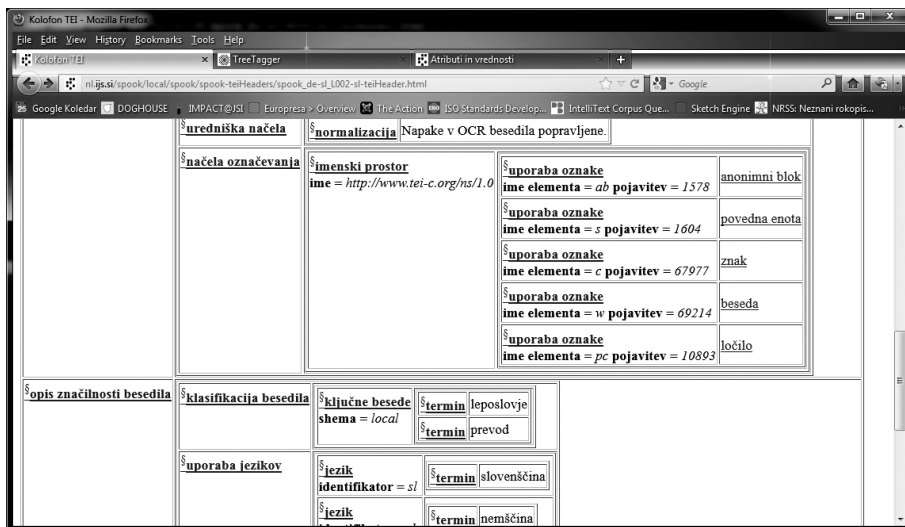
Korpus ima vsega skupaj 52 milijonov znakov besedila vključno s presledki, uporablja pa 270 različnih znakov, tudi npr. črke cirilične in drugih abeced, razne vrste navednic in vezajev ter matematične simbole.

3.3 SPOOK v zapisu TEI

V drugi fazi smo besedila, kot je bilo opisano v prejšnjem poglavju, avtomatsko jezikoslovno označili, obenem pa pretvorili format v takega, ki je skladen s priporočili

za zapis besedil TEI P5 (TEI P5: Guidelines for Electronic Text Encoding and Interchange), torej z de-facto standardom v digitalni humanistiki. Priporočila določajo zapis XML za razne vrste jezikovnega gradiva, ki se uporablja v znanstvene namene, kot so npr. slovarji, tekstnokritične izdaje, opisi rokopisov, itd., pa tudi za jezikoslovno označene korpuse in strukture lastnosti. V projektu SPOOK uporabljamo TEI tako za zapis korpusa kot tudi oblikoskladenjskih priporočil.

V korpusu SPOOK je vsako besedilo zapisano kot dokument TEI, ki se začne s kolofonom, nato pa vsebuje besedilo. TEI v svojem kolofonu ponuja bogat nabor meta-podatkov, od zapisa raznih vrst odgovornosti za besedilo in kodiranje, opis vira, uredniška načela, itd. V sliki 3 je kot primer dan del kolofona enega od besedil – kolofon tu ni neposredno zapisan v TEI/XML, pač pa v spletni obliki, ki je avtomatsko izvedena iz XML in tudi opisno prevede sicer angleška imena elementov v slovenske termine (Erjavec 2010b). Prikazani del kolofona vsebuje kratek opis uredniških načel normalizacije besedila, opis uporabe oznak v besedilu, klasifikacijo besedila kot prevod leposlovnega dela in definicije kod za jezike. Slednje so zanimive s stališča medsebojnega navezovanja standardov: TEI določa, da morajo biti kode jezikov zapisane po standardu ISO 639, s čimer že pravilna uporaba priporočil TEI zagotovi, da bodo tudi kode jezikov (in npr. zapis datumov, URLjev, itd.) sledili standardom in priporočilom, kar bistveno olajša strojno procesiranje takšnih jezikovnih virov.



Slika 3: Del kolofona TEI za eno od besedil korpusa SPOOK.

Strukturo besedila ilustrira slika 4. Besedilo dokumenta je sestavljeno iz prevodnih ustreznice, (*ab*) ki so povezane (preko *@xml:id* in *@corresp*) z istoležnimi ustreznici iz vzporednega besedila. Prevodne ustreznice so nato sestavljene iz

stavkov oz. povedi (*s*), pri čemer so bili stavki avtomatsko označeni (tako Tree-Tagger kot ToTaLe naredita vzporedno s tokenizacijo še segmentacijo na stavke), zato segmentacija ni vedno pravilna.

```
<text xml:id="spook_fr-sl_N064-fr:text" xml:lang="fr"
  corresp="spook_fr-sl_N064-sl.xml spook_fr-sl_N064-sl.text">
  <body>
    <ab n="1" xml:id="seg.1" corresp="spook_fr-sl_N064-sl.xml seg.1">
      <s>
        <w lemma="le" ana="T" ctag="DET:ART">Le</w>
        <c> </c>
        <w lemma="takfrisime" ana="Nc" ctag="NOM">takfrisime</w>
        <pc ctag=",">,</pc>
        <c> </c>
        <w lemma="un" ana="T" ctag="DET:ART">une</w>
        <c> </c>
        <w lemma="idéologie" ana="Nc" ctag="NOM">idéologie</w>
        <c> </c>
        <w lemma="messianique" ana="A" ctag="ADJ">messianique</w>
      </s>
    </ab>
```

Slika 4: Poravnano in jezikoslovno označeno besedilo v zapisu TEI.

Stavki so sestavljeni iz besed, ločil in presledkov. Besede (*w*) nosijo tri attribute, in sicer lemo (*@lemma*), oblikoskladenjsko analizo (*@ana* = oznaka SPOOK) in korpusno oznako (*@ctag* = oznaka TreeTagger). Za slovenska besedila pripišemo pregibne oznake JOS h korpusni oznaki, saj je ToTaLe z njimi označil korpus, leksikalne oznake IMP pa oblikoskladenjski analizi, ker so podobne analizam za ostale jezike.

Ločila (*pc*) seveda nimajo oblikoskladenjskih oznak ali lem, so pa označena s korpusno oznako (*@ctag*), pri čemer so te poenotene med jeziki in so kar enake kot samo ločilo.

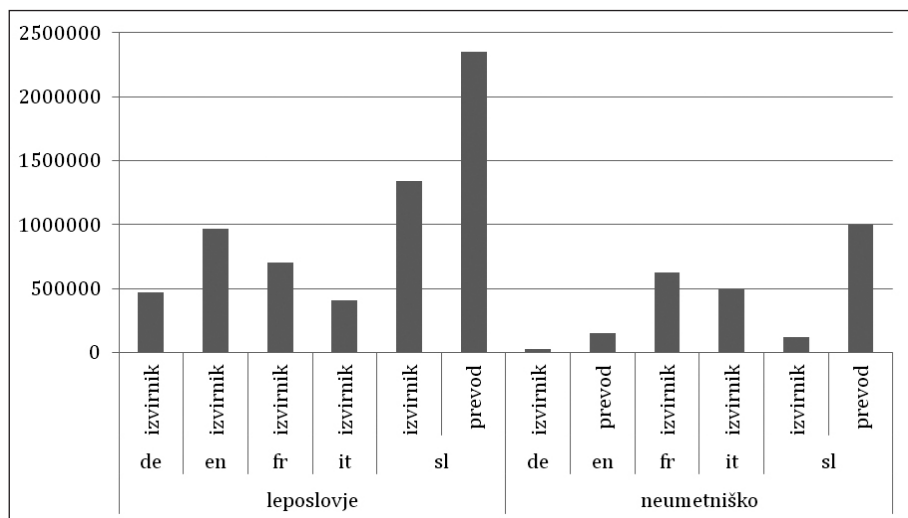
Vključevanje označenih presledkov je v računalniških korpusih prej posebnost kot pravilo, saj so pri jezikoslovnih obravnavah nepomembni. Vendar je, če drugega ne, vsaj lepo, če obdržimo stičnost pojavnic in s tem možnost pravilnega stavljenja prevodnih segmentov. Zato so v korpusu SPOOK označeni tudi presledki, in sicer kot presledek v elementu za posamezen znak (*c*, zapisano torej kot *<c> </c>*). Element za znak v korpusu SPOOK torej vsebuje vedno in samo presledek.

TEI služi kot kanonični format korpusa, pa tudi oblikoskladenjskih priporočil – ta zapis je dobro dokumentiran, izmenljiv in odporen na tehnološke spremembe. Za uporabo v aplikacijah je format v XML najprej potrebno transformirati v ciljni zapis, kar dosežemo s skriptami XSLT. Primer uporabe teh skript smo že videli na primeru slik spletne predstavitve oblikoskladenjskih priporočil in kolofona TEI. Podoben postopek uporabimo tudi pri konverziji v podatkovno bazo za konkordančnik; več o tem v naslednjem poglavju.

3.4 Velikost korpusa

Ko so besedila označena z metapodatki (npr. uporabljene oznake in njihovo število), ni težko prikazati medsebojnih razmerij med posameznimi jeziki oz. zvrstmi besedil. V tabeli 2 so podane porazdelitve nekaterih metapodatkov iz korpusa SPOOK, kjer so razmerja izražena s številom besed. Kot vidimo, ima SPOOK precej več leposlovnih kot neumetniških besedil, pri čemer so bila tu za namene projekta zbrana besedila s tujejezičnim izvirnikom in prevodom v slovenščino. Izvirna slovenska leposlovna besedila niso bila posebej zbrana za projekt SPOOK, temveč smo vključili romane iz korpusa Gigafida.

Tabela 2: Razmerja med zvrstmi besedil v korpusu glede na število besed.



Kar se velikosti datotek tiče, je posledica berljivosti XML in univerzalnosti TEI precejšnja potratnost s prostorom – celoten označeni korpus, ki vsebuje 8 milijonov besed, je velik več kot 800 MB. Vendar je prostor na disku poceni, datoteke XML pa se tudi učinkovito komprimirajo; stisnjen v formatu ZIP ima SPOOK tako samo 70MB.

4 KONKORDANČNIK CUWI

Za korpus SPOOK je bil razvit konkordančnik po imenu CUWI (Corpus Users' Web Interface), ki omogoča iskanje po besedilih v korpusih in raznovrstne prikaze rezultatov. Konkordančnik podpira iskanje tako po besedah oz. besednih zvezah kot po jezikoslovnih oznakah, podpira pa tudi omejitve iskanj glede na metapodatke besedil.

4.1 Implementacija konkordančnika

Konkordančnik je sestavljen iz zalednega dela, t.j. programske opreme, ki dejansko izvaja iskanje po korpusu, in čelnega dela sistema, torej spletnega vmesnika. Kot zaledni del uporabljamo CWB (Christ 1994; CWB: The IMS Open Corpus Workbench), ki je visokozmogljiv sistem v uporabi za številne korpusne tako pri nas kot po svetu. Za zasnovno iskanja v CWB je ilustrativno pogledati format datotek, ki jih sistem uporablja za posamezen korpus. En korpus je v CWB ena datoteka, ki jo iz kanoničnega formata TEI avtomatsko pretvorimo s skripto XSLT.

Datoteka CWB vsebuje strukturne elemente in označene pojavnice. Strukturni elementi so v resnici kar elementi XML z atributi, pojavnice pa so besede ali ločila, zapisane skupaj z oznakami v tabelarnem formatu, kot ilustrira slika 5.

```
<text id="FSL001" lang="fr" title="Fou de Vincent"
  title-sl="Nor na Vincenta" author="Hervé Guibert"
  translator="Brane Mozetič" type="leposlovje">
<seg id="FSL001.1" corresp="spook_fraslv-slv.cqp FSL001.1">
<s>
Dans      dans      dans      Ss      De      PRP
la         la         le        T        T        DET:ART
nuit       nuit       nuit      Nc      So      NOM
du         du         du        Sc      Dp      PRP:det
25         25         25        M        K        NUM
au         au         au        Sc      Dp      PRP:det
26         26         26        M        K        NUM
novembre   novembre   novembre  Nc      So      NOM
,           ,           ,         ,        ,        PUN
```

Slika 5: Format zapisa CWB.

CWB omogoča izpis metapodatkov (atributov elementa *text*) posameznih zadetkov v korpusu, pa tudi omejitev iskanja na tista besedila, kjer so metapodatki enaki kot zahtevana omejitev (npr. iskanje samo po prevedenih besedilih). Kot vidimo, je CWB korpus v zasnovi enojezičen, povezava z vzporednim korpusom pa je vzpostavljena na ravni segmentov. Tako v CWB vedno iščemo primarno po enem korpusu, lahko pa prikažemo tudi poravnani segment. Pojavnice in njihove oznake imajo v CWB enakovreden status, tako da je mogoče tako iskanje kot prikaz pojavnic oz. njihovih oznak, ali pa poljubne kombinacije obojih. Kar se tiče formata izpisa, velja še opomba, da CWB ne ohrani stičnosti pojavnic.

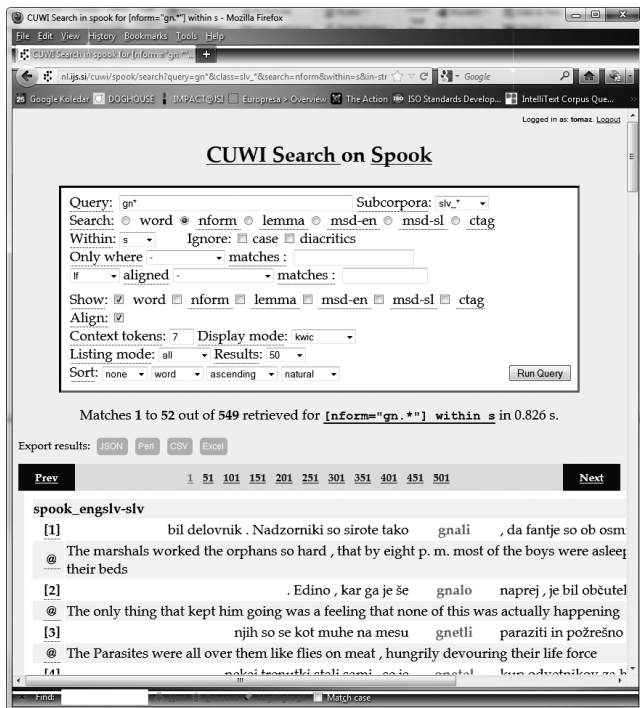
Spletni vmesnik CUWI smo razvili posebej za korpus SPOOK. CUWI je zasnovan na platformi Mojolicious (Mojolicious – Perl real-time web framework; McDaniel 2011, poglavje »Websockets«) za gradnjo spletnih aplikacij in je sestavljen iz zalednega dela CWB::Model, ki omogoča programski dostop (API) do korpusov v formatu CWB, in spletnega vmesnika CWB::CUWI, ki omogoča spletno poizvedovanje in prikaz. Platforma Mojolicious je napisana za jezik Perl, ki je tudi standardni skriptni vmesnik za korpuse CWB, zasnovana pa je po vzorcu reaktorskega demultiplekserja dogodkov (*reactor demultiplexing event loop*, Schmidt 1995), tako da lahko zelo majhno število spletnih servisov hkrati opravlja poizvedbe za več uporabnikov in predstavi rezultate, ko jih zaledni del CWB pripravi. Ta arhitektura omogoča veliko mero fleksibilnosti in zelo dober izkoristek sodobnih računalniških arhitektur z več procesorskimi jedri. Ker so korpusi SPOOK relativno kompleksni, je tako nastal konkordančnik, ki je splošno uporaben in ga je zelo lahko prilagoditi tudi za uporabo z drugimi korpusi in v drugačnih jezikovnih okoljih. Fleksibilna platforma, zgrajena po sodobnem modelu *Model-View-Controller* (MVC, Greer 2007; Burbeck 1992) omogoča tudi preprosto prilagajanje z dodatnimi predlogami za spletne strani (»templates«) in programskim vtičniki (»plugins«). Zaledni del CWB::Model prinaša dodatno možnost, da lahko uporabnik pripravi poizvedbe, ki bodo tekle samodejno, brez spletnega vmesnika, in jih z njegovo pomočjo strojno obdelala z lastnim programom. CUWI je prosto dostopen pod dvojno licenco *Perl Artistic Licence / GPL2*.

4.2 Uporaba konkordančnika

CUWI podpira raznovrstna iskanja in prikaze rezultatov iz korpusa SPOOK skozi masko, ki je podana v sliki 6. Iskanje je lahko enostavno ali pa z izrazi CWB. Pri enostavnem iskanju vtipkamo v iskalno okno niz ali zaporedje nizov (ki lahko vsebujejo tudi znaka ? in * za mehko ujemanje), in konkordančnik bo te nize iskal v polju, ki je bilo izbrano na iskalni maski. Polja so lahko sledeča:

1. *word* (beseda) je niz, kot se pojavi v besedilu korpusa, vključno z velikim začetnicami in pregibanjem, npr. »človeka«, »Človeka«, »Janeza«
2. *nform* (normalizirana oblika besede) je niz, ki je identičen besedi, samo, da je vedno zapisan z malimi črkami, npr. »človeka«, »janeza«
3. *lemma* (lema) je osnovna oblika besede, npr. »človek«, »janez«
4. *msd-en* (angleška oblikoskladenjska oznaka) je oznaka SPOOK, podana v angleškem jeziku, npr. »Ncm« (za »Noun common masculine«)
5. *msd-sl* (slovenska oblikoskladenjska oznaka) je oznaka SPOOK, podana v slovenskem jeziku, npr. »Som« (za »samostalnik občno_ime moški spol«)
6. *ctag* (korpusna oznaka) je oznaka programa TreeTagger, npr. »VVINF« za nemški nedoločnik, ki v sistemu SPOOK ustreza slovenskemu »Ggn« oz. »glagol glavni nedoločnik« in angleškemu »Vmn« oz. »Verb main infinitive«.

Z enostavnim načinom tako lahko npr. iščemo vse besede, ki se končajo na *-uje* (iskalni izraz »*uje«, polje »word«), leme, ki vsebujejo »gn« (iskalni izraz »*gn*«, polje »lemma«), ali pa vsa zaporedja pridevnika in samostalnika (iskalni izraz »P* S*«, polje »msd-sl«).



Slika 6: Vmesnik konkordančnika CUWI.

Uporaba skladnje CWB v iskalnem oknu je sicer bolj dolgovozna, vendar omogoča izražanje bistveno bolj kompleksnih iskalnih pogojev. Po eni strani lahko tu namesto mehkega ujemanja uporabljamo regularne izraze (npr. "[^aeiou]+" za vse pojavnice, ki ne vsebujejo samoglasnikov), po drugi pa lahko medsebojno kombiniramo več polj (npr. "[msd-sl=«P.*»] [lemma="konj" & ctag = "Somei"]" za nize, ki so sestavljeni iz pridevnika in besede z lemo »konj«, ki je v imenovalniku ednine.

Iskanje lahko dodatno omejimo, npr. tako, da mora biti celotno zaporedje v enem stavku ali pa, da iskanje poteka le znotraj določenega dela korpusa, npr. v posameznem besedilu. Te omejitve je mogoče vnesti s pomočjo maske programa CUWI ali pa neposredno z ukazi CWB, pri čemer CUWI vedno pokaže, kakšen ukaz je dejansko izvedel, kar uporabniku olajša sestavljanje kompleksnejših iskalnih zahtev.

CUWI podpira več načinov izpisa najdenih fraz v korpusu. Zadetki s kontekstom so lahko prikazani v vezanem besedilu, v standardnem formatu KWIC (»Key-Word In Context« oz. ključna beseda s kontekstom) ali pa brez konteksta, vendar s številom pojavitev v korpusu, torej kot frekvenčni seznam. Pri vseh načinih izpisa velja, da lahko izpišemo ne samo pojavnice (besedilo), temveč katerokoli kombinacijo oznak. Konkordančnik tudi ponuja razne možnosti sortiranja (po kontekstu in od zadaj) in filtriranja (vse, naključni vzorec).

Rezultate iskanja lahko shranimo na svoj računalnik kot razpredelnico v formatu CSV ali Excel, kar omogoča nadaljnje analize najdenih nizov v enem od programov za delo z razpredelnicami. Konkordančnik prav tako omogoča shranjevanje rezultatov v obliki podatkovnih struktur za skriptne jezike v formatu JSON (*JavaScript Object Notation*) ali Perl, če jih želimo programsko obdelovati. Kot je bilo že omenjeno, podpira konkordančnik prikaz poravnanih segmentov, če imamo izbran prevedeni slovenski korpus, tudi za več jezikov hkrati. Iskanje je mogoče tudi omejiti s postavitvijo dodatnih iskalnih zahtev nad poravnanimi segmenti. Tako lahko npr. iščemo vse pojavitve leme »mati«, kjer se v poravnanih italijanskih segmentih ne pojavi lema »madre«.

5 ZAKLJUČKI

Prispevek je predstavil bolj tehnične vidike korpusa SPOOK, in sicer oblikoskladenjsko označevanje in sistem oznak SPOOK, zapis korpusa v XML in TEI in spletni konkordančnik CUWI, s katerim je mogoče iskati po korpusu. Priporočila in vmesnik do konkordančnika so dostopni na spletni strani <<http://nl.ijs.si/spook/>>, pri čemer so priporočila prosto dostopna po licenci Creative Commons, konkordančnik pa zaradi avtorskih pravic nad izvornimi besedili zahteva za uporabo geslo – zaradi teh omejitev je tudi korpus kot celota dostopen samo partnerjem projekta SPOOK.

Zahvala

Avtor se zahvaljuje Jan Joni Javoršku za implementacijo konkordančnika CUWI in za pripombe na besedilo, za slednje tudi Darji Fišer. Delo predstavljeno v tem delu je financiral projekt ARRS J6-2009-0581 »Slovensko prevodoslovje – viri in raziskave«.

Literatura

- Arhar, Špela in Gorjanc, Vojko, 2007: Korpus FidaPLUS: nova generacija slovenskega referenčnega korpusa. *Jezik in slovnstvo* 52/2, 95–110.
- Burbeck, Steve, 1992: *Applications Programming in Smalltalk-80(TM): How to use Model-View-Controller (MVC)*. <<http://st-www.cs.illinois.edu/users/smarch/st-docs/mvc.html>> (Dostop 15. 5. 2012)
- Christ, Oliver, 1994: A modular and flexible architecture for an integrated corpus query system. *Proceedings of the Conference in Computational Lexicography, COMPLEX'94*, Hungarian Academy of Sciences, Budimpešta, 23–32.
- Creative Commons licence Attribution-ShareAlike 3.0* <<http://creativecommons.org/licenses/by-sa/3.0/>> (Dostop 1. 5. 2012)
- CWB: *The IMS Open Corpus Workbench*. <<http://cwb.sourceforge.net/>> (Dostop 1. 4. 2012)
- Erjavec, Tomaž, 2009: Odprtost jezikovnih virov za slovenščino. Stabej, Marko (ur.): *Infrastruktura slovenščine in slovenistike*. 28. Simpozij Obdobja, Ljubljana: Znanstvena založba Filozofske fakultete, 115–121.
- Erjavec, Tomaž (ur.), 2010a: *MULTEXT-East Resources Version 4* »MondiLex«. <<http://nl.ijs.si/ME/V4/>> (Dostop 1. 5. 2012)
- Erjavec, Tomaž, 2010b: Text Encoding Initiative Guidelines and their Localisation. *Infoteka*. 11/1, 3a–14a.
- Erjavec, Tomaž, 2012a: MULTEXT-East : morphosyntactic resources for Central and Eastern European languages. *Language resources and evaluation*, 46/1, 131–142.
- Erjavec, Tomaž, 2012b: The goo300k corpus of historical Slovene. *Proceedings of the Eight International Conference on Language Resources and Evaluation, LREC'12*, Pariz, ELRA.
- Erjavec, Tomaž, Ignat, Camelia, Pouliquen, Bruno in Steinberger, Ralf, 2005: Massive multi-lingual corpus compilation: Acquis Communautaire and ToTaLe. *Proceedings of the 2nd Language & Technology Conference*, Poznan, Poljska, 32–36.
- Erjavec, Tomaž, Krek, Simon, 2008: Oblikoskladenjska priporočila in označeni korpusi JOS. Erjavec, Tomaž in Žganec Gros, Jerneja (ur.). Zbornik Šeste konference Jezikovne tehnologije, zbornik 11. mednarodne multikonference Informacijska družba – IS 2008, zvezek C. Ljubljana: Institut Jožef Stefan, 49–53.

- Extensible Markup Language (XML)*, <<http://www.w3.org/XML/>> (Dostop 1. 4. 2012)
- Greer, Derek, 2007: *Interactive Application Architecture Patterns*. <<http://aspirin-graftsman.com/2007/08/25/interactive-application-architecture/>> (Dostop 15. 5. 2012)
- Ide, Nancy in Jean Veronis, 1994: Multext (multilingual tools and corpora). *Proceedings of the 15th International Conference on Computational Linguistics, Co-Ling'94*, Kyoto, 90–96.
- JavaScript Object Notation. <<http://www.json.org/>> (Dostop 1. 5. 2012)
- Jezikovni viri starejšega slovenskega jezika IMP*, <<http://nl.ijs.si/imp/>> (Dostop 1. 5. 2012)
- McDaniel, Adam, 2011: *Html5: Your Visual Blueprint for Designing Rich Web Pages and Applications*. *Visual Blueprint*, 7. zv., John Wiley & Sons.
- Mojolicious – Perl real-time web framework*. <<http://mojolicio.us/>> (Dostop 1. 5. 2012)
- Oblikoskladenjske specifikacije JOS V1*. <<http://nl.ijs.si/jos/josMSD-sl.html>> (Dostop 1. 5. 2012)
- Projekt JOS: jezikoslovno označevanje slovenskega jezika*. <<http://nl.ijs.si/jos/>> (Dostop 1. 5. 2012)
- Schmidt, Douglas C., 1995: Reactor: An Object Behavioral Pattern for Demultiplexing and Dispatching Handles for Synchronous Events. Coplien, Jim in Schmidt, Douglas C. (ur.) *Pattern Languages of Program Design*. Addison-Wesley. Novejša različica: <http://www.dre.vanderbilt.edu/~schmidt/PDF/reactor-siemens.pdf> (Dostop 15. 5. 2012)
- Schmid, Helmut, 1994: Probabilistic Part-of-Speech Tagging Using Decision Trees. *Proceedings of International Conference on New Methods in Language Processing*. Manchester, Velika Britanija.
- Schmid, Helmut, 1995: Improvements in Part-of-Speech Tagging with an Application to German. *Proceedings of the ACL SIGDAT-Workshop*. Dublin, Irska.
- TEI P5: Guidelines for Electronic Text Encoding and Interchange*. <<http://www.tei-c.org/Guidelines/P5/>> (Dostop 1. 4. 2012)
- The Unicode Consortium*, <<http://unicode.org/>> (Dostop 1. 4. 2012)
- TreeTagger – a language independent part-of-speech tagger*. <<http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>> (Dostop 1. 9. 2011)
- XSL Transformations (XSLT) Version 1.0*, <<http://www.w3.org/TR/xslt>> (Dostop 1. 4. 2012)

Pogostost vzročno-pojasnjevalnega *namreč* v izvirniku in prevodu: korpusni pristop

Nataša Hirci

Oddelek za prevajalstvo, Filozofska fakulteta Univerze v Ljubljani
 natasa.hirci@ff.uni-lj.si

Abstract

The aim of this paper is to address the issue of additions and omissions of connectives, which are a fairly common phenomenon in translation, and to elucidate the notion of explication of certain linguistic elements. The article reports on a study carried out on the first Slovene translational corpus, SPOOK, focusing on the additions of the Slovene expression *namreč* in translations from English into Slovene. According to the Slovene-English dictionary definition, the expression *namreč* has as a straightforward equivalent in the English adverb *namely*. A corpus approach was applied in the study to determine the frequency of translations with *namreč*. The results of the corpus analysis are compared with the equivalents proposed in the bilingual dictionary. Since *namely* is often overused in trainee translators' work from Slovene into English, the use of *namreč* and the role played by *namely* in Slovene translations from English were studied to shed more light on the typical use of the two expressions. The results of the study also contribute to the raising of awareness among the future translators, illustrating how inappropriate it is to rely purely on dictionary solutions and how necessary to become a habitual user of suitable corpora while translating.

Keywords: translation into mother tongue, translation into foreign language, text connectives, explication, corpus-based studies

Ključne besede: prevajanje v materni jezik, prevajanje v tuji jezik, besedilni povezovalci, eksplicitacija, korpusni pristop

1 UVOD

Utemeljitelj slovenskega knjižnega jezika in začetnik slovenske pisane besede, Primož Trubar, ni splošno znan le kot avtor prvih slovenskih knjig, ampak je poznan tudi kot prvi prevajalec svetopisemskih besedil v slovenski jezik. Tako je slovenski knjižni jezik že vse od svojih začetkov tesno prepleten s prevodom. Še dandanes so prevedena besedila, ki smo jim v Sloveniji izpostavljeni na vsakem koraku, ena najtesnejših stičnih točk slovenske kulture in jezika z ostalimi kulturami in jeziki po svetu. Vsakodnevno se srečujemo s prevedenimi besedili iz najrazličnejših jezikov, vendar verjetno najpogosteje s prevodi iz angleškega jezika. Medjezikovni in medkulturni prenos se vrši tudi v obratni smeri, vendar v znatno manjši meri. Ker je angleški jezik vse bolj prisoten v vseh sferah tako javnega kot privatnega življenja, je v Sloveniji prevajanje v ta jezik še posebej prisotno, pa naj je prevajalec tisti, ki mu je angleški jezik prvi jezik ali jezik A, ali tisti, ki mu je ta jezik drugi jezik ali jezik B.¹

Prav zaradi tega je raziskovanje značilnosti medjezikovnega in medkulturnega prenosa in umeščanje tujejezičnih besedil v slovenski prostor ključnega pomena, saj na ta način lahko pridemo do novih spoznanj, ki lahko pomembno prispevajo k razumevanju ne zgolj prevedenih besedil, temveč tudi prevajalčevih odločitev pri ubesedovanju iz enega v drug jezik. Še do nedavnega so bile značilnosti prevodov na Slovenskem opisovane na podlagi predvsem kvalitativnih analiz in še to zgolj posameznih besedil. Z razmahom korpusnega jezikoslovja pa smo tudi v prevodoslovju prišli do temeljnega preobrata v tovrstnih raziskavah. Še posebej v zadnjih letih lahko v Sloveniji opazimo živahen razmah jezikoslovnih in prevodoslovnih raziskav, ki so bile osnovane na korpusni metodologiji (prim. Fišer 2010, Gorjanc 2005, Gorjanc in Vintar 2007, Grad 2010, Hirci 2007, 2012, Mikolič Južnič 2010, 2011, Pisanski Peterlin 2009, 2010, Pokorn 2005, 2009, Schlamberger Brezar 2009, 2011, Vintar in Fišer 2009, Vintar 2010, 2011, če naštejemo zgolj nekatere). Vendar te raziskave niso bile osnovane na prevodoslovnih korpusih, temveč so večinoma vključevale enojezične referenčne korpeuse, ki so osvetljevali določene jezikovne strukture v enem jeziku, ali dvo- in večjezične vzporedne korpusne vire, ki so omogočali kontrastivno raziskovanje izvirnikov in njihovih prevodov.

Prvi slovenski prevodoslovni korpus SPOOK je večjezični vzporedno-primerljivi korpus angleško-, nemško-, francosko- in italijansko-slovenskih prevodov z izvirniki ter primerljivimi slovenskimi izvirnimi besedili, ki je nastajal od leta 2009 do 2011 v okviru projekta *Slovensko prevodoslovje – viri in raziskave*.² Korpus vsebuje

¹ Materni jezik v tem prispevku razumemo kot prvi jezik ali jezik A, medtem ko je prvi tuji jezik jezik B, drugi tuji jezik pa jezik C (prim. Pokorn 2005, Hirci 2007, 2012).

² Vodja projekta Špela Vintar.

sodobna leposlovna dela v slovenskem jeziku, bodisi izvirno napisana v slovenščini (korpus SLV-SLV) ali pa prevode tujih književnih del v slovenski jezik, in sicer iz angleščine (podkorpus ENG-SLV), francoščine (podkorpus FRA-SLV), nemščine (podkorpus GER-SLV) ali italijanščine (podkorpus ITA-SLV). Vsebuje tudi izvirnike romanov, ki so stavčno poravnani s prevodom v slovenščino. Strukturni atributi v korpusu SPOOK navajajo tip besedila, naslov dela, in sicer tako naslov izvirnika kot tudi naslov prevoda, avtorja in prevajalca dela, in seveda jezik izvirnika in prevoda. Besede v korpusih so označene z lemo in oblikoskladenjsko oznako.³ Za našo raziskavo sta relevantna dva podkorpusa korpusa SPOOK, vzporedni angleško-slovenski del korpusa in korpus slovenskih izvirnikov. Angleško-slovenski podkorpus, korpus SPOOK ENG-SLV, vsebuje šest angleških izvirnikov in šest prevodov teh književnih del v slovenski jezik, kar znaša 1.160.451 besed in 136.137 poravnanih segmentov. Slovenski podkorpus, korpus SPOOK SLV-SLV, pa vsebuje triindvajset izvirnikov, tj. slovenskih leposlovnih del, kar znaša 1.454.275 besed.

Korpus SPOOK odpira povsem nove možnosti za raziskovanje prevodoslovnih pojavov (prim. Vintar, 2009). S pomočjo prevodoslovnega korpusa lahko namreč pridemo do določenih zaključkov o značilnostih prevedenih besedil na osnovi zelo velikega števila besedil, ki so sicer za raziskovalca malodane neobvladljiva brez pomoči korpusne metodologije, medtem ko tak pristop omogoča hkratno primerjavo jezikovnih vzorcev v prevedenih besedilih z jezikom v izvirnikih. Rezultati raziskav, osnovanih na prevodoslovnem korpusu SPOOK, bodo zagotovo pripomogli pri razvoju prevajalskih strategij in različnih tehnik prevajanja, nena zadnje pa bodo prispevali tudi k boljšemu razumevanju in sprejemanju prevodov v slovenski jezik.

Pogosto si predstavljamo, da so prevajalski problemi iz prvega jezika, torej jezika A v tuji jezik, tj. jezik B, zelo specifični in povsem drugačni od tistih iz jezika B v jezik A. V nekaterih primerih pa se izkaže ravno obratno: prevajalski izziv, ki ga zaznavamo v smeri A-B nam lahko osvetli tudi raziskava določenega jezikovnega sredstva v smeri B-A. Na Oddelku za prevajalstvo na Filozofski fakulteti Univerze v Ljubljani namreč skušamo bodoče prevajalce ustrezno usposobiti za prevajanje tako v materni jezik kot tudi v tuje jezike. Pri prevajalskih vajah se nemalokrat zgodi, da študenti določene strukture iz svojega jezika (pre)pogosto vnašajo kot dobesedni prevod tudi v tuji jezik, medtem ko se slepo zanašajo na slovarske definicije in realnega stanja, torej dejanske rabe izraza v določenem kontekstu, ne preverjajo dodatno. Poglejmo primer angleškega prislova *namely*. Pogosto lahko opazimo, da se ta struktura znajde v študentskih prevodih v angleški jezik kljub temu, da se prislov *namely* v angleških izvirnih besedilih ne pojavlja tako pogosto kot se izraz *namreč* v slovenskih besedilih (prim. Klinar 1994, Jurko 2007). Zato

³ Več o označevanju korpusa: *Spook morphosyntactic specifications* na <http://nl.ijs.si/spook/msd/html-sl/>.

bomo v naši raziskavi skušali osvetliti rabo izraza *namreč* in vlogo angleškega prislova *namely* v slovenskih prevodih iz angleščine.

Sledi kratek pregled idej o prevodnih univerzalijah, še posebej se bomo posvetili pojmu eksplicitacije. Kot primer bomo analizirali vzročno-pojasnjevalni povezovalc *namreč* in njegovo slovarsko definicijo primerjali z dejansko rabo na primerih iz korpusa SPOOK. Becher (2011) trdi, da so povezovalci odličen način raziskovanja eksplicitacije pri prevajanju, saj prevajalci pri svojem delu pogosto posegajo po dodajanju ali izpuščanju teh jezikovnih sredstev. Predpostavljamo, da bi dodatna osvetlitev rabe te jezikovne strukture v slovenskem jeziku in primerjava z rabo v besedilih, prevedenih iz angleškega v slovenski jezik, lahko prispevala k večji ozaveščenosti glede rabe tudi v obratni smeri prevajanja. Obenem bi taka raziskava lahko prispevala tudi k potrditvi hipoteze o eksplicitaciji kot univerzaliji, še posebej, če bi se izkazalo, da prevajalci med prevajanjem v slovenski jezik izraz *namreč* uporabljajo tudi takrat, ko za to ni nekega eksplicitnega razloga v izvirni predlogi.

2 PREVODNE UNIVERZALIJE?

Pojem **prevodne univerzalije** je že od svojih začetkov nemalo kontroverzen, saj ima tako zagovornike kot tudi resne nasprotnike. Zadnji dve desetletji med prevodoslovci še vedno vzbuja veliko zanimanja, obenem pa raziskave, posvečene prevodnim univerzalijam, spremlja tudi kar nekaj razburjenja. O čem pravzaprav govorimo? O tem, da naj bi se jezik v prevedenih besedilih predvidoma pomembno razlikoval od jezika, ki ga najdemo v izvirnikih. Tu ne gre za pojave, povezane s samo kakovostjo prevoda, temveč za določene prevodne lastnosti ali posebnosti, ki naj bi se izkazovale na različnih jezikovnih ravneh. Osnovno razmejitev prevodnih univerzalij lahko pripisujemo razmišljanjem dveh prevodoslovcev, Gidoeona Tourya (1995), ki je predlagal dva t. i. *prevodna zakona*, tj. standardizacijo in interferenco, in Mone Baker (1993, 1996), ki je definirala štiri *prevodne univerzalije* in jih kategorizirala kot eksplicitacijo, presevanje, normalizacijo in zglaševanje (prim. Vintar 2009, 2011). Prav ta razmejitev je tudi najpogostejša osnova prevodoslovnih raziskav teh pojavov v zadnjih desetletjih in se je bomo deloma posluževali tudi v naši študiji.

Kljub temu, da so raziskave o prevodnih univerzalijah postale bolj intenzivne šele v zadnjih dveh desetletjih, lahko razmišljanja o eksplicitaciji pravzaprav najdemo že v šestdesetih letih prejšnjega stoletja. Vinay in Darbelnet sta že leta 1958 pojav eksplicitacije definirala kot vnašanje informacij v ciljni jezik, ki so v izvirnem jeziku le implicitne, vendar jih je mogoče razbrati iz konteksta ali situacije. Vanderauwera (1985) razume eksplicitacijo kot razširitev zgoščenega odlomka, dodajanje

modifikatorjev, kvalifikatorjev in veznikov, ali celo dodatnih informacij, da bi dosegli večjo jasnost besedila (prim. Claudiu 2010). Eksplicitacija se lahko odraža v daljšanju besedila, močnejših kohezivnih in logičnih vezeh, večji redundanci, večji berljivosti, zaznamovani rabi ločil in boljši organizaciji teme in reme (Pápai 2004: 144). Pym (2005a, 2008) svoja razmišljanja osredinja na predloge Kinge Klaudy (prim. Klaudy 1996, 2008) in v bistvu razpravlja o dveh različnih vrstah eksplicitacije: o obvezni, ki jo narekuje jezikovni sistem, in o namerni, ko prevajalec dodaja informacije, da bi se izognil dvoumnosti. Treba pa je dodati, da eksplicitacije ne razume v smislu prevodnih univerzalij in takšne hipoteze zavrača.

Domneve o prevodnih univerzalijah so nekateri raziskovalci oblikovali in opazovali s pomočjo introspekcije, intuicije in raziskav, ki so temeljile na kontrastivni analizi izvirnikov in prevodov, vendar doslej še ni bilo narejenih raziskav, ki bi v zadostni meri uspele potrditi, da resnično obstajajo. Becher (2010a) ugotavlja, da je sicer vrsta raziskav skušala potrditi hipotezo o eksplicitaciji Blum-Kulke (1986: 21); ta pravi, da lahko med izvirnikom in njegovim prevodom opazujemo nekažno kohezivno eksplicitnost, in to ne glede na število razlik med dvema jezikovnima in besedilnima sistemoma. Nadaljuje z razmišljanji Mone Baker in Juliane House. Baker (1993, 1996) trdi, da je eksplicitacija prevodna univerzalija, torej je tipična in razlikovalna značilnost prevedenih besedil, kar se vsekakor navezuje na razmišljanja Blum-Kulke (1986: 19). Po drugi strani pa se House (2004: 193) ne strinja s podano idejo o eksplicitaciji kot prevodni univerzaliji, saj trdi, da razlike v jezikovnih in stilističnih konvencijah med izvirnikom in prevodom lahko pripišemo zgolj določeni eksplicitaciji znotraj posameznega prevoda in ne univerzalni tendenci prevajalcev po eksplicitaciji. Vendar ta razmišljanja niso nujno izključujoča, saj ni razlogov za to, da bi verjeli v en sam tip eksplicitacije (prim. na primer Saldanha 2008, ali Klaudy 2008, ki razlikuje med štirimi različnimi tipi eksplicitacije). Becher (2010b: 4–5) je odločno nenaklonjen razlagam o eksplicitaciji kot prevodni univerzaliji, ki naj bi bila lastna vsem prevodom in je ni moč razložiti kot željo po približanju besedila ciljnemu občinstvu, torej zgolj v luči prevajalčevega prilagajanja izvirnika pričakovanjem ciljne publike, kar deluje zgolj znotraj določenega jezikovnega para med katerima je prevajalec prevajal. Kot primere navaja neustrezne in nepravilne zaključke raziskav, o katerih poročajo Øverås (1998), Pápai (2004) in Konšalová (2007), podobno pa velja tudi za Fabricius-Hansen (1998) ter Olohan in Baker (2000). Sprašuje se, če ne bi bilo bolje prepogosto rabo določene jezikovne strukture v prevedenih besedilih pripisovati interferenci izvirnega jezika, ne pa tega pojava razlagati kot eksplicitacijo, ki je lastna prevodu samemu. Zanimivo je tudi razmišljanje Baumgarten, Meyer in Özçetin (2008: 198); zaključili so, da eksplicitacija zagotovo ni univerzalni pojav, saj jo včasih lahko opazimo, včasih pa tudi ne. Ko je prisotna, se kaže kot eksplicitacija, ki jo narekujejo komunikacijske ali stilistične norme ciljnega jezika, zagotovo pa ni lastna prevajalskemu procesu in potemtakem izven prevajalčevega

nadzora. Že Puurtinen (2004: 170–175) je v svojih raziskavah o eksplicitaciji odkrila negativne učinke interference izvirnega jezika. Becher (2010b: 21) dodaja, da se v prihodnje morda lahko nadajamo bolj obetavnih rezultatov tudi na tem področju in pri tem posebej izpostavi delo Steinerja in kolegov iz Univerze v Saarbrücknu (Steiner 2005, 2008; Hansen-Schirra et al. 2007). Prav Hansen-Schirra in sodelavci (2007: 225) so v svoji raziskavi prišli do zaključka, da bi bistveno večjo leksikalno zgoščenost prevedenega besedila morebiti vseeno veljalo pripisati posledici prevajalskega procesa kot takega ne glede na obravnavan jezikovni par, kar spet govori v prid neki univerzalni težnji po eksplicitaciji.

Pym (2005a, 2005b, 2008) že od vsega začetka kritično izpostavlja težave s terminološko in konceptualno nedorečenostjo termina prevodne univerzalije. Trdi, da prevajalci radi posegajo po strategijah, ki zmanjšujejo tveganje, da bi bilo ciljno besedilo narobe razumljeno, in prav zaradi tega posegajo po eksplicitaciji. Tudi Becher (prim. 2011: 4) v svojih zadnjih prispevkih vse pogosteje predlaga odmik od pojma prevodu lastni eksplicitaciji; razlaga, da je različne primere eksplicitacije bolje razlagati s pomočjo leksikalno-slovničnih in pragmatičnih faktorjev. Trdi, da se prevajalci ponavadi zatekajo k eksplicitaciji takrat, ko skušajo bralcu olajšati procesiranje besedila in zmanjšati tveganje, da bi bralci besedilo narobe razumeli, ko skušajo upoštevati komunikacijske preference ciljnega jezika, medtem ko je povod za eksplicitacijo lahko tudi leksikalno ali skladenjsko neprekrivanje med izvirnim in ciljnim jezikom.

Že prej je bilo omenjeno, da je za zbiranje raziskovalnih dokazov na tem področju najprimernejša korpusna metoda, ki omogoča kvantitativne primerjave izvirnih in prevedenih besedil. Če bi s korpusnimi izsledki hipotezo o prevodnih univerzalijah uspeli dokazati, bi to pomenilo, da se pri prevajanju še posebej literarnih besedil določen del izvirne večplastnosti ali večpomenskosti ne ohranja. Večina starejših študij je namreč temeljila na nekih intuitivnih razmišljanjih in šele v zadnjih desetletjih je tovrstnim prevodoslovnim raziskavam dala nov zagon korpusna metodologija. Tako imamo zdaj s pomočjo prevodoslovnega korpusa SPOOK možnost dokazati ali ovreči hipotezo o prevodnih univerzalijah tudi v Sloveniji.

3 ČLENKI, VEZNIKI, KONEKTORJI ALI POVEZOVALCI?

Slovenska slovnica (Toporišič 1991, 2000) definira *členek* kot nepregibno besedno vrsto, ki izraža sporočevalčevo oceno, sodbo ali mnenje in po kateri se ne moremo vprašati z nobeno vprašalnico, lahko pa ga nadomestimo s stavkom. Kot besedno vrsto je členke že leta 1974 definiral Jože Toporišič, ki pravi, da

so »nepregibna besedna vrsta, z njimi vzpostavljamo zveze s sobesedilom, izražamo pomenske odtenke posameznih besed, delov stavkov, celih stavkov in povedi ali tvorimo skladenjske naklone« (Toporišič 1991: 384). Prej so namreč večino členkov prištevali med prislove. Tudi *Slovar slovenskega knjižnega jezika* (2000) označuje številne členke večinoma kot *prislove*. *Slovenska slovnica* (Toporišič 1991: 384–385 in 2000: 448–449) deli nepregibno besedno vrsto členkov na trinajst skupin, in sicer na skupine:

- navezovalnih členkov (*a, potem, sicer, saj*),
- členkov čustvovanja (*žal, dobro*),
- poudarnih členkov (*ravno, posebno, zlasti, predvsem*),
- izvzemalnih členkov (*le, samo, edino, vsaj*),
- presojevalnih členkov (*tako, približno, skoraj, malodane*),
- dodajalnih členkov (*tudi, prav tako, poleg tega*),
- členkov zadržka (*pravzaprav, pač, saj, komaj, sicer, resda*),
- členkov potrjevanja in soglašanja (*da, dejansko, gotovo, kajpak*),
- členkov možnosti in verjetnosti (*morda, mogoče, konec koncev, navsezadnje*),
- členkov mnenja in domneve (*baje, menda, denimo, recimo*),
- vprašalnih členkov (*ali, a kaj, kajne, mar*),
- spodbujalnih členkov (*morda, naj, res, vsajše, kaj, nikar, saj*) in
- skupino členkov zanikanja in nesoglašanja (*ne, nikar, nikakor, nasprotno*).

Vendar Smolej (2004) pravi, da že iz samih poimenovanj teh skupin lahko razberemo, da je temelj zgornje delitve predvsem preučitev besede kot nosilke pomena, torej pomensko razčlenjevanje besed in besednih zvez in ne toliko skladenjsko in besedilnofunkcijsko vedenje členkov. Dodaja, da če bi »v središče preučitve in analize členkov postavili /.../ skladenjski in besedilotvorni oz. besedilnopovezovalni vidik in dalje celo pragmatičnosporočevalni vidik členkov, bi lahko nekatere podskupine členkov združili« na manj skupin, »/.../ katerih meje bi izhajale predvsem iz skladenjskega, besedilnega in dalje pragmatičnosporočevalnega kriterija« (Smolej 2004: 47). Tako v skupino **besedilnih povezovalcev** avtorica uvršča členke kot so *kakorkoli že, kljub vsemu, konec koncev, kot (že) rečeno, na koncu koncev, namreč, navsezadnje, nena zadnje, po eni/drugi strani, poleg tega, potem, potemtakem, pravzaprav, se pravi, sicer, skratka, torej, vseeno, za nameček* in mnoge druge (ibid.).

Znotraj skupine **pojasnjevalnih povezovalcev** (povzeto po Toporišič 1991: 379–380) lahko ločimo med **pojasnjevalnimi členki** (na primer *namreč*) in

pojasnjevalnimi vezniki (na primer *in sicer*, *in to*, *kakor*, *namreč*, *sicer*, *to je*, *tj.*, *to se pravi*). Istovrstni pojasnjevalni vezniki so *in sicer*, *in to*, *kakor*, *namreč*, *sicer*, *to je*, *tj.*, *to se pravi* (prim. Tabela 1). Tu spremni del veznika ali vezniške zveze samo z drugimi besedami pove to, kar je bolj splošno povedano že v jedru:

- *Rojstni dan ima ta mesec, **in sicer** desetega.*
- *Leto ima štiri letne čase, **namreč** pomlad, poletje, jesen in zimo.*

Novi stavek – ali del stavka – prejšnjega osvetli tako, da ga pojasni s konkretnimi podrobnostmi; v takih primerih je razmerje sorodno apoziciji ali pojasnjevalnemu priredju. Prejšnji stavek (ali del stavka) torej osvetlimo s specifično spremenjeno, popravljeno ali povzeto vsebino, in sicer tako, da vsebino prejšnjega ubesedimo na boljši, ustrežnejši način (povzeto po Skubic 1999: 218, ki je svoja razmišljanja o členkih osnoval na Hallidayevi angleški funkcijski slovnici).

Tabela 1: Število vseh pojavnih pojasnjevalnih povezovalcev v korpusu SPOOK⁴ in njihova pogostost na 10.000 besed.

	SPOOK SLV-SLV	Pogostost na 10.000 besed	SPOOK ENG-SLV	Pogostost na 10.000 besed
in sicer	32	0,2200	49	0,4266
in to	318	2,1866	188	1,6370
namreč	253	1,7396	138	1,2016
sicer	580	3,9882	269	2,3423
to je	502	3,4518	453	3,9445
tj.	0	0	0	0
to se pravi	5	0,0343	0	0

Poglobljene, bolj teoretično obarvane slovenistične razprave o členkih, besedilnih konektorjih in besedilnih povezovalcih najdemo v Černelič (1991, 1993), Gorjanc (1998), Jakop (2000), Smolej (2004), Skubic (1999), Toporišič (1991, 2000), če naštejemo le nekaj avtorjev, ki so se v zadnjih dveh desetletjih intenzivneje ukvarjali s to problematiko. Gorjanc (1998) razpravlja o t. i. konektorjih, ki jih analizira v znanstvenih besedilih. Pravi, da so poseben »problem pri členkovnih konektorjih tisti, ki se izrazno prekrivajo z vezniki, npr. namreč, torej; sicer, saj /.../. Vsi z nestabilno pojavitvijo in – neglede na položaj – vzpostavljanjem enakega pomenskega razmerja v konektorski zvezi bi tako sodili med členkovne, npr. namreč, torej« (Gorjanc 1998: 380).

⁴ Pojasnjevalni veznik *tj.* smo izpustili iz nadaljnje analize, saj se v korpusu SPOOK ne pojavi niti enkrat.

Vendar za namene naše raziskave nadaljnja klasifikacija teh jezikovnih sredstev ni bistvenega pomena. Zaradi izredne obsežnosti te tematike se bomo v članku osredotočili zgolj na preučevanje vzročno-pojasnjevalnega besedilnega povezovalca *namreč*⁵, in sicer v vlogi tako člena kot veznika. Pojasnjevalni povezovalci lahko *namreč* vzročno pojasnjujejo jedro, torej delujejo vzročno ali kavzalno (prim. *zakaj, kajti, namreč, sicer*) ali pa jedro izrazijo z drugimi besedami, torej delujejo pojasnjevalno ali eksplikativno (predvsem *to je, tj., in sicer*, tudi *namreč*).

3.1 Primer: *namreč* kot vzročno-pojasnjevalni povezovalec

V prispevku se bomo osredotočili na rabo vzročno-pojasnjevalnega povezovalca *namreč*. Slovenska slovnica pravi, da »*namreč v nekem dodatku utemeljuje ali vzročno pojasnjuje jedro vezniške zveze*« (Toporišič 1991: 379). Opredeljuje različne vloge izraza *namreč*, in sicer:

1. v pojasnjevalnem priredju kot **pojasnjevalni veznik**:
 - *Midva, **namreč** brat in jaz, se dobro razumeva.*
2. v vzročnem priredju kot **vzročni veznik**:
 - *Vsi so vznemirjeni, zdaj gre **namreč** za pomembne stvari.*
3. kot **pojasnjevalni členek**, ki dopolnjuje in pojasnjuje, kar je bilo že prej povedano.
 - *Tam, **namreč** pri sosedovih, je vse po starem.*

Skušali bomo ugotoviti, ali bi pojasnjevalni povezovalec **namreč** v teh treh vlogah – v nadaljnji analizi jih *namreč* ne ločujemo posebej – bilo smiselno razlagati kot del teorije o prevodnih univerzalijah, in sicer kot primer eksplicitacije (po Baker 1996). Rabo tega jezikovnega sredstva bomo analizirali v luči predlagane slovarske definicije in pojavnosti v prevodih iz angleškega v slovenski jezik. V ospredju bo analiza rabe vzročno-pojasnjevalnega *namreč* in prevodov angleškega prislova *namely* v slovenski jezik, pri čemer bomo uporabili korpusni pristop.

4 ZASNOVA RAZISKAVE

V raziskavi se osredotočamo na pregled rabe vzročno-pojasnjevalnega izraza *namreč* v besedilih prevedenih iz angleškega v slovenski jezik, ki smo jih opazovali s pomočjo korpusa SPOOK. Uporabili bomo kvantitativne podatke, ki so osnovani na izsledkih korpusnih raziskav v prevodoslovju in podatke iz korpusa

⁵ Za namene naše študije, ki je prevajalske narave, smo poimenovanje poenostavili in izraz *namreč*, ki je sicer osrednja tema te raziskave, poimenovali **pojasnjevalni povezovalec**.

vzporednih angleško-slovenskih besedil SPOOK ENG-SLV primerjali s podatki iz korpusa slovenskih izvirnikov, ki jih nudi korpus SPOOK SLV-SLV.

Po pregledu pogostosti rabe *namreč* v izvirnikih in primerjavi s pogostostjo tega izraza v vzporednem korpusu bomo lahko analizirali dejansko rabo tega izraza. Zanima nas, ali stanje v korpusu odseva predloge o prevodnih ustreznicih v slovarskih virih. V prihodnje bi bilo smiselno podrobneje preveriti tudi senzitiviteto študentske populacije pri rabi korpusov in poudariti, da je bodoče prevajalce treba še naprej in celo še pogostejše spodbujati k redni rabi korpusnih virov. Takšna pedagoška angažiranost je nujna ne le zato, da bi prevajalci med prevajanjem bili uspešni pri iskanju potrditve svojih intuitivnih rešitev, temveč tudi, da jim pomagamo priti do novih spoznanj o rabi določenih jezikovnih sredstev.

Naša raziskava skuša najti odgovore na naslednja vprašanja: kaj je tipično za določene jezikovne elemente v prevodih iz angleškega v slovenski jezik? Se še vedno obnašajo enako, kot se obnašajo v slovenskih izvirnikih? Kakšna je pri tem vloga dvojezičnega slovarja? Chesterman (2004) *namreč* trdi, da je določene jezikovne posebnosti najlažje preučevati skozi prizmo slovnice, slovarjev, korpusov in kontrastivnih raziskav. Tako se bomo tudi v tem članku osredotočili na raziskovanje pojavnosti določenega jezikovnega elementa v korpusu, potem izsledke primerjali s predlaganimi prevodnimi ustreznici in predvideno rabo v slovarskih virih, čemur bo sledila še kvalitativna kontrastivna analiza raziskovanega izraza.

5 REZULTATI IN DISKUSIJA

5.1 Primerjava slovarskih definicij za *namreč* in dejansko stanje v korpusu SPOOK

V nadaljevanju bomo pogledali slovarske predloge za izraze, ki se prevajajo z vzročno-pojasnjevalnim povezovalcem *namreč* in stanje primerjali s pojavniciami tega jezikovnega sredstva v korpusu SPOOK (glej Tabelo 2).

Tabela 2: Pogostost pojasnjevalnega povezovalca *namreč* v korpusu SPOOK.

Korpus	Št. pojavníc v korpusu	Št. pojavníc na 10.000 besed
SPOOK SLV-SLV	253	1,7396
SPOOK ENG-SLV	138	1,2016

Najprej si bomo ogledali primer gesla *namreč* v slovensko-angleškem dvojezičnem slovarju in predlagane prevedke primerjali s predlogi za prevedke teh ustreznice v angleško-slovenskem slovarju. Potem bomo predlagane prevedke preverili v korpusu SPOOK, in sicer tako, da bomo preverili število pojavnic v slovenskem delu korpusa SPOOK, za primerjavo pa bomo uporabili rešitve s povezovalcem *namreč* v podkorpusu SPOOK ENG-SLV. Predvidevamo, da bo stanje v korpusu odražalo predloge v slovarskih virih.

Slovensko-angleški slovar (Grad, Leeming 1997) pod geslom *namreč* ponuja več različnih rešitev, in sicer:

námreč – namely; that is to say; videlicet (krajšava: viz.); (= t.j., to je); to wit; in other words; nisem mogel govoriti z njim, bil je ~ bolan I could not speak to him, as (ali seeing that) he was ill; eden mojih sosedov, ~ neki X.Y., ... one or my neighbours, to wit, one X.Y., ...

Veliki angleško-slovenski slovar (Grad et al 1997) za *namely*, ki ga najdemo kot prvo možnost prevedka za *namreč* v slovensko-angleškem slovarju poda zgolj prevedek z *namreč*. Najprej pogledajmo slovarske možnosti za prevedke za *namreč* v angleškem jeziku in jih potem primerjajmo še s slovenskimi prevedki, ki so v angleško-slovenskem slovarju ponujeni pod temi gesli:

- **namely** = namreč
- **that is (to say)** = to se pravi;
- **videlicet** ali **viz.** = namreč; to je
- **to wit** = to je, namreč; kot sledi; in to
- **in other words** = z drugimi besedami

Veliki angleško-slovenski slovar (Grad et al. 1997) zanimivo pod geslom *namely*, ki ga v slovensko-angleškem slovarju sicer najdemo na prvem mestu, obravnava *namely* kot prislov in za prevodno ustreznico ponuja *namreč*, medtem ko pod geslom *that is to say*, ki ga najdemo kot drugega pod tem geslom, ponuja večbesedni izraz *to se pravi*, izraza *namreč* pa ne. Za katere od predlaganih prevedkov lahko najdemo prevode z izrazom *namreč* v korpusu SPOOK? Poudariti je treba, da za izraze *namely*, *that is (to say)*, *videlicet/viz.*, *to wit* in *in other words*, ki naj bi v luči slovarske definicije bili ustreznice za *namreč*, v podkorpusu SPOOK ENG-SLV na prvi pogled presenetljivo ne najdemo niti ene same pojavnice. Vendar pa taka distribucija v pogostosti dejansko ne preseneča tako zelo, še posebej, če vzamemo v obzir dejstvo, da korpus SPOOK vsebuje leposlovna besedila, medtem ko se zgornji prevedki ponavadi uporabljajo v bolj strokovnih, znanstvenih in tehniških besedilih. Tako bi hipotezo o pogostosti teh prevedkov bilo smiselno dodatno preveriti

v drugih korpusih, na primer v slovenskem referenčnem korpusu *FidaPlus*⁶, in te izsledke primerjati s pojavnostjo angleškega prislova *namely* v angleških referenčnih korpusih kot sta *BNC* in *COCA*⁷ (prim. Jurko 2007: 35–36), kar pa žal presega okvir tega članka.

Kako pa je s pojavnicami za slovarske predloge za njihove prevedke v slovenščini? Poglejmo primere, ki so ilustrirani v Tabeli 3.

Tabela 3: Število pojavnic v podkorpusih SPOOK ENG-SLV in SPOOK SLV-SLV.

	Št. pojavnic v podkorpusu SPOOK ENG-SLV	Št. pojavnic v podkorpusu SPOOK SLV-SLV
namreč	138	253
to se pravi	0	5
to je	453*	502**
kot sledi	0	1
in to	188*	318**
z drugimi besedami	5	1

Primeri z * in ** niso vsi ustrezni, saj je kvalitativna ročna analiza pokazala, da modificirane povedi ne izražajo pomena v smislu pojasnjevalnega *namreč*.

Namely se v podkorpusu SPOOK ENG-SLV ne pojavi niti enkrat, čeprav lahko najdemo kar 138 pojavnic za *namreč*. Kaj to kaže? *Namreč* se torej pojavlja kot prevod nekih drugih angleških jezikovnih struktur ali pa ga prevajalci dodajajo, ko želijo nekaj eksplicitno poudariti. Zastavlja se vprašanje, katere jezikovne strukture v angleškem jeziku so prevajalci dejansko prevajali s pomočjo izraza *namreč*. V korpusu najdemo primere, ko je *namreč* uporabljen kot prevedek za različne angleškem jezikovne strukture, ki smo jih razdelili v 12 skupin (glej Tabelo 4):

⁶ Slovenski referenčni korpus *FidaPlus* je dostopen na <http://www.fidaplus.net/>.

⁷ Angleški korpus *BNC* najdemo na <http://www.natcorp.ox.ac.uk/>, medtem ko je korpus ameriške angleščine *COCA* dostopen na <http://corpus.byu.edu/coca/>.

TABELA 4: Primeri padajoče pogostosti različnih jezikovnih struktur v angleščini v podkorpusu SPOOK ENG-SLV, ki so jih prevajalci prevajali s pojasnjevalnim povezovalcem *namreč*; S = samostalnik, G = glagol.

<i>Namreč</i> kot prevedek za:	Število primerov v podkorpusu SPOOK ENG-SLV	Število primerov na 10.000 besed
1. 0 – primer dodajanja*	69	0,6095
2. for S + G	43	0,3705
3. because/as + (if) + S + (probably) + G	11	0,0957
4. I mean/Well, I've never/you see/I suppose/Now	7	0,0609
5. apparently + S + G	1	0,0086
6. by + S + G	1	0,0086
7. I think	1	0,0086
8. it's just (that) S + G	1	0,0086
9. on the grounds that	1	0,0086
10. such indeed is + S	1	0,0086
11. that is, for + S	1	0,0086
12. viz.	1	0,0086

*Primer dodajanja izraza *namreč*, ko v angleški predlogi ni podobnega jezikovnega sredstva.

V nadaljnji razpravi se bomo podrobneje osredotočili zgolj na tiste primere iz korpusa SPOOK, ki se v korpusu pojavijo vsaj petkrat.

Primeri različnih angleških jezikovnih struktur, ki so jih prevajalci prevajali z izrazom *namreč*:

1. 0 – primer dodajanja *

Harry, however, had never been less interested in Quidditch; he was rapidly becoming obsessed with Draco Malfoy.

*Spet ga je **namreč** obsedel Drecko Malfoy.*

That does not sound elvish to me.

*Ne zveni mi **namreč** kot vilinščina.*

They are not for all ears.

*Niso **namreč** za vsako uho.*

V teh primerih je dodajanje izraza *namreč* večinoma primer dodatnega pojasnjevanja oziroma eksplicitacije, saj bi ga lahko brez težav tudi izpustili in bi prevod pomensko še vedno ustrežal izvirniku. Ti primeri so izredno zanimivi, saj so lep prikaz prevoda, kjer se je prevajalec namerno odločil za dodajanje izraza *namreč*, ker je čutil, da je treba nekaj dodatno vzročno pojasniti, čeprav razlogov za to ni moč najti v angleški predlogi, saj v angleških stavkih ni izrazov, ki bi tudi sicer napeljevali na dodatno pojasnjevanje. Prav zaradi tega v teh primerih lahko govorimo o eksplicitaciji. V podkorpusu SPOOK ENG-SLV najdemo kar 69 primerov od skupaj 138 primerov prevodov z izrazom *namreč*, v katerih je prevajalec posegel po modifikaciji in se odločil dodati izraz *namreč*, čeprav ga na to niso napeljevale primerljive strukture v angleškem jeziku, torej se je za tako strategijo dodajanja oziroma eksplicitiranje odločal zavestno. Torej je natanko polovica primerov v tem korpusu takih, ko se je prevajalec odločal za eksplicitacijo. Je pa treba poudariti, da nekateri primeri vsebujejo časovne odvisnike, na primer *has ever been, at that moment, since*, ki so morda napeljevali prevajalca k dodatnemu pojasnjevanju z izrazom *namreč*, čeprav v luči slovarske definicije ta izraz ni njihova neposredna ustreznica.

2. for S + G

For coming unbidden to this place death is our law .

Naša postava namreč veleva smrt za nepovabljenca.

For it is plain that to you, at any rate, they are not strange .

Očitno je namreč, da vsaj tebi niso tuji.

Kot veznik se *for* v angleškem jeziku uporablja predvsem v bolj literarno obarvanih besedilih, zagotovo tudi v nekoliko zastarelem pomenu *zato ker/ zaradi*. Vsekakor angleški veznik *for* uporabljamo v primerih, ko želimo navesti ali pojasniti razloge za to, kar je navedeno pred tem, torej v prejšnji povedi. Ker se večina primerov v korpusu SPOOK s tem angleškim izrazom pojavlja v prevodih kultnega domišljjskega romana *The Lord of the Rings: The Two Towers*, torej *Gospodar prstanov: Stolpa* v prevodu Branka Gradišnika, kjer je avtor romana J. R. R. Tolkien namenoma uporabljal zelo svojstven, ponekod seveda nekako starinsko obarvan angleški jezik, da bi bolje pričaral domišljjski svet, je seveda lažje razumeti zelo pogosto rabo veznika *for* v smislu kavzalnega prislova *because*. Tako ne preseneča prevod z izrazom *namreč*, ki je sicer povsem ustrezen. Vendar pa bi ga prevajalec vsaj ponekod morda vseeno lahko prav tako uspešno nadomestil tudi s sicer markiranim knjižnim izrazom *zatorej* ali morebiti celo starinskim *zategadelj*, ki bi dodajal pridih nečesa nenavadnega, domišljjskega in obenem vzbudil občutek nekakšne starinskosti, torej nekega sveta tam nekje daleč vstran od navadnih smrtnikov.

3. **because / as + (if) + S + (probably) + G**

'You should record that and play it back every night, because you probably won't get another one.

*Po mojem tega **namreč** nikoli več ne boš doživel.*

For at such times, when material support is not enough, it is as if we are being judged

*...kadar nimamo zadostnih materialnih sredstev, se **namreč** počutimo, kakor da nas vsi ocenjujejo.*

Izraz *because*, še posebej v kombinaciji z modalnim *probably* vsekakor deluje kavzalno, torej nekaj vzročno pojasnjuje in v tem smislu se je prevajalec najverjetneje tudi odločil za prevod z vzročno-pojasnjevalnim izrazom *namreč*. Vendar takšno kombinacijo ponavadi uporabljamo pri razlagah v smislu *saj/zato/ker najverjetneje*, torej ni posebnega razloga za modifikacijo stavka z izrazom *namreč*. V drugem primeru lahko rabo *as if* razložimo v smislu *kot da* in tudi tu ni posebnega razloga za dodajanje pojasnjevalnega *namreč*. Tak primer bi morda tudi bilo moč uvrstiti med tiste, kjer v angleški predlogi ne najdemo izraza, ki naj bi ga prevajali z izrazom *namreč*, a se je prevajalec vseeno odločil za njegovo rabo in s tem za eksplcitacijo, saj je čutil potrebo po nekem dodatnem pojasnjevanju.

4. **I mean / Well, I've never**

To Stockholm, I mean.

*V Stockholm **namreč**.*

...'well, I've never mixed an antidote for a love potion..

*...**namreč**, še nikoli nisem namešal protizvarka za..*

V manj formalnih kontekstih lahko diskurzni označevalec *I mean* uporabljamo razlagalno oziroma kot jezikovno sredstvo za to, da pojasnimo ali popravimo, kar smo pravkar povedali. Lahko bi ga uporabili tudi s smislu *v bistvu* ali *torej*, morda pa nadomestili celo z bolj pogovornim *no*. V zgornjem primeru vsekakor vsebuje ta pomen, saj modificirana poved deluje v nekakšni razlagalni ali celo pojasnjevalni vlogi, zato mu prevedek z izrazom *namreč* ustreza. Podobno je v drugem primeru diskurzni označevalec *well* napeljal prevajalca, da je dodatno pojasnil stanje, tj. *da še nikoli ni namešal protizvarka*, in bi se ga dalo prevesti tudi z *v bistvu/pravzaprav* ali s *torej*, ki jih tukaj lahko razumemo v razlagalno-pojasnjevalni funkciji. Podobno velja tudi za primere z ostalimi diskurznimi označevalci, kot na primer *now/you see*, kjer so se prevajalci odločili za dodatno pojasnjevanje z izrazom *namreč*.

Morda bi veljalo kot zanimivost izpostaviti še primer z angleškim izrazom *viz.* in njegovim prevodom v slovenščino.

Instead of asking what the original author intended as clues, we should perhaps be asking an entirely different question, viz., what do the verses tell us about the person..

*..morali nemara postaviti povsem drugačno vprašanje, **namreč**, kaj nam verzi povedo o osebi..*

Zgornji primer je edini primer rabe prislova *viz.* (skrajšano za *vedelicet*), ki se v sodobnejših besedilih sicer ne pojavlja prav pogosto, vendar pa vsaj po slovarski definiciji v angleško-slovenskem slovarju sodeč ustreza prevedku z *namreč*, za kar je dokaz tudi naveden primer. To je tudi edini primer v celotnem korpusu SPOOK, kjer se izraz *namreč* pojavi kot prevedek zgolj enega izmed predlaganih ustreznih prevedkov v angleško-slovenskem dvojezičnem slovarju.

V analiziranih primerih lahko vidimo, da pojasnjevalne zveze v večini primerov z drugim delom prvega pojasnjujejo ali natančneje določajo. Glede vsebinske nadrejenosti je del zgradbe, ki pojasnjuje, največkrat podrejen, z besedilnega vidika je to ponavadi nekaj sekundarnega, vendar v smislu obvestilnosti ne redundantnega (prim. Gorjanc 1998). Vsekakor bi v določenih primerih to lahko ostalo tudi besedilno neizraženo. Besedilo brez pojasnjevanja se pravzaprav bistveno ne spreminja, zaradi dela, ki ostaja neizražen, pa postane zgolj informativno bolj zgoščeno (prim. Gorjanc 1998: 381–382). To potemtakem kaže, da je v zgoraj navedenih primerih modifikacija povedi z dodajanjem izraza *namreč* tako v vzročni kot v pojasnjevalno-vezniški vlogi primer eksplicitacije.

Najbolj presenetljiv zaključek te raziskave je, da se vzročno-pojasnjevalni povezovalc *namreč* pojavlja kot prevedek tudi tistih jezikovnih struktur, kjer ga ne bi nujno rabili, a ga prevajalci dodajajo eksplicitno, najverjetneje zato, da še dodatno podkrepijo in pojasnijo pomen določenega segmenta besedila ali pojasnijo vzrok za takšno stanje, medtem ko se kot prevedek za *namely*, ki ga kot ustreznico predlaga angleško-slovenski dvojezični slovar, ta izraz ne pojavi niti enkrat.

6 ZAKLJUČEK

Kot smo nakazali že v uvodnem delu, pojav hipoteze o eksplicitaciji kot prevodni univerzaliji že desetletja buri duhove med prevodoslovci in ima še vedno tako zagovornike kot tudi ostre kritike. Prevodoslovci se *namreč* še vedno ne strinjajo, kaj točno je eksplicitacija in v kolikšni meri bi ji lahko pripisovali neko univerzalno vrednost, ki je lastna samemu procesu prevajanja. Dosedanje raziskave vsekakor kažejo na to, da med prevajanjem prevajalci dejansko pogosto skušajo vzpostaviti ustrezno besedilno kohezivnost prevoda tako, da v besedilo vnašajo neke dodatne informacije, da bi prevod ciljna publika čim uspešneje sprejela. V kolikor je takšna modifikacija,

torej takšno vnašanje ali dodajanje del samega prevajalskega procesa in ne rezultat kontrastivnih razlik med nekim jezikovnim parom je še vedno težko določiti.

V članku smo najprej predstavili pojem prevodnih univerzalij in se natančneje posvetili eksplicitaciji. Potem smo kot primer za analizo izbrali vzročno-pojasnjevalni povezovalac *namreč* in skozi prizmo slovarske definicije in dejanskih primerov iz prevodoslovnega korpusa SPOOK skušali osvetliti rabo tega povezovalca in ugotoviti, kako se obnaša v prevodih iz angleškega jezika. Bi v analiziranih primerih lahko govorili o eksplicitaciji? V primerih, ko so prevajalci dodajali *namreč* takrat, ko za to ni bilo posebnega razloga v angleških predlogah, najverjetneje to drži, saj je videti, da so prevajalci posegali po takšni modifikaciji, torej dodajanju zato, da bi ciljnemu občinstvu olajšali branje in razumevanje prevoda.

Najbolj presenetljiva je skorajda popolna diskrepanca med predlogi slovarskih prevedkov za *namreč* in dejanskem stanju v korpusu SPOOK. Čeprav naj bi pri prevajanju iz slovenskega jezika v angleški jezik v luči slovarske definicije v slovensko-angleškem dvojezičnem slovarju pojasnjevalni povezovalac *namreč* najpogosteje prevajali z angleškim prislovom *namely*, saj ga pod tem geslom najdemo na prvem mestu, in obratno, se primeri s prevodi, ki vključujejo pojasnjevalni izraz *namreč* v slovenskem delu korpusa SPOOK zelo razlikujejo od tega, kar kot ustreznico navaja slovensko-angleški dvojezični slovar. Kot rečeno, v angleško-slovenskem podkorpusu SPOOK ENG-SLV ne najdemo niti enega primera s prevedkom *namreč*, ki naj bi bil prevod za angleški prislov *namely*. Vsi primeri, ki so prevajani s tem jezikovnim sredstvom, tj. *namreč*, so prevodi nekih drugih angleških izrazov. To je vsekakor zelo zanimiva informacija, še posebej, ker študenti prevajalstva kar preradi posegajo po izrazu *namely*, ko prevajajo v angleški jezik, najverjetneje prav zaradi tega, ker je v slovensko-angleškem slovarju pod geslom *namreč* kot prva prevodna ustreznica naveden *namely*. Morda ozaveščanje o rabi tega jezikovnega sredstva, kar je jasno nakazala tudi naša raziskava, lahko prispeva k temu, da bodo bodoči prevajalci pri prevajanju iz slovenskega v angleški jezik takrat, ko se v slovenskem besedilu pojavi *namreč*, zavestno manj pogosto posegali po prevodu angleškega prislova *namely*. Obenem pa raziskava nakazuje tudi nujnost nadgradnje in izboljšave do sedaj dostopnih dvojezičnih slovarskih virov za slovensko-angleški jezikovni par, še posebej za potrebe prevajanja iz jezika A v jezik B.

Pym (2005a: 32), podobno tudi Becher (2011), razume eksplicitacijo v povezavi z implicitacijo. Ko povzema razmišljanja Klaudy in Karoly (2005), pride do zaključka, da je eksplicitacija lahko obvezna ali pa ne, je torej izbirna ali namerna; prav ta druga možnost je izredno zanimiva za prevodoslovje, saj je specifična za prevajanje, ne pa nujno tudi za določen jezikovni par. Prevajalec se *namreč* lahko v vsakem primeru znova in znova posebej odloča za možnost eksplicitacije, če bo torej neko poved modificiral in nekaj dodatno pojasnjeval, ali pa ne bo posegel

po taki strategiji. Če bi torej uspeli potrditi tudi implicitacijo med prevajanjem iz slovenščine v angleščino, ki naj bi v luči hipoteze o asimetriji delovala v obratni smeri (prim. Klaudy, Karoly 2005), bi o takih primerih verjetno res lahko govorili kot o možnih kandidatih za eksplicitacijo v smislu prevodne univerzalije (prim. raziskavo nominalizacije v Mikolič Južnič 2010: 149). Žal je vzorec, ki smo ga izbrali za analizo, odločno premajhen, da bi lahko prišli do bolj splošnih zaključkov, kaj šele, da bi lahko govorili o potrditvi prevodne univerzalije. Pravzaprav se zdi, da gre v analiziranih primerih prej za neko prevajalčevo osebno strategijo, kot pa za splošno prevajalsko tendenco po eksplicitaciji.

Vendarle pa ta raziskava vseeno nakazuje, da je moč tudi s tovrstnimi 'majhnimi' raziskavami osvetliti določene jezikovne strukture in jih je zato primerno predstaviti tudi v pedagoškem okolju. Študenti so načeloma bolj naklonjeni eksperimentalno podprtim primerom kot pa učiteljevi intuiciji, zato bi jih s takšnimi rezultati morda lažje uspeli prepričali o tem, na kaj vse je treba biti pozoren, ko se med prevajanjem odločamo o neki prevodni ustreznici.

Literatura

- Bajec, Anton in sod., 2000: *Slovar slovenskega knjižnega jezika*. Ljubljana: DZS.
- Baker, Mona, 1993: Corpus Linguistics and Translation Studies: Implications and Applications. Baker, Mona, Gill Francis in Elena Tognini-Bonelli (ur.) *Text and Technology: In Honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins. 233–250.
- Baker, Mona, 1996: Corpus-based translation studies: The challenges that lie ahead. Somers, Harold (ur.): *Terminology, LSP and Translation. Studies in Language Engineering in Honour of Juan C. Sager*. Amsterdam/Philadelphia: John Benjamins. 175–186.
- Baumgarten, Nicole, Bernd Meyer in Demet Özçetin, 2008: Explicitness in Translation and Interpreting. A Review and Some Empirical Evidence (of an Elusive Concept). *Across Languages and Cultures* 9/2. 177–203.
- Becher, Viktor, 2010a: Abandoning the notion of “translation-inherent” explicitation. Against a dogma of translation studies. *Across Languages and Cultures* 11/1. 1–28.
- Becher, Viktor, 2010b: Towards a more rigorous treatment of the Explicitation Hypothesis in translation studies. *trans-kom* 3/1. 1–25.
- Becher, Viktor, 2011: When and why do translators add connectives? A corpus-based study. *Target* 23/1. 26–47. http://viktor.becher.bz/downloads/papers/2011/becher_2011__02bec.pdf. (Dostop 4. 5. 2012)
- Blum-Kulka, Shoshana, 1986: Shifts of Cohesion and Coherence in Translation. House, Juliane in Shoshana Blum-Kulka (ur.): *Interlingual and Intercultural Communication*. Tübingen: Narr, 17–35.

- Claudiu, Mihăilă, 2010: Translation studies: Simplification and Explication Universals. <http://www.slideshare.net/claudiumihaila/report-3832657>. (Dostop 4. 5. 2012)
- Chesterman, Andrew, 2004: Hypotheses about translation universals. Hansen, Gyde, Malmkjær Kirsten, in Gile, Daniel (ur.): *Claims, Changes and Challenges in Translation Studies*. Amsterdam: Benjamins. 1–13.
- Černelič Ivanka, 1991: Členek kot besedna vrsta v slovenskem knjižnem jeziku. *Jezikovni zapiski*. 73–85.
- Černelič Kozlevčar, Ivanka, 1993: O delitvi členkov. *Zbornik Slavističnega društva Slovenije* 3. 213–225.
- Englund Dimitrova, Brigitta, 2005: *Expertise and Explication in the Translation Process*. Amsterdam/Philadelphia: John Benjamins.
- Fabricsius-Hansen, Cathrine, 1998: Informational Density and Translation, with Special Reference to German – Norwegian – English. Johansson, Stig in Signe Oksefjell (ur.): *Corpora and Cross-linguistic Research. Theory, Method, and Case Studies*. Amsterdam etc.: Rodopi. 197–234.
- Fabricsius-Hansen, Cathrine, 2005: Elusive Connectives. A Case Study on the Explicitness Dimension of Discourse Coherence. *Linguistics* 43/1. 17–48.
- Fišer, Darja, 2010: Semantično označevanje korpusov. Vintar, Špela (ur.): *Slovenske korpusne raziskave*. Ljubljana: Znanstvena založba Filozofske fakultete. 110–130.
- Frankenberg-Garcia, Ana, 2004: *Are translations longer than source texts? A corpus-based study of explication*. <http://www.linguatca.pt/Repositorio/Frankenberg-Garcia2004.doc/>. (Dostop 2. 2. 2012.)
- Gorjanc, Vojko, 1998: Konektorji v slovničnem opisu znanstvenega besedila. *Slavistična revija* 46/4. 367–388.
- Gorjanc, Vojko, 2005: *Uvod v korpusno jezikoslovje*. Domžale: Izolit.
- Gorjanc, Vojko in Špela Vintar, 2007: Korpusna analiza vloge označevalcev medleksemskih razmerij v organizaciji besedila. *Jezik in slovstvo*. 52/3–4. 117–129.
- Grad, Martin, 2010: Izražanje osebnosti v akademskem diskurzu: primerjava rojenih in tujih govorcev angleščine. Vintar, Špela (ur.): *Slovenske korpusne raziskave*. Ljubljana: Znanstvena založba Filozofske fakultete. 72–89.
- Grad, Anton, Ružena Škerlj in Nada Vitorovič, 1997: *Veliki angleško-slovenski slovar*. Ljubljana: DZS
- Grad, Anton, Henry Leeming, 1997: *Slovensko-angleški slovar*. Ljubljana: DZS.
- Gumul, Ewa. 2006: Explication in Simultaneous Interpreting: A Strategy or a By-product of Language Mediation? *Across Languages and Cultures* 7/2. 171–190.
- Hansen-Schirra, Silvia, Stella Neumann in Erich Steiner, 2007: Cohesive Explicitness and Explication in an English-German Translation Corpus. *Languages in Contrast* 7/2. 205–229.
- Hirci, Nataša, 2007: *Učinkovitost uporabe sodobnih prevajalskih virov pri prevajanju v nematernalni jezik*. PhD dis., Ljubljana: Filozofska fakulteta.

- Hirci, Nataša, 2009: Možnosti uporabe sodobnih prevajalskih virov med študenti prevajalstva. Mikolič, Vesna (ur.): *Jezikovni korpusi v medkulturni komunikaciji*. Koper: Univerza na primorskem, Znanstveno-raziskovalno središče, Založba Annales. 117–132.
- Hirci, Nataša, 2012: Electronic Reference Resources for Translators: Implications for Productivity and Translation Quality. *Interpreter and Translator Trainer* 6/2, 219–235.
- House, Juliane, 2004: Explicitness in Discourse across Languages. Koller Werner, Juliane House in Klaus Schubert (ur.): *Neue Perspektiven in der Übersetzungs- und Dolmetsch-wissenschaft*. Bochum: AKS, 185–208.
- Jakop, Nataša, 2000: Vežanost poudarnih členkov na določeno besedno vrsto oziroma stavčni člen. *Jezikoslovni zapiski* 6, 67–80.
- Jurko, Primož, 2007: Divergent polysemy: the case of Slovene namreč vs. English namely. Komar, Smiljana in Uroš Mozetič (ur.): *ELOPE: Studies in the English language and literature in Slovenia*. Ljubljana: Filozofska fakulteta, 29–41.
- Klaudy, Kinga, 1996: Back Translation as a Tool for Detecting Explicitation Strategies in Translation. Klaudy, Kinga, Lambert, José in Sohár, Anikó (ur.): *Translation Studies in Hungary*. Budapest: Scholastica. 99–114.
- Klaudy, Kinga, 2008: Explicitation. Baker, Mona in Gabriela Sladanh (ur.): *Routledge Encyclopaedia of Translation Studies*. London in New York: Routledge. 104–108.
- Klaudy, Kinga in Krisztina Karoly, 2000: The Text-organizing Function of Lexical Repetition in Translation. Olohan, Maeve (ur.): *Intercultural Faultiness. Research Models in Translation Studies I: Textual and Cognitive Aspects*. Manchester: St. Jerome. 143–159.
- Klaudy, Kinga in Krisztina Karoly, 2005: Implication in Translation: Empirical evidence for operational asymmetry in translation. *Across Languages and Cultures* 6/1. 13–28.
- Klinar, Stanko, 1994: *Gradivo za vaje iz prevajanja v angleščino*. (3. Izdaja). Ljubljana: Filozofska fakulteta.
- Konšalová, Petra, 2007: Explicitation as a Universal in Syntactic De/condensation. *Across Languages and Cultures* 8/1. 17–32.
- Mauranen, Anna, 2000: Strange Strings in Translated Language: A Study on Corpora. Olohan, Maeve (ur.): *Intercultural Faultiness. Research Models in Translation Studies I: Textual and Cognitive Aspects*. Manchester: St. Jerome. 119–141.
- Mikolič Južnič, Tamara, 2010: Kontrastivni in prevodoslovni pogledi na nominalizacijo skozi italijansko-slovenski vzporedni korpus. Vintar, Špela (ur.): *Slovenske korpusne raziskave*. Ljubljana: Znanstvena založba Filozofske fakultete. 132–155.
- Mikolič Južnič, Tamara, 2011: Vpliv besedilnih tipov na pojavljanje nominalizacije v slovenščini: korpusna raziskava. Kranjc, Simona (ur.): *Meddisciplinarnost v slovenistiki*. Obdobja 30. Ljubljana: Znanstvena založba Filozofske fakultete. 321–327.

- Olohan, Maeve in Mona Baker, 2000: Reporting that in Translated English: Evidence for Subliminal Processes of Explication? *Across Languages and Cultures* 1/2. 141–158.
- Øverås, Linn, 1998: In Search of the Third Code: an Investigation of Norms in Literary Translation. *Meta* 43/4. 557–570.
- Pápai, Vilma, 2004: Explication: A Universal of Translated Text? Mauranen, Anna in Pekka Kujamäki (ur.): *Translation Universals. Do They Exist?* Amsterdam/Philadelphia: John Benjamins. 143–164.
- Pisanski Peterlin, Agnes, 2009: Izražanje svojilnosti v prevedeni angleščini: korpusna analiza. Mikolič, Vesna (ur.): *Jezikovni korpusi v medkulturni komunikaciji*. Koper: Univerza na primorskem, Znanstveno-raziskovalno središče, Založba Annales. 105–116.
- Pisanski Peterlin, Agnes, 2010: Se za strukturiranje besedila v prevodih uporabljajo drugačni elementi kot v izvirniku? Korpusna analiza medpovednega in mestavčnega in. Vintar, Špela (ur.): *Slovenske korpusne raziskave*. Ljubljana: Znanstvena založba Filozofske fakultete. 156–177.
- Pokorn, K. Nike, 2005: *Challenging the Traditional Axioms: Translation into a non-mother tongue*. Amsterdam/Philadelphia: John Benjamins Publishing.
- Pokorn, K. Nike, 2009: Proučevanje literarnih prevodov. Kocijančič Pokorn, Nike (ur.): *Sodobne metode v prevodoslovnem raziskovanju*. 40–58.
- Puurtinen, Tiina, 2003: Explicating and Implicitating Source Text Ideology. *Across languages and Cultures* 4/1. 53–62.
- Pym, Anthony, 2005a: *Explaining explication*. Krisztina Karoly in Ágata Fóris (ur.): *New Trends in Translation Studies. In Honour of Kinga Klauďy*. Budapest: Akadémia Kiadó. 29–34. http://www.tinet.org/~apym/online/translation/explication_web.pdf/. (Dostop 10. 3. 2012.)
- Pym, Anthony, 2005b: Text and Risk in Translation. Aijmer, Karin in Cecilia Alvstad (ur.): *New Tendencies in Translation Studies*. Göteborg: Göteborg University. 69–82.
- Pym, Anthony, 2008: On omission in simultaneous interpreting: Risk analysis of a hidden effort. Hansen, Gyde, Andrew Chesterman in Heidrun Gerzymisch-Arbogast (ur.): *Efforts and Models in Interpreting and Translation Research*. Amsterdam/Philadelphia: John Benjamins. 83–105.
- Saldanha, Gabriela, 2008: Explication Revisited: Bringing the Reader into the Picture. *trans-kom* 1/1. 20–35. http://www.trans-kom.eu/bd01nr01/trans-kom_01_01_03_Saldanha_Explication.20080707.pdf. (Dostop 30. 4. 2012).
- Schlamberger Brezar, Mojca, 2009: *Povezovalci v francoščini: od teoretičnih izhodišč do analize v diskurzu*. Ljubljana: Znanstvena založba Filozofske fakultete, Oddelek za prevajalstvo.
- Schlamberger Brezar, Mojca, 2011: Slovenski členki v luči Ducrotove integrirane pragmatike. Kranjc, Simona (ur.): *Meddisciplinarnost v slovenistiki. Obdobja 30*. Ljubljana: Znanstvena založba Filozofske fakultete. 405–411.

- Skubic, Andrej, 1999: Ogled kohezijske vloge slovenskega členka. *Slavistična revija* 47/2. 210–238.
- Smolej, Mojca, 2004: Členki kot besedilni povezovalci. *Jezik in slovstvo* 49/5. 45–57.
- Steiner, Erich, 2005: Explicitation, Its Lexicogrammatical Realization, and Its Determining (Independent) Variables – towards an Empirical and Corpus-based Methodology. http://www.hf.uio.no/forskningssprojekter/sprik/docs/pdf/Report_36_ESteiner.pdf. (Dostop 21. 4. 2012)
- Steiner, Erich, 2008: Empirical Studies of Translations as a Mode of Language Contact. “Explicitness” of Lexicogrammatical Encoding as a Relevant Dimension. Siemund, Peter in Noemi Kintana (ur.): *Language Contact and Contact Languages*. Amsterdam/Philadelphia: John Benjamins. 317–345.
- Toporišič, Jože, 1974: Kratko oblikoslovje slovenskega knjižnega jezika. *Seminar slovenskega jezika, literature in kulture*. Ljubljana: Oddelek za slovanske jezike in književnosti. 29–50.
- Toporišič, Jože, 1991 [1976]: *Slovenska slovnica*. 3. ponatis. Maribor: Obzorja.
- Toporišič, Jože, 2000: *Slovenska slovnica*. Maribor: Obzorja.
- Toury, Gideon, 1995: *Descriptive Translation Studies and Beyond*. Amsterdam/Philadelphia: John Benjamins.
- Vanderauwera, Ria, 1985: *Dutch Novels Translated into English: The Transformation of a ‘Minority’ Literature*. Amsterdam: Rodopi.
- Vinay, Jean-Paul in Jean Darbelnet, 1995 [1958]. *Comparative Stylistics of French and English: A Methodology for Translation*. Prevod Juan C. Sager in Marie-Josée Hamel. Amsterdam/Philadelphia: John Benjamins.
- Vintar, Špela in Darja Fišer, 2009: Gradnja in analiza korpusov za prevodoslovne raziskave. Kocijančič Pokorn, Nike (ur.): *Sodobne metode v prevodoslovnem raziskovanju*. 80–109.
- Vintar, Špela, 2009: Slovenski prevodoslovni korpus. Marko Stabej (ur.): *Infrastruktura slovenščine in slovenistike. Obdobja* 28. Ljubljana: Znanstvena založba Filozofske fakultete. 385–391.
- Vintar, Špela, 2010: Luščenje terminologije iz angleško-slovenskih vzporednih in primerljivih korpusov. Vintar, Špela (ur.): *Slovenske korpusne raziskave*. Ljubljana: Znanstvena založba Filozofske fakultete. 37–53.
- Vintar, Špela, 2011: Samodejno odkrivanje semantičnih premikov v prevodih. Kranjc, Simona (ur.): *Meddisciplinarnost v slovenistiki. Obdobja* 30. Ljubljana: Znanstvena založba Filozofske fakultete. 525–531.

Samostalnik *človek* med generično in specifično rabo v korpusu prevedenih in izvirnih besedil v slovenščini

Agnes Pisanski Peterlin

Oddelek za prevajalstvo, Filozofska fakulteta Univerze v Ljubljani
agnes.pisanski@guest.arnes.si

Abstract

In recent years, Translation Studies research has shown that the language of translations differs from that of comparable target language originals. Interference is generally accepted to be among the most common processes that contribute to the specific characteristics of translations. The aim of this paper is to establish whether it is possible to identify differences in the use of the Slovene noun *človek* (*human being/man*) between translated texts and comparable originals. The comparison focuses on the frequency and functions of the noun *človek* in three subcorpora of *Spook*, a multilingual parallel and comparable corpus of literary Slovene texts. By comparing the use of *človek* in texts translated from English with its use in texts originally written in Slovene, four functions of the noun in question are identified: indefinite generic, species-generic, specific and ascriptive. For all the functions, the corpus analysis shows that there are differences in the frequency of use between translated texts and comparable originals. An attempt is made to account for these differences by comparing the instances of *človek* in translations with the corresponding passages in the English originals. Interference is found to play an important role in accounting for the differences.

Keywords: generic use, specific use, the slovene noun *človek*, translated slovene, corpus-based studies

Ključne besede: generična raba, specifična raba, samostalnik *človek*, prevedena slovenščina, korpusna analiza

1 UVOD

V okviru prevodoslovja se je v zadnjih dveh desetletjih vrsta raziskovalcev ukvarjala z vprašanjem, ali je s sistematično primerjavo prevodov in primerljivih izvirnikov v ciljnem jeziku mogoče zaznati značilnosti, ki so tipične za jezik prevodov (npr. Mauranen 2000, Olohan in Baker 2000, Tirkkonen-Condit 2004, če naštejemo le nekatere). Več študij, ki so bile narejene na primerljivih besedilnih korpusih, je tako pokazalo, da je v prevodih v resnici mogoče zaznavati precejšnja odstopanja od primerljivih izvirnikov v ciljnem jeziku, kar so raziskovalci praviloma pripisali procesu nastanka prevodov, torej procesu prevajanja.

Verjetno je najpogostejše zaznani pojav v jeziku prevodov interferenca, to je vpliv izvirnega jezika skozi izvirno besedilo na ciljno besedilo (in posledično morda celo na ciljni jezik). Glede na pogostost interference ni presenetljivo, da prevodoslovci ta pojav opisujejo kot neločljivo povezan s prevajanjem: Toury (1995: 274–297) govori o interferenci kot o zakonitosti prevajanja, pogledi na interferenco pa so zelo različni: v preteklosti je imel izraz izrazito slabšalno konotacijo, s Touryjevim deskriptivnim pristopom pa postaja nevtralnejši (prim. tudi Mauranen 2007: 43).

Namen tega prispevka je s pomočjo korpusne analize, izvedene na delu korpusa *Spook*, ugotoviti, ali je samostalnik *človek* v besedilih, ki so v slovenščino prevedena iz angleščine, zaradi interference rabljen drugače kot v primerljivih besedilih, ki so bila izvirno napisana v slovenščini. Samostalnik *človek* je bil za to primerjavo izbran zato, ker predstavlja primer splošnega samostalnika¹, ki se v slovenščini pojavlja v različnih funkcijah, posamezne funkcije pa se ne pokrivajo povsem s funkcijami vzporednih angleških samostalnikov *human being* in *man*. Korpus *Spook* je za tovrstno primerjavo posebej primeren, saj vsebuje le literarna besedila, pri katerih kakovost prevoda ni vprašljiva, kar lahko pri tovrstnih analizah v nekaterih drugih žanrih predstavlja problem, saj se vprašanje interference v takih primerih tesno prepleta z vprašanjem kakovosti prevoda. Analiza korpusa literarnih besedil, kakršen je *Spook*, pa omogoča preučevanje interference brez vrednostnih sodb o kakovostnem standardu prevoda². V nadaljevanju prispevka so najprej podrobneje opisane različne funkcije samostalnika *človek*, nato sta predstavljena korpus *Spook* in metoda dela, sledi strnjena predstavitev rezultatov, v zadnjem delu prispevka pa diskusija in glavni sklepi.

¹ Giacalone Ramat in Sansò (2007: 95–96) izpostavita pomen splošnih samostalnikov z osnovnim pomenom *človek*, *oseba*, *posameznik*, *stvar*, *kraj* ali *način*, ki predstavljajo enega od virov za nedoločne zaimke v različnih jezikih. Primeri tovrstnih nedoločnih zaimkov so francoski *on*, angleška *nobody*, *nothing*, francoski *rien*, nemški *man*, latinski *nemo* itd. (prim. tudi Lehmann 2002: 44–50 za natančnejšo razlago razvoja posameznih zaimkov).

² Interferenca v okviru Touryjevega pristopa sama po sebi ni nujno negativna, tolerantnost do interference pa je v veliki meri odvisna od sociokulturnih dejavnikov (prim. Toury 1995: 278, Pym 2008: 321).

2 PREGLED LITERATURE

Samostalnik *človek* je zanimiv primer splošnega samostalnika, ki se uporablja tako v generičnem kot v specifičnem pomenu. Primerjava slovenskega samostalnika *človek* s sorodnimi samostalniki tipa *man* (*človek*) pokaže zanimive vzporednice med slovenščino in drugimi evropskimi jeziki.

O generični rabi nekega samostalnika govorimo v tistih primerih, ko se izbrani izraz očitno nanaša na množico, vrsto ali razred posameznikov, o specifični pa, kadar se nanaša na konkretnega posameznika ali predmet. V nekaterih primerih ni jasno, ali je samostalnik rabljen specifično ali generično, in takrat lahko prihaja do dvoumnosti. Stavek *Iščem svinčnik* je dvoumen, saj lahko pomeni, da govorec išče nek konkretni svinčnik (specifična raba), svinčnik se v tem primeru nanaša na nek konkretni predmet. Če pa govorec izraža željo, da bi našel kakršen koli primerek predmeta, ki bi ga lahko poimenovali svinčnik, pa tovrstna raba ni specifična, temveč generična: izraz svinčnik se v tem primeru nanaša na kateri koli predmet iz množice svinčnikov.

Samostalnik *človek* sodi med tiste splošne samostalnike, ki v številnih jezikih prek generične rabe skozi proces gramatikalizacije³ prehajajo v zaimke (prim. npr. Giacalone Ramat in Sansò 2007, Ragnarsdóttir in Strömquist 2005, Vi-berg 2010). V svojem pregledu širitve in upada nedoločnih oblik tipa *man* (*človek*) Giacalone Ramat in Sansò (2007) za evropske jezike ugotavljata, da je stopnja gramatikalizacije tovrstnih samostalnikov v različnih jezikih zelo različna. Močno so prisotne na tako imenovanem »področju Karla Velikega«, od koder prehajajo tudi v zahodno- in južnoslovanske jezike, medtem ko njihova raba na področju vzhodnoslovanskih jezikov ni izpričana. Giacalone Ramat in Sansò (2007) skozi zgodovinski pregled pokažeta, da je bila raba nedoločnih oblik tipa *man* (*človek*) zgodovinsko bolj razširjena, saj je bila prisotna v več germanskih in romanskih jezikih kot danes. Zelo znan primer germanskega jezika, v katerem je ta konstrukcija nekoč obstajala, danes pa je izginila, je angleščina. Od romanskih jezikov je ohranjena predvsem še v sodobni francoščini (*on*), čeprav je zgodovinsko izpričana v številnih starejših variantah drugih romanskih jezikov. Giacalone Ramat in Sansò (2007) v procesu gramatikalizacije konstrukcije tipa *man* (*človek*) v različnih jezikih opišeta več stopenj, ki jih natančneje povzemam spodaj, pri čemer za vsako posamezno rabo za ilustracijo navajam tudi po en njun primer; pri primerih, ki niso v slovenščini, je poleg angleškega prevoda naveden tudi slovenski.⁴ Prva stopnja je

³ Giacalone Ramat in Sansò (2007: 126–7) poudarjata, da polne razlage za ta proces gramatikalizacije še ni, podatki pa kažejo, da je vir za opisani proces pri samostalnikih tipa *man* (*človek*) najverjetneje pozna latinščina in stiki med jeziki, podrobnejše podatke o tem, kateri stiki, pa bo treba še pridobiti.

⁴ Kjer ni drugače navedeno, so slovenski prevodi moji. V zgledih, ki so povzeti po literaturi, so v primerih, kjer končna ločila niso bila uporabljena, le-ta dodana.

raba splošnega samostalnika tipa *man* (*človek*) za generični opis človeške vrste (torej človek v nasprotju z npr. drugimi živimi bitji, v nasprotju z Bogom itd.). Primer, ki ga navajata Giacalone Ramat in Sansò (2007), je iz latinščine, iz Matejevega evangelija (4,4):

(1) *Non in solo pane vivit **homo**.*

(1a) 'Man does not live by bread alone.'

(1b) 'Človek naj ne živi samo od kruha.' (Slovenski standardni prevod)

Druga stopnja je raba samostalnika tipa *man* (*človek*) v nedoločni generični rabi (v smislu *kdorkoli*). Ob drugih primerih te rabe Giacalone Ramat in Sansò (2007) za ilustracijo navajata tudi stavek v slovenščini in angleški prevod:

(2) *V taki situaciji bi **človek** iskal drugo službo.*

(2a) 'In such a situation **one** would look for another job.'

Tretjo stopnjo predstavlja raba konstrukcije tipa *man* (*človek*) v nedoločni referenčni rabi (v smislu *nekdo*). Takšno rabo poznajo nekateri germanski jeziki (švedščina, nemščina) in je v svojem bistvu zelo podobna trpniku. Primer, ki ga navajata Giacalone Ramat in Sansò (2007), je švedski in je povzet po Altenberg (2005: 94):

(3) ***Man** hat mordat Palme.*

(3a) 'Someone murdered Palme. / Palme has been murdered.'

(3b) 'Nekdo je umoril Palmeja. / Palme je bil umorjen/Palmeja so umorili.'

Četrta stopnja se pojavlja le v nekaterih jezikih. Predstavlja jo raba konstrukcije tipa *man* (*človek*) za določni zaimsek za prvo osebo (množine ali celo ednine), v nekaterih jezikih celo za druge osebe. Najbolj sistematična je tovrstna raba *on* v francoščini. Primer, ki ga navajata Giacalone Ramat in Sansò (2007), je povzet po Grafström (1969: 273):

(4) ***On** a du pain pour nos vieux jours.*

(4a) 'We've got bread for our old days.'

(4b) 'Imamo kruha za stare dni.'

Poleg generične in specifične rabe se samostalnik *človek* lahko pojavi tudi v predikativni rabi (za veznim glagolom, npr. *biti*), a v tem primeru ne označuje entitete, temveč lastnost. Takšna raba je pripisovalna (samostalnik osebkju pripisuje neko lastnost) in ni referenčna (prim. tudi Quirk in sod. 1992: 273), saj se tak samostalnik ne nanaša niti na vrsto (generična raba), niti na posameznika (specifična raba). Spodnji primeri iz Bond in sod. (1995: 2) v angleščini ponazarjajo razlike med generično (5), specifično (6) in pripisovalno (7) rabo za samostalnik *mammoth* (*mamut*); k vsakemu primeru je za pojasnilo dodan slovenski prevod.

- (5) *Mammoths are extinct.*
 (5a) 'Mamuti so izumrli.'
 (6) *There's a mammoth in my garden!*
 (6a) 'Na mojem vrtu je mamut.'
 (7) *That animal is a mammoth.*
 (7a) 'Ta žival je mamut.'

3 KORPUS IN METODA

3.1 Korpus

V analizi je zajet del korpusa *Spook*. Korpus *Spook* je večjezični korpus, ki obsega dva večja podkorpusa, in sicer vzporedni podkorpus izvirnih sodobnih literarnih besedil (romanov) in njihovih prevodov v slovenščino, ki ga sestavljajo štirje deli, angleško-slovenski, nemško-slovenski, francosko-slovenski in italijansko-slovenski, ter primerljivi podkorpus izvirnih slovenskih sodobnih literarnih besedil (romanov). Korpus *Spook* je nastal na Oddelku za prevajalstvo Filozofske fakultete Univerze v Ljubljani ter na Odseku za tehnologije znanja IJS med 2009 in 2012 v okviru projekta *Slovensko prevodoslovje – viri in raziskave*. Vsi deli korpusa so oblikoskladenjsko označeni, vzporedni deli korpusa so stavčno poravnani.

V pričujoči analizi je bilo v prvi fazi analizirano gradivo iz dveh delov korpusa *Spook*. Prvi del obsegajo slovenski prevodi literarnih besedil, izvirno napisanih v angleškem jeziku (v nadaljevanju bom za ta del uporabila oznako *Sl prevodi*), drugi del pa primerljivi slovenski izvirniki, prav tako literarna besedila (v nadaljevanju bom za ta del uporabila oznako *Sl izvirniki*). Vsa prevedena besedila so v slovenščino prevedli prevajalci, ki so rojeni govorci slovenščine, vsa izvirna besedila so napisali avtorji, ki so prav tako rojeni govorci slovenščine. V tabeli 1 so predstavljeni podatki o obsegu izbranih delov korpusa.

Tabela 1: Podatki o obsegu izbranih delov korpusa

	Sl izvirniki	Sl prevodi
Okvirno število besed v korpusu	1.454.275	1.148.419

V drugi fazi analize je bil ob prevodih vključen tudi pripadajoči del angleških izvirnikov, toda ker primerjava ni kvantitativna, podatkov o obsegu tega dela korpusa ne navajam posebej.

3.2 Metoda

V prvi fazi analize so bili s pomočjo korpusnega iskalnika identificirani vsi primeri izraza *človek* (v ednini) v sklopih *Sl prevodi* in *Sl izvirniki*. Vsi primeri so bili nato ročno pregledani in glede na pomen razvrščeni po funkcijah. V korpusu so bile identificirane štiri funkcije samostalnika *človek*, in sicer **specifična**, **pripisovalna**, **generična v smislu predstavnik človeške vrste** in **nedoločna generična**. Funkcije so opredeljene tako, kot so opisane v literaturi (prim. razdelek 2). Surovi kvantitativni rezultati so bili preračunani v obliko števila pojavitev na 10.000 besed, kar omogoča kvantitativno primerjavo različnih kategorij rabe.

V drugi fazi analize so bili s pomočjo poravnave prevodov in njihovih pripadajočih angleških izvirkov identificirani izvorni izrazi, ki so bili prevedeni z izrazom *človek*. Natančen pregled izvirkov je dal dodaten vpogled v razloge za razlike med sklopoma *Sl izvirniki* in *Sl prevodi*.

4 REZULTATI

Kvantitativni rezultati korpusne analize so predstavljeni v razdelku 4.1, rezultati primerjave prevodov in pripadajočih angleških izvirkov po posameznih funkcijah pa v 4.2.

4.1 Pogostost rabe posameznih funkcij

V tabeli 2 so predstavljeni kvantitativni rezultati korpusne analize za oba sklopa besedil, *Sl izvirniki* in *Sl prevodi*.

Tabela 2: Pogostost funkcij samostalnika *človek*

	Sl izvirniki		Sl prevodi	
	Absolutno število	/10.000 besed	Absolutno število	/10.000 besed
Nedoločna generična raba	311	2,1	189	1,6
Generični predstavnik človeške vrste	129	0,9	83	0,7
Pripisovalna raba	244	1,7	113	1,0
Specifična raba	123	0,8	110	1,0
Skupaj	807	5,5	495	4,3

Za vsak sklop besedil je v prvem stolpcu navedeno skupno število pojavitev in število pojavitev po posameznih rabah. V drugem stolpcu je za vsak sklop besedil preračunano povprečno število pojavitev na 10.000 besed v posameznih rabah, s čimer postanejo podatki primerljivi ne glede na razlike v dolžini med sklopoma besedil.

4.2 Primerjava prevodov in pripadajočih angleških izvirkov

V tabeli 3 je predstavljen pregled pripadajočih angleških izvirkov, ki so bili v slovenščino prevedeni s samostalnikom *človek*.

Tabela 3: Pregled pripadajočih angleških izvirkov po funkcijskih kategorijah

	Pogosto (10 primerov ali več)	Manj pogosto (manj kot 10 primerov)	Posamezni primeri (ne več kot en primer)
Nedol. gen. raba	0*, one, you, man	person, people, we, I, anyone	men, fellow, he, everyone
Gen. predstav. čl. v.	man/Man	person, human, human beings, people, Men, human being	anthropomorphic form, mortal, manmade, one
Pripisovalna	man, person, 0*	type, one	people, bloke, soul, mortal, someone, character
Specifična	man	someone, person, fellow, 0*, one	body, human being, men, human, he, guy

* V kategorijo, označeno z 0, so zajeti tisti primeri, za katere v izvirkih ni mogoče določiti konkretnega samostalnika oz. zaimka, ki bi bil izhodišče za prevod s *človek*, to pa so neosebne strukture, nominalizacije, neosebne glagolske oblike, trpniki ipd., pa tudi primeri, v katerih je bil v prevodu samostalnik *človek* dodan.

Pripadajoči angleški izvirki so razvrščeni v tri skupine, ki so v tabeli predstavljene v obliki stolpcev. Za vsako rabo so v prvem stolpcu navedeni pogosti

pripadajoči angleški izvirniki, to so tisti izrazi, ki se pojavijo vsaj v desetih primerih. V drugem stolpcu so ločeno po rabah navedeni manj pogosto rabljeni pripadajoči izvirniki, ki se pojavljajo več kot enkrat, vendar manj kot desetkrat. V tretjem stolpcu so za vsako rabo posebej navedeni še tisti pripadajoči angleški izvirniki, ki se v posamezni rabi pojavijo le po enkrat. Vsak sklop primerov je razvrščen po pogostosti, pogostejši izrazi so navedeni pred manj pogostimi.

5 DISKUSIJA

Iz tabele 2 je razvidno, da se samostalnik *človek* v podkorpusu prevodov v celoti pojavlja nekoliko manj pogosto (4,3 primera na 10.000 besed) kot v podkorpusu izvirnikov (5,5 primera na 10.000 besed). V prevodih je torej primerov samostalnika *človek* v celoti približno 20 % manj kot v izvirnikih, kar ni zanemarljiva razlika. V nadaljevanju so podrobneje razčlenjeni rezultati za posamezne rabe samostalnika *človek*.

5.1 Nedoločna generična raba

Primerjava pogostosti posameznih funkcij pokaže, da je samostalnik *človek* v obeh podkorpusih najpogostejše rabljen v nedoločni generični rabi, torej v smislu *kdorkoli*. V podkorpusu izvirnikov je takšna raba izrazito pogosta, saj se pojavljata 2,1 primera tovrstne rabe na 10.000 besed, kar je malo manj kot polovica vseh primerov, v podkorpusu prevodov pa je tovrstna raba nekoliko manj prevladujoča (1,6 primera na 10.000 besed), a prav tako predstavlja najpogostejšo funkcijo tega samostalnika. V prevodih je torej primerov samostalnika *človek* v nedoločni generični rabi približno za eno četrtno manj kot v izvirnikih. Za boljše razumevanje razlike v pogostosti med podkorpusoma je smiselno analizirati, v katerih primerih se kot prevod angleškega izvirnika pojavi slovenski samostalnik *človek* v nedoločni generični rabi.

Primerjava prevodov in pripadajočih angleških izvirnikov pokaže, da kategorija nedoločne generične rabe v prevodih pravzaprav izhaja iz dveh tipov angleških ustreznice, ki pa jih v prevodih ni več mogoče ločevati. Velika večina ustreznice (150 primerov oziroma 79 % vseh primerov nedoločne generične rabe) v celoti ustreza nedoločni generični rabi splošnega samostalnika *človek*, kakršno opredeljujeta Giacalone Ramat in Sansò (2007), torej v smislu *kdorkoli* (prim. razdelek 2). V korpusu *Spook* so najpogostejši pripadajoči angleški izvirniki kategorija, označena z 0, generični *you* ali generični *one*. Ker predstavlja kategorija, označena z 0, heterogeno

skupino oblik, v rabi katerih se slovenščina in angleščina močno razlikujeta (načeloma pa velja, da so vsaj nekatere med njimi, npr. neosebne glagolske oblike, v slovenščini rabljene manj pogosto, prim. npr. Pisanski Peterlin 2007: 259), zagotovo ni presenetljivo, da se te oblike pogosto prevajajo z osebnimi konstrukcijami s *človek* v nedoločni generični rabi. Oba preostala pogosta izvirnika, generični *you* in *one*, sta pričakovana (prim. npr. van der Auwera 2012); *you* je seveda bolj pogovoren in manj formalen, *one* pa izrazito formalen in v literaturi pogosto naveden kot angleška ustreznica generičnemu samostalniku tipa *man* (*človek*) v drugih jezikih (prim. npr. Altenberg 2005). Tudi za kombinacijo slovenščina-angleščina Kocijančič Pokorn⁵ (1997: 13) na podlagi kontrastivne analize pokaže, da sta za izražanje generične osebe najpogostejši prevodni ustreznici angleški zaimsek *one* in slovenski generični *človek*. S primerjavo angleškega izvirnika s slovenskim prevodom identificira generično rabo samostalnika *človek* kot najpogostejši slovenski prevod za angleški generični *one* (kot možni prevodni ustreznici v gradivu identificira še neosebno obliko *se* ter generično drugo osebo ednine, *ti*). Tudi obratna primerjava slovenskega izvirnika in angleškega prevoda pokaže tesno povezavo med *one* in *človek*: v prevodu v angleščino je zaimsek *one* edina prevodna ustreznica za generični *človek*.

Spodnji primeri (8–10) ilustrirajo to vrsto rabe samostalnika *človek* v podkorpusu prevodov (korpus *Spook*), ob vsakem primeru je dodan še pripadajoči izvirnik.

(8) *Naravno je, da **človek** čuti nekaj radovednosti glede teh reči... čarovnike določenega kalibra ta plat magije od nekdanj privlači...*

(8a) *It is natural **to feel** some curiosity about these things... Wizards of a certain caliber have always been drawn to that aspect of magic...*

(9) *In ne pozabi na čelado: **človek** nikoli ne ve.*

(9a) *And wear your fuzz plates; **you** never know.*

(10) *Seveda ni nobenega zagotovila o bona fides in v bistvu se mora **človek** zanašati na instinkt; vendar pa ima podjetje malo težav.*

(10a) *There was no guarantee of bona fides, of course, and basically **one** had to rely on gut-reaction; but the firm experienced very few problems, really.*

V 39 primerih (to je 21 % vseh primerov nedoločne generične rabe) je pripadajoča ustreznica slovenskemu nedoločnemu generičnemu samostalniku *človek* angleški samostalnik *man* (ali *person* oz. *fellow*), ki se v pravi nedoločni generični rabi sicer v angleščini ne pojavlja. Pri nekaterih od teh primerov je mogoče zaznati premik v prevodu: ne v smislu napake, temveč v smislu drugega poudarka (primer 11).

(11) *In če **človek** začne razmišljati o tistem delčku svojih možganov, ki gleda homunkulusa na zaslonu, se na zaslonu pojavi ta del možganov, zaslon pa začne opazovati spet kak drug delček možganov.*

⁵ Omenjena raziskava se sicer osredotoča na angleški izraz *one* in se s problematiko generičnih samostalnikov tipa *man* (*človek*) ne ukvarja.

(11a) *And when the **person** thinks about this part of the brain (the bit that is watching the homunculus on the screen) they put this bit of the brain on the screen and there is another bit of the brain watching the screen.*

V primeru (11) je slovenski prevod mogoče razumeti generično: če **človek** začne razmišljati v smislu če začneš ali začnemo razmišljati, torej če kdorkoli začne razmišljati, angleški izvirnik pa ni generičen, saj je govora o specifični (vendar hipotetični) osebi (prim. t.i. psevdogenerično rabo samostalnika *man*, kot jo opiše Fuertes-Olivera 2007: 226). V takšnih primerih se rabi zelo približata. V drugih primerih angleški *man*, ki sicer sam po sebi ne izraža generičnega pomena, takšen pomen pridobi, ker je rabljen v stavku, ki izraža generičnost (primer 12), ali z zaimkom, ki nakazuje generičnost (v smislu *kdorkoli*, kot v primeru 13).

(12) *Kadar je mraz, **človek** vidi svojo sapo, kadar je vroče, je pa ne.*

(12a) *On cold days **a man** can see his breath, on a hot day he can't.*

(13) *Bilo je jasno: kar se je njega tikalo, **človek**, ki lahko pogleda Harryja in reče »imenitno«, pač ne sodi v njegovo družbo.*

(13a) *It was clear that as far as he was concerned, **any man** who could look at Harry and say "excellent" was a man with whom he could never see eye to eye.*

Tako slovenščina kot angleščina imata bogat nabor sredstev za izražanje generičnosti, vendar pa se ta dva nabora med seboj močno razlikujeta. Toporišič (2000: 503) govori o »splošnih vršilcih dejanja«, med katere uvršča tretjo osebo množine, zaimenski *človek* v funkciji osebka (pri čemer kot vzporedne navaja nemški *man*, angleški *one* in francoski *on*), splošnoosebni *se*, trpnik, prvo osebo množine, drugo osebo ednine itd. Za angleščino Johansson (2007: 175) izpostavlja v prvi vrsti trpnik, osebne zaimke (*we*, *you*, *they*) in generični zaimek *one*. Za pričujočo primerjavo je zlasti pomembno dejstvo, da angleščina nima generičnega samostalnika tipa *man* (*človek*) v nedoločni rabi, ki bi bil najbolj očitna ustreznica slovenskemu nedoločnemu generičnemu samostalniku *človek*. Zaimek *one*, ki je, kot je bilo omenjeno zgoraj, najbližja angleška ustreznica generičnemu samostalniku tipa *man* (*človek*) v drugih jezikih, pa je v nasprotju s slovenskim generičnim samostalnikom *človek* mnogo bolj formalen in zato v svoji rabi bolj omejen. Glede na povedano je bolj omejena raba samostalnika *človek* v nedoločni generični funkciji v prevodih v primerjavi z izvirniki pravzaprav pričakovana, in je v skladu s *hipotezo unikatnih elementov* (*Unique Items Hypothesis*), ki jo je postavila Sonja Tirkkonen-Condit (2004). Ta predpostavlja, da se tista sredstva, ki so lastna le ciljnemu jeziku, ne pa tudi izvirnemu jeziku, v prevodih pojavljajo precej manj pogosto kot v izvirnikih, napisanih v istem jeziku. Laviosa (2008: 125) meni, da je mogoče to hipotezo uvrstiti v okvir interference, kot Touryjeve zakonitosti prevajanja.

5.2 Pripisovalna raba

Druga najpogostejša funkcija samostalnika *človek* je v obeh podkorpusedih pripisovalna, a razlika v pogostosti rabe je med podkorpuseda precejšnja. V slovenskih izvirnikih je tovrstna raba mnogo pogostejša (na 10.000 besed se pojavi okrog 1,7 primera takšne rabe) kot v prevedenih besedilih (približno en primer na 10.000 besed); to pomeni, da je v prevodih primerov samostalnika *človek* v pripisovalni funkciji 40 % manj kot v primerljivih izvirnikih.

Primerjava prevodov in pripadajočih angleških izvirnikov pokaže, da je najpogostejše izhodišče za prevod s samostalnikom *človek* v pripisovalni rabi samostalnik *man* (npr. kot v primeru 14, ki je naveden spodaj), na drugem mestu je samostalnik *person*, razmeroma pogost je tudi tovrsten prevod kategorije, označene z 0. Redko, a več kot enkrat, se v izvirniku pojavita še *type* in *one*, v korpusu pa se pojavi tudi nekaj posameznih angleških izhodišč za takšen prevod.

(14) *Kdo bi vedel? Jih je že tako ali drugače razjezil. Dober človek je bil, ta Florean.*

(14a) *Who knows? He must've upset them somehow. He was a good man, Florean.*

Zdi se, da prav na podlagi angleških izvirnikov morda lahko sklepamo o razlogih za razmeroma veliko razliko med slovenskimi izvirniki in prevodi. Dejstvo, da so razmeroma pogosti prevodi, v katerih se v slovenščini pojavi *človek* v pripisovalni rabi, v angleščini pa je namesto tega rabljena neka drugačna struktura ali pa ustrezne sploh ni, priča o tem, da so prevajalci samostalnik *človek* v pripisovalni rabi v prevodih dodajali. To nakazuje, da je morda v slovenščini na splošno pripisovalna raba samostalnika *človek* pogostejša, v angleščini pa se enako razmerje lahko izraža še z drugimi sredstvi (primer 17), ali pa se samostalnik *človek* enostavno niti ne izrazi (primera 15 in 16). Kljub dodajanju samostalnika *človek* v pripisovalni rabi v prevodih je tovrstne funkcije v prevedenih besedilih morda manj prav zato, ker se s tem odraža interferenca iz angleščine.

(15) *Sploh pa so ljudje dobesedno stali v vrsti, da bi mi izdali razne packarije iz njegovega življenja. Niso vsi tako prepričani, da je bil prekrasen človek, veš – stopil je na kar nekaj pomembnih prstov.*

(15a) *People were queuing to dish the dirt on Dumbledore anyway. Not everyone thought he was so wonderful, you know – he trod on an awful lot of important toes.*

(16) *David je imel rad tega moža, Fredericka Burkeja, ki je bil prijazen, čeprav zadržan človek, kakor je bila tudi Davidova mati prijazna in zadržana.*

(16a) *David liked this man, Frederick Burke, who was kind, if remote, like his mother, who was kind and remote.*

(17) *McHudurra resda ni bila **človek**, ki bi znal prisluhnuti; Dumbledore, though in many ways more intimidating, still seemed less likely to scorn a theory, however wild.*

(17a) *Professor McGonagall did not invite confidences; in čeprav je Dumbledore v mnogih ozirih zbujal več strahospoštovanja, bi bilo pri njem manj verjetno, da bi se posmihal še tako divji teoriji.*

5.3 Generična raba v smislu predstavnika človeške vrste

V podkorpusu slovenskih izvirnikov je tretja najpogostejša funkcija samostalnika *človek* generična raba tega samostalnika v smislu predstavnika človeške vrste. Tovrstna raba je v tem podkorpusu manj pogosta in obsega nekaj manj kot en primer (0,9) na 10.000 besed. V podkorpusu prevodov je takšna funkcija še redkejša (0,7 primera na 10.000 besed), pravzaprav je v tem podkorpusu to najredkejša funkcija samostalnika *človek*. V prevodih je tako primerov samostalnika *človek* v smislu predstavnika človeške vrste okrog 20 % manj kot v izvirnikih.

Primerjava prevodov in pripadajočih angleških izvirnikov pokaže, da je daleč najpogostejši pripadajoči izvirnik samostalnika *človek* v smislu predstavnika človeške vrste angleški izraz *man* (občasno zapisano tudi z veliko začetnico, *Man*), kot npr. v primeru (18). Razmeroma redki so primeri, ko so s *človek* prevedeni izrazi *person*, *human*, *human beings*, *people* in *Men* (zapisano z veliko začetnico), vsak od njih se pojavi manj kot desetkrat (kot npr. primer (19)), ob njih se pojavlja še peščica izrazov, ki se najdejo le po enkrat.

(18) *Sveto pismo je napisal **človek**, draga moja, ne Bog.*

(18a) *The Bible is a product of **man**, my dear. Not of God.*

(19) *Je sveti gral **človek**?*

(19a) *The Holy Grail is **a person**?*

Dejstvo, da je prav ta funkcija samostalnika *človek* v prevodih najredkejša, je zanimivo. Gotovo ga ni mogoče pripisati izboru romanov v posameznem podkorpusu, saj podkorpus prevodov vsebuje fantazijski roman J. R. R. Tolkiena, ki se v podrobnosti ukvarja z vprašanjem človeške vrste v primerjavi z drugimi imaginarnimi bitji, kot so vilini in orki, mladinski fantazijski roman J. K. Rowling, ki prav tako pogosto primerja človeško vrsto s fantastičnimi bitji, ter uspešnico Dana Browna, ki se ukvarja z vprašanjem človeka in Boga, in bi že na podlagi tega dejstva pričakovali, da bo tovrstna raba v podkorpusu prevodov bogato zastopana.

Prav tako razmeroma redke tovrstne rabe v prevodih ni mogoče neposredno pripisovati leksikogramatikalnim razlikam med angleščino in slovenščino, saj generični *man* v smislu predstavnika človeške vrste v angleščini seveda obstaja. Vendar pa drži, da je v sodobni angleščini takšna raba samostalnika *man* nekoliko zaznamovana, medtem ko to za slovenski samostalnik *človek* ne velja, razlika je torej pragmatična. Ker v angleščini samostalnik *man* lahko pomeni tako *človek* kot *moški*, sodobni slogovni priročniki v izogib seksizmu pogosto priporočajo nadomeščanje izraza *man* s kakšnim drugim nevtralnijšim izrazom (npr. *human*, *human being*, ali z množinsko obliko, *people*, ali z drugo parafrazo), prim. npr. (20) spodaj. Posledično je raba *man* v smislu predstavnika človeške vrste v sodobnih angleških besedilih vse redkejša, kar so že pokazali rezultati različnih študij, ki so zaznale močno upadanje pogostosti rabe androcentričnih generikov (Fuertes-Olivera 2007: 222).

(20) *Ko se **človek** okuži s tem virusom, samo sedi v naslanjaču in se sploh ne more premakniti, nič ne pije in ne je, zato kmalu umre.*

(20a) *And when **people** get the virus they just sit on the sofa and do nothing and they don't eat or drink and so they die.*

Takšnega upada v rabi samostalnika *človek* v smislu predstavnika človeške vrste v slovenščini seveda ni. Vendar pa prav trend v angleščini pomeni, da se tudi v prevodih večkrat pojavi kakšna druga oblika, ne *človek*, pod vplivom angleškega izvirnika. To je npr. razvidno iz primera (21), v katerem se kot prevedek za *human beings* pojavita izraza *človek* in *ljudje*.

(21) *Ljudje, ki verjamejo v boga, mislijo, da je bog ustvaril **človeka**, ker mislijo, da je **človek** najpopolnejše živo bitje, ampak tudi **človek** je samo živalska vrsta in se bo razvil v drugo živalsko vrsto, ki bo pametnejša od **človeka** in bo **ljudi** zaprla v živalske vrtove, kot smo mi zaprli v živalske vrtove šimpanze in gorile.*

(21a) *And people who believe in God think God has put **human beings** on the earth because they think **human beings** are the best animal, but **human beings** are just an animal and they will evolve into another animal, and that animal will be cleverer and it will put **human beings** into a zoo, like we put chimpanzees and gorillas into a zoo.*

5.4 Specifična raba

Najredkejša funkcija samostalnika *človek* v podkorpusu slovenskih izvirnikov je specifična funkcija, v tej funkciji se samostalnik *človek* pojavi v 0,8 primera na 10.000 besed. V podkorpusu prevodov je ista funkcija zastopana nekoliko pogostejše (približno en primer na 10.000 besed), kar je enako pogosto, kot se v tem

korpusu pojavlja pripisovalna raba, in pogostejše od generične rabe tega samostalnika v smislu predstavnika človeške vrste. V prevodih se torej samostalnik *človek* v specifični rabi v povprečju pojavlja za četrtno pogostejše kot v izvirnikih.

Primerjava prevodov in pripadajočih angleških izvirnikov pokaže, da je najpogostejše izhodišče, ki se prevede v *človek* v specifični rabi, izraz *man*. V manj kot desetih primerih se pojavijo še *someone* (nedoločna specifična raba), *person* in *fellow*, ostali izvirniki se pojavijo po enkrat. Primer (22) ilustrira prevod specifičnega *man* v slovenščino s *človek*:

(22) *Ta **človek** je bil okrutno prebičan!*

(22a) *This **man** has been mercilessly whipped!*

Razlog za opaženo razhajanje med podkorpusoma ni povsem očiten. Znova bi utegnili prihajati do sprememb zaradi dvojnega pomena angleškega *man* (*človek* in *moški*). Prevajalci se morda v specifičnem pomenu večkrat odločijo za samostalnik s širšim pomenom, *človek*, tudi v primerih, ko bi bil v slovenščini lahko uporabljen (in morda tudi pogostejša izbira) *moški*. Tvrsten primer je (23):

(23) *In res, za Percyjevimi koraki je pošepaval tisti **človek**, ki ga je Harry videval v Preroških novicah, in grivo sivečih las in črni plašč mu je pikljal sneg.*

(23a) *And sure enough, the **man** Harry had seen in the Daily Prophet was following along in Percy's wake, limping slightly, his mane of graying hair and his black cloak flecked with snow.*

V nekaterih primerih je interferenca iz angleščine še bolj očitna. Raba samostalnika *človek*, kakršna se pojavi v primeru (24), je v slovenščini seveda možna, a izrazito redkejša kot v angleščini.

(24) *Brani pred čim, **človek**?*

(24a) *Defend himself against what, **man**?*

6 SKLEP

Namen pričujočega prispevka je bil s pomočjo korpusne analize ugotoviti, ali je samostalnik *človek* v literarnih besedilih, ki so v slovenščino prevedena iz angleščine, pod vplivom interference rabljen drugače kot v primerljivih literarnih besedilih, ki so bila izvirno napisana v slovenščini. Razčlemba rezultatov korpusne analize je takšne razlike v resnici identificirala, in sicer tako v smislu splošne pogostosti (v prevodih je samostalnik *človek* rabljen redkeje kot v slovenskih izvirnikih) kot tudi v pogostosti rabe posameznih funkcij (v prevodih je samostalnik *človek* v vseh funkcijah razen v specifični rabljen redkeje kot v primerljivih izvirnikih). Ob tem se pojavlja vprašanje o razlogih za takšna odstopanja: na podlagi primerjave

prevodov s pripadajočimi angleškimi izvirniki je bilo mogoče sklepati o nekaterih možnih vzrokih, ki so pogosto izvirali v interferenci, v nekaterih primerih je bilo mogoče ta sklepanja dopolniti tudi z drugimi podatki, pridobljenimi iz literature.

Pričujoča študija predstavlja nujen prvi korak k natančnejšemu razumevanju rabe samostalnika *človek* v izvirni in prevedeni slovenščini, saj prinaša razmero-ma natančen popis posameznih rab tega samostalnika v slovenščini, na podlagi katerega bo mogoče primerjavo še nadgraditi. Treba je namreč upoštevati, da v analizi zaradi obsega »ročnega« razvrščanja po kategorijah ni bilo mogoče izkoristiti celotnega potenciala korpusa *Spook*, primerjava je zato obsegala le prevode iz angleščine. Za pridobitev bolj tehtnih podatkov o rabi samostalnika *človek* v prevedenih besedilih je logičen nadaljnji korak primerjava prevodov iz različnih jezikov. Prav takšna primerjava bo namreč omogočila pridobivanje dodatnih podatkov o razlogih za razlike v rabi in boljši vpogled v to, v kolikšni meri je mogoče zaznane razlike pripisovati interferenci. S primerjavo prevodov iz različnih jezikov bi bilo namreč vsaj teoretično mogoče zaznati tudi primere, v katerih se pojavljajo podobna odstopanja od primerljivih izvirnikov, ki očitno niso vezana na izvirni jezik in niso povezana z interferenco.

Literatura

- Altenberg, Bengt, 2005: The generic person in English and Swedish. A contrastive study of *one* and *man*. *Languages in Contrast* 5(1). 93–120.
- Bond, Francis, Kentara Ogura in Tsukasa Kawaoka, 1995: Noun phrase reference in Japanese-to-English machine translation. *Proceedings of the Sixth International Conference on Theoretical and Methodological Issues in Machine Translation (TMI '95)*. 1–14. <http://xxx.lanl.gov/abs/cmp-lg/9601008> (Dostop 22. 4. 2012.)
- Giacalone Ramat, Anna, in Andrea Sansò, 2007: The spread and decline of indefinite man -constructions in European languages: An areal perspective. Paolo Ramat in Elisa Roma (ur.) *Europe and the Mediterranean as Linguistic Areas: Convergencies from a historical and typological perspective*. Amsterdam / Philadelphia: John Benjamins Publishing Company. 95–131.
- Grafström, Åke, 1969: On remplaçant nous en français. *Revue de Linguistique Romane* 33. 270–298.
- Fuertes-Olivera, Pedro A., 2007: A corpus-based view of lexical gender in written Business English. *English for Specific Purposes* 26(2). 219–234.
- Johansson, Stig, 2007: *Seeing through multilingual corpora: On the use of corpora in contrastive studies*. Amsterdam / Philadelphia: John Benjamins Publishing Company.
- Kocijančič Pokorn, Nike, 1997: A Slovene-English contrastive analysis of *one*. *Slovenski jezik/Slovene Linguistic Studies* 1. 5–23.

- Laviosa, Sara, 2008: Description in the translation classroom. Anthony Pym, Miriam Shlesinger in Daniel Simeoni (ur.) *Beyond Descriptive Translation Studies*. Amsterdam / Philadelphia: John Benjamins Publishing Company. 119–132.
- Lehmann, Christian, 2002: *Thoughts on grammaticalization*. Druga izdaja. Erfurt: Seminar für Sprachwissenschaft der Universität (ASSidUE, 9).
- Mauranen, Anna, 2000: Strange strings in translated language. A study on corpora. Maeve Olohan (ur.) *Intercultural Faultlines. Research Models in Translation Studies 1: Textual and Cognitive Aspects*. Manchester, UK in Northampton, MA: St Jerome. 119–142.
- Mauranen, Anna, 2007: Universal tendencies in translation. Gunilla Anderman in Margaret Rogers (ur.) *Incorporating Corpora: The Linguist and the Translator*. Multilingual Matters Ltd. : Clevedon. 32–48.
- Olohan, Maeve, in Mona Baker, 2000: Reporting *that* in translated English: Evidence for subconscious processes of explicitation? *Across Languages and Cultures* 1(2). 141–158.
- Pisanski Peterlin, Agnes, 2007: Grammatical forms of text-organising metatext: A Slovene-English contrastive analysis. *Slovenski jezik/Slovene Linguistic Studies* 6. 251–265.
- Pym, Anthony, 2008: On Toury's laws of how translators translate. Anthony Pym, Miriam Shlesinger in Daniel Simeoni, (ur.) *Beyond Descriptive Translation Studies*. Amsterdam / Philadelphia: John Benjamins Publishing Company. 311–328.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, in Jan Svartvik, 1992: *A Comprehensive Grammar of the English Language*. Deseti ponatis. (1. izdaja 1985.) London in New York: Longman.
- Ragnarsdóttir, Hrafnhildur, in Sven Strömqvist, 2005: The development of generic *maður/man* for the construction of discourse stance in Icelandic and Swedish. *Journal of Pragmatics* 37(2). 143–155.
- Tirkkonen-Condit, Sonja, 2004: Unique items – over- or under-represented in translated language? Anna Mauranen and Pekka Kujamäki (ur.) *Translation Universals: Do they exist?* Amsterdam / Philadelphia: John Benjamins Publishing Company. 177–184.
- Toporišič, Jože, 2000: *Slovenska slovnica*. (1. izdaja 1976.) Maribor: Obzorja.
- Toury, Gideon, 1995: *Descriptive translation studies and beyond*. Amsterdam/Philadelphia: John Benjamins.
- van der Auwera, Johan, 2012: From contrastive linguistics to linguistic typology. *Languages in Contrast*. 12 (1). 69–86.
- Viberg, Åke, 2010: Swedish Impersonal Constructions from a Crosslinguistic Perspective: An Exploratory Corpus-based Study. *Orientalia Suecana* 59. 122–158.

Neosebne glagolske oblike v prevodni in izvirni slovenščini: primer deležja

Tamara Mikolič Južnič

Oddelek za prevajalstvo, Filozofska fakulteta Univerze v Ljubljani
tamara.mikolicjuznic@ff.uni-lj.si

Abstract

The paper focuses on a particular type of Slovene non-finite verbal forms, i.e. the adverbial participle called *deležje* and its frequency in original and in translated Slovene literary works. Although it has been often stated that the adverbial participle is a rare feature in contemporary Slovene and although Slovene native speakers often feel a certain degree of awkwardness towards it and seem to tend to avoid it, the assumptions mentioned are mostly based on intuition rather than real evidence, as no comprehensive study about the use of the Slovene participle has yet been done. In this paper, the analysis focuses on the frequency of the adverbial participle in literary texts, differentiating among original and translated texts. The data was collected from the Spook corpus, which comprises a set of original Slovene literary texts and a set of subcorpora of translations from several languages into Slovene. Since we shall be concentrating on the translations from Italian, a comparison between the examples of adverbial participles in translated texts and their original counterparts will also be presented in order to establish whether there are any links between the parts of speech or syntactical structures of the original and the occurrence of the participle in the translations.

Keywords: Slovene adverbial participle, translation corpus, translated Slovene, original Slovene, Italian

Ključne besede: deležje, prevodni korpus, prevodna slovenščina, izvirna slovenščina, italijanščina

1 UVOD

Ena od najzanimivejših (in najpomembnejših) možnosti, ki nam jih ponujajo korpusi, je preverjanje tako širše uveljavljenih kot novonastalih intuitivnih prepričanj in hipotez o jeziku. V nasprotju s tradicionalnim jezikoslovnim pristopom se korpusni pristop (in še bolj radikalno korpusno jezikoslovje nasploh) namreč osredotoča na sistematično proučevanje avtentičnih primerov jezika v rabi (Sinclair 1997: 31).

Ta prispevek je namenjen preverjanju hipoteze omenjene vrste: predvsem upamo, da bomo lahko vsaj za obravnavani besedilni tip dobili svež vpogled v vprašanje, koliko se deležja v sodobni slovenščini dejansko uporabljajo in ali so res tako redka, kot to intuitivno trdi marsikateri govorec slovenskega jezika, poleg tega pa nameravamo empirično preveriti, ali je pri uporabi deležij kaka razlika med izvirnimi in prevedenimi slovenskimi besedili.

Raziskava je bila izvedena v okviru projekta Slovensko prevodoslovje: raziskave in viri (prim. Vintar 2009) na korpusu Spook, ki združuje večjezični vzporedni korpus ter primerljivi korpus slovenskega jezika. Vzporedni korpusi, če povzamemo Erjavca (1997: 83), so primerljivi korpusi, torej korpusi primerljivih besedil (določenega žanra ali področja), ki vključujejo tako izvirna besedila kot njihove prevode. Prednosti vzporednih korpusov se pokažejo zlasti pri kontrastivnih in prevodoslovnih študijah, saj ponujajo nezamenljiv pogled v izvirno in prevedeno besedilo hkrati. Primerljivi korpusi pa so zbirke primerljivih (torej istovrstnih) besedil v dveh ali več jezikih.

Ker želimo preveriti rabo neosebne glagolske oblike v vlogi nosilnih elementov stavčne strukture, je korpus Spook idealno orodje za tako zasnovano raziskavo: njegova zgradba (ki bo natančneje opisana v razdelku Korpus in metoda) nam omogoča, da preverimo na eni strani pojavljanje deležja v izvirnih, na drugi strani pa v prevedenih slovenskih besedilih. Dodatna možnost, ki smo jo uporabili pri analizi, pa je sočasno opazovanje prevedenih izsekov besedila in njihovih izvirnikov, saj smo preverili tudi, ali je pojavljanje deležja v prevedenih besedilih morebiti povezano s pojavljanjem specifičnih struktur v izvirnih besedilih.

2 NEOSEBNE GLAGOLSKÉ OBLIKE V SLOVENŠČINI IN DELEŽJE

Tako neosebne glagolske oblike nasploh kot deležje specifično niso bili pogosto predmet jezikoslovnih raziskav o slovenskem jeziku. Poleg Breznikovih zapisov

(zbranih v Breznik 1967) z začetka 20. stoletja in Jesenškovih (npr. 1998) diahronih razprav o rabi deležnikov in deležij je zaslediti nekaj kontrastivnih študij o neosebni glagolskih oblikah oz. njihovi polstavčni rabi v povezavi z različnimi evropskimi jeziki (za primerjavo z angleščino gl. npr. Milojević Sheppard 1992, Kocijančič Pokorn in Šuštaršič 2001, 1999 ter Šuštaršič in Kocijančič Pokorn 1998; s francoščino npr. Schlamberger Brezar 2005a in 2005b, nazadnje pa Mezeg 2011). Na splošno bi lahko iz omenjenih del sklepali, da je raba polstavkov v slovenščini dokaj omejena, saj se ob prevajanju iz tujih jezikov (npr. angleščine in francoščine), v katerih so tovrstni polstavki zelo pogosti, prevajalci pogosto zatekajo k alternativnim ubeseditvam.

V skladu z Jesenškovimi raziskavami je pogostnost deležnikov in deležij na –č in –ši v knjižni besedi zrasla tekom 19. stoletja pod vplivom prekmurskega prevoda nove zaveze (Jesenšek 1998: 12), nato pa proti koncu stoletja začela upadati in praktično izginila iz sodobnega jezika Ivana Cankarja in drugih pisateljev, zlasti zaradi kritičnega pogleda na rabo polstavkov v slovenščini Antona Breznika. Breznik (1967: 82–94) je skoraj vse vrste polstavčnih struktur odprto kritiziral z navedbo primerov iz sodobnih avtorjev, saj jih je razumel kot francosko sintaktično interferenco¹ ter je svetoval rabo odvisnih stavkov z osebnimi glagolskimi oblikami. Kot takratno stanje povzema Jesenšek (1998: 15), so sodobni pisatelji njegovo mnenje v veliki meri sprejeli in *»oblika je skoraj povsem izginila; njena raba se je omejila le na posamezne glagole, in to predvsem za stilno zaznamovano izražanje«*.

Da se je deležniško-deležijska tipologija ohranila v slovenskih slovnica, gre velika zasluga Jožetu Toporišiču² (1991), ki je slovenske polstavčne tvorbe v svojih delih opredelil in opisal, ni pa posvetil izrecne pozornosti njihovi dejanski rabi, pogostnosti in sprejemljivosti. Deležje je v skladu z Žagarjevo (2001: 290) definicijo nepregibna (prislovna) oblika glagola; Toporišič (1991: 339) specificira, da deležja na –č in –aje izražajo *»glagolsko dejanje, ki se vrši hkrati s kakim drugim dejanjem«*, deležje na –e pa se *»večinoma rabi kot prislov načina«*. Deležijski polstavek je struktura, ki vsebuje deležje in *»izraža glagolsko dejanje, ki poteka hkrati z drugim dejanjem«* (Žagar 2001: 291). Nadalje Jesenšek (1998: 41) tudi ugotavlja, da se deležja pojavljajo zlasti v *»umetnostnih in strokovnih besedilih predvsem pri glagolih govorjenja, mišljenja, čustvenega zaznavanja in premikanja«*, poleg tega pa (ibid.) njihova raba peša in *»je omejena predvsem na pisno izražanje istodobnosti dogajanja glavnega stavka in deležijskega polstavka«*. Schlamberger Brezar (2005b: 127) gre še dlje in ugotavlja, da je raba deležja (primer, ki ga navaja, je *Sedeč na vrtu, je bral časopis*) *»daleč od naravne rabe v*

¹ Pojem interferenca razumemo kot pojav, pri katerem se površinske značilnosti izvirnika prenesejo v ciljno besedilo (prim. Toury 1995: 275).

² Toporišičevo izvirno tipologijo so pozneje prevzeli drugi avtorji, med katerimi je tudi Žagar (2001), ki ga navajamo v nadaljevanju.

govorjenem jeziku in da deluje starinsko, nerodno.» V svoji raziskavi političnih besedil ugotavlja, da se tovrstne strukture relativno pogostejše pojavljajo v sodobnih prevedenih besedilih, manj pa v izvirnih (ibid: 126–128).

Priponska obrazila, ki se pojavljajo pri deležjih v deležijskih stavkih, so *–oč/–eč*, *–aje* in *–(v)ši*. Zadnji v skladu z obstoječo strokovno literaturo praktično ni več del sodobnega jezika, saj je ta oblika »zelo redka, omejena na pisni jezik, s predznakom knjižno zastarelo« (Jesenšek 1998). Poleg omenjenih priponskih obrazil srečamo tudi deležja na *–e*, ki se jih pogosto upošteva kot prislove (ali tudi predloge) z glagolsko etimologijo.³ V opisano raziskavo so namenoma vključeni tudi primeri s tem zadnjim obrazilom, saj se zdi zanimivo vprašanje, ali se take oblike pojavljajo pogostejše od ostalih oz. ali je njihovo obnašanje v besedilih drugačno od ostalih deležij. Spodaj je za boljšo orientacijo navedenih nekaj primerov, ki so bili naključno izbrani iz korpusa Gigafida:

- deležje na *–eč*: *Sicer niso bili vsi dosegljivi za komentar, toda sodeč po izjavah mnogih bi še.*
- deležje na *–oč*: *To se na fotografiji tudi vidi: gledam torto, komaj čakajoč, da jo bom jedel.*
- deležje na *–aje*: *Živeti je treba, ne oziraje se na to, da padajo granate ...*
- deležje na *–ši*: *... na drugem pa je tedanja opozicija formirala alternativo politiko, začenši z Ruplovo demokrasko in Omanovo kmečko zvezo...*
- deležje na *–e*: *Snažilka je miže pokimala in molčala.*

Kot ugotavlja tudi Mezeg (2011: 54, 73), ki nam ponuja nadroben pregled slovenskega deležja v okviru kontrastivne raziskave o francoskih in slovenskih polstavkih, celovite raziskave, ki bi pokazala, kako pogosto se deležja in deležijski polstavki v sodobni slovenščini dejansko uporabljajo, ter s katero bi lahko posameznikov jezikovni čut nadomestili z empiričnimi podatki iz večjih korpusov, zaenkrat še ni. Avtorica sicer tudi ugotavlja, da prisotnost deležijskih polstavkov v analiziranih prevodih ni zanemarljivo majhna, zlasti v leposlovnih besedilih (ibid: 322–323).

S tem prispevkom želimo doprinesiti k omenjenim obsotoječim delnim raziskvam in preveriti, kako je s pojavljanjem deležja v leposlovnih besedilih, ki so zbrana v korpusu Spook (opisanem v naslednjem razdelku).

Raziskava, ki jo bomo predstavili, se torej osredotoča na analizo pogostnosti deležja v slovenskih leposlovnih besedilih. Ugotoviti bomo skušali, kolikšna je prisotnost deležja v tem besedilnem tipu na splošno, poleg tega pa bomo tudi

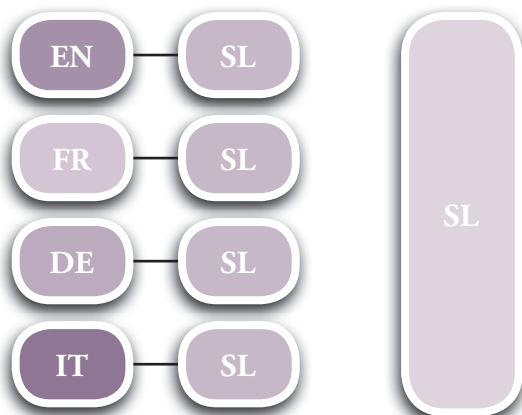
³ V zvezi s tem vprašanjem prim. tudi Schlamberger Brezar (2011), kjer je izpostavljena hipoteza, da gre pri tovrstnih oblikah za proces gramatikalizacije oz. prehoda od glagolskih oblik k prislovom oz. predlogom.

preverjali, ali med izvirnimi besedili v slovenskem jeziku in prevodi iz drugega jezika – v danem primeru italijanščine – prihaja do razlik v pogostnosti deležja. Italijanščina je namreč jezik, kjer so strukture z neosebniimi glagolskimi oblikami raznovrstne in zelo pogoste (prim. Lenassi 2004), zaradi česar bi lahko pričakovali določeno mero interference v smeri povišanja prisotnosti deležijskih stavkov, kot se to dogaja npr. v političnih besedilih, ki jih je analizirala Schlamberger Brezar (2005b).

3 KORPUS IN METODE

3.1 Korpus

Kot smo že omenili, je bila raziskava opravljena s pomočjo korpusa Spook, ki je nastal v sklopu projekta Slovensko prevodoslovje: viri in raziskave (Vintar 2009). Zgradba samega korpusa je kompleksnejša, kot so potrebe te raziskave, kot vidimo na sliki 1. Razdeljen je na dva dela, od katerih eden vsebuje izvirna slovenska leposlovna dela (1.454.275 pojavnic), drugi pa slovenske prevode iz različnih jezikov (angleščine, francoščine, nemščine in italijanščine; skupno 2.708.740 pojavnic) ter njihove izvirnike. Ker smo se v raziskavi omejili na prevode iz italijanskega jezika, smo iskanja opravili le na tem delu celote, ki obsega skupno 478.591 pojavnic.⁴ V pripadajočih izvirnih italijanskih besedilih je skupno 486.843 pojavnic. Korpus je označen in lematiziran, zato omogoča tudi avtomatsko iskanje po besednih vrstah in podvrstah. Besedila, zbrana v korpusu, so daljša in krajša prozna dela sodobnih avtorjev, izdana v zadnjih 25 letih in prevedena v zadnjih 15 letih. Prevodi so delo različnih prevajalcev.



Slika 1: Zgradba korpusa Spook (vir: spletna stran korpusa Spook)

⁴ V osnovi gre za leposlovni del korpusa ISPAC, ki je nastal v okviru doktorske naloge Mikolič Južnič (2007).

Iskalnik (slika 2), ki je bil pripravljen namensko na Institutu Jožef Štefan, ponuja številne različne možnosti iskanja, od preprostega besednega iskanja do uporabe kod MSD, in sicer tako na izvirnih slovenskih besedilih, kot tudi na vseh prevodnih podkorpusih hkrati ali pa na katerem izmed podkorpusov posebej. Seveda je prav tako mogoče iskati po izvirnih besedilih v tujih jezikih ali pa pogledati izvirni odsek in vzporedno njegov prevod, bodisi če želimo naš iskalni pogoj postaviti za besedila v izvirnem jeziku bodisi za njihove slovenske prevode. Korpus je omejeno (z uporabniškim imenom in geslom) dostopen preko spleta.

Slika 2: Iskalnik po korpusu Spook

3.2 Metode

V skladu z željo, da bi preverili tako pogostnost deležja v slovenskih izvirnih in prevedenih leposlovnih besedilih kot tudi morebitno povezavo med pojavljanjem deležja v slovenskih prevodih in prisotnostjo neosebne glagolske oblike v izvirnih italijanskih besedilih, je analitično delo po predhodnem nabiranju primerov rabe potekalo v več fazah.

Preden smo se lahko lotili analize, je bilo potrebno iz korpusa izluščiti primere. Pri tem sta se pokazali dve težavi. Po eni strani je bilo kmalu opaziti, da pri samodejnem označevanju deležij prihaja do relativno številnih napak zaradi dvoumnosti oblik deležja oz. deležnika. Velik del deležijskih oblik je namreč identičen kot deležniške, ob kategorizaciji pa so bile take oblike avtomatsko označene kot deležniški pridevniki. V nekaterih primerih sicer samodejno označevanje deluje odlično in pravilno loči oblikovno identične oblike deležnikov in deležij ter jih pravilno kategorizira, kot to prikazuje tabela 1, v kateri sta primera pojavnice *sopeč* iz podkorpusa slovenskih izvirnih besedil izpisana skupaj s kodami MSD,⁵ ki kažejo, da gre v prvem primeru za deležniški pridevnik, v drugem pa za dele-

⁵ Na spletni strani korpusa Spook je dostopna tudi povezava do razlag kod MSD ter seznama uporabljenih kod in njihovih pomenov.

žijski prislov. Drugače je pa v primeru pojavnice *ležech*, ki je vselej kategorizirana kot deležnik (in torej pridevnik), v resnici pa se v tej vlogi izmed štirih primerov pojavi samo enkrat. Primer je prikazan v tabeli 2, kjer tako kot v tabeli 1 vidimo tako zadetke v obliki kwic,⁶ kot jih je moč izvoziti v datoteko za program Microsoft Excel. Kljub (v tem primeru zaradi prostorske stiske namenoma izbranemu) skromnemu, 5-besednemu sobesedilu ob ključni besedi lahko hitro ocenimo, da gre v tabeli 2 le pri drugem od primerov zares za deležnik.

Tabela 1: Primer rezultatov iz korpusa Spook za iskano besedo *sopeč* (s prikazanimi kodami MSD)

usta	,	da	je	slišati	<< sopeč >>	glas	,	in	oba	za	trenutek
Sosmt	,	Vd	Gp-ste-n	Ggyn	<< Pdnmein >>	Somei	,	Vp	Zc-mdt	Dt	Sometn
moral	biti	hvaležen	.	Težko	<< sopeč >>	je	zlezla	z	njega	.	“
Ggnd-em	Gp-n	Ppnmein	.	Rsn	<< Rd >>	Gp-ste-n	Ggdd-ez	Dr	Zotmer	.	& 34;

Tabela 2: Primer rezultatov iz korpusa Spook za iskano besedo *ležech* (s prikazanimi kodami MSD)

na	fotelju	.	Mama	je	<< ležech >>	na	kavču	strmela	v	zrak
Dm	Sometn	.	Sozei	Gp-ste-n	<< Pdnmein >>	Dm	Sometn	Ggnd-ez	Dt	Sometn
vrečko	tobaka	in	poleg	nje	<< ležech >>	zavitek	papirčkov	,	potem	pa
Sozet	Somer	Vp	Dr	Zotzer	<< Pdnmetn >>	Somei	Sommr	,	Rsn	Vp
vse	več	časa	je	prebil	<< ležech >>	ob	televizorju	.	Na	srečo
Rsn	Rsr	Somer	Gp-ste-n	Ggdd-em	<< Pdnmein >>	Dm	Sometn	.	Dt	Sozet
In	tako	sem	se	,	<< ležech >>	v	vroči	vodi	,	rahlo
Vp	Rsn	Gp-spe-n	Zp-----k	,	<< Pdnmein >>	Dm	Ppnzem	Sozem	,	Rsn

Zaradi tovrstnih netočnosti smo se odločili, da bomo sicer preverili, kakšni rezultati se pojavijo z avtomatskim iskanjem po besednovrstni kategorizaciji s kodami MSD, vendar bomo poleg tega pojavljanje deležja preverili tudi s preprostim iskanjem glede na obrazila. To delo nam je olajšalo dejstvo, da so deležja nepregibna, je pa po drugi strani zelo zamudno, saj kakršnokoli iskanje po zaporedju črk prinaša zelo veliko 'šuma' oz. nepravilnih primerov, pri katerih je naključno zadnji del besede enak kot pri dani deležijski obliki. Priponska obrazila, ki smo jih preverili, smo našli že zgoraj: *-eč*, *-oč*, *-aje*, *-ši*; poleg teh pa smo preverili še nekatera druga obrazila, s katerimi se prav tako tvori deležja, vendar se taka deležja danes pogosto ne zaznavajo kot glagolska, temveč bolj kot načinovni prislovi; med temi so *-če*,⁷ *-te*, *-de* in *-eje* (zadnja se sicer pojavlja pri eni sami besedi, *smeje*). V raziskavo so te oblike

⁶ Kwic = Key Word in Context, oz. ključna beseda s sobesedilom.

⁷ Pravzaprav gre pri *-če* za združitev primerov z obrazili *-eče*, *-oče* ter takimi, kjer je samoglasnik pred *-če* izpuščen (npr. pri *molče*).

vključene zato, ker se jih vseeno da razvezati v odvisne stavke z osebnimi glagolskimi oblikami, tako kot je to mogoče narediti z ostalimi deležji, poleg tega pa bi se v prevodih lahko pojavljali kot odziv na italijanske neosebne glagolske oblike (zlasti na gerundij, ki ga je pogosto moč interpretirati načinovno).

Druga vrsta metodološke težave, s katero smo se srečali, je prav tako odvisna od nezmožnosti razdvoumljanja med deležniki in deležji pri iskanju po preprostem zaporedju črk. Ko smo izluščili relevantne primere iz korpusa in odstranili tiste, ki imajo naključno enake zadnje črke, nam je preostalo še razdvoumljanje med enakozvočnimi deležji oz. deležniki, kar je bilo kljub relativno zajetnemu številu primerov (gl. rezultate) mogoče narediti le ročno.

Ko smo zbrali vse primere rabe deležja tako v izvirnih kot v prevedenih slovenskih besedilih, smo se lotili še analize izvirnih italijanskih odsekov, katerih prevod je vseboval deležje. Za te primere je bilo ročno opredeljeno, v kateri besedni vrsti se je v izvirniku skrival pomen slovenskega deležja ter kakšno funkcijo je dana besedna vrsta opravljala v stavku. V zgledu (1),⁸ na primer, je v slovenskem delu uporabljeno deležje *mijavkajoč* prevod italijanskega *miagolando*, ki je sedANJI gerundij glagola *miagolare* (it. mijavkati), njegovo funkcijo v povedi pa je moč razumeti bodisi kot časovni bodisi načinovni odvisnik (kar je tipično za italijanski gerundij, vendar se bomo temu posvetili v diskusiji).

(1) Mačka je skočila s fotelja in šla *mijavkajoč* pred njo na verando.

Miagolando, la gatta la precedette sulla veranda.

Podobno smo analizirali vse najdene primere, dobljene rezultate pa smo na koncu navzkrižno primerjali in iskali morebitne povezave. Izsledki so predstavljeni v naslednjem razdelku.

4 REZULTATI IN DISKUSIJA

Kot smo videli zgoraj, je analiza potekala v različnih fazah, zato bo tudi predstavljena v treh razdelkih. Najprej se bomo posvetili pojavljanju deležja v podkorpusu izvirnih slovenskih leposlovnih besedil, kjer si bomo ogledali, koliko je bilo dejanskih primerov deležja in kakšna je njihova pogostnost ob upoštevanju velikosti celotnega podkorpusa. Nato bomo enako naredili za podkorpus besedil, prevedenih iz italijanščine. Ob pregledovanju številčnosti deležja v omenjenih podkorpusih se bomo ustavili tudi pri podatkih o tem, koliko je bilo med prvotnimi rezultati iskalnih pogojev dejanskih deležij, kolikšen del pa je pripadal deležnikom oz. drugim besednim vrstam. Nazadnje bodo prikazani rezultati analize izvirnih italijanskih primerov ter njihova primerjava s pripadajočimi slovenskimi prevodi.

⁸ Ta in vsi primeri v nadaljevanju so vzeti iz korpusa Spook.

4.1 Pogostnost deležja v izvirnih slovenskih leposlovnih besedilih

Po zgoraj opisani metodi smo iz podkorporusa izvirnih slovenskih besedil najprej poskusili izluščiti primere s pomočjo kod MSD, pri čemer smo uporabili iskalni ukaz za deležijske prislove (tako so namreč kategorizirana deležja v danem sistemu). Tak iskalni pogoj je vrnil 387 primerov, od katerih je le en bil napačno kategoriziran.⁹ Zaradi relativno nizke številke – nenavadno se nam je zdelo, da bi v korpusu, ki šteje skoraj 1,5 milijona pojavnih, bilo deležij tako malo – smo nato želeli preveriti njihovo prisotnost tudi iščoč po priponskih obrazilih. Z osmimi zgoraj navedenimi obrazili smo pred izločanjem šuma nabrali 20.490 primerov. Ko pa smo izločili vse tiste primere, kjer je šlo za popolnoma drugačne besedne vrste (npr. *daleč* pri priponi *-eč* ali *tisoč* pri priponi *-oč* ipd.), nam je ostalo 1.985 primerov, ki so glede na različna priponska obrazila razporejeni, kot je prikazano v tabeli 3. V tem skupnem številu so sicer všteti tako primeri deležij kot deležnikov, ki smo jih naknadno prav tako razločili. Končna številka vseh najdenih deležij v izvirnih slovenskih besedilih je torej 1.350.

Tabela 3: Rezultati iskanja deležij v izvirnih slovenskih besedilih po priponskih obrazilih po izločitvi šuma

Priponsko obrazilo	Prečiščeni rezultati (deležja + deležniki)	Samo deležja
-oč	423	337
-eč	233	137
-aje	70	70
-ši	4	4
-če	854	401
-de	345	345
-te	52	52
-eje	4	4
Skupno	1.985	1.350

Če zaradi primerljivosti z drugim analiziranim podkorpustom (prevodne slovenščine), ki je po obsegu precej manjši, preračunamo pogostnost na 500.000 pojavnih, je torej v izvirnih slovenskih besedilih moč naleteti na deležje 464,1-krat vsakih pol milijona pojavnih, kar znaša približno 0,1 odstotka.

⁹ Primer je naslednji: ... *vzela zalet za svoj znani, tako žuboreč smeh* ...; beseda *žuboreč* v tem primeru ni deležje, temveč deležnik.

V primerjavi s popolnoma samodejnim iskanjem po korpusu smo se torej s precejšnjo mero ročnega dela dokopali do bistveno višje in bolj realistične¹⁰ številke deležij (387 s kodami MSD, 1350 – kar 3,5-krat več – s priponskimi obrazili).¹¹

Če natančneje pogledamo še pogostnost različnih priponskih obrazil, ugotovimo, da je 'pravih deležij' (torej tistih, ki so tvorjena s priponami *-oč*, *-eč*, *-aje* in *-ši*) 518, ostalih pa 802. Med prvimi se najpogosteje pojavlja priponsko obrazilo *-oč*, s 337 primeri, dokaj pogosto je tudi *-eč*, ki ga najdemo v 137 primerih, še redkeje pa je priponsko obrazilo *-aje*, ki se pojavi le v 70 primerih. Čeprav se priponsko obrazilo *-ši* vedno kategorizira kot arhaično, pa je vseeno prisotno v obravnavanem korpusu (sodobnih besedil!) z nezanemarljivimi 4 primeri (dveh različnih sodobnih avtorjev). Med ostalimi priponskimi obrazili, s katerimi se tvori deležja, ki so normalno v stavkih rabljena kot prislovna določila načina, je daleč najpogostejše obrazilo *-če* (*-oče* se pojavi 224-krat, *-eče* 137-krat, ostalih 40 primerov pa se razporedi med različne možnosti z izpuščenim samoglasnikom), velikokrat se pojavi tudi *-de* (345 pojavitev), dokaj redek je *-te* (52 primerov), *-eje* pa se pojavi samo štirikrat v pojavitvah deležja *smeje*. Za boljšo predstavbo je v spodnjih primerih za vsako vrsto omenjenih priponskih obrazil podan primer (po vrsti, tako kot so našeta obrazila v tabeli 3 zgoraj, le pri obrazilu *-če* so navedeni trije primeri za tri različne možnosti).

- (2) Čakajoč v vrsti za vozovnice sem v steklu ...
- (3) Ob njem sta, *držeč* ga za roko, stopala ...
- (4) In on ji je *dobrikaje* poljubljaj roko ...
- (5) ... povprečna masa sto kil, *vštevši* otroke in shirane starce ...
- (6) a. Večer je romantičen in ona ga *ljubeče* gleda ...
 b. Pri tem je *mljaskajoče* zagrizel v sirovo konico ..
 c. Julija je povedila pogled in *molče* prikimala ...
- (7) ... se je dvignil in obstal napol *sede* v postelji ...
- (8) Je nekaj, kar sem *hote* skril ...
- (9) ... kjer se je politična smetana *smeje* naslajala nad žepnimi diplomati ...

¹⁰ Povsem mogoče je, da smo tudi pri ročnem iskanju kak primer izgubili, vendar ne kake večje skupine, ki bi bistveno spremenila končni rezultat. Poleg tega je med primeri, ki so bili samodejno označeni kot deležja, le 7 takih pojavitev, ki jih z iskanjem s priponskimi obrazili nismo zajeli, in sicer gre za deležja *miže* (štiri pojavitve), *diše* in *drve*, ostali pa so vsi 'ujeti' med različna priponska obrazila.

¹¹ Ob tem naj poudarimo, da je seveda nujno izraziti spoštovanje do velikega truda, vloženega v razvijanje orodij za samodejno razdvajanje besed morfološko tako zahtevnega jezika, kot je slovenščina; za prihodnost lahko samo izrazimo željo, da bi se ta orodja še naprej razvijala in izpopolnjevala.

4.2 Pogostnost deležja v prevedenih slovenskih leposlovnih besedilih

Podobno kot podkorpusa izvirnih slovenskih besedil smo se lotili tudi podkorpusa prevedenih slovenskih besedil. Iskalni pogoj, postavljen s pomočjo kot MSD, nam je izluščil borih 46 primerov, izmed katerih je prav tako kot pri izvirnih slovenskih besedilih en primer bil kategoriziran narobe.¹² Z iskanjem po priponskih obrazili se nam je pred izločanjem šuma nabralo 8.577 primerov, po prečiščenju rezultatov pa je ostalo 329 primerov, od katerih jih je bilo 230 dejansko rabljenih v funkciji deležja. Ponovno smo torej z ročno metodo izluščili bistveno večje število primerov – tokrat kar 5-krat toliko kot z avtomatsko metodo.

Preračunano v z zgornjimi rezultati primerljive številke, je v podkorpusu slovenskih prevodov iz italijanskega jezika torej pogostnost deležja 240,3 pojavitev na 500.000 besed oziroma 0,05 odstotka vseh pojavnic. V primerjavi s pogostnostjo v izvirnih besedilih (kjer je bila pogostnost 464,1 na 500.000 besed oziroma 0,1 odstotka) opazimo veliko razliko, saj se v prevedenih besedilih deležja pojavljajo pol manjkrat. Zakaj prihaja do tako izrazitega odstopanja, težko z gotovostjo trdimo, vendar je taka situacija vsekakor presenetljiva, saj bi, kot smo povedali v uvodnih poglavjih, pričakovali pri prevodih iz italijanščine, ki je bogata z neosebniimi glagolskimi oblikami, ravno obratno. V vsakem primeru ugotovljenega razmerja med izvirno in prevodno slovenščino zaradi omejenosti analiziranega korpusa ne moremo upoštevati kot popolnoma zanesljivo glede pogostnosti deležja, saj smo med drugim v prevodnem delu opazovali le prevode iz italijanskega jezika. Vendar bo ravno zaradi te izbire zanimivo opazovati, ali se pojavljajo morebitni vzorci glede situacij, v katerih se deležje pojavlja kot prevod določenih izvirnih struktur, kar bomo skušali osvetliti v naslednjem razdelku.

Pred tem si pogledajmo поблиže še pogostnost posameznih priponskih obrazil v podkorpusu prevedenih slovenskih besedil, prikazano v tabeli 4.

12 Primer, ki ga omenjamo, je naslednji: *Ogorek je izginil v noči kot žareč kosček lave*. Beseda *žareč* je v tem primeru rabljena kot deležnik, ne deležje.

Tabela 4: Rezultati iskanja deležij v prevedenih slovenskih besedilih po priponskih obrazilih po izločitvi šuma

Priponsko obrazilo	Prečiščeni rezultati (deležja + deležniki)	Samo deležja
-aje	35	35
-oč	42	34
-eč	33	14
-ši	0	0
-de	87	87
-če	119	47
-eje	7	7
-te	6	6
Skupno	329	230

Deležij, tvorjenih s prvimi štirimi priponami, je skupno 83, tistih, ki se jih upošteva kot deležijske prislove, pa 147. V primerjavi s pojavljanjem v podkorpusu izvirnih slovenskih besedil opazimo razlike v pogostnosti različnih priponskih obrazil: če sta priponski obrazili *-aje* in *-oč* v prevedenih slovenskih besedilih skorajda izenačeni in je *-aje* na prvem mestu, je v izvirnih besedilih daleč najpogostejše obrazilo *-oč*: pojavlja se skoraj petkrat pogostejše kot *-aje*, ki je šele na tretjem mestu, za *-eč*. Priponsko obrazilo *-ši*, ki je bilo zelo redko že v izvirnih slovenskih besedilih, je popolnoma odsotno v prevodih.

Prav tako se razlikuje pogostnost ostalih obravnavanih priponskih obrazil: v prevedenih besedilih je najpogostejše obrazilo *-de*, na drugem mestu pa je *-če* (v 24 primerih se pojavi *-eče*, 13-krat *-oče*, 11-krat pa z izpuščenim samoglasnikom); v izvirnih besedilih je ravno obratno; podobno se zgodi tudi za ostali dve obrazili, kjer je v prevedenih besedilih *-eje* pogostejši od *-te*, v izvirnih besedilih pa je *-te* veliko pogostejši.

Tako kot pri podkorpusu izvirnih slovenskih besedil, tudi za podkorpus (iz italijanščine) prevedenih slovenskih besedil navajamo primere za vsak tip priponskega obrazila v enakem vrstnem redu, kot so navedena v tabeli 4 (pri čemer za obrazilo *-ši* seveda ni primera).

(10) ... in zdaj je tekel čez travnike, *skakaje* čez grmičevje ...

(11) Stopil je k vratom *opirajoč* se na palico ...

(12) *Ležeč* v babičini postelji bo zdaj zdaj skočil ...

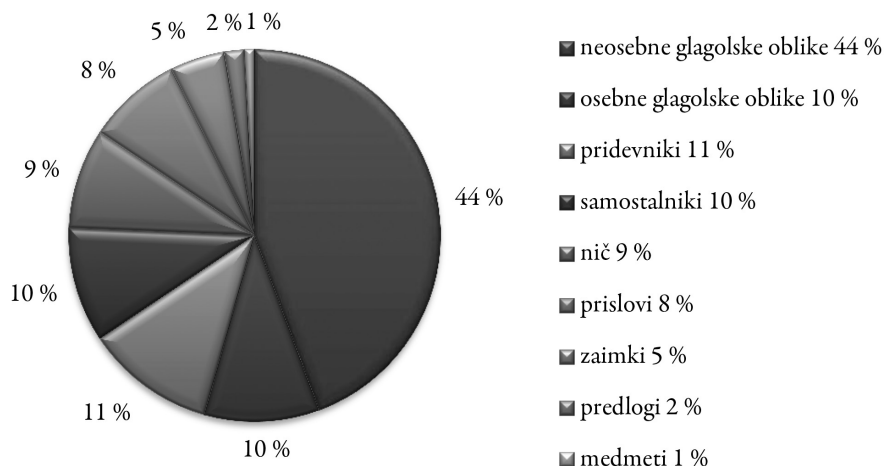
- (13) *Sede* drug ob drugem sta prebila več ur ...
- (14) a. ... pa je *kleče* na tleh privzdignil Friderikovo glavo ...
 b. Medtem ko so se sušili, so *lesketajoče* spreminjali barvo ...
 c. *Molče* je pobožala sončnico ...
- (15) Rus je *smeje* se nekaj dejal ...
- (16) Andrea je *hote* preslišal njun prepir ...

4.3 Primerjava primerov iz prevedenih slovenskih leposlovnih besedil z njihovimi izvirniki

Pri primerjavi slovenskih prevodov z njihovimi izvirniki smo se osredotočili na dve vprašanji: po eni strani smo ugotavljali, katere besedne vrste v izvirniku najpogosteje nosijo pomen, ki ga v slovenščini izraža deležje; po drugi strani pa nas je zanimalo, v kakšnih sintaktičnih strukturah se te besedne vrste nahajajo, saj se deležja (še zlasti tista, ki jih ne čutimo več kot glagolska) navadno razume kot sredstvo za izražanje načina.

4.3.1 Besedne vrste izvirnikov

Poglejmo si najprej besedne vrste izvirnikov, ki so v slovenščini pripeljale do deležja. Glede na to, da je deležje ena od neosebne glagolske oblike, ki v slovenščini na splošno niso tako pogoste kot na primer v italijanščini, in tudi v skladu z ugotovitvami Schlamberger Brezar (2005b), ki smo jih omenili zgoraj, bi torej pričakovali da bomo v izvirnih italijanskih stavkih, katerih prevodi vsebujejo deležje, večinoma prav tako našli neosebne glagolske oblike (gerundij, infinitiv ali particip). Dejansko se je izkazalo, da je neosebne glagolske oblike res največ, kot vidimo na sliki 3, kjer so prikazani združeni podatki za vse vrste neosebne glagolske oblike in za druge besedne vrste, vendar njihova prevlada le ni tako zelo izrazita. Slovensko deležje je prevod italijanske neosebne glagolske strukture v 44 odstotkih vseh primerov (natančno gre skupno za 102 primera, ostali absolutni podatki pa so v tabeli 5).



Slika 3: Deleži izvirnih besednih vrst, prevedenih v slovenščino z deležjem

Tabela 5: Podrobni številčni podatki o neosebni glagolskih oblikah, prevedenih z deležjem

Besedna vrsta	Število primerov
gerundij	52
pretekli particip	33
nedoločnik	8
sedanji particip	7
gerundij + nedoločnik	2

Ob izraziti prisotnosti neosebni glagolskih oblik v izvirnih primerih pa opazimo tudi precejšnje število drugih besednih vrst: poleg neosebni glagolskih oblik najdemo tudi osebne glagole, zatem pa še pridevnike, samostalnike, prislove, redko pa tudi zaimke, medmete in predloge.¹³ V 9 odstotkih primerov se je celo zgodilo, da pravzaprav v izvirnem stavku ni bilo elementa, katerega neposredni prevod bi bilo obravnavano deležje, temveč je v slovenščini prišlo do (bolj ali manj natančne in potrebne) eksplicitacije pomena, ki je bil v izvirniku kvečjemu impliciran. V nadaljevanju je predstavljen po en primer (italijanski izvirnik in slovenski prevod) za vsako besedno vrsto posebej.

¹³ Pri zadnjih treh kategorijah gre za mejne klasifikacije, kjer se pravzaprav v izvirnih primerih najdejo širše strukture, katerih del je zaimke, medmet ali predlog.

- (17) **Gerundij**
ital. Vi si diresse *caracollando* sulle sue gambette corte ...
slov. *Pozibavajoč* se na kratkih nožicah se je napotil ...
- (18) **Pretekli particip**
ital. L' imperatore lo aspettava *seduto* sul suo trono e beveva un boccale di vino
slov. Cesar ga je pričakal *sedeč* na svojem tronu in pil bokal vina
- (19) **Nedoločnik**
ital. ... specie, continuava a *mormorare* il Poeta, fortunatamente in teutonico ...
slov. ... in to še posebno, je *šepetaje* nadaljeval Pesnik, k sreči v tevtonščini ...
- (20) **Sedanji particip**
ital. ... e dal retro dell' iconostasi apparve una figura *brancolante* ...
slov. ... in nekdo je *tipaje* prikrevsal izza ikonostasa ...
- (21) **Gerundij + nedoločnik**
ital. ... ed entrò in casa *riprendendo a cantare* in sincronia con Di Stefano ...
slov. ... in odšel proti hiši, *prepevaje* svojo ljubo operno arijo ...
- (22) **Osebna glagolska oblika**
ital. La scienza si *dimostrava* incapace di comprendere ...
slov. Znanost po vsem *sodeč* ni bila sposobna dognati ...
- (23) **Pridevnik**
ital. ... poi si fermavano *incerti* sul da farsi ...
slov. ... nato pa prenehali, ne *vedoč*, kaj bi ...
- (24) **Samostalni**
ital. "Ascolti, figliola", disse in un *sussurro* ...
slov. "Poslušajte, draga" je rekla *šepetaje* ...
- (25) **Nič**
ital. ... sono io che preparo la colazione", osservò.
slov. ... jaz pripravim zajtrk," je navrgla *mimogrede*.
- (26) **Prislov**
ital. Collodi cominciò a camminare a *tentoni*, inciampò in qualcosa ...
slov. Collodi je *tipaje* hodil naokrog, se spotaknil ob nekaj ...
- (27) **Zaimek**
ital. Ma quanto al veleno, come puoi dire ...
slov. *Glede* strupa pa, le kako lahko izjaviš ...

(28) **Medmet**

ital. Dove loro avevano fallito, ebbero successo, *grazie* alla forza delle armi, gli americani ...

slov. Kjer je njim spodletelo, je *zahvaljujoč* orožju uspelo Američanom ...

(29) **Predlog**

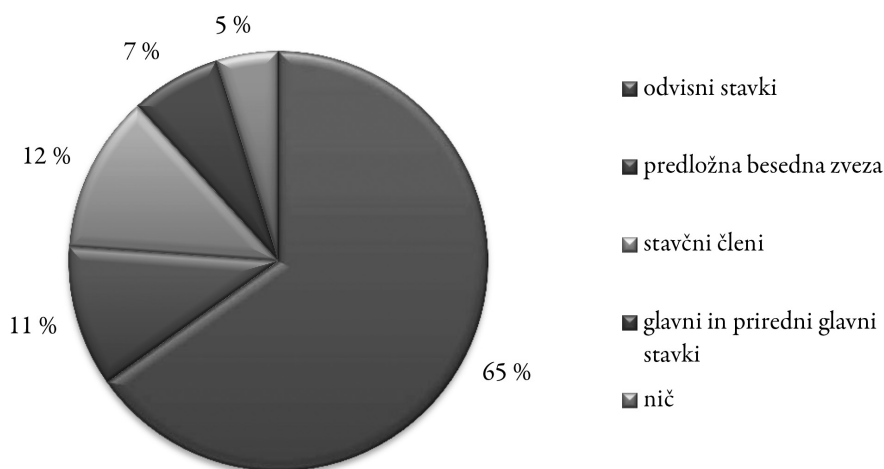
ital. ... che lo aveva assicurato *sulle* condizioni generali della madre ...

slov. Pomirila ga je *glede* maminega zdravstvenega stanja ...

Iz zgoraj povedanega lahko torej zaključimo, da večja ali manjša pogostnost deležja v slovenščini, prevedeni iz italijanščine, ni izrazito pogojena s prisotnostjo neosebne glagolske oblike v izvirnih besedilih, saj so le malo več kot štiri od desetih deležij v slovenščini odlikava katere od italijanskih neosebne glagolske oblike. Deležijski prevodi ostalih besednih vrst, ki se pojavljajo poleg glagola (predvsem pridevniki, samostalniki in prislovi), so pogosto posledica sintaktične strukture, v kateri se nahajajo te besede v izvirnih besedilih. V zgornjih primerih tako opazimo, da je pridevnik uporabljen kot določilo vezalnemu glagolu *fermarsi* (primer 23), kar je v slovenščino prevedeno s polnomenomskim glagolom v deležijski obliki (*vedeti*); samostalnik v primeru 24 pa je v izvirnem besedilu uporabljen kot del prislovnega določila načina (*in un sussurro*), ki je smiselno prevedeno v slovenščino z deležjem (*šepetaje*), ki prav tako izraža način. Podobno je prislov v primeru 26 (*tentoni*) uporabljen v vlogi prislovnega določila načina in tako tudi preveden (*tipaje*). Primeri, v katerih se pojavljajo zaimki, vezniki ali medmeti, so vedno sobesedilno zelo pogojeni in opazovane besede so vedno del širše besedne zveze; tudi po številčnosti gre za marginalne rabe, omejene na posebne okoliščine. Prav zaradi pogoste umestitve posamezne besede, ki v izvirniku nosi pomen, izražen v prevodu z deležjem, smo se odločili, da preverimo, v kakšnih sintaktičnih strukturah se pojavljajo te besede (prim. razdelek 4.3.2.). Med zgoraj navedenimi možnostmi izpostavimo še primere (kar 9 odstotkov jih je), ko v prevodu navadno prihaja do eksplicitacije pomena, ki je v izvirniku le nakazan (bodisi v neposrednem bodisi v širšem sobesedilu): glede na relativno nizko pogostnost deležja v prevodih literarnih besedil na splošno (kar je bilo omenjeno zgoraj v poglavju 4.2.) se zdi nenavadno in zanimivo, da prevajalci posegajo po deležjih, ko želijo eksplicirati del izvirnega pomena. Iz tega bi lahko sklepali, da uporaba deležja (vsaj v obravnavanem besedilnem tipu) le ni tako tuja in nenaravna, saj vsaj za te primere lahko z gotovostjo trdimo (za večino ostalih analiziranih primerov pa verjetno pav tako), da rabe deležja ne moremo pripisati pojavi interference z izvirnim jezikom (kot je za sodobne prevode (zlasti iz francoščine) trdil zgoraj omenjeni Breznik (1967: 82–94)).

4.3.2 Sintaktične funkcije izvirnih struktur

Kot je bilo omenjeno zgoraj, smo se analize sintaktičnih struktur, ki se v italijanskih besedilih pojavljajo tam, kjer smo v slovenščini zasledili deležja, lotili z namenom, da bi odkrili morebitne povezave med funkcijo, ki jo določen element opravlja v izvirnem besedilu, in verjetnostjo, da bo tak element preveden z deležjem. Izkazalo se je, da je paleta možnih italijanskih sintaktičnih struktur, ki so v slovenščino bile prevedene z deležjem, precej obsežna: med najdenimi strukturami, ki so navedene na sliki 4 s pripadajočimi deleži, najdemo odvisne stavke, glavne stavke in priredne stavčne zveze glavnih stavkov, stavčne člene, predložne besedne zveze ter nič (v primerih, ko je slovenski prevod ekspliciral informacijo, ki v izvirniku ni bila ubesedena). Podrobnejša razvrstitev najdenih sintaktičnih struktur je navedena v tabeli 6.



Slika 4: Deleži izvirnih sintaktičnih struktur, prevedenih v slovenščino z deležji

Tabela 6: Podrobni prikaz vrst sintaktičnih struktur, prevedenih v slovenščino z deležjem

Vrsta strukture		Število primerov	Delež (v %)
Odvisni stavki	načinovni odvisnik	83	36,10
	vzročni odvisnik	25	10,87
	časovni/načinovni odvisnik	13	5,65
	prilastkov (oziralni) odvisnik	7	3,05
	časovni odvisnik	4	1,74
	pogojni odvisnik	4	1,74
	dopustni odvisnik	4	1,74
	predmetni odvisnik	3	1,30
	krajevni odvisnik	2	0,87
	namerni odvisnik	2	0,87
	izvzemalni odvisnik	1	0,43
	odvisnik ozira	1	0,43
	primerjalni odvisnik	1	0,43
Glavni in priredni glavni stavki	glavni stavek	13	5,65
	ločno priredje	3	1,30
Predložne besedne zveze		25	10,87
Stavčni členi	prilastek	13	5,65
	povedkovo določilo	10	4,35
	prislovna določila	5	2,18
Nič		11	4,78

Na sliki 4 vidimo, da med izvirnimi strukturami, ki so bile prevedene v slovenščino z deležjem, izrazito prevladujejo odvisni stavki, med temi pa, kot je razvidno iz tabele 6, z veliko prednostjo prevladujejo načinovni odvisniki (primer 30), ki skupaj s primeri, ki jih lahko interpretiramo načinovno ali časovno (primer 31), obsegajo kar 65-odstotni delež vseh odvisnih stavkov (oz. dobrih 40 % vseh struktur). Druga najbolj množična skupina so vzročni odvisniki (primer 32), ki pa se pojavljajo bistveno manj pogosto (17 % vseh odvisnikov oz. slabih 11 % vseh struktur). Vsi ostali tipi odvisnikov skupaj (primeri 33–42) obsegajo komaj 12,6 odstotkov vseh najdenih struktur. Relativno redki – skupno 7 odstotkov – so tudi primeri, kjer je z deležjskim stavkom bil preveden glavni stavek ali priredni glavni stavek (primeri 43–44), nekoliko večje število pa je takih primerov, kjer prevod z deležjem izvira iz različnih stavčnih členov: prilastka (primer 45), povedkovega določila (46) in raznih prislovnih določil (primer 47). V 11 odstotkih vseh

najdenih primerov je slovensko deležje nastalo iz predložne besedne zveze (primer 48): v teh primerih gre za ustaljene zveze v italijanščini, ki so tudi v slovenščini izražene z ustaljenimi oblikami (deležje *glede* ima v primeru 48 prav tako vlogo predloga, ki uvaja predložno zvezo). Na zadnjem mestu (čeprav ne po pogostnosti) so v tabeli 6 omenjeni primeri, kjer v originalu de facto ni elementa, ki je v prevodu ekspliciran z deležjem (tak primer je že naveden zgoraj pod številko 25).

(30) **Načinovni odvisnik**

ital. ... ragazzino prese qualcosa da sotto gli stracci e *tremando di paura* si avvicinò a Hervé ...

slov. ... deček izpod cap nekaj izvlekel, se *trepetajoč od strahu* približal Hervéju ...

(31) **Časovni/načinovni odvisnik**

ital. ... che stringendo il pugno e *sventolando la bandierina* ruggivano ...

slov. ... ki so s stisnjenimi pestmi in *mahajoč z zastavicami* vpili ...

(32) **Vzročni odvisnik**

ital. ... io che la portai al fiume, *credendo* che fosse ragazza ...

slov. ... sem jo hotel peljati k reki, *misleč*, da je še dekle ...

(33) **Prilastkov (oziralni) odvisnik**

ital. ... Baudolino, *inginocchiato presso al corpo di Federico*, gli sollevava il capo ...

slov. ... Baudolino pa je *kleče na tleh* privzdignil Friderikovo glavo ...

(34) **Časovni odvisnik**

ital. ... chiese, *mentre le passava accanto* e le sfiorava la guancia con un bacio ...

slov. .. je vprašal in jo *mimogrede* rahlo poljubil na lice ...

(35) **Pogojni odvisnik**

ital. ... già una bella impresa studiare a Parigi, *se si pensava* che era nato ...

slov. ... lep dosežek že sam študij v Parizu *glede na to*, da se je rodil ...

(36) **Dopustni odvisnik**

ital. ... sentiva il rumore delle stoviglie e, *da qualunque parte del giardino fosse* ...

slov. ... zmerom zaslišala žvenket jedilnega pribora, *ne glede na katerem koncu vrta je bila* ...

(37) **Namerni odvisnik**

ital. ... sbatteva l'uscio della cucina e si rifugiava nell' orto *a piangere* ...

slov. ... je zaloputnila s kuhinjskimi vrati in se *jokajoč* skrila na zelenjadni vrt ...

- (38) **Krajevni odvisnik**
 ital. ... chino sul tavolo, una mano dietro la schiena, l'altra *a stringere la stecca*, incredulo ...
 slov. ... eno roko za hrbtom in *v drugi stiskajoč palico*, otrpnil in ni mogel verjeti ...
- (39) **Predmetni odvisnik**
 ital. ... aprì gli occhi lui sentì la propria voce *dire piano* ...
 slov. ... oči, je zaslišal svoj glas, *mrmljajoč* ...
- (40) **Izvezalni odvisnik**
 ital. ... il male che *senza rendermi conto* nel 1980 (intervista a Khomeini) feci ...
 slov. ... leta 1980 (intervju s Homeinijem) *nevede* povzročila ...
- (41) **Odvisnik ozira**
 ital. ... tuttavia avesse qualcosa sulla coscienza *per quanto riguardava la morte di Federico* ...
 slov. ... kljub temu pa ni imel mirne vesti *glede Friderikove smrti* ...
- (42) **Primerjalni odvisnik**
 ital. ... e lo stringeva *come se volesse spezzarlo* ...
 slov. ... ga je, kot da bi ga *hote* zdrobil ...
- (43) **Glavni stavek**
 ital. *Ho brancolato sino al fondo del corridoio* ...
 slov. *Tipaje* sem se odvrlekel do konca hodnika ...
- (44) **Ločno priredje**
 ital. *O Ardrouni gli aveva confidato qualcosa*, o l'aveva ...
 slov. Ne *glede* na to, ali mu je Ardrouni kaj razkril ali pa je sam dojel ...
- (45) **Prilastek**
 ital. E tuttavia parlava un greco *scorrevole* ...
 slov. Kljub temu pa je *tekoče* govoril grško ...
- (46) **Povedkovo določilo**
 ital. *Testi Plinio e Isidoro*, si decise invece di porre in quella terra ...
 slov. *Sklicevaje se na Plinija in Izidorja Seviljskega* so se zato raje odločili umestiti ...
- (47) **Prislovno določilo**
 ital. ... Penelope non aveva mai fatto commenti *in proposito*...
 slov. ... Penelopa ni *glede tega* nič rekla ...
- (48) **Predložna besedna zveza**
 ital. *Quanto all'ammirevole capacità di unirsi* ...
 slov. *Glede* občudovanja vredne sposobnosti združevanja moči ...

Iz zgoraj navedenih podatkov lahko torej sklepamo, da se deležje v analiziranih leposlovnih prevodih, kadar je sploh uporabljeno (prim. zgoraj opisano razliko v pogostnosti v prevedeni in v izvirni slovenščini, poglavji 4.1. in 4.2.), dejansko daleč najpogosteje uporablja kot prevod odvisnih stavkov, zlasti načinovnih odvisnikov. Čeprav v izvirnih besedilih deli, ki so bili prevedeni z uporabo deležja, lahko vsebujejo zelo različne vrste sintaktičnih struktur (in besednih vrst, kot smo videli v poglavju 4.3.1.), se deležje najmanj štirikrat pogosteje¹⁴ najde v vlogi prevodne ustreznice načinovnega odvisnika kot v vlogi prevoda katerekoli druge strukture. To dejstvo gotovo ni presenetljivo, saj je interpretacija deležja na splošno v slovenščini zelo pogosto prav tako načinovna (zlasti deležja na –e, prim. Toporišič 1991: 340).

S tega vidika je sicer bolj zanimivo, da se lahko deležje pojavlja kot prevod tako zelo raznolikih izhodiščnih struktur in pomenov, kar nakazuje na veliko fleksibilnost jezika kot takega in tudi na ne tako zelo omejeno uporabnost slovenskega deležja.

Med ostalimi izvirnimi primeri, v katerih je prisoten odvisni stavek, se relativno pogosto pojavljajo še vzročni in časovni odvisniki (ki so, kot smo že omenili, pogosto neločljivi od načinovnih), kar prav tako ustreza ugotovitvam Toporišiča (1991: 514 glede vzročnih odvisnikov, 339 glede izražanja istodobnosti oz. časovnih odvisnikov) o vlogah slovenskega deležja. Izmed ostalih struktur se v več kot 10 odstotkih primerov v izvirnikih pojavljajo le še predložne besedne zveze, pri katerih gre v vseh primerih za ustaljene besedne zveze (ali rabe posameznih besed) tako v italijanskem jeziku kot v slovenščini (kar nazorno prikazuje primer 48) in torej za posebne okoliščine, kjer jezikovni čut slovenskih prevajalcev pri rabi deležja ni igral nobene vloge. Druge pojavitve, ki posamezno ne presegajo 6 odstotkov vseh najdenih primerov (tej številki se približujeta kategoriji glavnih in prirednih stavkov ter prilastkov), večinoma pa se gibljejo okrog odstotka ali dveh, razumemo predvsem kot ad hoc pojavitve, uporabne v danih okoliščinah (kot prikazujejo navedeni primeri), vendar ne tudi širše posplošljive prevodne možnosti.

Posebej velja še enkrat omeniti slabih 5 odstotkov primerov, pri katerih je v slovenskem prevodu prišlo do eksplicitacije pomena, ki v izvirniku ni bil ubeseden. Čeprav bi del takih primerov lahko pripisali prevajalčevemu slabšemu razumevanju besedila (primer 49, kjer je po a. objavljeni prevod, pod b. izvirnik in pod c. naš dobessedni prevod), gotovo velja, da gre pri večini primerov dejansko za eksplicitacijo, kot jo številni avtorji (med prvimi Vinay in Darbelnet 1995) razumejo v prevodoslovju, kar pa sicer ni predmet tokratne raziskave.

¹⁴ Preko štirikrat pogosteje se pojavlja v primerjavi z drugima najpogostejšima izvirnima italijanskima strukturama, vzročnimi odvisniki in predložnimi besednimi zvezami (vsaka se pojavlja v slabih 11 odstotkih primerov); vse ostale strukture se pojavljajo še bistveno redkeje, kot je razvidno iz tabele 6.

- (49) a. Hipatija, ti govoriš o napaki stvaritve in zlu, a le čemu – *glede* na to, da te vse to ne zadeva in živiš v tem gozdu, kjer je vse tako lepo, kot si ti?¹⁵
 b. Tu parli dell'errore della creazione, Ipazia, e del male, ma come se non ti toccasse, e vivi in questo bosco come se tutto fosse bello come te.
 c. Ti govoriš o napaki stvaritve, Hipatija, in o zlu, a kot da se te ne bi dotikalo in živiš v tem gozdu, kot da bi bilo vse tako lepo kot ti.

5 SKLEP

Kljub temu, da slovensko deležje materni govorniki navadno dojemajo kot fenomen naše jezikovne preteklosti, se ta glagolska oblika še vedno pojavlja tako v opisih jezika (Toporišič 1991) kot v realnih (pisnih)¹⁶ besedilih. Vprašanje, ki smo si ga postavili v pričujočem prispevku, je, kako pogosto se dejansko deležje uporablja v slovenskih leposlovnih besedilih in ali se enako pogosto pojavlja v izvirnih slovenskih in prevedenih leposlovnih besedilih. Analiza korpusa Spook je pokazala, da je pogostnost deležja v izvirnih besedilih približno 464-krat na 500.000 pojavnic, kar ni zanemarljivo malo. Pri tem je nujno opozoriti, da gre v približno 30 % za z vidika sodobne slovenščine izrazito predložne ali prislovične oblike (*glede; mimogrede, nehote*),¹⁷ v ostalih primerih pa za raznoliko rabo številnih avtorjev. Zelo zanimivo je poleg tega dejstvo, da se v iz italijanščine prevedenih leposlovnih besedilih deležje pojavlja znatno redkeje, in sicer približno pol manjkrat kot v izvirnih besedilih. Če ob tem upoštevamo še, da je v tem primeru predložnih oz. prislovičnih deležij kar 59 %, je še bolj jasno, da je živa raba deležja v prevedenih leposlovnih besedilih bistveno bolj siromašna kot v izvirnih besedilih. V obeh primerih gre za isti besedilni tip in približno enako obdobje nastajanja besedil; prav tako imamo v obeh primerih opravka s številnimi različnimi avtorji besedil (bodisi pisatelji bodisi prevajalci), tako da razlik ne moremo pripisati idiosinkratičnim izbiram posameznikov. Tako skromno uporabo deležja v prevodih lahko razložimo z vplivom hiperkorektnosti, ki so ji prevajalci podlegli v strahu, da ne bi pretirano prenašali izvirnih (italijanskih) brezosebne glagolske strukture v slovenski jezik. Vsi so se najbrž tekom svojega izobraževanja v materinščini vsaj na osnovnošolskem in srednješolskem nivoju srečali s preskriptivnim pogledom na rabo deležja, ki ga je izrazil že Breznik (1967), za njim pa številni drugi avtorji, kot smo omenili v uvodnem delu, in posledično tudi ključni akterji izobraževanja. Prevajalci, ki se ukvarjajo z jezikom, kot je italijanščina, se gotovo (vsaj intuitivno) zavedajo razlike v pogostnosti in splošni uporabnosti neosebne glagolske strukture med obema obravnavanima jezikoma, zato se bodo še skrbneje

¹⁵ V objavljenem prevodu pride do napačnega prenosa pomena, saj v izvirniku ni implicirana namernost, ki izhaja iz (retoričnega) vprašanja 'le čemu' v slovenskem prevodu. Ker avtor izvirni stavek interpretira nekoliko po svoje, čuti potrebo po ekspliciranju vzročnosti z zvezo 'glede na to'.

¹⁶ Študij o pojavljanju deležja v govorjenem jeziku zaenkrat ni, za druge žanre pisnega jezika pa prim. tudi Mikolič Južnič (2012).

¹⁷ Prim. tudi Schlamberger Brezar (2011).

obvarovali pasti interference na tem področju. Posledica takega razmišljanja je relativno konservativna, torej redkejša raba deležja v analiziranih prevedenih besedilih.

Poleg splošne pogostnosti deležja nas je zanimalo, ali je morda med strukturami, prisotnimi v izvirnih besedilih, ter prisotnostjo deležja v slovenskih prevodih mogoče zaznati kako povezavo. Izkazalo se je, da je dejansko deležje najpogostejše uporabljeno, ko imamo v izvirnem stavku neosebne glagolske oblike, uporabljene v načinovnih odvisnih stavkih (kar je v skladu z eno pogostejših pomenskih interpretacij deležja), vendar se razmeroma pogosto pojavi tudi kot prevod drugih besednih vrst in sintaktičnih struktur.

Ob zbiranju primerov deležja v korpusu Spook smo imeli možnost preveriti tudi, kako natančno je samodejno oblikoskladenjsko označevanje besednih vrst: izkazalo se je, da se pri deležjih program ni izkazal kot najbolj zanesljiv, saj je bilo s kodami MSD pravilno kategoriziranih le med 29 in 41 odstotki deležij (odvisno od tega, kako opredelimo deležje oz. kaj vključimo v to kategorijo). Slovnična kategorija zaradi svoje podobnosti z drugimi oblikami (npr. deležniki) gotovo predstavlja izziv za samodejno označevanje in potrebnega bo še precej dela, preden bodo rezultati samodejnega označevanja dovolj natančni in uporabni za jezikoslovne analize.

Prihodnje raziskave bodo osredotočene na preverjanje prisotnosti deležja v slovenskih prevodih leposlovnih besedil iz drugih jezikov, ki so na voljo v korpusu Spook (angleščina, francoščina, nemščina), kajti zanimivo bo opazovati, ali je hiperkorektna težnja prevajalcev iz italijanščine splošen fenomen slovenske prevodne literature ali je bolj vezana na specifičen jezikovni par.

Drugi vidik, ki ga nameravamo v prihodnje raziskati, zadeva druge besedilne tipe: ali je pogostnost deležja, ugotovljena na primeru analiziranega korpusa izvirnih leposlovnih besedil, podobna pogostnosti v splošnih slovenskih besedilih? Kaj pa v različnih vrstah strokovnih jezikov? Ali pa v prevodih drugih besedilnih tipov? Na ta in številna druga vprašanja bo treba odgovoriti, preden si bomo lahko ustvarili resnejšo sliko o rabi deležja v modernem slovenskem jeziku.

Bibliografija

- Breznik, Anton, 1967: *Življenje besed*. Maribor: Obzorja.
 Erjavec, Tomaž, 1997: Računalniške zbirke besedil. *Jezik in slovstvo* 42/2–3. 81–96.
 Jesenšek, Marko, 1998: *Deležniki in deležja na -č in -ši. Razširjenost oblik v slovenskem knjižnem jeziku 19. stoletja*. Maribor: Slavistično društvo Maribor.
 Kocijancič Pokorn, Nike in Rastislav Šuštaršič. 1999: Slovensko-angleška protistavna analiza nedoločnika v vlogi premega predmeta. *Vestnik* 33/1–2. 267–281.

- Kocijančič Pokorn, Nike in Rastislav Šuštaršič, 2001: Slovensko-angleška protistavna analiza nedoločnika v vlogi osebkovega oz. povedkovnega določila. *Slovenski jezik/Slovene Linguistic Studies* 3. 37–41.
- Korpus Gigafida <<http://www.gigafida.net>> (Dostop 13. 05. 2012.)
- Korpus Spook. Slovensko prevodoslovje: viri in raziskave: <<http://lojze.lugos.si/spook/index.html>> (Dostop 13. 05 2012.)
- Lenassi, Nives. 2004: Implicitne glagolske konstrukcije: problem njihovega poučevanja v okviru poslovnega italijanskega jezika. *Vestnik* 38/1–2. 173–195.
- Mezeg, Adriana, 2011: *Korpusna podprta analiza francoskih polstavkov in njihovih prevedkov v slovenščino*. PhD dis., Univerza v Ljubljani.
- Mikolič Južnič, Tamara, 2007: *Nominalne strukture v italijanščini in slovenščini : pogostnost, tipi in prevodne ustreznice*. PhD dis., Univerza v Ljubljani.
- Mikolič Južnič, Tamara, 2012: Korpusna analiza slovenskega deležja v različnih besedilnih tipih. Erjavec, Tomaž in Žganec Gros, Jerneja (ur.): *Zbornik 8. konference Jezikovne tehnologije, Ljubljana, 8.–9. oktober 2012*. Ljubljana: Institut Jožef Stefan. 147–152.
- Milojevič Sheppard, Milena, 1992: Polstavki v angleščini in slovenščini. *Vestnik* 26/1–2. 224–237.
- Schlamberger Brezar, Mojca, 2005a: Nekatere skladenjske značilnosti prevodov iz francoščine v slovenščino – lekcija iz kontrastivne slovnice. Mikolič, Vesna in Karin Marc Bratina (ur.): *Slovenščina in njeni uporabniki v luči evropske integracije*. Koper: Založba Annales. 261–268.
- Schlamberger Brezar, Mojca, 2005b: Politična besedila kot tip besedil in postopki prevajanja stalnih formul. Kocijančič Pokorn, Nike, Erich Prunč in Alessandra Riccardi (ur.): *Beyond Equivalence – Jenseits der Äquivalenz – Oltre l'equivalenza – Onkraj ekvivalence*. Gradec: Institut für Theoretische und Angewandte Translationswissenschaft. 121–135.
- Schlamberger Brezar, Mojca, 2011: Le gérondif et le participe présent et leur évolution vers la grammaticalisation: étude contrastive du slovène et du français. *Linguistica* 51. 333–348.
- Sinclair, John, 1997: Corpus Evidence in Language Description. Wichmann, Anne, Steven Fligelstone, Tony McEnery in Gerry Knowles (ur.): *Teaching and Language Corpora*. London/New York: Longman. 27–39.
- Šuštaršič, Rastislav in Nike Kocijančič Pokorn, 1998: Slovensko-angleška protistavna analiza angleškega nedoločnika v funkciji osebkva. *Vestnik* 32/1–2. 299–320.
- Toporišič, Jože, 1991: *Slovenska slovnica*. Maribor: Obzorja.
- Toury, Gideon, 1995: *Descriptive translation studies and beyond*. Amsterdam/Philadelphia: John Benjamins.
- Vinay, Jean-Paul in Jean Darbelnet, 1995: *Comparative Stylistics of French and English: A Methodology for Translation*. Amsterdam/Philadelphia: John Benjamins.
- Vintar, Špela, 2009: Slovenski prevodoslovni korpus. Stabej, Marko (ur.): *Infrastruktura slovenščine in slovenistike. Simpozij Obdobja 28*. Ljubljana: Znanstvena založba Filozofske fakultete. 385–91.
- Žagar, France, 2001: *Slovenska slovnica in vadnica*. Maribor: Obzorja.

Pridevniški polstavki v francoščini in prevedeni slovenščini

Adriana Mezeg

Oddelek za prevajalstvo, Filozofska fakulteta Univerze v Ljubljani
adriana.mezeg@gmail.com

Abstract

The article focuses on a contrastive corpus-based analysis of French sentence-initial adjectival detached constructions and their Slovenian translations, extracted from a 2.5-million word French-Slovenian parallel corpus of contemporary literary and journalistic texts (*FraSloK*). Syntactically and semantically opaque in the surface structure, detached constructions are commonly used in French written texts instead of semantically equivalent subordinate clauses, thus making the writing more fluid and dynamic. Even though such structures exist in Slovenian, little is known about their actual use and frequency, particularly in authentic contemporary written texts. When translating them from a foreign language, however, contrastive studies show that detached constructions are predominantly subject to morpho-syntactic expansions and are thus seldom retained in translations. This is also the case in the present study which confirms the proposed explicitation hypothesis in that around 80% of Slovenian translations of French sentence-initial adjectival detached constructions are syntactically and semantically more explicit than their French counterparts. The analysis also reveals some important differences in the strategies used in the two subcorpora and raises questions which need to be further explored.

Keywords: explicitation, translation strategies, adjectival detached constructions, automatic extraction of detached constructions, parallel corpus

Ključne besede: eksplicitacija, prevodne strategije, pridevniški polstavke, računalniško luščenje polstavkov, vzporedni korpus

1 UVOD

V francoskih pisnih besedilih so namesto stavkov z osebno glagolsko obliko iz stilističnih razlogov pogosto rabljene polstavčne strukture, s katerimi avtorji dosegajo večjo konciznost oziroma jedrnatost besedil (Rossi-Gensane 2009: 188). Polstavki zaradi okrnjenosti na skladenjski (npr. neizraženost osebkoma in, pri nekaterih tipih, povedka na površinski ravni) in na semantični ravni (implicitnost logične povezave z glavnim stavkom zaradi neizraženih stavčnih povezovalcev) predstavljajo težave, povezane s kognitivnimi dejavnostmi tvorjenja in recepcije besedil (Combettes 1998: 5–6), še posebno kadar se nahajajo na začetku povedi pred glavnim stavkom (prav tam). Omenjene strukture so lahko še toliko težje razumljive za govorce jezikov, v katerih ne obstajajo ali pa so redko rabljene, kakor npr. v slovenščini. Težave pri dekodiranju in prenosu vsebine francoskih polstavčnih struktur v slovenščino med drugim opažamo pri študentih francoščine z Oddelka za prevajalstvo ljubljanske Filozofske fakultete, zato želimo raziskati načine prevajanja francoskih polstavkov v slovenščino in na osnovi analize določiti dejavnike, ki bi olajšali skladenjsko in semantično interpretacijo teh francoskih struktur. V članku se bomo omejili samo na francoske pridevniške polstavke z začetka povedi pred glavnim stavkom (t. i. levi polstavki) in na njihove slovenske prevode, pridobljene z metodo računalniškega luščenja iz francosko-slovenskega korpusa publicističnih in literarnih besedil (*FraSloK*).

2 ZNAČILNOSTI FRANCOŠKIH PRIDEVNIŠKIH POLSTAVKOV IN HIPOTEZA

2.1 Splošne značilnosti polstavkov v francoščini

V francoščini poznamo več vrst polstavkov (fr. *constructions détachées*), ki se med seboj ločijo glede na jedro (deležje, sedanji in pretekli deležnik, pridevnik, samostalnik). Njihove skupne značilnosti so:

- a) v pisnem jeziku ločenost od preostalega besedila z nekončnim ločilom (vejico, dvopičjem, podpičjem, pomišljajem);
- b) stavčna mobilnost: nahajajo se lahko na začetku povedi pred glavnim stavkom, med osebkom in povedkom ali za osebkom in povedkom;
- c) sekundarna predikacija: polstavčna struktura na površinski ravni nima eksplicitno izraženega osebkoma, temveč samo predikat, tj. povedek oziroma del povedke, ki z osebkom (lahko tudi s predmetom oziroma z drugimi udeleženci, odvisnimi od povedja) glavnega stavka vzpostavi povezavo, identično tisti v popolni predikaciji;

- d) skriti nanosnik: polstavčna struktura obvezno implicira nanosnik, ki ga v glavnem stavku po navadi predstavlja osebek (lahko tudi predmet) (prirejeno po Combettes 1998: 10–13).

Posamezne značilnosti lahko ponazorimo z naslednjim primerom pridevniškega polstavka (primer (1)):

- (1) ***Fatiguée**, Marie s'est endormie tout de suite.*
 [****Utrujena**, je Marie takoj zaspala.*]¹

Podčrtani pridevniški polstavek sestoji iz pridevnika *fatiguée* (*utrujena*), ki stoji na začetku povedi (kriterij b)) in je od glavnega stavka ločen z vejico (kriterij a)). V strukturi s pridevnikom *fatiguée* osebek ni površinsko izražen, je pa razviden iz pridevniške končnice *-ée*, ki ustreza ženskemu spolu ednine in ga v glavnem stavku predstavlja lastno ime *Marie*. Polstavčna pridevniška struktura torej vsebuje skriti nanosnik (*Marie* était *fatiguée*/*Marie* je bila *utrujena*; kriterij d)) in vzpostavlja koreferenco med tem skritim nanosnikom in osebkom glavnega stavka. Ostane nam še kriterij sekundarne predikacije (c)), ki ga je mogoče razložiti s parafrazo *Marie était fatiguée, Marie s'est endormie tout de suite.* (*Marie je bila utrujena, Marie je takoj zaspala.*), pri čemer drugi del (glavni stavek) vsebuje glavno oziroma primarno predikacijo, prvi del v neparafraziranem primeru (1) zgoraj, ki je semantično podrejen drugemu, pa sekundarno predikacijo zaradi neizraženosti povedka na površinski ravni.

Po zgornjih kriterijih sodeč izpostavljene francoske strukture v slovenščini ustrezajo kategoriji *polstavek*, ki ga Toporišič (2004: 632) definira kot »tak del povedi (ali poved), ki mu je jedro neosebna glagolska oblika ali kak pridevnik ali samostalnik iz povedka«, kar lahko ponazorimo s primerom (2), ki izpolnjuje vse Combettesove kriterije (gre za primer pridevniškega polstavka v treh različnih stavčnih pozicijah):

- (2) ***Vsa rdeča v obraz od sramu**, je Eva odvihrala iz šole.*
*Eva je, **vsa rdeča v obraz od sramu**, odvihrala iz šole.*
*Eva je odvihrala iz šole, **vsa rdeča v obraz od sramu**.*

2.2 Francoski pridevniški polstavki

Jedro pridevniškega polstavka predstavlja pridevnik, ki se z nanosnikom (navadno z osebkom, redkeje s predmetom) glavnega stavka ujema v spolu in številu. Pridevnik v polstavku lahko uvaja nikalni prislov (*non*) ali kakšen drug prislov (npr. *plus*, *enfin*, *souvent*, *trop*), pa tudi veznik (npr. *quoique*). Poznavanje elementov, ki lahko tvorijo polstavek, je ključno za uspešno računalniško luščenje primerov.

¹ V članku so vsi slovenski prevodi v oglatih oklepajih dobesedni zaradi strukturne primerjave z izhodiščnimi francoskimi primeri in niso vedno slovnično ustrezni.

Predikacija pridevniškega polstavka je drugotna tako s skladenjskega kot s semantičnega vidika, zasidra se namreč v primarno predikacijo, ki je informacijsko pomembnejša (Noailly 1999: 124). Informacija, ki jo nosi levi polstavek, lahko okrepi koherenco z neposrednim predhodnim kontekstom in s tem zagotovi tematsko kontinuiteto (Neveu 1998; Noailly 1999: 125; Combettes 1998). Sekundarna predikacija pridevniškega polstavka lahko s primarno predikacijo vzpostavi opisno ali prislovno vrednost:

Opisna vrednost:

Skladenjsko gledano pridevniški polstavek izhaja iz oziralnega stavka (Dubois et al. 2007: 140), npr.:

- (3) *Le chat, (qui est) satisfait, ronronne.* (Dubois in Lagane 1973: 182)
[*Maček, (ki je) zadovoljen, prede.*]

Oziralni odvisnik samo prinaša dodatno informacijo, neko podrobnost, ki se nanaša na osebek glavnega stavka (Dubois in Lagane 1973: 182). O opisni vrednosti torej govorimo, kadar je med obema predikacijama težko vzpostaviti kakršno koli prislovno vrednost (Boch, Tutin in Laurent 2009: 103). V tem primeru lahko poveza med pridevniškim polstavkom in glavnim stavkom ustreza preprosti sopostavitvi (fr. *juxtaposition*) sestavin izreka (Combettes 1988: 115), npr.:

- (4) *Immobile, il regardait par la fenêtre. (Il regardait par la fenêtre et il était immobile.)* (Combettes 1988: 115)
[**Nepremičen*, je gledal skozi okno. (Gledal je skozi okno in bil (pri tem) nepremičen.) → *Nepremičen* je gledal skozi okno./*Nepremično* je gledal skozi okno.]]²

Kadar pridevniški polstavek lahko parafraziramo z vrinjenim vezalnim ali prislovnim stavkom, ima torej dopolnjevalno oziroma pojasnjevalno funkcijo (Riegel, Pellat in Rioul 1999: 150, 179), saj so vse lastnosti vezane na dopolnjevalni značaj informacije, ki jo prenašajo (prav tam: 179).

Prislovna vrednost:

Od prislovnih vrednosti pridevniško jedro polstavka z glavnim stavkom lahko vzpostavlja vzročno, časovno, dopustno/protivno in pogojno razmerje. Logično razmerje vzročnosti se pogosto prekriva s časovno vrednostjo (gl. primer (5)), ki izraža preddobnost ali sočasnost z glavno predikacijo (Boch, Tutin in Laurent 2009: 101), lahko pa tudi z več drugimi prislovnimi vrednostmi (npr. pogojnostjo, dopustnostjo). V takem primeru je pravilna določitev vrednosti odvisna od konteksta.

- (5) *Jeune, je ne comprenais pas de quoi il s'agissait.* (Havu 2002: 5)
[**Mlad*, nisem razumel, za kaj je šlo. → ?*Ker sem bil mlad*, nisem razumel, za kaj je šlo.]/?*Ko sem bil mlad*, nisem razumel, za kaj je šlo.]

² V slovenščini vsebino pridevniškega polstavka lahko izrazimo s povedkovim prilastkom ali s prislovnim določilom načina. V francoščini se prislov *immoblement* (*nepremično*) uporablja redko in deluje nerodno.

Na pravilno interpretacijo semantične vezi lahko vpliva glagolski čas povedka glavnega stavka (če je v pogojniku, je vez med polstavkom in glavnim stavkom pogojna). Kadar polstavek uvaja veznik (npr. *quoique* v primeru (6), ki eksplicira dopustnost (Wagner in Pinchon 1962: 149)), ta poudari semantično vrednost, ki je tako ne moremo zamenjati z drugimi vrednostmi:

- (6) *Quoique riche, il aidait toujours les pauvres.*
 [Čeprav bogat, je vedno pomagal revnim. → Čeprav je bil bogat, je vedno pomagal revnim.]

Francoskemu pridevniškemu polstavku v slovenščini ustreza polstavek s pridevniškim jedrom, če ima pridevnik ob sebi dopolnilo (gl. primer (2)). V prevodih francoskih del so namreč pridevniške besede, ki stojijo levo od odnosnice in niso skladenjsko razvite, včasih od preostalega besedila ločene z vejico, kar ni ustrezno glede na sprejeto pravopisno normo. Breznik (1967: 103) navaja primer (7) iz Flaubertovih *Treh povesti*, v katerem je francoski besedni red ohranjen v slovenskem prevodu, in ugotavlja, da je v slovenskem prevodu raba vejice napačna in »bi bila opravičena le, če bi povedkovo določilo nadomeščalo samostojen stavek, toda taka raba je romanska« (prav tam). Brez rabe vejice bi imela pridevniška zveza *usa bela* funkcijo povedkovnega prilastka. Pri analizi bomo morebitne takšne prevode, kakršen je ta pri (7), označili kot slovnično neustrezne.

- (7) *Toute blanche, elle éblouissait comme une nappe de neige.*
 **Vsa bela, se je blestela kakor sneženo polje.*

2.3 Raba polstavkov v slovenščini in hipoteza

V slovenščini nimamo na voljo veliko študij, ki bi se ukvarjale z vprašanjem pogostnosti rabe polstavčnih struktur. To velja predvsem za rabo v 20. in 21. stoletju, medtem ko je za 19. stoletje pogostnost nekaterih tovrstnih struktur podrobno raziskal Marko Jesenšek (1998). Pravopis in slovnice v povezavi s polstavki navajajo tipologijo in njihove glavne značilnosti, ne pomagajo pa nam odgovoriti na vprašanje, koliko se danes sploh še uporabljajo, tako da je govorec oziroma pisec glede tega prepuščen lastnemu jezikovnemu občutku. Do rabe polstavkov v slovenščini je bil kritičen Anton Breznik (1967), saj so po njegovih besedah polstavki »romanska posebnost in se jih moramo ogibati« (Breznik 1967: 88). Namesto njih je treba uporabljati stavke z osebno glagolsko obliko, kajti po Brezniku (1967: 94) slovenski stavek mora vsebovati glagol.

Obstoječe kontrastivne raziskave, ki so proučevale prevode polstavkov (npr. Lujčić 2006; Milojević Sheppard 1992, 1993), ugotavljajo redko ohranitev polstavčnih struktur in težnjo po rabi osebnih glagolskih oblik ter po eksplicitaciji logičnih povezav med izvirnimi polstavki in glavnimi stavki. Pri analizi

prevodov francoskih pridevniških polstavkov v slovenščino lahko pričakujemo enako strategijo, saj ohranitev francoske skladnje v slovenskih prevodih pogosto slovnično ni mogoča ali pa je slogovno neustrezna, kot kažejo dobesedni prevodi francoskih primerov na prejšnjih straneh (tj. primerov (1), (4), (5) in (7)). Na osnovi tega postavljamo hipotezo o eksplicitaciji skladenjsko-semantičnih prvin pri prevajanju francoskih pridevniških polstavkov v slovenščino. Pričakujemo, da bo hipoteza, ki jo bomo preverili v četrtem poglavju, večinsko potrjena v obeh podkorpuseh *FraSloK*.

V pričujoči raziskavi eksplicitacijo razumemo kot prevajalsko tehniko, s katero v ciljnem jeziku eksplicitno izrazimo tisto, kar je bilo v izhodiščnem jeziku implicitno, in sicer na podlagi sobesedila (Vinay in Darbelnet 1958, 1995: 342) in/ali (zunaj)jezikovnega znanja. To pripelje do večje eksplicitnosti ciljnega besedila. Pričakujemo lahko tudi, da bodo prevajalci polstavke ohranili v slovenskih prevodih, kar imamo za znak implicitacije, ki jo definiramo kot neubeseditev informacije (skladenjske, semantične, pragmatične itn.), ki jo naslovnik lahko razbere na podlagi sobesedila in/ali (zunaj)jezikovnega znanja (podobno Becher 2010: 2).

3 FRANCOSKO-SLOVENSKI KORPUS *FRASLOK*

Analitični del članka temelji na gradivu iz francosko-slovenskega korpusa *FraSloK*, ki je bil zgrajen v okviru doktorskega dela A. Mezeg (2011) in je prvi samostojni³ vzporedni korpus za ta jezikovni par. Njegova gradnja je potekala od novembra 2007 do januarja 2010, cilj pa je bil zgraditi elektronski vzporedni korpus sodobnih pisnih besedil velikosti približno milijon besed na jezik, kar naj bi omogočalo dovolj število podatkov za proučevanje raznovrstnih jezikovnih pojavov, predvsem slovničnih (Sinclair 1991: 100; Hunston 2002: 26).

Na končno podobo korpusa *FraSloK* sta poleg časovno-finančnih omejitev in zgoraj omenjenih vnaprej določenih kriterijev zasnove ključno vplivala dva dejavnika, in sicer razpoložljivost slovenskih prevodov izvirnih francoskih besedil in pridobivanje avtorskih pravic. Iz francoščine se v slovenščino prevajajo različni tipi besedil, npr. umetnostna, humanistična, poljudnoznanstvena, promocijska, publicistična, tudi pravna in politična besedila (več o tem v Mezeg 2011: 103–106). Pregled razpoložljivih jezikovnih virov za omenjeni jezikovni par je pokazal, da so za vključitev v korpus najprimernejša prozna besedila in časopisni članki iz mesečnika *Le Monde diplomatique* oziroma *Le Monde diplomatique v slovenščini*, in sicer zaradi treh ključnih pogojev, ki jih drugi tipi besedil ne

³ Z izjemo francosko-slovenske komponente besedil EU znotraj *Evrokorpora* za ta jezikovni par za zdaj ni na voljo drugih javno dostopnih korpusov, ki bi vsebovali druge vrste besedil.

izpolnjujejo v celoti: a) zadostno število besedil različnih avtorjev, prevajalcev in založb, da se izognemo prevelikemu vplivu posameznega avtorja ali prevajalca na rezultate analize, b) natančen podatek o jeziku izvirnika in c) popolna vzporednost, ki je nujna za stavčno poravnavanje besedil (Mezeg 2011: 107).

Pridobivanje korpusnega gradiva je bilo v veliki meri odvisno od volje potencialnih besedilodajalcev. Uredništvu *Le Monde diplomatique* in *Le Monde diplomatique v slovenščini* sta brez zadržkov dovolili vključitev časopisnih člankov v korpus in njihovo uporabo v raziskovalne namene; slovensko uredništvo nam je članke posredovalo v elektronski obliki, francoske pa smo pridobili iz spletnega arhiva *Le Monde diplomatique*. Zatem smo začeli z zbiranjem besedil za literarni podkorpus, in sicer najprej za slovenski del, kar je povezano z reševanjem avtorskih pravic. Glede na izkušnje drugih graditeljev korpusov, da predvsem nosilci avtorskih pravic literarnih del niso naklonjeni temu, da bi bila njihova dela vključena v (javno dostopne) elektronske korpuse, smo menili, da literarnega podkorpusa ne bo mogoče oblikovati, če bo že odziv slovenskih založb negativen. Večina slovenskih založb, na katere smo se obrnili, se je pozitivno odzvala naši prošnji in nam želena besedila tudi poslala po elektronski pošti. Drugačno izkušnjo smo imeli s francoskimi založbami, ki besedila zelo nerade dajo iz rok, pa čeprav zgolj v raziskovalne namene. Po dolgotrajnih pogajanjih nam je uspelo dobiti dovoljenje za vključitev večine zelenih francoskih romanov v korpus, medtem ko z nekaterimi nosilci avtorskih pravic pogajanja še potekajo. Popoln seznam besedil, vključenih v *FraSloK*, je objavljen v Mezeg 2011: 345–370.

Korpus *FraSloK* torej sestavljata dva podkorpusa, v katera so vključena celotna besedila, in sicer:

- a) podkorpus publicističnih besedil v skupnem obsegu 1 164 074 besed oziroma pojavnic, v katerega je vključenih 300 francoskih člankov iz časnika *Le Monde diplomatique* in njihovih prevodov iz slovenske različice *Le Monde diplomatique v slovenščini*, objavljenih med leti 2005 in 2009;
- b) podkorpus literarnih besedil v skupnem obsegu 1 302 911 pojavnic, v katerega je vključenih dvanajst sodobnih francoskih romanov in njihovih slovenskih prevodov, objavljenih med leti 1995 in 2008.

Kot je razvidno iz tabele 1 na naslednji strani, korpus vsebuje nekaj manj kakor 2,5 milijona pojavnic. Literarni podkorpus je le slabih 6 % večji od publicističnega podkorpusa, tako da še vedno lahko rečemo, da je korpus *FraSloK* uravnotežen z vidika zajetih tipov besedil. Razlika v velikosti je pogojena z dejstvom, da smo želeli v korpus vključiti celotna besedila, zato bi bilo težko doseči popolno sorazmerje, saj dolžina romanov ni primerljiva z dolžino časopisnih člankov. Razlika med podkorpusoma je vidna tudi na jezikovni ravni, saj sta slovenska dela v

posameznih podkorpused nekoliko manjša od vzporednih francoskih delov. Pripišemo jo lahko dejstvu, da v nasprotju s francoščino samostalniških besednih zvez v slovenščini ne uvajajo členi in da so osebni zaimki zajeti v glagolskem obrazilu, ne pa eksplicitno izraženi v površinski strukturi,⁴ seveda pa bi bila za natančnejše vzroke potrebna podrobnejša analiza.

Tabela 1: Velikost korpusa *FraSloK*.

	Publicistični podkorpus	Literarni podkorpus	Število pojavnice glede na jezik
Francoski del	637 297	701 715	1 339 012
Slovenski del	526 777	601 196	1 127 973
Število pojavnice v podkorpusu	1 164 074	1 302 911	
Število pojavnice v korpusu <i>FraSloK</i>	2 466 985		2 466 985

Po končanem zbiranju besedil sledi tehnični vidik gradnje, katerega cilj je priprava besedil za poznejše raziskovanje in analiziranje korpusa s konkordančnimi orodji. Z izjemo francoskih romanov, ki jih je bilo treba digitalizirati in zaradi mogočih napak pri preslikavanju posledično ročno pregledati ter popraviti, smo vsa druga besedila, kot že rečeno, pridobili v elektronski obliki bodisi od besedilodajalcev bodisi s svetovnega spleta. Besedila so bila nato pretvorjena v enoten elektronski zapis, in sicer v navadne besedilne datoteke (.txt) oziroma format ANSI, ki je predpogoj za obdelavo besedil z orodjem *ParaConc* (Barlow 2005). Zanj smo se odločili, ker od dostopnih vzporednih komercialnih konkordančnikov omogoča obdelavo besedil v slovenskem jeziku in ima dobre iskalne možnosti. To orodje smo uporabili tudi za stavčno poravnavo izvirnih in prevedenih besedil. Vse besedilne datoteke so bile tudi avtomatsko označene, in sicer francoske besednovrstno z označevalnikom *TreeTagger* (Schmid 1994), slovenske pa oblikoskladenjsko z orodjem *ToTale* (Erjavec et al. 2005).⁵ Do korpusa *FraSloK* je mogoče dostopati prek konkordančnika *ParaConc*, ki omogoča osnovno iskanje po pojavnica ali napredno iskanje na osnovi regularnih izrazov in/ali besednovrstnih oznak. pa Zahvaljujoč projektu *Slovensko prevodoslovje – viri in raziskave* (Vintar 2009) je korpus dostopen tudi na spletu, in sicer literarni podkorpus prek iskalnika na spletnem naslovu omenjenega projekta (<http://lojze.lugos.si/spook/index.html>), publicistični podkorpus pa prek korpusne orodjarne *Sketch Engine* (http://nl.ijs.si/noske/all.cgi/first_form; LeMonde: slovensko in LeMonde: francosko).

⁴ V slovenščini je slovnični osebek razviden iz glagolskega obrazila (npr. *je strgal* → on), v francoščini pa mora biti osebni zaimek izražen, ko se prvič pojavi v povedi (npr. *il arrachait*).

⁵ Označevanje je opravil T. Erjavec z Instituta Jožef Stefan.

4 RAČUNALNIŠKO LUŠČENJE FRANCOSKIH PRIDEVNIŠKIH POLSTAVKOV

Dosedanje raziskave polstavkov so večinoma temeljile na ročno izpisanih primerih iz tiskanih besedil, kar je po eni strani mogoče pripisati zapletenosti in zamudnosti gradnje primerne korpusa, če nam od že obstoječih nobeden ne ustreza ali nam ni dostopen, po drugi strani pa zapletenosti oblikovanja ukazov za avtomatsko iskanje polstavkov zaradi omejevalnih kriterijev, ki določajo njihovo zgradbo (npr. položaj v povedi, neizraženost osebka in osebne glagolske oblike na površinski ravni, vrsta jedra in upoštevanje vseh besednih vrst, ki se lahko nahajajo okrog jedra). Metodologija avtomatskega luščenja polstavkov iz elektronskega korpusa, ki jo bomo uporabili v raziskavi, tako predstavlja novost na področju proučevanja polstavkov.

Luščenja francoskih pridevniških polstavkov iz korpusa *FraSloK* smo se lotili prek naprednega iskanja orodja *ParaConc* na osnovi ukazov, sestavljenih iz besednovrstnih oznak in regularnih izrazov, ki omogočajo iskanje kompleksnih slovničnih vzorcev. Ob upoštevanju zgradbe pridevniških polstavkov oziroma besednih vrst, ki se lahko pojavljajo okrog pridevniškega jedra, smo za iskanje tovrstnih polstavkov uporabili naslednja dva ukaza:

Ukaz 1:

$$(<w\ R>[A-Z]\backslash w+(\backslash W<w\ R>\backslash w+)?\backslash W<w\ P>\backslash w+|<w\ P>[A-Z]\backslash w+(\backslash W?<w\ V>\backslash w+(\backslash W<w\ R>\backslash w+)\{0,2\}\backslash W<w\ P>\backslash w+)?$$

S prvim ukazom smo želeli poiskati vse leve polstavke, ki jih neposredno uvaja pridevnik (*P*), poleg tega pa tudi primere, ko polstavke s pridevniškim jedrom uvajajo druge besedne vrste, in sicer v naslednjih kombinacijah: od 0 do 2 prislova (*R*) + obvezni pridevnik (*P*) / od 0 do 2 prislova (*R*) + veznik (*V*) + od 0 do 2 prislova (*R*) + obvezni pridevnik (*P*). Primer (8) kaže kombinacijo dveh prislovov pred pridevniškim polstavčnim jedrom:

- (8) ***Bien moins solide qu'il ne le paraissait**, le groupe a brusquement cessé ses versements en 2007, contraignant le gouvernement à solliciter des banques locales un prêt de 2 milliards d'euros.* (LMD156-FRA)⁶
[Družba, ki ni tako stabilna, kot je sprva kazalo, je leta 2007 nenadoma ustavila svoje transakcije in vlado prisilila v najem posojila v vrednosti 2 milijard € pri domači banki.] (LMD156-SLV)

Ta ukaz se je izkazal za najproduktivnejšega, saj smo z njim izluščili 103 oziroma 82 % vseh pridevniških polstavkov iz publicističnega podkorpusa (LMD) in

⁶ Oznaka v oklepaju je identifikacijska številka besedila, iz katerega je vzet dani primer. LMD156-FRA se nanaša na besedilo 156 iz francoskega dela publicističnega korpusa (LMD). Za seznam korpusnih besedil gl. Mezeg 2011, Priloga 1.

169 oziroma 64 % vseh pridevniških polstavkov iz literarnega podkorpusa (LIT). Neuporabnih zadetkov je bilo v obeh podkorpusih več kakor 80 % (LMD 87 %, LIT 84 %), predvsem iz dveh razlogov: a) zaradi napačno pripisanih besednovrstnih oznak nekaterim besedam (npr. nekatere pretekle deležniške oblike in samostalniki (primer (9)) so v korpusu po pomoti označeni kot pridevniki); b) kadar besednovrstno pravilno označene besede ne nastopajo v polstavčni strukturi (primer (10)):

- (9) **Collégienne*, je venais de changer de ville d'études, comme souvent. (LIT008-FRA)⁷
 [**Ko sem se vpisala v nižjo gimnazijo, sem se morala, kot se pogosto zgodi, preseliti.* (LIT008-SLV)]
- (10) *Nouvelle soirée catastrophique avec Vincent.* **Toujours fou de désir pour lui. Il arrive à sept heures et dit qu'il repartira à neuf.* (LIT001-FRA)
 [Ponovno katastrofalen večer z Vincentom. **Še vedno nor* od želje po njem. *Pride ob sedmih in reče, da bo odšel ob devetih.* (LIT001-SLV)]

Ukaz 2:

[;“-,](\W)?(<w R>\w+(\W<w R>\w+)?\W<w P>\w+|<w P>\w+)(\W?<w V>\w+(\W<w R>\w+){0,2}\W<w P>\w+)?

Drugi ukaz se od prvega razlikuje le v tem, da išče enake kombinacije besednih vrst za nekončnim ločilom ([;“-,]). Primer (11) kaže kombinacijo prislova in pridevnika v polstavčni strukturi:

- (11) *En l'absence des forces de l'ordre, je le savais, j'aurais constitué une proie facile, quoique peu intéressante ; très conventionnel, mon habilement de cadre moyen n'avait rien qui puisse les séduire.* (LIT009-FRA)
 [Vedel sem, da bi bil v odsotnosti organov reda lahek, čeprav nezanimiv plen; zelo nevpadljiva obleka uradnika srednjega razreda, ki sem jo nosil, jih ni mogla pritegniti. (LIT009-SLV)]

S tem ukazom smo izluščili 23 polstavkov (18 % vseh polstavkov) iz publicističnega podkorpusa in 94 polstavkov (36 % vseh polstavkov) iz literarnega podkorpusa, kot kaže tabela 2 na naslednji strani. Neustreznih zadetkov je bilo v obeh korpusih veliko, deloma tudi zato, ker je ukaz v osnovi namenjen luščenju desnih polstavkov.

Kot je razvidno iz tabela 2, smo z danima ukazoma izluščili 126 levih pridevniških polstavkov iz publicističnega podkorpusa in 263 iz literarnega, skupno torej 389 levih pridevniških polstavkov.

⁷ V (9) je *Collégienne* jedro samostalniškega polstavka.

Tabela 2: Število ustreznih zadetkov pridevniških polstavkov glede na dane ukaze.

ukaz	Literarni podkorpus (LIT)		Publicistični podkorpus (LMD)	
	uporabni zadetki	%	uporabni zadetki	%
ukaz 1	169	64 %	103	82 %
ukaz 2	94	36 %	23	18 %
skupaj	263	100 %	126	100 %

Dane številke kažejo, da literarni podkorpus vsebuje 35 % več primerov pridevniških polstavkov od publicističnega podkorpusa. Ob upoštevanju velikosti posameznih podkorpusov (LMD 637 297 besed, LIT 701 715 besed) normalizacija (gl. Biber, Conrad in Reppen 1998: 263) izluščenih polstavkov pokaže, da na 2000 besed v publicističnem podkorpusu najdemo 0,4 polstavke, v literarnem pa 0,7 polstavkov; razlika v distribuciji polstavkov je torej približno 1 proti 1,75.

Ker se polstavki ne pojavljajo enakomerno v vsakem besedilu, si je smiselno ogledati njihovo distribucijo v posameznih korpusnih besedilih. Iz tabele 3 na naslednji strani je razvidno, da tretjino izluščenih polstavkov iz literarnega podkorpusa najdemo v romanu *Impératrice* (*Cesarica*, Shan Sa, 32,7 %), slabo četrtno v romanu *Le ventre de l'Atlantique* (*Trebuh Atlantika*, Diome Fatou, 23,6 %), približno desetino pa v romanih *Le testament français* (*Francoski testament*, Andrei Makine, 11,4 %) in *Je m'en vais* (*Grem*, Jean Echenoz, 10,3 %); drugi romani vsebujejo manj kakor 10 % vseh polstavkov, pri čemer nekateri ne vsebujejo niti enega pridevniškega polstavka (romana *Truismes* (*Svinjarije*, Marie Darrieussecq) in *Mammifères* (*Sesalci*, Pierre Mérot)). Glede na dolžino besedila po številu polstavkov na 2000 besed izstopa roman *Le ventre de l'Atlantique* z dvema polstavkoma, iz povprečja pa z več kakor enim pridevniškim polstavkom na 2000 besed izstopajo še romani *Impératrice* (1,6), *Un secret* (*Skrivnost*, Philippe Grimbert, 1,2) in *Je m'en vais* (1,1), medtem ko je Makinov *Le testament français* ravno v povprečju (0,7).

Tabela 3: Porazdelitev pridevniških polstavkov v posameznih besedilih francoskega dela literarnega podkorpusa.

Besedilo	Število polstavkov	Normalizacija (na 2000 besed)	Delež glede na skupno št. polstavkov
LIT001-FRA-Guibert	1	0,1	0,4 %
LIT002-FRA-Grimbert	15	1,2	5,7 %
LIT003-FRA-Echenoz	27	1,1	10,3 %
LIT004-FRA-Sijie	7	0,3	2,7 %
LIT005-FRA-Darrieussecq	0	0	0,0 %
LIT006-FRA-Gaude	1	0,04	0,4 %
LIT007-FRA-Merot	0	0	0,0 %
LIT008-FRA-Diome	62	2,0	23,6 %
LIT009-FRA-Houellebecq	13	0,3	4,9 %
LIT010-FRA-Makine	30	0,7	11,4 %
LIT011-FRA-ShanSa	86	1,6	32,7 %
LIT012-FRA-Bruckner	21	0,5	8,0 %
skupaj	263	/	100 %

Glede na tabelo 4 v publicističnem podkorpusu slabih 70 % časopisnih člankov iz *Le Monde diplomatique* ne vsebuje nobenega levega pridevniškega polstavka, po en polstavek najdemo v dobri petini (22,7 %) člankov, dva v 24 člankih (8,0 %), štiri oziroma šest pridevniških polstavkov pa vsebuje le po en članek.

Tabela 4: Porazdelitev pridevniških polstavkov v francoskem delu publicističnega podkorpusa.

Pogostnost	Število člankov	Delež glede na celoto
0	206	68,7 %
1	68	22,7 %
2	24	8,0 %
3	0	0,0 %
4	1	0,3 %
5	0	0,0 %
6	1	0,3 %
skupaj	300	100 %

5 STRATEGIJE PREVAJANJA FRANCOSKIH PRIDEVNIŠKIH POLSTAVKOV V SLOVENŠČINO

V tem poglavju bodo predstavljeni rezultati skladenjsko-semantične analize slovenskih prevodov francoskih pridevniških polstavkov, izluščenih iz korpusa *FraSloK*.

5.1 Skladenjska analiza slovenskih prevodov

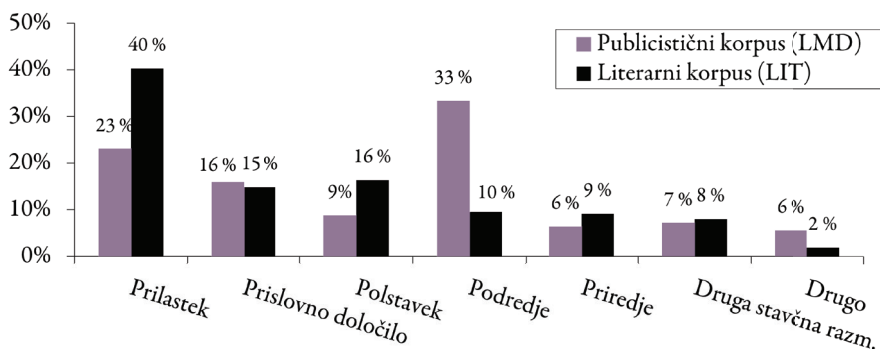
Prevode francoskih pridevniških polstavkov smo najprej analizirali glede na stavčne vloge in stavčna razmerja, da bi ugotovili, kako so prevajalci francoskih časopisnih člankov in romanov iz posameznih podkorpusov *FraSloK* prenesli v slovenščino njihovo vsebino.

Kot je razvidno iz grafa 1 spodaj, se razporeditev prevodnih strategij med podkorpusoma nekoliko razlikuje, predvsem pri najbolj zastopanih kategorijah. V slovenskem delu publicističnega podkorpusa je tretjina (33,3 % oziroma 42 primerov) pridevniških jeder francoskih polstavkov izražena z osebno glagolsko obliko ali s povedkovim določilom v odvisnem stavku (kategorija 'Podredje'), medtem ko je v literarnem podkorpusu ta strategija četrta po pogostosti in zajema približno desetino (9,5 % oziroma 25 primerov) prevodov.

Pridevniška funkcija francoskih polstavčnih jeder, v slovenščini izražena s povedkovim določilom (pridevniškega izvora) v odvisnem stavku, je v publicističnem podkorpusu ohranjena v 81 % prevodov (34 primerov) (gl. primer (12)–(12a)), v literarnem pa v 44 % prevodov (11 primerov):

(12) *Habituellement intense et chaotique, le trafic entre les deux pays n'a pas encore repris son cours.* (LMD170-FRA)

(12a) *Promet med obema državama, ki je ponavadi intenziven in kaotičen, se še ni obnovil.* (LMD170-SLV)



Slika 1: Porazdelitev slovenskih skladenjskih ustreznih francoskih pridevniških polstavkov.

Vsebina francoskih pridevniških jeder je torej v okviru podredja z osebno glagolsko obliko pogostejše izražena v literarnem podkorpusu (14 primerov, LMD 8 primerov), poglagoljenje oziroma verbalizacija pa je opaziti predvsem v primeru naslednjih francoskih pridevnikov: *capable/incapable* (sposoben/nesposoben → znati, močil/ne znati, ne moči), *scient/inscient* (zavedati selne zavedati se), *célèbre* (slaven → sloveti), *impatient* (nepotrpežljiv, neučakan → komaj čakati), *inquiet* (zaskrbljen → skrbeti), *las* (naveličan → naveličati se), *soucieux* (zaskrbljen → skrbeti) in *vivant* (živ → živeti, preživeti (primer (13)–(13a))):

(13) *Vivante*, je renaîtrais plus forte que jamais ! (LIT011-FRA)

(13a) *Če preživim*, se bom znova rodila, močnejša kot kdajkoli prej! (LIT011-SLV)

V obeh danih primerih ((12a) in (13a)) je skladenjska eksplicitacija obvezna, saj francoska pridevniška jedra nimajo desnega dopolnila, zato ohranitev polstavčne strukture v slovenščini ni mogoča (npr. (13) **Živa*, se bom znova rodila, močnejša kot kdajkoli prej!).

Najpogostejša prevodna strategija v slovenskem delu literarnega podkorpusa je izraženost vsebine francoskih pridevnikov iz polstavčne strukture s pridevnikom v funkciji prilastka ali povedkovnega prilastka (40,3 %), in sicer znotraj enostavne povedi ali glavnega stavka zložene povedi (kategorija 'Prilastek'), medtem ko je v publicističnem podkorpusu ta strategija druga po pogostosti in zajema 23,0 % prevodov. V publicističnem podkorpusu v tej kategoriji prevladuje raba levega ujemalnega pridevniškega prilastka (75,9 %; primer (14)–(14a)) pred povedkovim prilastkom pred odnosnico (20,7 %). V literarnem podkorpusu ima dobra polovica francoskih pridevniških jeder funkcijo povedkovnega prilastka (51,9 %; primer (15)–(15a)), medtem ko je raba levega prilastka opaziti v dobrih dveh petinah prevodov (42,5 %). Raba povedkovnega prilastka za pomožnim glagolom oziroma povedkom je v obeh korpusih redka (LMD 1 primer, LIT 6 primerov).

(14) *Illégale* - bien que le Conseil de sécurité ait refusé de la condamner -, l'intervention fait émerger de nouvelles questions. (LMD071-FRA)

(14a) *Ta nezakonita* intervencija – čeprav se Varnostni svet ni odločil za njeno obsodbo – je sprožila celo kopico novih vprašanj. (LMD071-SLV)

(15) *Le sable de la plage respirait la miséricorde. Plat, blanc, fin et poreux*, il laissait les vagues venir timidement sucer son âme. (LIT008-FRA)

(15a) *Pesek na plaži je puhtel od usmiljenja. Ves raven, bel, droben in prosojen* je puščal valovom, da so mu plaho sesali dušo. (LIT008-SLV)

V obeh danih primerih je skladenjska eksplicitacija zaradi nerazvitih francoskih pridevniških jeder (so brez desnega dopolnila) v slovenščini obvezna⁸ in

⁸ Ohranitev direktne strukture v slovenščini slovnično ni ustrezna, kadar pridevnik v polstavku nima desnega dopolnila (podobno Breznik 1967: 103): (14) **Nezakonita* – čeprav se Varnostni svet ni odločil za njeno obsodbo –, je ta intervencija sprožila celo kopico novih vprašanj; (15) **Raven, bel, droben in prosojen*, je puščal valovom, da so mu plaho sesali dušo.

vsebina pridevnikov, ki opisujejo nanosnika glavnega stavka, bolj integrirana v slovensko poved kakor v francoščini. Polstavke iz primera (15) bi v slovenščini lahko izrazili tudi s povedkovim določilom znotraj vezalnega priredja (*Bil je ves raven, bel, droben in prosojen in je puščal valovom, da so ...*), medtem ko raba prilastkovega odvisnika v konkretnem primeru, kakor tudi v večini podobnih primerov iz korpusa, ni mogoča, saj je nanosnik glavnega stavka pozajmljen (v (15) *il laissait ...*), s ponovitvijo samostalniške besedne zveze (npr. *Pesek v (15a)*) pa bi kršili medpovedno kohezijo (*Pesek na plaži je puhtel od usmiljenja. ?Pesek, ki je bil ves raven, bel, droben in prosojen, je puščal ...*).⁹

Ohranitev polstavčne strukture je druga najpogostejša prevajalska strategija v literarnem podkorpusu (16,3 % oziroma 43 primerov), v publicističnem podkorpusu pa zajema manj kakor desetino vseh prevodov (8,7 % oziroma 11 primerov) in je četrta po pogostnosti. V publicističnem podkorpusu so vsi polstavki ohranili pridevniško jedro, glede na izvirnik pa se je v slabih treh četrтинah primerov (8 oziroma 73 %) spremenilo njihovo mesto, saj so bili prestavljeni za odnosnico (primer (16a)); v tem položaju ustrezajo reduciranemu prilastkovemu odvisniku (v (16a) ..., *ki je odgovoren* ...) in so skladenjsko manj implicitni od izhodiščnih francoskih struktur:

(16) *Responsable d'énormes opérations de blanchiment à travers des maisons de change amies, son principal assistant était un ancien cadre de la Banque mexicaine.* (LMD260-FRA)

(16a) *Njegov glavni pomočnik, odgovoren za izjemno obsežne operacije pranja denarja preko prijateljskih menjalniških družb, je nekdanj delal v vodstvu Mehške banke.* (LMD260-SLV)

V literarnem podkorpusu je pridevniško jedro ohranjeno v 91 % prevodov (39 primerov): od tega je v približno 60 % primerov mesto polstavka ohranjeno pred glavnim stavkom, v drugih primerih pa se polstavki nahajajo za odnosnico. V slovenščini je raba pridevniškega polstavka pred odnosnico mogoča, kadar je pridevniško jedro skladenjsko razvito oziroma ima desno dopolnilo (primer (17)–(17a)):

(17) *Honteux comme celui que l'on surprend en train de parler avec lui-même, je me voyais tel que j'étais.* (LIT010-FRA)

(17a) *Osramočen kot nekdo, ki ga zalotimo, da se pogovarja sam s sabo, sem se zagledal tak, kakršen sem.* (LIT010-SLV)

Analiza kaže, da so nekateri levi polstavki ohranjeni tudi v primerih, ko francosko pridevniško jedro nima desnega dopolnila, npr.:

⁹ Sama poved (*Pesek, ki je bil ves raven, ...*) je sicer povsem ustrezna, vendar iz že omenjenega razloga ne v danem kontekstu oziroma v dani navezavi.

- (18) *Hier soir, **légèrement ivre**, en voyant ma main droite posée sur la nappe du dîner, éclairée par les guirlandes de Noël dispersées dans le laurier-rose, il m'a semblé évident que cette main n'avait été créée que pour pouvoir caresser Vincent ; ce matin, cette émotion est devenue abusive.* (LIT001-FRA)

- (18a) *Sinoči, **rahlo pijan**, ko sem gledal svojo desno roko, položeno na namizni prt, osvetljeno z božičnimi kitami, ovitimi po oleandru, mi je bilo očitno, da je bila ta roka ustvarjena samo zato, da lahko boža Vincenta; zjutraj se zdi to čustvo pretirano.* (LIT001-SLV)

V (18a) je prevajalec direktno prenesel v slovenščino francoske strukture brez spremembe besednega reda. Najbolj izstopa prav ohranitev levega pridevniškega polstavka, ki deluje popolnoma nepovezan s preostalim besedilom, predvsem zaradi prislovnega določila časa pred njim in časovnega odvisnega stavka, ki mu sledi. Prevod bi bil v slovenščini ustrezen, če bi polstavek prestavili za pomožni glagol časovnega odvisnika (*Sinoči, ko sem, **rahlo pijan**, gledal svojo desno roko, ...*) ali ga pretvorili v povedkov prilastek in (skupaj s prislovnim določilom) integrirali v odvisni stavek (npr. *Ko sem sinoči **rahlo pijan** gledal svojo desno roko, ...*).

Predvsem v romanu *Impératrice* (*Cesarica*) v nekaj primerih (npr. (19)) opazimo naštevanje golih pridevnikov v polstavčni strukturi, ki jih je prevajalka najbrž zaradi želje po poustvaritvi slogovnega učinka direktno prenesla v slovenščino (npr. (19a)):

- (19) ***Rouge, bleue, verte, scintillante, évanescence**, chaque étoile était une écriture mystérieuse et le ciel un livre sacré.* (LIT011-FRA)

- (19a) ***Rdeča, modra, zelena, iskriva, minljiva**, vsaka zvezda je bila skrivnostna pisava in nebo sveta knjiga.* (LIT011-SLV)

Kljub temu da je takšna »romanska« (Breznik 1967: 103) raba v slovenščini ne navadna, jo prav iz stilističnih razlogov (kopičenje ali asindeton) v literarnem besedilu lahko upravičimo, medtem ko bi jo v drugih vrstah besedil označili kot slovnično neustrezno. Gole pridevnike sicer lahko integriramo v slovensko poved z rabo prilastka, povedkovnega prilastka ali povedkovnega določila. V danem primeru ((19a)) bi bila npr. ustrežna raba povedkovnega določila znotraj dopustnega odvisnika z veznikom *najsi* za odnosnico (*Vsaka zvezda, **najsi bo rdeča, modra, zelena, iskriva, minljiva**, je bila skrivnostna pisava ...*), sprejemljiva pa bi bila tudi poved s spremenjenim besednim redom (prestavitev pridevnikov za odnosnico) brez skladenjsko-semantične eksplicitacije (*Vsaka zvezda, **rdeča, modra, zelena, iskriva, minljiva**, je bila skrivnostna pisava ...*).

Poleg večinske ohranitve pridevniških polstavkov je bila v treh primerih iz literarnega korpusa vsebina francoskih pridevniških jeder verbalizirana z deležjem na -č

(vedno na osnovi pridevnika *conscient* → dvakrat *zavedajoč se* za odnosnico, enkrat *vedoč* pred odnosnico), v enem pa z deležnikom na -č z ohranjenim mestom pred odnosnico (primer (20)–(20a)):

(20) *Dégoulinante de sueur dans ses longues robes de basin, les talons enfoncés dans le sable, elle jouait de son regard et intercalait quelques mots français.* (LIT008-FRA)

(20a) *Cedeča se od znoja, je v dolgih žametnih oblekah in s petkami, ugreznjenimi v pesek, uživala v njegovem pogledu in v stavke vpletala posamezne francoske besede.* (LIT008-SLV)

V (20a) verbalizacija ni presenetljiva, saj je *dégoulinante* pravzaprav izglagolski pridevnik (iz glagola *dégouliner* – *curljati, kapljati, cediti se*), ki ga v slovenščini lahko izrazimo le z deležjem (*cedeč se*) ali deležnikom (*cedeča se*). Primer bi pravzaprav sodil k francoskim polstavkom s sedanjo deležniško obliko, če bi se avtorica romana odločila za normativno rabo (tj. *Dégoulinant de sueur..., elle jouait...*).

Po pogostnosti tretja strategija, ki je v obeh podkorpusih uporabljena v približno 15 % vseh prevodov (LMD 15,9 %, LIT 14,8 %), je poprislovljenje pridevniških jeder francoskih polstavkov (kategorija 'Prislovno določilo'). V tej kategoriji je vsebina francoskega pridevniškega jedra iz polstavčne strukture v slovenskih prevodih iz korpusa *FraSloK* najpogostejše izražena s prislovno obliko (primer (21)–(21a)) ali s predložno samostalniško besedno zvezo (primer (22)–(22a)):

(21) *Immobile, je laisse les bruits et les odeurs m'envahir.* (LIT006-FRA)

(21a) *Negibno sedim in pustim, da me preplavljajo zvoki in vonjave.* (LIT006-SLV)

(22) *Silencieux, deux pêcheurs débarquèrent leur cargaison.* (LIT008-FRA)

(22a) *Dvoje ribičev je v tišini raztovorilo svoj ulov.* (LIT008-SLV)

V danih primerih zaradi golega pridevniškega jedra direktna ohranitev levega polstavka v slovenščini ni mogoča.¹⁰ V obeh primerih je pridevniško obliko sicer mogoče ohraniti v funkciji povedkovega prilastka (npr. *Negiben sedim in pustim, ...; Molčeca sta ribiča raztovorila ...*), vendar sta se prevajalca odločila za rabo prislova (v (21a)) oziroma predložne samostalniške zveze (v (22a))¹¹ v funkciji prislovnega določila.

Če izvezemo izpuste prevodov francoskih pridevniških polstavkov (kategorija 'Drugo'; LMD 5,6 %, LIT 1,9 %), nam ostaneta še kategoriji 'Priredje' (LMD

¹⁰ Ne moremo reči: **Negiben, sedim in pustim, da me preplavljajo zvoki in vonjave*. (nanosnik glavnega stavka je moškega spola) ali **Molčeca, sta ribiča raztovorila svoj ulov*., temveč je treba francoske pridevnike iz polstavčne strukture v slovenščini integrirati v poved z glagolom (npr. *Negiben/Negibno sedim in pustim, da me preplavljajo ...; Molčeca/Molče/V tišini sta ribiča raztovorila svoj ulov*).

¹¹ Mogoča je tudi raba prislova: *Molče sta ribiča raztovorila svoj ulov*.

6,3 %, LIT 9,1 %) in 'Druga stavčna razmerja' (LMD 7,1 %, LIT 8,0 %), ki sta v obeh podkorpused zastopani v manj kakor desetini prevodov.

V okviru priredja je vsebina večine izluščenih francoskih pridevniških jeder iz obeh podkorpused *FraSloK* v slovenskih prevodih izražena s pridevnikom v funkciji povedkovega določila (LMD 6 primerov od 8 (75 %), LIT 16 primerov od 24 oziroma 67 %) (primer (23)–(23a)), medtem ko je verbalizacijo francoskih pridevnikov opaziti v tretjini prevodov iz literarnega podkorpused (8 primerov) in v četrtini prevodov (2 primera) iz publicističnega podkorpused (primer (24)–(24a)):

(23) *Plus jeune que Maxime, Robert a été mobilisé sur le front de l'Est.*
(LIT002-FRA)

(23a) *Robert je mlajši od Maxima in vpokličejo ga v vojsko na vzhodno fronto.*
(LIT002-SLV)

(24) *Originaires de l'Etat de Sinaloa, le creuset du narcotrafic au Mexique, ils travaillaient dans le passé comme seconds couteaux des grandes organisations.* (LMD260-FRA)

(24a) *Slednji izvirajo iz zvezne države Sinaloa, »talilnega lonca« mehiške trgovine z mamili, in so v preteklosti delovali kot podaljšane roke velikih organizacij.* (LMD260-SLV)

Od drugih stavčnih razmerij v obeh podkorpused izstopa protistava (LMD 6 primerov od 9 oziroma 67 %, LIT 15 primerov od 21 oziroma 71 %), primeri zapleteno zloženih povedi so redki. Pri obeh tipih razmerij je sicer z izjemo enega primera pridevniška oblika v slovenskih prevodih ohranjena in ima funkcijo povedkovega določila:

(25) *Tranquille, vertueuse, soumise et désintéressée, elle doit attirer l'admiration des quatre mers et satisfaire la volonté des dieux.* (LIT011-FRA)

(25a) *Biti mora mirna, krepostna, pokorna in nesebična; vzbujati mora spoštovanje štirih mórij in izpolnjevati voljo bogov.* (LIT011-SLV)

V (25a) je skladenjska eksplicitacija obvezna, pomensko razmerje med polstavčnima deloma in glavnim stavkom pa ni eksplicitno izraženo. Ker se označena pridevnika iz (25) nanašata na pozajmljenega nanosnika (*elle/ona* = *cesarica*, omenjena v predhodnem kontekstu), izhodiščne opisne vrednosti zaradi ohranitve zaimenskega navezovanja ni mogoče eksplicitirati z rabo prilastkovega odvisnika ali prilastka (gl. razpravo pri primeru (15)–(15a)), vprašljiva pa bi bila tudi raba vezalnega priredja (npr. *Biti mora mirna, krepostna, pokorna in nesebična ter vzbujati spoštovanje štirih in izpolnjevati voljo bogov.*) zaradi kopičenja veznikov in enakovrednosti vsebine polstavčne strukture drugemu delu priredne povedi – polstavek ima namreč v francoščini pojasnjevalno oziroma dopolnjevalno funkcijo in prinaša informacije, ki so drugotnega pomena.

V nekaterih primerih protistave ugotavljamo, da bi bilo pomensko razmerje med deli povedi mogoče eksplicitirati na več načinov. V (26) je npr. med francoskim polstavkom in glavnim stavkom očitna vzročno-posledična povezava, ki se v slovenskem prevodu (primer (26a)) izgubi, čeprav bi jo lahko ohranili z rabo vzročnega odvisnega (*Ker je zelo nevarna, jo navadno zaupajo ...*) ali prirednega stavka (*Zelo je nevarna, zato jo navadno zaupajo ...*). Takšni primeri zmanjšujejo stopnjo semantične eksplicitacije¹² v slovenskem delu korpusa *FraSloK*.

(26) *Ce n'est plus du tout la même tâche ! Très dangereuse, elle est normalement confiée à un chalumiste-monteur, c'est-à-dire un ouvrier qualifié et expérimenté, capable d'assurer sa propre sécurité.* (LMD038-FRA)

(26a) *To sploh ni več ista naloga! Zelo je nevarna, navadno jo zaupajo varilcu-plezalcu, torej kvalificiranemu in izkušenemu delavcu, ki se je sposoben zavarovati.* (LMD038-SLV)

Če povzamemo, je analiza slovenskih skladenjskih ustreznih francoskih pridevniških polstavkov pokazala skladenjsko eksplicitacijo v 86 % prevodov iz publicističnega podkorpusa in v 82 % prevodov iz literarnega podkorpusa. Ob tem lahko zaključimo, da je bila pridevniška oblika francoskih pridevniških polstavčnih jeder ohranjena v dobrih 70 % prevodov iz obeh podkorpusov (LMD 71 %, LIT 73 %), v publicističnem podkorpusu najpogosteje v okviru povedkovnega določila in prilastka (redkeje pridevniškega polstavka in povedkovnega prilastka), v literarnem podkorpusu pa v okviru povedkovnega prilastka, ki mu sicer tesno sledijo tudi povedkovo določilo, prilastek in pridevniški polstavek. Verbalizacijo francoskega pridevniškega jedra iz polstavčne strukture smo ugotovili v približno 8 % prevodov iz obeh podkorpusov, adverbializacijo pa v 16 % (LMD) oziroma v 15 % (LIT) vseh prevodov iz korpusa *FraSloK*.

5.2 Eksplicitacija pomenskih vrednosti v prevodih francoskih pridevniških polstavkov

Glede na rezultate skladenjske analize je eksplicitacijo pomenskih razmerij med francoskimi pridevniškimi polstavki in glavnimi stavki mogoče opazovati znotraj kategorij 'Prilastek', 'Prislovno določilo', 'Podredje' in 'Priredje', v primerih iz drugih kategorij pa je razmerje tako kot v izvirniku implicitno ('Polstavek' in 'Druga stavčna razmerja') ali ni izraženo (izpusti prevodov pridevniških jeder iz francoskih polstavčnih struktur znotraj kategorije 'Drugo'). Semantična eksplicitacija je tako potrjena v 79 % primerov iz publicističnega podkorpusa in v 74 % primerov iz literarnega podkorpusa.

¹² Po teoriji relevance (Sperber in Wilson 1986) je taka eksplicitacija sicer nepotrebna, če jo naslovnik lahko sam izpelje iz konteksta.

Z vidika pomenskih razmerij je analiza prevodov francoskih pridevniških polstavkov, ki jo povzema graf 2 na strani 115, pokazala, da približno polovica francoskih pridevniških jeder iz obeh podkorpusev nosi opisno vrednost (LMD 54,8 %, LIT 47,9 %), ki je v slovenščini izražena s povedkovim določilom znotraj prilastkovega odvisnika ali vezalnega priredja, s prilastkom ali povedkovim prilastkom. Ta vrednost je najpogostejša, kadar med polstavkom in glavnim stavkom ni mogoče vzpostaviti nobene posebne semantične povezave (Combettes 2005: 34); pridevniki s takšno vrednostjo po navadi prinašajo dodatno informacijo o nanosniku glavnega stavka (npr. njegovo lastnost, značilnost) ali izražajo spremno okoliščino (npr. duševno stanje v (27)–(27a)) dejanja glavnega stavka, kar je razvidno iz konteksta:

(27) ***Curieux**, j'observais ses allées et venues de par les rues de Moscou.* (LIT010-FRA)

(27a) ***Začuden** sem opazoval njegovo kroženje po moskovskih ulicah.* (LIT010-SLV)

Prislovna vrednost je v publicističnem podkorpusu (24,6 %) za štiri odstotke višja od deleža neizraženih pomenskih razmerij (20,6 %), medtem ko v literarnem podkorpusu razlika med njima znaša le 0,3 % (prislovna vrednost: 26,2 %, razmerje ni izraženo: 25,9 %). Od prislovnih vrednosti v obeh podkorpuseh izstopa vrednost načina (LMD 29,0 % oziroma 9 primerov, LIT 39,1 % oziroma 27 primerov), v literarnem tudi vzročnost (29,0 % oziroma 20 primerov; LMD 16,1 % oziroma 5 primerov). Vrednost načina v slovenščini najpogosteje nosijo prislovi, redkeje predložne samostalniške zveze, s katerimi je izražena vsebina po navadi golih francoskih pridevniških jeder, ki pojasnjujejo spremno okoliščino osrednjega glagolskega dogodka (kako je nekdo nekaj naredil/počel), npr.:

(28) ***Muette**, j'admirais la patience de l'instituteur.* (LIT008-FRA)

(28a) ***Nemo** sem občudovala učiteljevo potrpežljivost.* (LIT008-SLV)

Na interpretacijo vzročno-posledične povezave med vsebino polstavka in glavnega stavka v večini primerov vplivata besedišče polstavka in glavnega stavka ter širši kontekst. V literarnem podkorpusu v polovici prevodov znotraj podredja vzročno vrednost nosi pridevnik *incapable* (*nesposoben*), ki je v slovenščini v vseh primerih verbaliziran:

(29) ***Incapable de se libérer des démons lovés dans son cœur**, il préfère se recroqueviller dans les ténèbres de la Cité interdite.* (LIT011-FRA)

(29a) ***Ker se ni mogel otresti zlih duhov v svojem srcu**, se je raje zvil v klobčič v mraku Prepovedanega mesta.* (LIT011-SLV)

Od drugih prislovnih vrednosti smo v publicističnem podkorpusu našli po štiri primere vrednosti namena, dopustnosti in protivnosti (12,9 % prislovnih vrednosti), v literarnem pa po šest primerov vrednosti časa in namena (8,7 %

prislovnih vrednosti) ter pet primerov protivnosti (7,2 %). Analiza je pokazala, da poleg konteksta določitev nekaterih omenjenih vrednosti, in sicer namena in časa, omogočajo naslednji dejavniki:

Vrednost namena: nosijo jo nekateri francoski pridevniki (npr. *désireux* – v želji (primer (30)–(30a)), *soucieux* – v prizadevanju ali da bi, *curieux* – iz radovednosti), ki pojasnjujejo, čemu oziroma s kakšnim namenom se je zgodilo ali se dogaja to, kar izraža osrednji glagolski dogodek; v prevodih iz korpusa *FraSloK* so ti najpogostejše izraženi s predložno samostalniško zvezo v vlogi prislovnega določila namena (npr. (30a)), redko z veznikom *da bi* znotraj namernega odvisnika:

(30) *Désireux de rassurer les industriels*, M. Prodi a évoqué une « méprise » et assuré que le TGV sera construit. (LMD281-FRA)

(30a) Prodi je v želji, da bi pomiril podjetnike, govoril o »napaki« in zagotovil, da bo hitra železnica zgrajena. (LMD281-SLV)

Vrednost časa: imajo jo nekateri francoski pridevniki (npr. *petit* – majhen, *jeune* – mlad, *consécutif* – tak, ki sledi, *vivante* – živa (gl. primer (31)–(31a))), ki jih je treba v slovenščini skladenjsko razviti; slovenski prevajalci so njihovo vsebino najpogostejše izrazili s predlogom ali s predložno samostalniško zvezo v funkciji prislovnega določila časa, redko s povedkovim določilom znotraj časovnega odvisnika (primer (31a)); v enem primeru je pridevniško polstavčno jedro v francoščini uvajal časovni veznik, ki je bil ohranjen v slovenskem prevodu (primer (32)–(32a)):

(31) *Vivante*, elle ne m'avait jamais totalement comprise. (LIT011-FRA)

(31a) Ko je bila še živa, me ni nikoli popolnoma razumela. (LIT011-SLV)

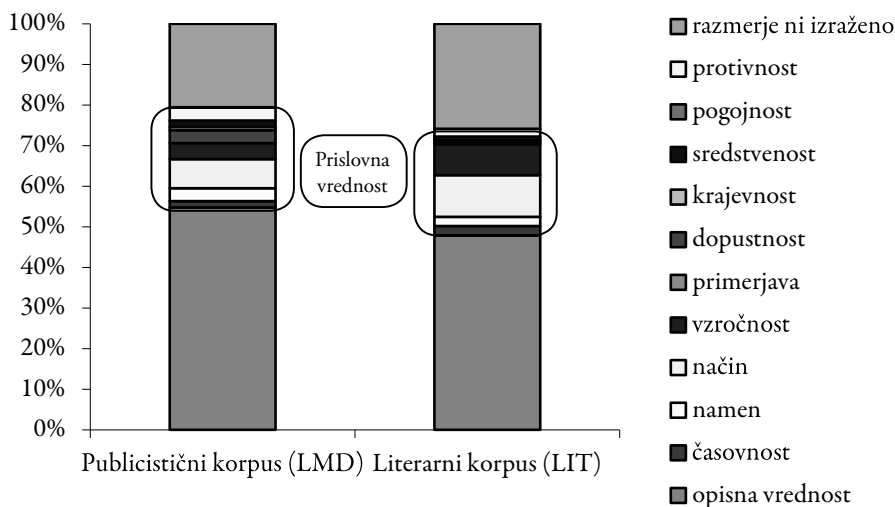
(32) *Lorsque, lassés d'être plongées dans l'opaque repos nocturne, les pupilles désirent enfin les nuances du jour, le soleil se lève, inlassablement, sur des couleurs volées à la douceur de l'art pour borner le monde.* (LIT008-FRA)

(32a) Ko se zenice naveličajo motnega nočnega počitka in si znova zaželejo odtenkov dneva, takrat vzide neutrudno sonce in razkrije barve, ukradene nič hudega sluteči umetnosti, da bi zamejile svet. (LIT008-SLV)

Druge vrednosti, navedene v legendi grafa 2, se v obeh podkorpusih pojavijo le v primeru ali dveh ali pa sploh ne. Razen pri primerih pogojne vrednosti, ki jo v francoščini določa povedek glavnega stavka v pogojniku (primer (33)), drugod zaradi (pre)majhnega vzorca ni bilo mogoče razbrati dejavnikov, ki bi poleg konteksta olajšali določitev posameznih pomenskih razmerij.

(33) *Vivante*, Mere aurait été terrifiée d'avoir accouché d'un empereur, d'un dieu. A présent, tous deux se taisaient de fierté. (LIT011-FRA)

(33a) Če bi mati še živela, bi jo obšla groza ob spoznanju, da je rodila cesarja in božanstvo. (LIT011-SLV)



Slika 2: Povzetek semantičnih vrednosti prevodov francoskih pridevniških polstavkov.

6 ZAKLJUČNE UGOTOVITVE

V članku smo predstavili strategije prevajanja francoskih pridevniških polstavkov, ki zaradi skladenjske reduciranosti in pomenske nedoločnosti na površinski ravni nematernim govorcem francoščine povzročajo težave pri njihovem razumevanju in/ali dekodiranju, predvsem kadar imajo v povedi začetni položaj oziroma se nahajajo pred odnosnico. Empirično gradivo smo pridobili iz francosko-slovenskega korpusa publicističnih in literarnih besedil z metodo računalniškega luščenja na osnovi ukazov, sestavljenih iz besednovrstnih oznak in regularnih izrazov. Ta metoda se je izkazala manj primerna za pridobivanje kompleksnih struktur, kot so polstavki, saj so bili rezultati zaradi nezmožnosti vključitve kriterijev polstavčnosti v ukaze slabi, ročno analiziranje in razvrščanje zadetkov pa precej zamudno. Rezultati so bili slabši tudi zaradi napačno pripisanih besednovrstnih oznak nekaterim pojavnicam, tako da so izboljšave potrebne tudi na področju avtomatskega označevanja naravnega jezika, saj je uspešnost slovnice raziskave, katere računalniško luščenje primerov temelji na besednovrstnih oznakah, odvisna od kakovostne označenosti korpusa. V prihodnosti bi bilo sicer zanimivo opraviti tudi ročno analizo pogostnosti rabe polstavkov v korpusnih datotekah in jo primerjati z rezultati avtomatskega luščenja, s čimer bi preverili uspešnost oziroma natančnost metode, uporabljene v tej raziskavi.

Število izluščenih francoskih pridevniških polstavkov kaže, da so ti nekoliko pogostejši v romanah kakor v časopisnih člankih, pogostnost njihove rabe v francoščini

pa ni odvisna le od tipa besedila, temveč tudi od pisca oziroma od njegovega sloga pisanja, saj so polstavki v manjšini francoskih besedil iz obeh podkorpusov *FraS-loK* rabljeni pogosto, v večjem delu pa redko ali sploh ne.

Analiza prevodov francoskih pridevniških polstavkov je potrdila pričakovano skladenjsko eksplcitacijo v 82 % (LIT) oziroma 86 % (LMD) vseh primerov, medtem ko je bila eksplcitacija pomenskih razmerij nekoliko nižja (LIT 74 %, LMD 79 %), predvsem zaradi števila ohranjenih polstavkov. Primerjava rezultatov luščenja in kontrastivne analize torej kaže, da je raba polstavkov pogostejša v izvirni pisni francoščini kakor v prevedeni pisni slovenščini, kjer izstopa razvezovanje francoskih polstavčnih jeder v stavke z osebno glagolsko obliko in dodajanje logičnih povezovalcev med stavčne dele.

Ohranitev polstavkov v slovenščini je glede na rezultate analize sicer višja v prevodih iz literarnega podkorpusa in je torej žanrsko pogojena, pogojena pa je tudi z njihovo pogostostjo v izvirniku: večje je njihovo število, večja je verjetnost, da bo kakšen ohranjen v prevodu. Poleg tega njihova raba lahko izhaja iz prevajalčevega sloga ali vpliva jezika izvirnika na jezik prevoda. Raba polstavkov v sodobni pisni slovenščini sicer ni natančno raziskana, je pa slogovno zaznamovana (gl. npr. Milojević Sheppard 1992: 235). Opisi v slovenskih slovnicih so skromni in pomanjkljivi (raba ločil za levimi polstavki npr. ni natančno določena) in po navadi ne vsebujejo podatka o normi, zato se mora prevajalec ali pisec glede njihove (ne)rabe zanašati na svoj občutek za jezik. V prihodnosti bi veljalo dopolniti opise polstavkov v slovnici in natančneje raziskati njihovo rabo v izvirni slovenščini, kar bi omogočilo zaključke, ali njihova ohranitev v prevedeni slovenščini, predvsem v literarnem diskurzu, kjer ni nezanemarljiva, odstopa od norme in torej izhaja iz vpliva jezika izvirnika, ali pa je samo slogovno pogojena.

Bibliografija

- Barlow, Michael, 2005: *ParaConc*. Houston: Athelstan.
- Becher, Viktor, 2010: Abandoning the notion of “translation-inherent” explicitation: against a dogma of translation studies. *Across Languages and Cultures* 11/1. 1–28.
- Biber, Douglas, Susan Conrad in Randi Reppen, 1998: *Corpus linguistics. Investigating language structure and use*. Cambridge: Cambridge University Press.
- Boch, Françoise, Agnès Tutin in Damien Laurent, 2009: Construction détachée et adjectifs d’affects. Apothélos, Denis, Bernard Combettes in Franck Neveu (ur.): *Les linguistiques du détachement. Actes du colloque international de Nancy (7.–9. junij 2006)*. Bern: Peter Lang. 99–115.
- Breznik, Anton, 1967: *Življenje besed* (priređil Jakob Šolar). Maribor: Založba Obzorja.

- Combettes, Bernard, 1988: Fonctionnement des nominalisations et des appositions dans le texte explicatif. *Pratiques* 58. 107–119.
- Combettes, Bernard, 1998: *Les constructions détachées en français*. Pariz: Ophrys.
- Combettes, Bernard, 2005: Les constructions détachées comme cadres de discours. *Langue française* 148 (Tematska številka *Les adverbiaux cadratifs*). 31–44.
- Dubois, Jean, in René Lagane, 1973: *La nouvelle grammaire du français*. Pariz: Larousse.
- Dubois, Jean, Mathée Giacomo, Louis Guespin, Christiane Marcellesi, Jean-Baptiste Marcellesi in Jean-Pierre Mével, 2007: *Grand dictionnaire : Linguistique & Sciences du langage*. Pariz: Larousse.
- Erjavec, Tomaž, Camelia Ignat, Bruno Pouliquen in Ralf Steinberger, 2005: Massive multi-lingual corpus compilation: Acquis Communautaire and totale. Vetulani, Zygmunt (ur.): *Proceedings of the 2nd Language and Technology Conference, Poznan, 21.–23. april*. Poznan: Wydawnictwo Poznanskie Sp. z o.o. 32–36.
- Havu, Eva, 2002: L'interprétation des constructions détachées. *Cercles de linguistique appliquée à la communication* 10, <http://www.ucm.es/info/circulo/no10/havu.htm>. (Dostop: 14. 5. 2012)
- Hunston, Susan, 2002: *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Jesenšek, Marko, 1998: *Deležniki in deležja na -č in -ši. Razširjenost oblik v slovenskem knjižnem jeziku 19. stoletja*. Maribor: Slavistično društvo Maribor.
- Lujić, Andra, 2006: *Slovenske prevodne ustreznice gerundija (gérondif) in sedanjega deležnika (participe présent) v frankofonskem afriškem romanu*. Diplomsko delo, Univerza v Ljubljani, Filozofska fakulteta, Oddelek za prevajalstvo.
- Mezeg, Adriana, 2011: *Korpusno podprta analiza francoskih polstavkov in njihovih prevedkov v slovenščini*. Doktorska disertacija, Univerza v Ljubljani, Filozofska fakulteta, Oddelek za romanske jezike in književnosti.
- Milojević Sheppard, Milena, 1992: Polstavki v angleščini in slovenščini. *Vestnik* 26/1–2. 224–237.
- Milojević Sheppard, Milena, 1993: *Morpho-syntactic expansions in translation from English into Slovenian as a prototypical response to the complexity of the original*. München: Verlag Otto Sagner.
- Neveu, Franck, 1998: *Études sur l'apposition. Aspects du détachement nominal et adjectival en français contemporain, dans un corpus de textes de J.-P. Sartre*. Pariz: Honoré Champion Éditeur.
- Noailly, Michèle, 1999: *L'adjectif en français (Collection L'essentiel français)*. Pariz: Ophrys.
- Riegel, Martin, Jean-Christophe Pellat in René Rioul, 1999: *Grammaire méthodique du français*. Pariz: Presses Universitaires de France.

- Rossi-Gensane, Nathalie, 2009: Register variation in the non-standard use of non-finite forms. Beeching, Kate, Nigel Armstrong in Françoise Gadet (ur.): *Sociolinguistic variation in contemporary French*. Amsterdam, Filadelfija: John Benjamins Publishing Company. 177–191.
- Schmid, Helmut, 1994: *TreeTagger*. Univerza v Stuttgartu, <http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger>. (Dostop: 14. 5. 2012)
- Sinclair, John, 1991: *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sperber, Dan, in Deirdre Wilson, 1986: *Relevance – Communication and cognition*. Oxford: Basil Blackwell.
- Toporišič, Jože, 2004: *Slovenska slovnica* (Četrta, prenovljena in razširjena izdaja). Maribor: Obzorja.
- Vinay, Jean-Paul, in Jean Darbelnet, 1958: *Stylistique comparée du français et de l'anglais*. Pariz: Les éditions Didier.
- Vinay, Jean-Paul, in Jean Darbelnet, 1995 [1958]: *Comparative stylistics of French and English. A methodology for translation*. Prevedla Juan C. Sager in M.-J. Hamel. Amsterdam, Filadelfija: John Benjamins Publishing Company.
- Vintar, Špela, 2009: Slovenski prevodoslovni korpus. Stabej, Marko (ur.): *Infrastruktura slovenščine in slovenistike (Simpozij Obdobja)*. Ljubljana: Znanstvena založba Filozofske fakultete. 385–391.
- Wagner, Robert-Léon, in Jacqueline Pinchon, 1962: *Grammaire du français classique et moderne*. Pariz: Hachette.

Primerjava slovenskih prevodnih variant povezovalca *mais* v francosko-slovenskem vzporednem korpusu

Mojca Schlamberger Brezar

Oddelek za prevajalstvo, Filozofska fakulteta Univerze v Ljubljani
mojca.brezar1@guest.arnes.si

Abstract

The paper analyses the translation possibilities of French connective *mais* in Slovene on the basis of the French-Slovene parallel corpus FraSloK, a part of the larger translation corpus SPOOK, available on the website <http://nl.ijs.si/spook/>, which includes literary texts and newspaper articles as well as their Slovene translations. Based both on qualitative and quantitative methodology, the paper gives an overall view of a sample of 2500 phrases containing a connective *mais* in different contexts and its Slovene translations. First, different values of the connective *mais* are pointed out that can give different results in Slovene translations. Then, the Slovene equivalents are counted and discussed. In the second part, translation choices from the FraSloK and their stylistic values are analysed and evaluated from the stylistic point of view. It becomes obvious that the choice of the connective in the translation depends more on the translator than on the value of the context: some of the translators have a clear predilection for a certain connective whereas others try to base their choice of the connective on the context and type of text - in narration other connectives would be used than in dialogues. The most natural translations are those that make no preference for one single connective. The influence of the norm, confining certain connectives to oral use, is visible in different translations of connectives in dialogue or narration.

Keywords: French connective *mais*, translation equivalents of connective in Slovene, parallel corpus analysis, translation equivalents in different contexts

Ključne besede: francoski povezovalec *mais*, prevodne ustreznice povezovalcev, analiza vzporednega korpusa, prevodne ustreznice v različnih kontekstih

1 TEORETIČNA IZHODIŠČA

V članku se ukvarjamo z raziskavo prevodnih ustreznih francoskega veznika in povezovalca *mais* v vzporednem prevodoslovnem korpusu SPOOK, natančneje v njegovem literarnem slovensko-francoskem delu, poimenovanem tudi FraSloK, ki je dostopen na spletnem naslovu <http://nl.ijs.si/spook/>. Povezovalac *mais* v francoščini nastopa v različnih kontekstih, ki zajemajo govorjeni jezik z različnimi izrekanjskimi položaji od pogovornih do zbornih, kakor tudi pisni jezik, predvsem v pripovedi. Na podlagi korpusnega materiala smo preverjali, kakšne prevodne variante se za ta povezovalac v različnih kontekstih uporabljajo v slovenščini.

Večjezični prevodoslovni korpus SPOOK je nastal v okviru projekta Špele Vintar J6-2009, *Slovensko prevodoslovje – viri in raziskave*, njegov francoski del FraSloK pa je izdelala Adriana Mezeg za potrebe svoje doktorske disertacije *Korpusno podprta analiza francoskih polstavkov in njihovih prevedkov v slovenščini* (Mezeg 2011). Vsebuje literarna dela in publicistična besedila v obsegu skoraj 2,5 milijona besed (2.466.985) (Mezeg 2010); publicistična besedila izhajajo iz 300 besedil in njihovih prevodov iz *Le Monde diplomatique* med leti 2006 in 2009 (1.164.074 pojavnic), literarna besedila pa predstavlja 12 modernih francoskih in frankofonih romanov, ki so v zadnjih 15 letih izšli v Franciji in v slovenskem prevodu (Mezeg prav tam). Količinsko literarni korpus predstavlja malo več kot polovico pojavnic (1.302.911). Natančnejši pregled naslovov zajetih literarnih del ter njihovih avtorjev in prevajalcev je podan ob analizi v nadaljevanju. Naša raziskava je potekala na podkorpusu literarnih besedil, in sicer v omejenem obsegu 2500 pojavnic, ki smo jih poiskali z vmesnikom na spletni strani korpusa SPOOK (glej zgoraj). Na podlagi analize pojavnic iz korpusa FraSloK smo pregledali slovenske veznike oziroma povezovalce,¹ za katere se prevajalci največkrat odločajo pri prevajanju francoskega veznika in povezovalca *mais*, ter skušali definirati kontekste, ki so v prevodu bolj naklonjeni rabi določenih veznikov.

Teoretični okvir za korpusne raziskave v opisnem prevodoslovju so bili dani kmalu po razcvetu korpusnega jezikoslovja. Korpusne raziskave v translatologiji so se razmahnile z Mono Baker v devetdesetih letih prejšnjega stoletja (cf. Baker 1993, 1995), ko je začela graditi prvi angleško-francoski vzporedni korpus besedil iz kanadskega parlamenta Canadian Hansard (www.tsrali.com) in so od takrat postale tako rekoč zakonitost v raziskovanju prevodov (Malmkjaer 2004: 91). V prevodoslovju veliko primerjalnih študij predvsem zadnjega desetletja izhaja iz predpostavke, da se jezik prevoda prav zaradi postopka prevajanja razlikuje od jezika primerljivega izvirnika (Pisanski Peterlin 2009: 105). Preverjali bomo, ali to velja tudi v primeru naših prevodov.

¹ O pojmihi veznik in povezovalac na tem nestu ne bi odpirali širše razprave; veznik nam pomeni enoto skladenjske, povezovalac pa diskurzivne analize. Za podrobnejšo razmejitev pojmov glej Schlamberger Brezar 2009.

Raziskave konektorjev-povezovalcev kot diskurzivnih dejavnikov za slovensko-francosko kombinacijo v slovenskem prostoru segajo v 90 leta dvajsetega stoletja in so bile večinoma opravljene na transkripcijah govorjenih besedil s poudarkom na njihovem pomenu (Schlamberger Brezar 1998, 2000, 2009). V prevodoslovne namene za kombinacijo teh dveh jezikov širša raziskava še ni bila izvedena z izjemo kratkega pregleda slovenskih ustreznih pojavnic v pretolmačenih besedilih (Schlamberger Brezar 2011), se pa v okviru računalniških raziskav pripravljajo poskusi avtomatičnega označevanja pomena in razdvoumljanja prevodov konektorjev (npr. Balažic Bulc, Gorjanc 2009).

Po opredelitvah vrednosti povezovalca *mais* v francoščini in opisu njegovih prevodnih variant v slovenščini bomo v luči kontrastivne metode predstavili še kvantitativni in kvalitativni pristop k analizi izbranih povezovalcev, ki jih v našem prostoru občasno nazivajo tudi konektorji.² Po eni strani nas zanima kvantiteta povezovalcev v posameznem delu in številčnost posameznih prevodnih variant, po drugi strani pa je pomembna kvalitativna raziskava vrednosti povezovalcev glede na kontekst in ustreznost prevoda in seveda dejstvo, ali v prevodu zasledimo kake posebnosti, po katerih se razlikuje od izvirnih slovenskih besedil.

2 VREDNOSTI POVEZOVALCA IN VEZNIKA *MAIS* V FRANCOŠČINI IN NJEGOVE MOŽNE PREVODNE VARIANTE V SLOVENŠČINI

2.1 *Mais* v različnih besedilnih položajih

Mais v francoščini je povezovalec, ki uteleša protivno-dopustno logično razmerje tako v govoru kot v pisnem izražanju (prim. Schlamberger Brezar 2009). Oblikoslovno je protivni priredni veznik, katerega pomen bi lahko pojasnili kot »in + nasprotno od pričakovanj«.

Slovarji (Petit Robert in TLF (www.atilf.atilf.fr)) najprej ločijo med prislovno rabo *mais* iz lat. *magis* (v pomenu *več*) kot prislova, in *mais*, prirednega veznika oziroma povezovalca. Poleg *mais* v francoščini nastopa še več povezovalcev vezniškega ali prislovnega izvora, kot na primer *or*, *pourtant*, *cependant*, *néanmoins*, *toutefois*, ipd., ki jih v tem prispevku ne obravnavamo. Vrednosti *mais* se do določene mere prekrivajo z njimi, ima pa tudi posamezne posebne vloge, ki pridejo do izraza predvsem v okoliščinah izražanja v govorjenem jeziku. Skupno slovarja navajata vsaj devet različnih vlog. Že sam pregled slovarskega članka nam torej pokaže, da delovanje *mais* še zdaleč ni enotno. Ducrot (Ducrot et al., 1980: 93–131) poskuša

² Pojem povezovalci ali konektorji zaznamuje vsa sredstva, ki v besedilih zagotavljajo vezavo na ravni izrekov, v slovenščini v tej funkciji lahko nastopajo besede, ki pripadajo različnim besednim vrstam, poleg veznikov še prislovi in členki.

strniti njegove vrednosti z upoštevanjem izvenjezikovnih³ dejavnikov, kot so kontekst in nameni govorcev, kakor tudi njihove sodbe o situaciji in njihovi medsebojni odnosi. V tem kontekstu *mais* kot protivni veznik uvaja nasprotovanje, ki je lahko treh vrst glede na to, ali se nanaša na govorno dejanje govorca *G*, ko izreka propozicijo *P*, sklepe, ki jih *G* potegne iz *P* (kljub temu da *X* dopušča resničnost *P*) ali samo resničnost *P*.

Po tej poti je raziskoval tudi J. M. Adam (1990) in podal opise, po katerih v diskurzu povezovalc *mais* prevzema mnoge funkcije, od katerih jih je najbolj izpostavljenih pet: argumentacijski *mais*, dopustni *mais*, *mais* spodbijanja, stopenjevalni *mais* in fatični *mais*.

Argumentacijski *mais* vsebuje dva eksplicitna argumenta in dva na splošno implicitna in vedno nasprotujoča si sklepa (Adam 1990: 206, Schlamberger Brezar 2009: 203–207). Izraz *P mais Q* predpostavlja, da lahko propozicija *P* služi za argument sklepu *r* in da je propozicija *Q* argument, ki ta sklep *r* izniči. Argumentacijsko gibanje, ki ga implicira izrek *P mais Q*, bi lahko prikazali takole: *P* je resničen in na podlagi *P* bi lahko sklepali *r*, pa ne bomo, ker velja *Q*. *Q* je močnejši argument za sklep *ne-r* kot *P* za sklep *r*. Princip lahko razložimo na izreku *Cet étudiant est intelligent, mais paresseux* (*Ta študent je inteligenčen, ampak len*), kjer na osnovi prvega dela lahko izpeljemo sklep »naredil bo izpit«, iz drugega dela pa lahko sklepamo le nasprotno. Gre za argumentacijski *mais*, ki temelji tudi na predpostavkah govorcev o stanju stvari in njihovem odnosu do konteksta in širše predmetnosti, kar Ducrot in Anscombe (1983) poimenujeta topos. V tem primeru lastnost *P* (*biti inteligenčen*) ne more implicirati *ne-Q* (*ne biti len*): ta implikacija ne temelji na splošnih normah, pač pa na govorcevih predstavah in presupozicijah, ki jih ima o svetu, ki ga obdaja. Argumentacijski *mais* je odvisen od diskurzivnega prostora govorca, kjer je *P* argument za sklep *C* in *Q* argument za sklep *ne-C*, in kjer *mais* kaže, da je sklep argumentacijsko usmerjen k *ne-C*.

Dopustni *mais* se pojavlja vedno v zvezi z nikalnico, ki je lahko tudi implicitna, npr. *Cet étudiant est intelligent, mais il échoue à tous les examens* (*Ta študent je inteligenčen, ampak ne naredi nobenega izpita*). Dopustni *mais* torej vodi k direktnemu sklepu. *Mais* spodbijanja se tvori z zanikano prvo propozicijo, ki se kombinira z *non pas* (*plus, point*), po shemi *non pas P, mais Q* (Adam, 1990: 194), kot na primer v *Il n'était pas malade, mais paresseux* (*Ni bil bolan, ampak len*). Tu je navezava na nasprotovanje prvemu delu povedi direktna in do nje ni treba dostopati prek implicitnih ali eksplicitnih sklepov. Razlike med temi tremi *mais* so malenkostne in vključujejo večinoma iste slovenske prevodne ustreznice, najbližje je vsekakor *ampak*, več podatkov o vrednostih bomo podali v nadaljevanju.

³ Ducrot uporablja termin "extra-linguistiques" v narekovajih; gre namreč za eno prvih študij, ki preučuje povezovalce s pragmatičnega vidika.

Nekoliko drugačen je stopnjevalni *mais* – *non seulement mais aussi* – morda najprepoznavnejši, ki kot tak obstaja tudi v slovenščini v različicah *ne le, ne samo, ampak, marveč, temveč tudi*.

Še bolj tipičen, predvsem za govorjeni diskurz, pa je t. i. fatični *mais* (Adam, 1990, Schlamberger Brezar 2009: 205–207). Tega lahko razložimo kot *mais* govorjenega diskurza, ki še ohranja svojo argumentacijsko vrednost, vendar nastopa lahko tudi v povsem okrnjenih kontekstih in se kombinira z drugimi zaznamovalci diskurza. Ta kategorija nam omogoča »enotno obravnavo« vseh *mais*, ki odstopajo od običajne, predpisane, zborne, normativne rabe. Vrednosti fatičnega *mais* se razlikujejo, če uvaja pripovedne, vprašalne ali vzklične izreke, lahko pa ima tudi metajezikovno vrednost (Schlamberger Brezar 2009: 206). Fatični *mais* je kot zaznamovalec besedilnih segmentov najznačilnejši za govorjeni jezik – najobičajneje je postavljen na začetek izreka oziroma odstavka, ki po M. A. Morel tvori enoto analize govorjenega jezika (Morel in Danon-Boileau 1998, Schlamberger Brezar 2009). V korpusu bi ga pričakovali v dialogih. Od vseh vrednosti *mais* edino ta vstopa v kombinacijo z drugimi diskurzivnimi zaznamovalci, kot na primer *mais oui, mais bien sur*, kjer ima tako rekoč nasprotno vrednost od pričakovane (Schlamberger Brezar 2005), ali *mais quand même* in podobno. Po funkciji je najbolj podoben slovenskemu *ampak*, pri primerih, ko gre v argumentaciji v nasprotno smer od običajnega *mais*, pa med prevodne možnosti vključujemo tudi *seveda, jasno* in druge prislove oziroma členke zatrjevanja.

Vidimo torej, da je *mais* povezovalac z različnimi slogovnimi vrednostmi, ki pa se jih da opredeliti tudi glede na tip besedila. Če vzamemo Adamovo tipološko besedilno razdelitev (Adam 1994), bo povezovalac *mais* vključeval drugačna navodila za interpretacijo, če bo nastopal v argumentacijskem ali pripovednem besedilnem tipu oziroma v dialogu. Slogovnim vrednostim slovenskih povezovalcev, ki se lahko pojavljajo v podobnih kontekstih in ki bodo posledično najbrž vplivale na prevajalske odločitve, se bomo posvetili v nadaljevanju.

2.2 Raznolikost slovenskih protivnih veznikov in njihova slogovna vrednost

Kot je bilo rečeno, *mais* pretežno interpretiramo kot veznik oziroma povezovalac, ki izraža dopušcanje ali nasprotovanje. V slovenščini ga bomo torej iskali predvsem v okviru protivnih razmerij.

Po Toporišiču (1976: 523–23, 2000: 650–652) protivno priredje lahko zaznamuje kontrast ali razliko, nasprotje, zamenjavo ali izvzemanje. V *Slovenski slovnici* (Toporišič, 2000, 1976) so navedeni naslednji protivni vezniki: *toda, ampak, a,*

vendar, drugače, namreč, samo, le. Pojem povezovalca oziroma konektorja v Toporišičevi slovnici namreč še ni opredeljen. Ne govori o specializaciji veznikov za katero koli od funkcij, podaja samo primere, tako na primer *pa, a, toda, ali, vendar, ampak* »izražajo nekako zanikano nasprotje prvemu delu vezniške zveze« (Toporišič, 1976: 378, 2000: 439–440), *vendar, ali, ampak* čustveno poudarjajo nasprotje, *a, toda* sta njihovi predvsem zborni varianti – *a* je bližje *vendar*, *toda* je bližje *pa*. *Samo, le* vključujeta izvzemanje. Če sklenemo, same stilistične oznake prevajalcu razen informacij zborni/pogovorno ne dajo veliko priporočil za prevajanje, sodbe o tem, kaj je čemu bližje, pa se v *Slovenski slovnici* žal ne opirajo na kakšne s primeri podprte razlage.

Claude Vincenot v svoji slovenski slovnici *Essai de grammaire slovene* (1975: 297), namenjeni francoskim govorcem, obravnava »jonctifs adversatifs simples« (enostavne protivne veznike) in *pa/toda/ampak/marveč/temveč* = *mais, alali (littéraire)* = *mais*, *vendar* = *pourtant, cependant*, *samolle* = *seulement*, »jonctifs adversatifs doubles« (dvojne protivne veznike – v smislu dopuščanja): *sicer, vendar* – *certes, pourtant* ali pa te, ki zaznamujejo »renchérissement positif« (stopnjevanje): *ne samolle ... ampak, marveč, temveč tudi* (*non seulement, mais encore*). Glede na Vincenotove opise so kot prevodne ustreznice *mais* najprimernejši *pa, toda, ampak, marveč, temveč*, manj ustrezna pa *vendar* in *samo*. Tudi te sodbe niso z ničemer utemeljene.

V *Francosko-slovenskem slovarju* (Grad 1984) so dani naslednji pomeni, kar za prevajanje ni nepomembno: veznik *toda, ampak; temveč*; starinsko *več*; *non seulement mais encore* = *ne le, temveč tudi*; *mais non!* = *toda ne! vendar ne!*, *mais oui* = *seveda, gotovo*; *mais si* = *pač*; *je n'en puis mais* = *ne morem za to*. Slovarska opredelitev do določene mere prevajalca usmerja v večpomenskost in upoštevanje konteksta. Torej že površen pregled tematike obeta raznolike prevodne možnosti. Vidimo, da Gradov slovar in Vincenotova slovnica ponujata določeno kategorizacijo protivnih veznikov, kjer pa so kriteriji bolj implicitni kot eksplicitni, pa vendar v očeh prevajalca oprijemljivejši kot v *Slovenski slovnici*.

Tudi po *Slovarju slovenskega knjižnega jezika* (v nadaljevanju *SSKJ* (Bajec et al. 1994)) so protivni vezniki medsebojno zamenljivi, vendar ima vsak od njih sebi lastno slogovno vrednost. Ta je do neke mere označena v *SSKJ*, še bolj pa zakoreninjena v stilistiki prenosnika, ki je bodisi pisni, bodisi govorni, ter posameznika in njegovega odnosa do jezika. Vsi – *a, ampak, toda, vendar* – imajo med opredelitvami na prvem mestu po oznaki besedne vrste veznik razlago »nasprotje s prej povedanim«, v nadaljevanju pa smo se osredotočili še na dodatne slogovne oznake, ki jih srečamo v *SSKJ* in bi za prevajalce utegnile biti vodilo za slogovno oziroma pomensko razlikovanje.

Veznik in povezovalc *a* v *SSKJ* ima že pri prvi vrednosti oznako *knjižno*, slovar opozarja tudi na možno rabo v vezalnem priredju.

Veznik in povezovalc *ampak* je v *SSKJ* posebej izpostavljen v medmetni rabi, kjer povsem ustreza francoskemu faticnemu *mais*, kot veznik pa se uporablja: 1) v protivnem priredju, 2) ekspresivno v *ne samo ... ampak tudi*, torej v stopnjevalnem priredju, 3) ekspresivno za izražanje nasprotja in 4) z oznako *pogovorno ekspresivno* na začetku stavka. Tudi ta zadnja raba ga uvršča med najboljše možne različice faticnega *mais*. Zanimivo je, da je med slovenisti in lektorji okvalificiran kot pogovorni (prim. M. Križaj Ortar, Slovenski knjižni jezik 2 1998/1999, <http://kozjestezece.tripod.com>).

Veznik in povezovalc *pa* je opisan kot prislov ali veznik, med drugim tudi: 1) veznik v protivnem priredju, 2) za izražanje rahlega nasprotja, 3) vzročno-posledično, vzročno-sklepalno razmerje.

Veznik in povezovalc *toda* se navaja z naslednjimi atributi: v protivnem priredju, izražanje nasprotja s prej povedanim, pa tudi izražanje nepričakovanega (»mlad, *toda* že izkušen«); na začetku (od)stavka za prehod k drugi misli; ekspresivno »*toda* čast komur čast, govoriti pa zna« – gre za dopustno gibanje. Ta druga vrednost je konektorska oziroma povezovalna.

Veznik in povezovalc *vendar* po *SSKJ* nastopa v protivnem priredju: 1) izražanje nasprotja s prej povedanim, 2) oslavljen pomen, izraža nejevoljnost, nestrpnost, 3) oslavljen pomen, poudarja samoumevnost – v vezniški rabi v protivnem priredju in za omejevanje prej povedanega.

Tudi členka *samo* in *le* sta opisana kot kazalca protivnega priredja. Po *SSKJ* sta *le* in *samo* tako rekoč enakovredna, spremljajo ju naslednje navedbe: *le* – prisl. izraža omejenost na navedeno in *le* (vezn.) kot posebno geslo – metaoznaka v protivnem priredju, za omejevanje »Govori lepo, le malo pretiho.« *Samo* – prislov – izraža omejenost na navedeno. *Samo* v vezniški rabi za izražanje nasprotja s prej povedanim »Govori lepo, samo malo pretiho. Vlaki prihajajo, samo z zamudo.« Pa še primer – ekspr. nav. v zvezi *ne samo, ampak tudi*. Opis v *SSKJ* ne vsebuje oznake, da bi bil *le* zbornejši od *samo*, *vendar* je taka tendenca prisotna v govoru in pisavi, posledično pa tudi v prevodih, kot bomo videli v nadaljevanju.

Poseben primer sta *marveč* in *temveč*, ki z besedico *več* v korenu nakazujeta na posebno jakost dogajanja. V *SSKJ* sta oba pospremljena z oznako »v protivnem priredju«, pri tem da je kot sinonim *marveč* podan *ampak*, pri vezniku *temveč* pa je poudarjeno, da gre za utemeljevanje nove trditve namesto prej zanikane, kar predstavlja poseben pogoj glede na druge veznike.

Glede na pomensko analizo (Schlamberger Brezar 2009; Žagar, Schlamberger Brezar 2009) je *ampak* v slovenščini glede na funkcijsko raznolikost najpopolnejši približek francoskega *mais* in edini, ki ga lahko nadomesti v praktično vseh njegovih položajih. Skupaj s *pa* sta glede na *SSKJ* najbolj stilno zaznamovana, predvsem kot rabljena v govorjenem diskurzu. *Ampak* je kot zaznamovalec pripovednega registra po »ustnem izročilu« tudi manj zaželen med lektorji (glej tudi M. Križaj Ortar), ki zato najraje predlagajo nadomeščanje s katerim drugim, ki naj bi imel enako vrednost. V korpusu, ki nam je služil za analizo, ni podatka, v kolikšni meri so uporabljeni vezniki oziroma povezovalci rezultat izbora lektorja oziroma prevajalca. Kakšen je status veznika *ampak* v prevodih, bomo pa videli v nadaljevanju v empirični raziskavi.

Glede na opredelitev *SSKJ* se v glavnih vrednostih slovenski protivni vezniki med seboj le malo razlikujejo. Raba enega ali drugega veznika za izražanje protivnih razmerij je torej stvar sloga in izbire pisca oziroma prevajalca. Za govorjeni diskurz so glede na slogovne oznake primernejši *ampak*, *pa*, *samo*, *le*, medtem ko so *marveč*, *temveč*, *toda*, *a* bolj knjižni, *vendar* pa je do neke mere zaznamovan tudi s svojimi členkovnimi vrednostmi samoumevnosti, nestrpnosti. Kako se s posameznimi vrednostmi veznika *mais* v francoščini in njegovimi prevedki v slovenščino spopadajo slovenski prevajalci, bomo pogledali v nadaljevanju.

3 REZULTATI ANALIZE

Po korpusu FraSloK (gl. A. Mezeg 2010, 2011) na povezavi <http://nl.ijs.si/spook/konkor/> smo z ukazom za enostavno iskanje dobili 2500 zadetkov po vzporednem francosko-slovenskem podkorpusu oziroma 1250 pojavnic *mais* v francoščini in 1250 različic prevoda v slovenščini. Kvantitativna analiza je potekala avtomatsko prek vmesnika, dostopnega na spletni strani, kvalitativna pa ročno, ob upoštevanju konteksta pojavljanja posameznih veznikov in njihovih prevodnih ustreznih. V naboru smo zajeli 7 romanov, rezultati glede na vsakega od njih pa so podrobneje predstavljeni v nadaljevanju.

Ker se že na prvi pogled opazi, da se za prevod enakovrednih vrednosti *mais* v primerljivih kontekstih pojavljajo različni slovenski prevedki, ki so bili bolj odvisni od prevajalčeve izbire kot od (ne)enakovrednosti konteksta, smo lahko postavili hipotezo, da se prevajalci pri izbiri veznikov, prevodnih ustreznih *mais*, ravna predvsem po svojem lastnem slogu pisanja oziroma prevajanja, ki pa je odvisen tudi od besedilnega tipa, ki na tem mestu nastopa v romanu. Pri opredelitvah besedilnega tipa se sklicujemo na Adama (1994) in v danih besedilih lahko ločujemo predvsem med naracijo, opisom in dialogom. Za vse prevajalce je bilo namreč značilno, da so se v dialogih kot posnetkih

govorjenega diskurza večkrat odločali za uporabo ustreznih *ampak* ali *pa*, medtem ko so v delih, kjer gre za pripoved oziroma opis, uporabljali glede na SSKJ »zbornejše« povezovalce.

V delu *Fou de Vincent, Nor na Vincenta*, katerega avtor je Hervé Guibert, prevod pa je prispeval Brane Mozetič, smo s korpusnim pregledom na 15 straneh analize dobili 49 primerov *mais* z vzporednimi prevodi. Prevladoval je veznik-povezovallec *toda*, ki je bil uporabljen 25-krat. 10-krat je bil *mais* preveden kot *a*, petkrat kot *pa*, trikrat z *ampak*, po dvakrat pa s *samo* ali *temveč* oziroma je bil izpuščen.

Glede na naravo besedila, gre za prvoosebno pripoved, v kateri prevladuje naračija, ki jo sem in tja prekinjajo dialogi, so se nam za osvetlitev načinov prevajanja zdeli zanimivi naslednji primeri. Primer 1⁴ spodaj je tipičen primer prevajanja *mais* v govoru, in sicer v vprašalnih povedih:

1)

Hector me prend à part et me demande : « **Mais** c' est qui ? »

Hector me povleče na stran in me vpraša: »**Pa** kdo je to?«

Gre torej za tipičen primer dialoga, kjer predvidevamo, da bo prevajalec posegal po povezovalcih, ki so značilnejši za govorni jezik. V delu *Nor na Vincenta* je večina *mais*, ki nastopajo v dialogih, prevedena na ta način (gl. tudi primer 2 spodaj),

2)

c'est du brown sugar, c'est de moins bonne qualité que la blanche, **mais** c'est moins cher.

To je zate, brown sugar, ni tako dober kakor beli, je **pa** cenejši.

Primeri s *toda* in *a* so uporabljeni, kadar gre za pripoved (primera 3 in 4 spodaj):

3)

Il faudrait pouvoir témoigner de ce moment, car il fut immense, **mais** il faudrait que son récit soit aussi cristallin qu' un poème de Cavafy...

Moral bi imeti moč, da bi pričal o tem trenutku, kajti bil je neizmeren, **toda** njegov opis bi moral biti tako kristalno čist kot Kavafisova pesem ...

4)

A cette époque j'avais peu d'argent, **mais** toujours un flacon de parfum précieux.

V tistem času sem imel malo denarja, **a** vedno stekleničko dragocenega parfuma.

Veznik in povezovallec *toda* se pojavlja tudi na začetku izreka, kar je v skladu z opredelitvijo v SSKJ:

⁴ Vsi primeri so vzeti iz korpusa Spook, podkorpus FraSloK. Zadržali smo izvorno obliko, čeprav zapis ločil ni v skladu s pravili slovenskega pravopisa.

5)

Mais mon humeur guillerette me fait croire que je peux aller au-devant d'une ...**Toda** v svoji živahnosti verjamem, da grem lahko naproti pričakovanemu razočaranju ...

Prevodi kombinacij fatičnega *mais* v pripovednih oziroma vprašalnih povedih niso enakovredno obravnavani, kar je razvidno iz primerov 6 in 7 spodaj, kjer prvič nastopa *toda*, ki v tem primeru ni najbolj ustrezen, drugič pa *pa*:

6)

Si, **mais** de temps en temps.Kako da ne, **toda** kdaj pa kdaj.

7)

Mais c'est qui ?**Pa** kdo je to?

Prevajalec Brane Mozetič se sicer za izbiro povezovalca večinoma odloča na osnovi besedilnega tipa v skladu s slovarsko oznako o literarnem/zbornem in nevtralnem/pogovornem. Za literarno/zborna varianto prevoda *mais* prevladuje *toda*, sicer je nabor veznikov precej razgiban. Raznoliko v skladu z registrom izkorišča možnosti, ki jih daje slovenščina, pa tudi upošteva normo in običajni kontekst pojavljanja.

V romanu Philippa Grimberta *Un secret, Skrivnost*, ki ga je prevedla Jožica Grum, je shema uporabljenih veznikov oziroma povezovalcev za francoski *mais* naslednja: od 71 pojavitev *mais* v izvirniku jih je 47 prevedenih s *toda*, 9 z *ampak*, 6 s *pa*, 4 z *vendar*, po enkrat nastopijo *pa*, *temveč*, *a*, ali izpusti, en *mais* ostane brez poravnave, tako da ne dobi prevodne ustreznice.

Del romana predstavlja prvoosebna pripoved. Prvoosebni *mais* spominja na govorjeni kontekst in je preveden tudi z *ampak*, kot se to zgodi v primeru 8.

8)

J'en étais le fruit, **mais** avec une jouissance morbide je me plantais devant le miroir pour inventorier.Bil sem njun plod, **ampak** pred ogledalo sem se postavljal z boleznim veseljem, da bi naredil spisek svojih pomanjkljivosti.

Vendar *mais* v prvoosebni pripovedi ni vedno preveden z *ampak*, pač pa tudi s *toda*, *pa*, *vendar* (primeri 9, 10 in 11), ne da bi bile za to dane razlike v strukturi izvirne povedi ali v odnosu do konteksta:

9)

Je tentais de le faire fléchir **mais** je sortais rarement vainqueur de nos empoignades.Poskušal sem upogniti njegovo moč, **toda** redkokdaj sem bil zmagovalac v najinih rokoborbah.

10)

La tache jaune les désignait au regard des autres *mais* leur permettait aussi de se reconnaître, soudant une communauté qui, à...

Rumena lisa jih je zaznamovala v pogledih drugih, hkrati *pa* nenadoma tudi njim samim omogočila, da so se prepoznali med seboj,

11)

Maxime aurait aimé se dispenser de ces manifestations traditionnelles, *mais* il s' y prête de bonne grâce.

Maxime bi se sicer raje izognil takšnim običajem, *vendar* se jim rade volje prepusti.

Tudi fatični *mais* na začetku vprašalne povedi je preveden ne z *ampak*, pač pa s *toda*, kar je morda malo preveč »privzdignjeno«, papirnato, ko se nam ponuja možnost prevoda s *pa* »kaj pa, če«:

12)

Mais que se passera -t-il si Hannah et Simon reviennent de leur exil ?

Toda kaj, če se bosta Hannah in Simon vrnila iz pregnanstva?

V neki drugi poziciji pa se ta *ampak* pojavi v prvoosebnem govoru – tudi tu se *pa* ponuja kot naravnejši.

13)

Mais avait-il des soucis ?

Apak, ali je sploh imel skrbi?

V omenjenem prevodu med slovenskimi ustreznici *mais* prevladuje *toda* kot zborni veznik, raba veznika in povezovalca *ampak* ni vezana na dialoge, saj z izjemo prvoosebnih pasusov v začetku večinoma pripoveduje tretjeosebni vsevedni pripovedovalec. Za izjemo bi lahko označili primere, kot je 12, kjer gre za polpremi govor, kar bi lahko približali oznaki pogovorno.

V prevodu romana *Je m'en vais, Grem*, Jeana Echnoza, ki je delo Suzane Koncut, je od 130 prevodov *mais* 59-krat uporabljen veznik *a*, 36-krat *ampak*, 13-krat *pa*, devetkrat *vendar* (od tega trikrat *vendar pa*), dvakrat *toda*, ki jim sledijo prevodi predvsem v dialogih prisotnih kombinacij povezovalca *mais* fatične narave z drugimi zaznamovalci diskurza: *mais aussi* – *pa tudi* se pojavi enkrat, *mais quand même* dvakrat kot *a vseeno*, po enkrat tudi *mais sans* v prevodu *ne da bi*, *mais nom de dieu* kot *zaboga no*, *mais pourtant* – *a vendarle*, *mais enfin* kot *kaj pravzaprav*, *mais oui* je *seveda*, 2 *mais* v dialogih sta podana kot *no*. Enkrat srečamo v naboru tudi *maïs*, *koruza*, kar je napaka, ker vmesnik ne loči besed glede na naglasna znamenja. Ob seštevku vseh variant v slovenskem prevodu prevladuje veznik in povezovalac *a*, vendar bi strategijo prevajalke lahko poimenovali tudi kot mnogo-vrstnost. O tem pričajo naslednji primeri, kjer *mais* nastopa kot *ampak* predvsem v dialogih. Primer 14 podaja začetni stavek v knjigi:

14)

Je te laisse tout *mais* je pars.Vse lahko obdržiš, *ampak* jaz odhajam.

Pripoved v knjigi je tretjeosebna, od časa do časa jo prekinjajo dialogi. Najbolj tipičen prevod za *mais* v pripovedi je *a* (primer 15):

15)

Ce soir il n'y pensait même pas, distrait *mais* moins préoccupé que prévu par la scène qui venait de se jouer avec Suzanne.Nocoj na to ni niti pomislil, saj mu je misli odvrčal, *a* ga *vendar*le manj vznemirjal, kot bi lahko pričakoval, prizor, ki se je pravkar odigral s Suzanne.

Povezovalc *a* se pojavlja tudi na začetku povedi oziroma izreka (kot povezovalc, prim. 16):

16)

Mais une fois remonté sur la passerelle, il ne percevrait plus qu'un léger...*A* ko se je spet vzpel na najvišjo palubo, je zaznaval samo še rahlo, neprekinjeno pokanje,

Pri tipičnem *mais* spodbijanja se izmenjujejo prevodi z *ampak* ali s *toda* (primera 17, kjer glavni junak našteva, kaj ima v svoji galeriji, in 18)

17)

Pas de sculpteurs vu ses antécédents *mais* des peintres, bien sûr, comme Beucier, Spontini, Gourdel et surtout...Glede na svojo preteklost seveda ne kiparjev, *ampak* slikarjev, na primer Beucierja, Spontinija, Gourdela in predvsem ...

18)

Etiques, atones, discrets *mais* tenaces, ils ont, ils savent qu'ils n'ont qu'un petit rôle dans la vie.Jetične, nemočne, skromne, *toda* žilave se zavedajo, da jim v življenju pripada le neznatna vloga.

V naraciji nastopa kot možna ustreznica *mais* tudi *vendar*, ki mu njegova členkovna narava s pomenom vztrajanja oziroma občutka, da je skrajni čas, da se bo nekaj zgodilo, dajeta večjo težo (primer 19):

19)

Ferrer se lève, le monde le regarde, Ferrer est bien embarrassé *mais* il s'avance.Ferrer vstane, ljudje ga opazujejo, Ferrer je v precejšnji zadregi, *vendar* stopi naprej.

Naslednji primeri so predvsem značilni za prevajanje *mais* v dialogih. Fatični *mais* dobi, kadar nastopa v povezavi z drugimi vezniki, kot na primer *mais enfin*, *mais*

quand même ipd., različne prevodne variante v slovenščini, seveda vedno kontekstu ustrezne (primera 20 in 21):

20)

Je sais, je vois bien, dit Ferrer. **Mais** enfin ce n'est pas très bon, franchement, c'est loin d'être le meilleur.

Ampak pravzaprav ni tako zelo dober, še zdaleč ni najboljši iz serije.

21)

Mais c'est quand même un peu coûteux pour ce que c'est.

A vseeno je nekam drago za to, kar je.

V nadaljevanju podajamo še primera, kjer *mais* zaradi kombinacije z *oui* in *si* dobi nasprotno vrednost: *mais oui* v primeru 22 in *mais si* v primeru 23:

22)

Il avait dit **mais** oui, laissez donc votre sac.

Rekel je, **seveda**, pustite torbico tukaj.

23)

Allez, on va faire une exposition. Non ? fit l'autre. **Mais si**, dit Ferrer, bien sûr, bien sûr.

Ja, ja, je rekel Ferrer, seveda, seveda.

Kadar se pojavita dva *mais* eden za drugim, sta prevedena različno v skladu z zapovedmi slovenske stilistike, ki je proti ponavljanju – v primer 24 spodaj prvič kot *a*, drugič *pa*:

24)

Mais d'abord on lui fit visiter la maison : bien isolée, téléphone et cuisine moderne, mobilier de bois blanc bon marché de genre nordique **mais** qu'on trouve jusqu'en banlieue parisienne.

A najprej so mu razkazali hišo: dobro izolirana, telefon in televizija, velika peč in moderna kuhinja, ceneno pohištvo iz belega lesa v nordijskem slogu, ki **pa** ga je mogoče najti vse do pariških predmestij.

Prisoten je tudi en izpust veznika *mais* oziroma gre za zamenjavo z medmetom (primer 25):

25)

Mais nom de Dieu, se demandait-il cependant, *mais* enfin qu'est-ce qui me prend.

Zaboga, no, se je medtem spraševal, kaj me je pravzaprav nocoj prijelo, da takole preklinjam.

Suzana Koncut predstavlja živ prevod z raznolikim naborom veznikov, vedno ustrezen glede na kontekst. Tvegamo lahko stilistično oznako: glede na Mozetiča se kaže, da prevajalka raje uporablja *a* kot *toda*. Torej gre za osebni izbor, ki deloma spominja na avtomatsko prevajanje: *a* za pripovedni del, *ampak* in *pa* za dialoge.

V delu *Balzac et la petite tailleur chinoise*, *Balzac in kitajska šiviljica*, avtorja Daia Sijeja, kitajskega avtorja, ki piše v francoščini, prevajalka Alenka Moder Saje za 235 primerov povezovalca in veznika *mais* 78-krat uporabi *pa*, 57-krat *vendar*, 42-krat *ampak*, 34-krat se odloči za izpust veznika, ki ga označi z ločilom oziroma ga zamenja z *in*, desetkrat ga prevede z *a*, osemkrat s *samo*, petkrat s *toda*, enkrat *kakor koli že*, enkrat *čeprav*. Ker vmesnik ne loči med *i* in *ï* (*i* + trema), je bilo v tem naboru tudi 9 neustreznih pojavnic *maïs* (koruza).

Prevod s *pa* srečamo med drugim v primeru 26:

26)

Rarement des orages ou des averses, **mais** des pluies fines, constantes et sournaises, des pluies dont on eût dit...

Redkokdaj nevihte ali nalivi, pač **pa** droben dežek, stalen in potuhnjnjen, kakor da ne bo nikoli nehal ...

Izpust veznika je lahko popoln, kot je to v primeru 27, lahko pa je *mais* nadomeščen z *in*, kar je pogost primer prevajanja predvsem pri *mais* »spodbijanja«, tudi v primeru 28:

27)

Son poids etait un peu etonnant par rapport a sa taille **mais** il me fut impossible de savoir...

Glede na velikost je bil presenetljivo težek, bogve, kaj je notri ...

28)

Il était à peine gros comme la paume d'une main, **mais** avec une jolie sonnerie, pleine de douceur.

Velika je bila komaj kot pest **in** je zelo prijetno, zelo milo zvonila.

Po tem vzoru so prevedene tudi sintagme kot *belle mais inculte*, *sale mais solide*, protivno razmerje prevajalka izrazi z veznikom *in*, ki pa ne daje istega učinka kot je v izvirniku.

Formula *pa* in izpust se izkaže za najpogostejšo (primer 29):

29)

Mais Luo se montra un conteur de génie: il racontait peu, **mais** jouait tour à chaque personnage, en changeant sa voix et ses gestes.

Luo **pa** se je izkazal kot genialen pripovedovalec : bolj malo je opisoval, raje je igral različne osebe in pri tem spreminjal glas in vedenje.

Precej primerov je prevedenih z *vendar* (prim. 30 spodaj):

30)

une page de livre depuis une éternité. Je m'en approchai tout de suite, **mais** le résultat fut plutôt décevant : c' était un catalogue de couleurs.

Takoj sem stopil zraven, *vendar* sem bil kar malo razočaran: bil je tovarniški katalog barv blaga.

Veznik in povezovalc *samo* je bistveno redkejši, pa *vendar* ga zasledimo v primeru 31:

31)

Mais comment une valise de livres aurait-elle pu échapper aux Gardes rouges ?

Samo kako se je lahko kovček s knjigami izmuznil rdeči gardi?

Glede na primere ne moremo trditi, da prevajalka striktno razlikuje med pisnim in ustnim prenosnikom, saj *a* in *pa* nastopata tako v dialogih kot v pripovedi (primer 32):

32)

Écoute, on est de bons amis, on s'amuse bien ensemble, *mais* si tu commences à raconter des conneries sur les livres interdits, merde.

Poslušaj, prijatelji smo, dobro se razumemo, *a* če boš govoril traparije o prepovedanih knjigah, presneto.

Tudi v primeru fatičnega *mais* prihaja do izpustov (primer 33):

33)

- Si, c'est moi qui l'ai fait ce soir. - *Mais* regarde !

Poglej!

Na splošno pri A. Moder Saje lahko opazujemo celo paletu različnih rešitev, večino ma ustreznih glede na izrekanjsko situacijo v romanu, poleg tega pa je veliko izpustov protivnega veznika (34 od 235), kar kaže na večjo implicitacijo medpovednih razmerij, kar je dejansko obraten proces, kot bi ga pri prevajanju pričakovali.

Precej je primerov *mais* »spodbijanja«, ki so prevedeni z *ampak*, *vendar* ali *pa*, sledi izpust (na primer zgoraj navedeni pod številko 28, 29, a takih je vsaj še 20): gre torej za pisateljevo slogovno naravnost po predstavljanju predmetnosti s pomočjo veznika *mais*, ki se v slovenščini razveže na različne načine.

V delu *Les Truismes, Svinjarije*, avtorice Marie Darrieussecque v prevodu Marka Crnkoviča se srečamo s prvoosebno pripovedjo, ki se pred nami odvija kot monolog glavne osebe, pripovedovalke. Med 235 prevodi povezovalca *mais* prevladuje *ampak* s 163 pojavnicami, *pa* mu sledi 36-krat, naslednja strategija je izpust (19-krat), potem pa si po frekvenci še sledijo *čeprav* šestkrat, *toda* štirikrat, *samo* dvakrat in po enkrat *pravzaprav*, *kljub temu*, *drugače* (*pa*) ter 2-krat *seveda* kot varianta *mais en fait*, *mais oui*. Prevajalčeva strategija je mnogovrstnost, ki je vezana na govorjeni diskurz – monolog.

Tipični primeri prevajalčevega pristopa k vezniku *mais* s *pa* ali *ampak* so razvidni v primerih 34 in 35 spodaj:

34)

Il pleuvait dehors **mais** à l'Aqualand il fait toujours beau et chaud.

Zunaj je deževalo, v Aqualandu **pa** je vedno lepo in toplo vreme.

35)

On peut bien gagner sa vie à l'Aqualand, **mais** je m'y suis toujours refusée, même dans les moments où ma mère menaçait...

V Aqualandu ni težko kaj zaslužiti, **ampak** jaz nisem nikoli tega počela, niti takrat ne, ko mi je mama grozila, da me bo vrgla iz stanovanja.

Pogosto se kot prevodna varianta *mais* pojavlja *samo* (prim. 36):

36)

comme dans les magazines **mais** en plus appétissante.

kot iz revije, **samo** še veliko bolj privlačna.

Toda kot zborni veznik oziroma povezovalc nastopa tudi v začetku (prim. 37), kar predvideva tudi SSKJ, v primeru 38 pa navajamo izpust:

37)

Mais ce n'était pas ça non plus qui me donnait ce moral terrible, cette...

Toda te grozne morale mi ni dvigoval vznemirljivi občutek, da začenjam življenje znova ...

38)

se rendaient compte de rien, trop occupés d'eux-mêmes et de leur plaisir, **mais** le lit de massage devenait, sous leurs nouvelles envies, une sorte de meule

Nič niso razumeli, preveč so bili zaposleni sami s sabo in s svojimi užitki.

Masažna postelja je glede na njihove nove želje postajala kot kopica sena na polju.

V primeru izpusta povezovalca v primeru 38 zgoraj se protivno razmerje izgubi. V delu se pojavlja tudi *mais* v tipičnih primerih govorjenega konteksta (prim. 39, 40 in 41), kjer je fatični *mais* bodisi izpuščen (prim. 39), bodisi preveden v skladu s spremljevalnim kontekstom:

39)

Le gendarme qui était avec eux a dit : « **Mais** c'est la SPA qu'il faut appeler ! »

Z njimi je bil žandar, rekel je: »Ne, veterinarje je treba poklicati!«

40)

Supérieur s'est illuminé, il m'a regardée dans les yeux et il a dit que **mais** oui, qu'il me reconnaissait très bien même si je ne m'étais pas...

Takrat se je višjemu posvetilo, pogledal me je v oči in rekel, saj res, **seveda**, dobro se me še spominja, čeprav od takrat nisem najboljše skrbela zase ...

41)

La dame a dit en bégayant : « **Mais** elle pleure ! *mais* elle pleure ! »

Gospa je zajecjlala: »**Ampak** saj *vendar* joče! Ona joče!«

Predvsem v teh zadnjih primerih do izraza pridejo besedilne značilnosti romana. Gre za prvoosebno pripoved, ki je napisana knjižno, *vendar* ne v zelo visokem registru, kar tudi vpliva na izbor veznikov oziroma povezovalcev, ki jih je uporabil prevajalec. Prevladovanje *ampak* in *pa* v prevodu je vsekakor ustrezno.

Sledi delo *Eldorado* Laurenta Gaudéja v prevodu Tjaše Mohar. Tu bi strategijo lahko poimenovali z »*mais* je enako *vendar* ali izpust«. Glede na druge prevode to besedilo izstopa, slika postane precej drugačna kot pri prejšnjih analizah, ker močno prevladuje veznik in povezovalec *vendar*, sicer ne kot edina, a prevladujoča prevodna varianta za *mais*. Prehod med prevajalcema Crnkovičem in Moharjevo je razviden že po prevajanju veznikov. Od 255 *mais* jih Tjaša Mohar 190 prevede z *vendar*, na drugem mestu je izpust (19-krat), sledi *pa* s 15 pojavitvami, *ampak* z 12 pojavitvami, devetkrat je uporabljen *a*, dosledno vedno v primerih, ko gre za zloženi veznik: *a ko*, *a če*, *a še preden*. *Marveč* je zapisan 3, po 2-krat *samo* in *temveč* ter enkrat kombinacija veznikov *pa vendar*, *le* in zaznamovalec dopuščanja *že ... ko*.

Delo je tretjeosebna pripoved, dialogov praktično ni. Strategija *mais* je enako *vendar* na trenutke ni najbolj posrečena. Glede na razpoložljivost in slogovne vrednosti *vendar* se bojimo, da je raba nekoliko pretirana. Dva primera rabe *vendar* v vezniški vlogi sta prikazana spodaj (42 in 43), v primeru 44 *vendar* nastopa kot povezovalec na začetku povedi:

42)

Il se retourna d'un coup *mais* ne vit rien d'autre, dans la foule, que les badauds qui avançaient à...

Nenadoma se je obrnil, *vendar* je v množici videl samo radovedneže, ki so se s počasnimi koraki pomikali naprej ...

43)

Il essaya de sourire *mais* n'y parvint pas tout à fait.

Skušal se je nasmehniti, *vendar* mu ni povsem uspelo.

44)

Mais il se dit que si elle l'avait reconnu, c'est qu'en deux ans, au moins,

Vendar si je rekel, da ga vsaj v dveh letih čas ni preveč popačil, glede na to, da ga je prepoznala.

V primerih 45 in 46 imamo opravka z izpustom veznika *mais*, v primeru 45 je zamenjan z *in*, kjer se v prevodu v odnosu do izvirnika izgubi informacija, saj je protivno priredje nadomeščeno z vezalnim, v primeru 46 pa gre za *mais* spodbijanja, ki vnaša nasprotje v razmerje, in je tudi izpuščen. Implicitno v francoskem izreku je, da biti majhen in mršav ne pomeni nič dobrega, ta informacija pa se v slovenskem prevodu izgubi. Takih primerov je v celotnem korpusu prevoda *Eldorado* še več.

45)

Nous restons dans la pénombre du café **mais** nous jouissons de la vue sur la place.

Sediva v polmraku kavarne in opazujeva dogajanje na trgu.

46)

C'était un homme petit, au corps maigre **mais** au visage racé.

Bil je majhne rasti, mršavega telesa in plemenitega obraza.

Veznik *a* se vedno pojavlja v kombinaciji s časovnim veznikom v primerih kot so *mais lorsque – a ko, mais avant que – a še preden* (drugega srečamo v primeru 47 spodaj), podobno je tudi s *pa* (primer 48 spodaj):

47)

Mais avant qu'il n'ait pu le faire, elle reprit la parole et sa voix le gifla.

A še preden je lahko spregovoril, je spet povzela besedo in njen glas je zvenel kot klofuta.

48)

Mais lorsqu'il croisa son regard, il fut frappé par cette tristesse noire qui...

Ko pa je ujel njen pogled, ga je presunila črna žalost, zaradi katere se je z vso močjo oklepala zaščitne ograje.

Veznik in povezovalac *pa* nastopa tudi sam (prim. 49):

49)

Plus riche que moi, bien sûr, **mais** qui ne l'est pas ?

Seveda je bogatejši od mene, kdo *pa* ni?

Precej je primerov stopnjevanja, kjer se kot ustreznica *mais* pojavlja *marveč* (prim. 50) oziroma *temveč* (primer 51), kar deluje privzdignjeno:

50)

Et ce ne fut pas pour demander quoi que ce soit, **mais** pour raconter.

Ne zato, da bi kaj vprašala, *marveč* zato, da bi mu povedala svojo zgodbo.

51)

Non pas pour se noyer **mais** pour être encore davantage au cœur de la tourmente.

Ne zato, da bi utonil, *temveč* da bi se potopil še globlje v osrčje neurja.

Veznik in povezovalac *ampak* je rezerviran za tipične govorne položaje, kjer sam zase predstavlja argumentacijo (primer 52 spodaj):

52)

... – **Mais** ... – Je ne sais pas et, au fond, cela n'a aucune importance.

»**Ampak** ...

Ob primerih, ki prikazujejo na splošno ustrezno in raznoliko rabo različic *mais* v prevodu dela Eldorado moramo vseeno še enkrat omeniti klasifikacijo veznika

in povezovalca *vendar*. V pomenu vključuje tudi zadržana čustva oziroma možno izražanje nestrpnosti; morda se s tem vpisuje v poetiko dela Eldorado, a vseeno se raba zdi na trenutke pretirana.

Zadnje delo je *Les Mamifères, Sesalci*, Pierra Mérota v prevodu Jana Jone Javorška. *Mais* se pojavi 107-krat, v prevodu pa prevladuje *ampak* z 71 pojavitvami, sledijo *pa* (devetkrat), *vendar* (osemkrat), *samo* (petkrat), *toda* (trikrat), *čeprav* in *a* (vsak po dva-krat) in izpust protivnega veznika oziroma njegova zamenjava (petkrat), ter *pa vendar-le* (enkrat) za sestavljeni *mais malgré tout*. Tudi tu gre za tretjeosebno pripoved, vendar je jezik gibek in blizu pogovornim frazam, naracijo pa pogosto prekinjajo dialogi.

Povezovalc *mais* na začetku je preveden kot *toda* ali *ampak*, prevladuje pa slednji (prim. 53 in 54):

53)

Mais l'oncle n'a aucun honneur, ses résolutions du matin s'effondrent au coucher du soleil.

Toda kaj ko stric nima nobene časti, njegovi jutranji sklepi se s sončnim zahodom razblinijo.

54)

Mais le grand-père se mit à rugir : « Il ne faut plus toucher à l'alcool ! »

Ampak dedek je ves zaripnil: »Alkohola se ne smeš niti dotakniti!«

Pogosta je tudi pogovorna raba v dialogih, kjer prevladuje fatični *mais*, tudi v kombinaciji z drugimi vezniki in prislovi (primeri 55, 56 in 57 spodaj):

55)

– Ah **mais** ..., fit l'oncle, et mes symptômes ? Et leur cause ? – Ne touchez plus.

– **Ja, ampak** ..., se je oglasil stric, kaj pa moji simptomi?

56)

Mais soit ... Admettons que votre école soit dans cette fameuse rue.

Pa naj bo ...

57)

Mais, en définitive, quelle que soit cette quantité, vous êtes alcoolique.

Ampak dejansko ste ne glede na količino alkoholik takrat, ko drugim rečete, da je alkohol vaš glavni problem.

Med tipične veznike, ki se kot prevodna ustreznica *mais* pojavljajo v dialogu, spada tudi *pa* (primer (58) spodaj):

58)

Il est composé d'une grande variété d'écorchés vifs, **mais** surtout de professionnels de la nuit : serveuses, barmen, portiers, patrons.

Noč sestavlja velik izbor živih okostnjakov, predvsem **pa** njeni profesionalni prebivalci: natararice, strežaji, vratarji, lastniki barov in diskotek.

Tudi pri Javoršku bi lahko prevodno strategijo poimenovali od konteksta odvisna mnogovrstnost, kot prevodna ustreznica *mais* prevladuje *ampak*. Nabor veznikov in povezovalcev za *mais* je povsem ustrezen.

4 SKLEP

Glede na opravljeno analizo lahko potegnemo nekaj zaključkov: na analizi vzorca 2500 pojavnih *mais* in njegovih prevodov se je hipoteza potrdila. V primerljivih kontekstih se za veznik in povezovalac *mais* pojavljajo različni slovenski prevedki, ki so bolj odvisni od prevajalčeve izbire kot od (ne)enakovrednosti konteksta, torej se prevajalci ravnaajo predvsem po svojem lastnem slogu pisanja oziroma prevajanja, ki pa je odvisen tudi od besedilnega tipa, ki na tem mestu nastopa v romanu.

Ne bi mogli trditi, da je kateri koli prevod neustrezen, lahko pa kvantitativno dokažemo preference posameznih prevajalcev za določen veznik: pri Mozetiču je to *toda*, pri Koncutovi *a*, pri Moharjevi *vendar*, pri Moder-Sajetovi *pa*, pri Crnkoviču in Javoršku *ampak*. Boljši prevodi variirajo povezovalce, manj posrečeni so tisti, kjer najdemo več avtomatizma in morda več izpustov, ki posredno vplivajo na interpretacijo. Ugotavljanje, ali so prevodi eksplicitnejši od izvirnika, že zaradi načina raziskave ni moglo priti v poštev, saj smo iskali primere protivnega razmerja, ki jih uvaaja veznik in povezovalac *mais*, bilo pa je precej primerov implicitacije v prevodih, ki so bili odvisni predvsem od odločitve prevajalca.

Glede na prevajanje različnih vrednosti *mais* prevod argumentacijskega *mais* in dopustnega *mais* ne kaže pri prevajalcih nobenih posebnosti ali možnih težav, prav tako ne stopnjevalni *mais*. Pri *mais* spodbijanja so se prevajalci pogosto odločali za izpust, pri čemer so se določene informacije izgubile.

Pokazalo se je tudi, da je prevajanje povezovalca *mais* kontekstualno vezano in se pri posameznih prevajalcih večinoma razlikuje tudi glede na tip besedila. Besedilne značilnosti, ki vplivajo na jezikovne registre in posledično na izbiro povezovalca, so dialog ali prvoosebna pripoved na eni strani, kjer je bil nabor povezovalcev in veznikov, ki so jih prevajalci smatrali za ustrezne, drugačen kot v tretjeosebni pripovedi, kjer sta prevladovala opis in naracija. Dialog in prvoosebna pripoved sta bila bolj naklonjena prevodu *mais* z *ampak* ali *pa* kot tretjeosebna pripoved in naracija.

V dialogih, torej okoljih, ki oponašajo govorjeni jezik, prevladuje fatični *mais*. To je še bolj očitno v primerih, kjer je *mais* dopolnjen še s kakšnim drugim pragmatičnim zaznamovalcem.

Po naših predhodnih kontrastivnih raziskavah (Schlamberger Brezar 2009) je v resnici najbolj prekriven z *mais* slovenski *ampak*, pa tudi *pa*, kar je za rešitev izbrala tudi večina prevajalcev.

Ob že zgoraj omenjeni normativnosti, ki naj bi omejevala rabo *ampak* na govoreni diskurz, se vseeno pokaže, da se prevajalci na eni strani izogibajo rabi *ampak* v naraciji, kjer raje izberejo *a* ali *toda*, rabe *ampak* pa nikakor ne omejujejo v dialogih.

Kaj nam je pri pregledu nabora prevodnih ustreznic veznika in povezovalca *mais* omogočila korpusna analiza? Predvsem nam je omogočila kvantitativno analizo z dostopom do velike količine besedila naenkrat. Kvalitativna obdelava pa se ne glede na korpusni pristop ne razlikuje veliko od »predkorpusnih« metodologij, saj je bilo za natančnejšo opredelitev pojavljanja veznika oziroma povezovalca in njegovih specifičnih prevodnih ustreznic treba analizirati širši kontekst in tip besedila, v katerem se ta pojavlja.

Pokaže se, da je prevajanje povezovalcev pri nekaterih prevajalcih povsem avtomatizirano, kar lahko dokažemo s ponavljanjem enega in istega veznika oz. povezovalca, ki se mestoma izpriča tudi kot manj ustrezen glede na jezikovni opis. Ta način pa po drugi strani predstavlja nevarnost zaradi polifonije veznika *mais*, ki je bila predstavljena v začetku. Vendar kljub temu, da prevajalci najbrž ne delajo poglobljene slovnične analize veznikov glede na vzorce, ki so bili predstavljeni v začetku, večinoma le-te ustrezno prevedejo glede na kontekst in besedilni tip.

Bibliografija

- Adam, Jean-Michel, 1990: *Éléments de linguistique textuelle*. Bruselj: Mardaga.
- Adam, Jean-Michel, 1994: *Les textes: types et prototypes*. Paris: Nathan.
- Anscombre, Jean-Claude in Oswald Ducrot, 1983: *L'Argumentation dans la langue*. Bruselj: Mardaga.
- Bajec, Anton et al., 1994: *Slovar slovenskega knjižnega jezika*. Ljubljana: DZS.
- Baker, Mona, 1993: Corpus linguistics and translation studies. Implications and Applications. Baker, M., Francis, G. in Tognini-Bonelli, E. (ur.): *Text and Technology: In honour of John Sinclair*. Amsterdam, John Benjamins. 17–45.
- Baker, Mona, 1995: Corpora in translation studies. An overview and some suggestions for future research. *Target* 7, 1995. 223–243.
- Balažic Bulc, Tatjana, Gorjanc, Vojko, 2009: Corpus tagging of connectors: the case of Slovenian and Croatia academic discourse. *Corpus linguistics 2009: [abstracts]*. Liverpool: University of Liverpool, str. 57.
- Ducrot, Oswald et al., 1980: *Les mots du discours*. Pariz: Minuit.

- Grad, Anton, 1984: *Francosko-slovenski slovar*. Ljubljana: DZS.
- Malmkjaer, Kirsten (ur.), 2004: *The Linguistic Encyclopedia*. London-New York: Routledge.
- Mezeg, Adriana, 2010: Compiling and using a French-Slovenian parallel corpus. V: Xiao, Richard (ur.). *Proceedings of The International Symposium on Using Corpora in Contrastive and Translation Studies : 2010 conference (UCC-TS2010)*. Ormskirk: Edge Hill University, <http://www.lancs.ac.uk/fass/projects/corpus/UCCTS2010Proceedings/papers/Mezeg.pdf>, dostop marec-maj 2012.
- Mezeg, Adriana, 2011: *Korpusno podprta analiza francoskih polstavkov in njihovih prevedkov v slovenščini: doktorska disertacija*. Ljubljana: Univerza v Ljubljani.
- Morel, Mary-Annick, Danon-Boileau, Laurent, 1998: *Grammaire de l'intonation*. Pariz, Gap: Ophrys.
- Olohan, Maeve, 2004: *The use of corpora in translation studies*. London-New York: Routledge.
- Pisanski Peterlin, Agnes, 2009: Kontrastivna retorika. Kocijančič-Pokorn, Nike (ur.). *Sodobne metode v prevodoslovnem raziskovanju* (Zbirka Prevodoslovje in uporabno jezikoslovje). Ljubljana: Znanstvena založba Filozofske fakultete, 146–162.
- Schlamberger Brezar, Mojca, 1998: Vloga povezovalcev v diskurzu. Štrukelj, Inka (ur.). *Jezik za danes in jutri: zbornik referatov na II. Kongresu Društva za uporabno jezikoslovje Slovenije*. Ljubljana: Društvo za uporabno jezikoslovje Slovenije, Inštitut za narodnostna vprašanja 194–202.
- Schlamberger Brezar, Mojca, 2000: Les connecteurs en combinaison avec les marqueurs modaux : l'exemple du français et du slovène. *Linguistica* 40, 2, 273–282.
- Schlamberger Brezar, Mojca, 2002: Le rôle syntaxique et pragmatique des connecteurs dans le discours argumentatif français. *Linguistica* 42 89–110.
- Schlamberger Brezar, Mojca, 2005: Mais - inverseur d'orientation argumentative. Mais justement? Vilkou-Poustovaia, Irina (ur.): *L'exception entre les théories linguistiques et l'expérience, Faits de langues*. 25, Gap, Pariz: Ophrys 146–150.
- Schlamberger Brezar, Mojca 2009: *Povezovalci v francoščini: od teoretičnih izhodišč do vloge v diskurzu* (Zbirka Prevodoslovje in uporabno jezikoslovje). Ljubljana: Znanstvena založba Filozofske fakultete.
- Schlamberger Brezar, Mojca, 2010: La préparation des corpus parallèles et comparables et leur utilité pour le traducteur. *Cahiers du GEPE Traduction et domaines de spécialités* (Actes des Journées d'études de Strasbourg « Outils de traduction, outils du traducteur ? ». Université Marc-Bloch- Strasbourg 2, 18-19 septembre 2008.

- Schlamberger Brezar, Mojca, 2011: Les marqueurs de l'oral en français et en slovène et la justification de leur utilisation en interprétation. Pavelin Lešić, Bogdanka (ur.). *Francontraste. 2, La francophonie comme vecteur du transculturel*. Mons: CIPA, 255–262.
- Toporišič, Jože, 1976, 2000: *Slovenska slovnica*. Maribor: Obzorja.
- Vincenot, Claude, 1975: *Essai de grammaire slovène*. Ljubljana: Mladinska knjiga.
- Žagar, Igor Ž., Schlamberger Brezar, Mojca, 2009: *Argumentacija v jeziku*, (Digitalna knjižnica, Dissertationes, 4). Ljubljana: Pedagoški inštitut.

Internetne strani

- Korpus Canadian Hansard <<http://www.tsrali.com/>> (dostop maj 2012)
- Korpus Spook <<http://nl.ijs.si/spook/>> (dostop julij-september 2012)
- Korpus Spook – FraSloK <<http://nl.ijs.si/spook/konkor/>> (dostop februar-julij 2012)
- Križaj Ortar, Martina: Slovenski knjižni jezik 2, 1998–1999, <<http://kozjestezi-ce.tripod.com>> (dostop maj, september 2012)
- Trésor de la langue française TLF <<http://www.atilf.atilf>> (dostop februar-julij 2012)

Pogled na pojme med jeziki in jezikovnimi viri

Darja Fišer, Kristina Bizjak

Darja Fišer, Kristina Bizjak

Oddelek za prevajalstvo, Filozofska fakulteta Univerze v Ljubljani

darja.fiser@ff.uni-lj.si, kristabizjak@gmail.com

Abstract

In this paper we compare the lexico-semantic inventory of two different types of language resources, namely the semantic lexicons Princeton WordNet and sloWNet with the English-Slovene parallel corpus of literary texts that is part of the SPOOK multilingual translation corpus. A lexical analysis was performed in order to establish the degree of lexical overlap between the wordnets and the corpus. A further analysis of the corpus that was manually annotated with wordnet senses shows the coverage of the conceptualizations in the Slovene semantic lexicon, constructed from a foreign-language resource, that are represented in the corpus and therefore relevant for Slovene. The results of both analyses show the greatest advantages and limitations of the developed resource.

Key words: lexical semantics, semantic annotation, wordnet, polysemy, word-sense disambiguation

Ključne besede: leksikalna semantika, semantično označevanje, wordnet, večpomenskost, razdvoumljanje

1 OZADJE IN SORODNE RAZISKAVE

Besedni pomen je zelo izmuzljiva kategorija, ki kljub poskusom številnih avtorjev, da ga ukrotijo, ostaja ohlapna. V praksi ga zato obravnavamo zelo pragmatično, odvisno pač od namena oz. konkretne aplikacije, v kateri ga uporabljamo. Na pogosto zabrisanost meja med posameznimi pomeni in subjektivno razlikovanje med njimi je pokazal že Lakoff (1987), s katerim se strinjajo številni leksikografi, ki opozarjajo, da so pomeni besed izpeljani, prilagojeni ali celo ustvarjeni s konkretnim kontekstom, v katerem je beseda uporabljena. Trdijo, da jih zaradi tega tudi ni mogoče dokončno in vnaprej naštet (Kilgariff 1997, Hanks 2000). Če privzamemo, da je pomene vsaj do neke mere mogoče popisati, je v leksikografiji in leksikalni semantiki še vedno ena osrednjih tem, kako določiti njihovo število in jih klasificirati (Atkins 1991).

Z besednim pomenom se ukvarjamo tudi v tej raziskavi. Z njo želimo ovrednotiti enega najpopularnejših načinov gradnje leksikalno-semantičnih virov, ki iz praktičnih razlogov pogosto temeljijo na tujejezičnih podatkovnih zbirkah (Vossen 1998, Tufiş, Cristea in Stamou 2004). Na podlagi angleškega predhodnika je bil izdelan tudi sloWNet (glej razdelek 2.2), prvi semantični leksikon za slovenščino (Fišer 2009). Cilj pričujoče raziskave je preveriti posledice, ki jih imajo tovrstni pristopi na uporabno vrednost dobljenih virov v praksi.

Ker bi za celovit in rigorozen preizkus prevodnega pristopa potrebovali tudi semantični leksikon, ki je bil izdelan posebej za ta jezik, kar pa je seveda povsem nerealistično, tovrstnih primerjav, vsaj v večjem obsegu, v literaturi nismo našli. Zato se pri testiranju avtomatsko zgrajenega sloWNeta naslanjamo na vir, ki realistične leksikalno-semantične informacije ponuja v izobilju, in sicer korpus. Primerjava leksikalne zbirke s korpusom ima dodatno pomembno prednost, da daje jasno sliko o (ne)ustreznosti vsebovanega leksikalnega inventarja glede na izpričano jezikovno rabo, kar je ključno, če želimo, da je izdelana zbirka tudi uporabna.

Raziskava je sestavljena iz dveh delov. V prvem s preprostimi metodami korpusne analize preverimo prekrivanje besedišča v leksikonu in v korpusu, s katerim želimo razkriti največje pomanjkljivosti v sloWNetu, v drugem delu pa si z opazovanjem pojmov na stičišču med dvema jezikoma prizadevamo ovrednotiti vpliv prenosa tujejezične kategorizacije pomenov v drug jezik in identificirati posledice, ki jih ta pristop prinaša.

Opazovanje pojmov v korpusu opravimo na podlagi semantičnih oznak v korpusu. Čeprav za slovenščino semantično označen korpus že obstaja (Fišer 2010), je označen korpus enojezičen, mi pa v tej raziskavi potrebujemo oznake za jezikovni par angleščina-slovenščina, zato smo se označevanja lotili sami, pri čemer se

naslanjamo na izkušnje, pridobljene v sorodnih projektih. Prvi vzporedni korpus, označen s pomeni iz wordneta, je MultiSemCor (Bentivogli, Forner in Pianta 2004), ki temelji na predpostavki, da se med prevajanjem izvirnika semantične informacije v veliki meri ohranijo. Avtorji raziskave zato uporabijo angleški korpus SemCor (Miller idr. 1994), ki že vsebuje semantične oznake iz Princeton WordNeta (Fellbaum 1998), ga prevedejo v italijanščino, avtomatsko poravnajo na besedni ravni in tako dobljene semantične oznake uporabijo tudi na italijanski strani. Recikliranje semantičnih oznak je mogoče, ker italijanski wordnet (Artale, Magnini in Strapparava 1997) vsebuje identične pojmovne kode.

Čeprav smo v naši raziskavi v grobem uporabili zelo podoben pristop, je projekt, kot so ga izvedli italijanski kolegi, problematičen iz različnih razlogov:

- (1) Pri delu so uporabili angleški korpus, ki so ga prevedli v italijanščino, vendar so za čim lažje avtomatsko vzporejanje korpusov na besedni ravni v naslednjem koraku prevajalcem naročili, naj angleške izvirne stavke prevajajo čim bolj dobesedno. S tem so bistveno vplivali na podobo italijanščine v prevodu, kar ima brez dvoma posledice tudi za vsebovan leksikalno-semantični inventar. Naš pristop je v primerjavi s tem boljši, ker uporabljamo vzporedni korpus, ki vsebuje prevode profesionalnih prevajalcev, namenjene objavi, zaradi česar ne podlegajo pragmatičnim potrebam konkretne raziskave in bolj verodostojno odražajo leksikalne in konceptualne fenomene, tako izvirnika kot prevoda.
- (2) Po prevajanju so nato oba dela korpusa avtomatsko poravnali na besedni ravni, s čimer so originalne semantične oznake prenesli v italijanščino. Kljub temu, da ta način pridobivanja semantičnih oznak deluje zelo privlačno, lahko glede na izkušnje iz predhodnih raziskav, v katerem smo tudi sami uporabili orodja za besedno poravnavo vzporednih besedil, zatrdimo, da zaenkrat takšne poravnave vsebujejo toliko napak, da pristop brez natančnega ročnega popravljanja rezultatov, ni uporaben. Besedna poravnava je pogosto problematična pri literarnih besedilih in drugih podobnih zvrsteh, kjer so prevodi nekoliko svobodnejši. Še slabše pa se avtomatska poravnava na besedni ravni odreže pri večjih strukturnih razlikah med jezikoma, npr. ko ima večbesedna zveza v izvirniku enobesedni prevod in obratno, kar je za angleško-slovensko kombinacijo zelo pogost pojav.
- (3) V želji po avtomatizaciji označevanja korpusa, ki je zelo zamudno in drago, različni avtorji pogosto posegajo tudi po semantičnih označevalnikih, kot sta na primer UKB (Agirre in Soroa 2009) in SenseRelate (Pedersen in Kolhatkar 2009), ki večpomenskim besedam glede na sobesedilo pripišeta najverjetnejši pomen iz wordneta. Vendar so tudi tovrstna orodja zaenkrat še premalo natančna za raziskave, kot je naša, za

katere je visoka zanesljivost oznak ključna, zato smo se v našem primeru označevanja lotili ročno.

V naslednjem razdelku predstavimo vire, ki smo jih v raziskavi uporabili, nato opišemo, kako je raziskava potekala. Temu sledi predstavitev rezultatov in diskusija, prispevek pa sklenemo s ključnimi ugotovitvami in načrti za prihodnje delo.

2 UPORABLJENI VIRI

V pričujoči raziskavi uporabljamo tri jezikovne vire: semantična leksikona za angleščino in slovenščino Princeton WordNet in sloWNet ter angleško-slovenski del vzporednega korpusa SPOOK. Tako leksikona kot korpus vsebujeta bogate leksikalno-semantične informacije, ki pa so v semantičnem leksikonu strukturirane in eksplicitno izražene, medtem ko jih je iz korpusa potrebno šele izluščiti. Korpus je poravnan na stavčni ravni, leksikona pa temeljita na skupnem inventarju pomenov, tako da je slovensko-angleške prevodne ustreznice za nek pojem zelo preprosto identificirati na podlagi njegove identifikacijske kode. V nadaljevanju razdelka na kratko predstavljamo glavne značilnosti virov, uporabljenih v raziskavi.

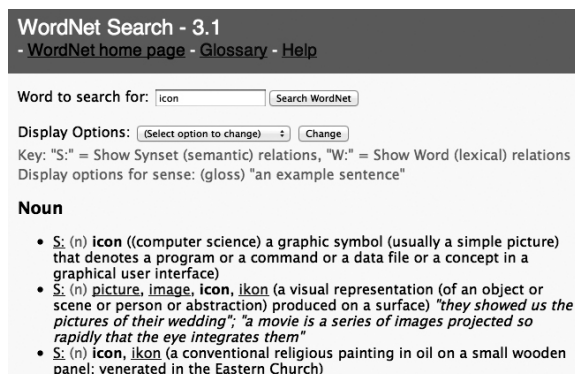
2.1 Princeton WordNet

Princeton WordNet (PWN) je obsežna leksikalna zbirka za angleški jezik, ki je začela nastajati v 80. letih prejšnjega stoletja na Oddelku za psihologijo na Univerzi v Princetonu in je kmalu postala zelo priljubljen pripomoček pri najrazličnejših nalogah računalniške obdelave naravnega jezika. V njej so samostalniki, glagoli, pridevniki in prislovi razvrščeni v t.i. sinsete, nize kognitivnih sinonimov oz. literalov, ki se uporabljajo za izražanje istega pojma (npr. *sick, ill*, slo. *bolan*). Sinsetom je dodana razlaga, pogosto tudi primer rabe in domenska oznaka, posamezni sinseti pa s semantičnimi in leksikalnimi relacijami (npr. *antonym*, slo. *protipomenka*) povezani v pojmovno mrežo. Wordnet vsebuje tako enobesedne kot večbesedne nize, pri čemer je upoštevana tudi metaforična in idiomatska raba (Fellbaum 1998: 3–17).

Čeprav je bila pred kratkim objavljena različica WordNet 3.1, je zanjo omogočeno samo iskanje v spletnem vmesniku, ne pa tudi datotečni prenos, zato v tej raziskavi uporabljamo različico 3.0, ki vsebuje 155.327 različnih besed. Te so razvrščene v 117.597 sinsetov, od katerih je slabih 70 odstotkov samostalniških. Enopomenskih besed v WordNetu je 128.321, večpomenskih pa 27.006, povprečna večpomenskost je tako 1,23 za samostalnike, 2,16 za glagole, 1,41 za pridevnike in 1,24 za prislove¹. Sinsete, v katerih se v spletnem pregledovalniku

¹ <http://wordnet.princeton.edu/man/wnstats.7WN> [15. 5. 2012]

Princeton WordNeta pojavlja večpomenska beseda *icon* (slo. *ikona*), prikazuje Slika 1.



Slika 1: Prikaz pomenov besede *icon* v Princeton WordNetu.

2.2 sloWNet

Po vzoru Princeton WordNeta je bil izdelan tudi wordnet za slovenščino (sloWNet), ki je ohranil strukturo PWN, slovenskim literalom pa so ponekod v manjšem številu dodane tudi že razlage pojmov in primeri rabe v slovenskem jeziku. Gradnja sloWNeta, ki je sicer še vedno v razvoju, je doslej potekala v treh fazah:

(1) **Avtomatska indukcija slovenskih sinsetov na podlagi različnih že obstoječih dvo- in večjezičnih jezikovnih virov (Fišer in Sagot, 2008)**

V tej fazi smo uporabili različne tipe že obstoječih leksikalnih virov, kot so dvojezični slovarji, večjezični vzporedni korpusi in Wikipedija, iz katerih smo izluščili večjezične leksikone, ki so vsebovali vse variante slovenskih prevodnih ustreznice za angleške iztočnice. Ustreden pomen smo jim pripisali z iskanjem preseka med leksikonom in wordneti v petih jezikih. Vse prevodne ustreznice, ki jim je bil pripisan isti pomen, smo združili v isti sinset (npr. *vojska*, *armada* za ang. *army*), prevodne ustreznice večpomenskih iztočnic, ki so jim bili pripisani različni pomeni, pa smo zakodirali v ločene sinsete (npr. *stranka* in *zabava* za ang. *party*).

(2) **Širjenje sloWNeta s pomočjo metod strojnega učenja (Sagot in Fišer, 2011)**

Ker smo v prvi fazi izdelave sloWNeta želeli zagotoviti osnovni nabor visoko zanesljivih sinsetov, je precejšnji delež gesel iz večjezičnega leksikona, ki smo ga izluščili iz različnih tipov jezikovnih virov, ostal neizkoriščen. Zaradi želje po čim večji natančnosti tako recimo nismo upoštevali

številnih večpomenskih iztočnic iz Wikipedije, ki pa bi bile za slovenski wordnet zelo dragocene. Zato smo osnovni wordnet uporabili kot model, s pomočjo katerega smo probabilistični klasifikator naučili razvrščati v najustreznejši sinset tudi preostala, težja gesla, s čimer smo sloWNet za dvakrat povečali.

(3) Identifikacija nezanesljivih literalov z uporabo referenčnega korpusa in distribucijske semantike (Sagot in Fišer 2012).

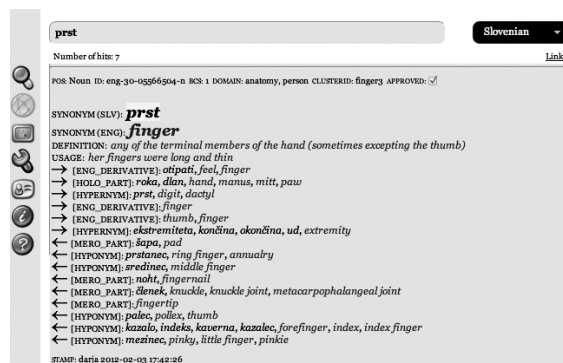
Razširjanje wordneta z metodami strojnega učenja je zahtevna naloga, ki ne prinaša popolnih rezultatov, zato je bilo v novo generiranih sinsetih precej šuma, ki bistveno zmanjšuje uporabno vrednost na ta način izdelanega semantičnega leksikona. Ker bi bilo ročno popravljanje razširjenega wordneta preveč zamudno in predrago, smo si v zadnji fazi projekta prizadevali identifikacijo napak v sloWNetu avtomatizirati. Pri tem smo uporabili referenčni korpus FidaPLUS², iz katerega smo izluščili kontekstualne informacije za literalne iz sloWNeta. V skladu z načeli distribucijske semantike smo nato kontekstualne informacije, pridobljene iz korpusa, primerjali z neposredno okolico literala v sloWNetovi semantični mreži. Kandidate z najslabšim rezultatom smo označili kot potencialne napake, ki jih bomo po ročnem pregledu po potrebi izbrisali.

Zadnja različica sloWNeta vsebuje 82.721 literalov, ki so razvrščeni v 42.919 sinsetov, kar predstavlja 36 % vseh sinsetov v Princeton WordNetu. Poleg enobesednih sloWNet vsebuje tudi številne večbesedne literalne in lastna imena. Samostalniki so daleč najbolj zastopani, saj predstavljajo več kot 70 % vseh sinsetov. 66 % literalov v sloWNet je enopomenskih, povprečna stopnja večpomenskosti pa je 2,07, kar je nekoliko več kot v Princeton WordNetu. Glede na to, da je bil sloWNet izdelan avtomatsko, višja stopnja večpomenskosti ne pomeni, da je slovensko besedišče bolj večpomensko, temveč nakazuje na napake v generiranih sinsetih.

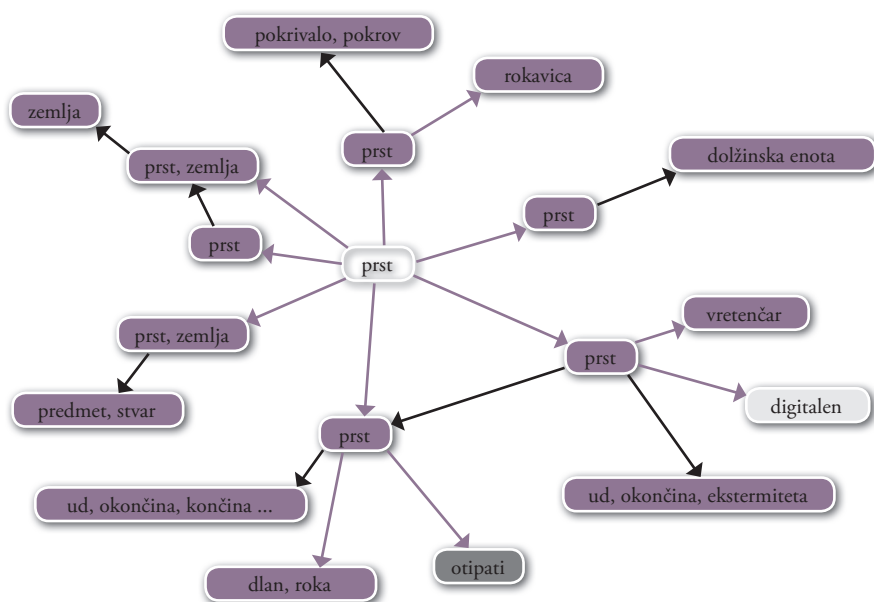
Z izdelavo sloWNeta se je kmalu pojavila potreba po orodju, ki bi omogočalo brskanje, vizualizacijo in popravljanje sinsetov. Ker nobeno obstoječe orodje iz različnih razlogov ni ustrezalo našim potrebam, smo razvili sloWTool³ (Fišer in Novak 2011), ki omogoča primerjavo sinsetov v različnih jezikih, osnovno in napredno iskanje po wordnetu, vizualizacijo semantične mreže v obliki hiperboličnih grafov in popravljanje sinsetov v najrazličnejših brskalnikih, s preprostim postopkom registracije in brez kakršnega koli predhodnega nameščanja programskih komponent. Slika 2 prikazuje primer osnovnega iskanja slovenske večpomenske besede *prst* v sloWToolu, Slika 3 pa vizualizacijo rezultatov za isti iskalni pogoj.

² <http://fidaplus.net/> [15. 5. 2012]

³ <http://nl.ijs.si/slowtool/> [15. 5. 2012]



Slika 2: Prikaz enega od sinsetov, ki vsebuje večpomensko besedo *prst* z angleško ustreznico, razlago, primerom rabe in semantičnimi relacijami.



Slika 3: Vizualizacija vseh pomenov besede prst v sloWNetu s semantičnimi relacijami, ki so ob njih zabeležene.

2.3 Korpus SPOOK

Za semantično označevanje besed in primerjavo ujemanja pojmov med jezikoma smo izbrali del korpusa SPOOK (Vintar 2009), ki je bil razvit v okviru projekta Slovensko prevodoslovje – viri in raziskave (2009–2012). Korpus v

vzporednem delu vsebuje izvirna besedila v angleščini, nemščini, francoščini in italijanščini ter njihove prevode v slovenščino, v primerljivem pa podobna besedila, ki so bila izvirno napisana v slovenščini in tako omogočajo primerjavo izvirne in prevedene slovenščine. Korpus, ki vsebuje 8 milijonov besed, je sestavljen iz pretežno literarnih besedil in je bil tokeniziran, oblikoskladenjsko označen, lematiziran in zapisan v XML v skladu z načeli TEI P5 (glej Erjavec 2012).

V pričujoči raziskavi smo uporabili del angleško-slovenskega vzporednega podkorpusa, iz katerega smo izbrali 5 knjižnih del čim bolj različnih avtorjev in žanrov, s čimer upamo, da smo zagotovili čim bogatejše besedišče in čim večjo raznolikost zastopanih pomenov. Izdelan podkorpus šteje nekaj več kot pol milijona besed na jezik, pri čemer je najkrajše besedilo znanstveno-fantastični roman *The Supernaturalist* avtorja Eoina Colferja, ki vsebuje veliko tehničnih opisov izmišljenega sveta v bližnji prihodnosti, najdaljši pa prvenec britanske pisateljice Zadie Smith *White Teeth*, ki je bogat s pogovornim jezikom. *The Way through the Woods* Colina Dexterja je kriminalka, *Harry Potter and the Deathly Hallows* izpod peresa J. K. Rowling je zadnji od njenih sedmih fantazijskih romanov, Tolkienov *The Two Towers* pa drugi del epske fantazijske pripovedi *Lord of the Rings*. Celoten seznam del, ki smo jih pri raziskavi uporabili, in njihovo velikost prikazuje Tabela 1.

Tabela 1: Seznam in velikost del, zajetih v raziskavo

Naslov in avtor dela	Št. besed - izvirnik	Št. besed - prevod
<i>The Supernaturalist</i> (Eoin Colfer)	62.235	58.775
<i>The Way through the Woods</i> (Colin Dexter)	87.024	76.270
<i>Harry Potter and the Deathly Hallows</i> (J. K. Rowling)	56.078	58.778
<i>Lord of the Rings: The Two Towers</i> (J. R. R. Tolkien)	146.771	150.367
<i>White Teeth</i> (Zadie Smith)	169.099	171.548
Skupaj	521.207	515.738

3 PREDSTAVITEV PRISTOPA

Raziskava je sestavljena iz dveh delov. V prvem opravimo celovito primerjavo leksikalnega inventarja, ki ga vsebujejo angleški in slovenski wordnet ter vzporedni korpus. Cilj leksikalne analize je preveriti, v kolikšni meri leksikalni zbirki pokrivata besedišče, ki se pojavlja v korpusu, oceniti njuno uporabno vrednost

pri delu s tematsko in slogovno raznolikimi besedili, kot je literatura, ter identificirati morebitne vrzeli v wordnetih, ki bi jih za zagotavljanje čim boljše pokritosti ključnega besedišča kazalo čim prej zapolniti.

V drugem delu raziskave pa za izbrane večpomenske besede v obeh jezikih ročno označimo naključni vzorec stavkov, v katerih se le-te pojavljajo. Z analizo pomenov, ki so za isti stavek uporabljeni v enem in drugem jeziku, prav tako preverimo, v kolikšni meri se pojmi v njih prekrivajo. Cilj semantične analize je ugotoviti jezikovno (ne)odvisnost pojmov, uporabljenih v korpusu, in na podlagi rezultatov semantičnega označevanja utemeljiti (ne)primernost pristopov gradnje leksikalno-semantičnih virov, ki izhajajo iz tujejezične semantične mreže. Nadalje si s semantično analizo korpusnih podatkov prizadevamo preveriti, v kolikšni meri je sloWNet, ki je še v razvoju, že zrel za uporabo v praksi, ter pridobiti vpogled v njegove trenutne največje pomanjkljivosti.

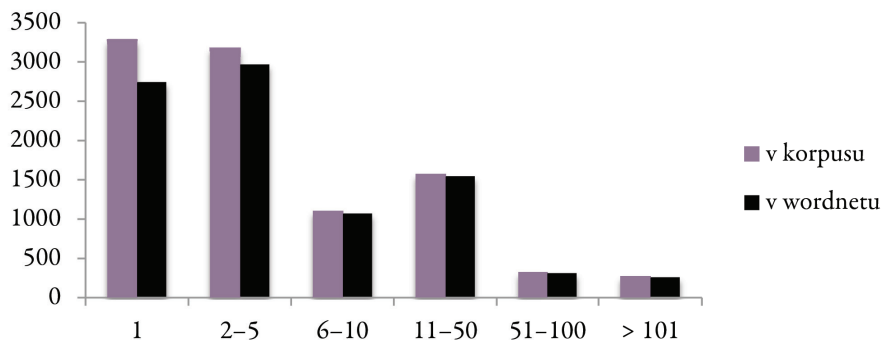
3.1 Leksikalna analiza

Pri leksikalni analizi smo se omejili na občne samostalnike, saj so le-ti po eni strani osnovni nosilci leksikalnega pomena in kot taki najbolj zanimivi za medjezikovno primerjavo, po drugi pa trenutno najboljše zastopani v slovenskem wordnetu, glede na to, da je bil njegov dosednji razvoj usmerjen prav nanje.

3.1.1 Primerjava angleškega wordneta in korpusa

Angleški del korpusa, ki šteje 521.207 besed, vsebuje 9.715 različnih lem, ki so označene kot občni samostalniki. Najpogostejši je *time*, ki se pojavlja 1.766-krat. Z več kot tisoč pojavitvami se ponašajo še: *man* (1.670), *eye* (1.409), *hand* (1.310), *thing* (1.209), *way* (1.177) in *people* (1.019). Dobra tretjina oz. 3.281 samostalnikov s tega seznama je enopojavnic, ki so z ozko specializiranega besedišča (npr. *lynx*), kratice (npr. *OAP*), izražajo avtorjevo kreativnost (npr. *toughie*), napake avtomatskega oblikoskladenjskega označevanja (npr. *Audi*), ali pa sicer splošne besede, ki se zaradi razmeroma majhnega obsega korpusa v njem slučajno ne pojavljajo pogosteje (npr. *burglar*).

Angleški wordnet vsebuje skoraj vse občne samostalnike iz korpusa, natančneje 8.909 oz. 91,7 %, kar je zelo visoka stopnja prekrivanja, še posebej če izvzamemo enopojavnice, s čimer stopnja prekrivanja med obema viroma naraste na 95,8 %. Stopnjo prekrivanja med viroma glede na pogostost, s katero se angleški samostalniki pojavljajo v korpusu, ponazarja slika 4.



Slika 4: Rezultati prekrivanja občnih samostalnikov v angleškem wordnetu in korpusu glede na njihovo pogostost v korpusu

39,5 % samostalnikov, ki se pojavljajo tako v korpusu kot v wordnetu, ima v wordnetu samo en pomen, ostali so večpomenski. Približno enak delež jih ima dva ali tri pomena (39,2 %), 4–10 pomenov ima 20,14 % samostalnikov, več kot deset pomenov pa je zelo redkih (1,19 %). Od vseh občnih samostalnikov v korpusu ima v wordnetu največ pomenov samostalnik *head*, in sicer 33.

Od besed, ki imajo v angleškem delu korpusa več kot 100 pojavitev, jih v angleškem wordnetu manjka le sedem: *something*, *anything*, *Cosmo*, *everything*, *other*, *anyone* in *everyone*, ki so vse, razen lastnega samostalnika *Cosmo*, ki je napačno označen kot občni, v wordnetu obravnavane kot zaimki in jih zaradi tega med samostalniškimi sinseti ni. Ob tej ugotovitvi, ki kaže, da pri neprekrivanju ne gre za pomanjkljivost wordneta, temveč za različno obravnavo besednih vrst med viroma, smo se odločili analizirati vseh 66 samostalnikov, ki se v korpusu pojavijo 10-krat ali pogosteje, v Wordnetu pa jih nismo našli. Izkaže se, da gre pri večini za diskrepance pri pripisovanju besedne vrste, napake v tokenizaciji oz. lematizaciji (npr. *ty* namesto *tie*), navajanju kanonične oblike v ednini oz. množini (npr. *goggle* v korpusu in *goggles* v wordnetu) ter zapisu z veliko oz. malo začetnico (npr. *mister* v korpusu in *Mister* v wordnetu). Resnično manjkajočih samostalnikov, ki so v korpusu pravilno označeni in lematizirani, pa jih v wordnetu vseeno ni, je samo šest, zato smo njihovo pogostost preverili še v obsežnih referenčnih korpusih BNC (Burnard in Aston 1998) in ukWaC (Ferraresi idr. 2008). Kot je razvidno iz Tabele 2, so tudi v teh virih z izjemo samostalnikov *orc* in *jus* zelo redki, zaradi česar jih upravičeno ni v wordnetu. Poleg tega je razmeroma visoko število zadetkov za lemo *jus* zgolj navidezno, saj se poleg latinskega izraza za pravo med zadetki pogosto pojavlja tudi napačno označena in lematizirana pogovorna različica besede *just*.

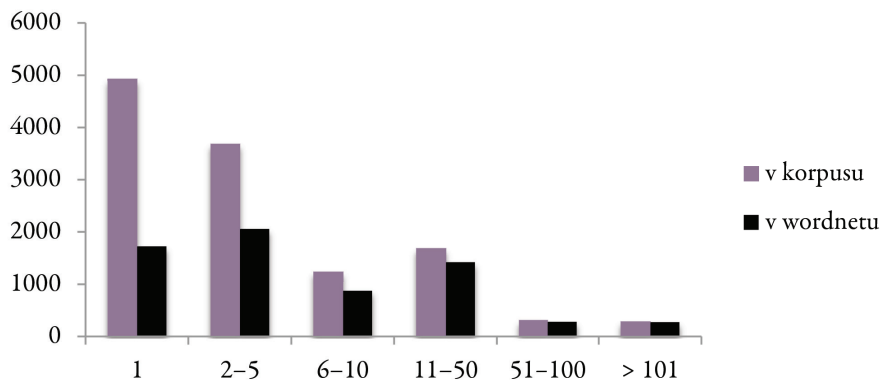
Tabela 2: Primerjava pogostosti pojavljanja samostalnikov, ki manjkajo v angleškem wordnetu in korpusih SPOOK, BNC ter ukWaC

Samostalni	SPOOK	BNC	ukWaC
supernaturalist	58 (111,3)	0 (0 na mio)	69 (0 na mio)
orc	56 (107,4)	173 (1,5 na mio)	782 (0,5 na mio)
jus	24 (46,1)	106 (0,9 na mio)	1519 (1 na mio)
bezoar	15 (28,8)	0 (0 na mio)	11 (0 na mio)
cilice	14 (26,9)	0 (0 na mio)	25 (0 na mio)
symbolologist	13 (24,9)	0 (0 na mio)	26 (0 na mio)

3.1.2 Primerjava slovenskega wordneta in korpusa

V slovenskem delu korpusa je 515.738 besed in 12.119 različnih lem, ki so označene kot občni samostalniki. Razlog za precej večje število samostalniških lem, kot jih najdemo v angleškem delu korpusa, je najverjetneje težja in zato tudi nekoliko slabša lematizacija slovenskega dela korpusa (npr. *kucelj* namesto *kucelj*), kar je razvidno že iz precej večjega števila enopojavnic v tem delu korpusa, in sicer 4.925 oz. 40,6 %. Preostala distribucija besed po pogostosti je zelo podobna angleški. Tudi v slovenskem delu korpusa se najpogosteje pojavlja samostalni *čas*, in sicer 1.490-krat. Več kot tisoč pojavitev imajo še *roka* (1.474), *človek* (1.429), *oči* (1.231), enako kot v angleščini, poleg njih pa še *dan* (1.152), *leto* (1133), *vrata* (1066) in *glava* (1.002). Poleg napak v tokenizaciji, oblikoskladenjskem označevanju in lematizaciji je med enopojavnicami tudi v slovenskem delu korpusa precej ozko specializiranih, pogovornih in kreativnih izrazov (npr. *relikviarij*, *bejbika*, *primitivščina*) ter kar nekaj takih, ki so jih prevajalci pustili v angleščini.

Za razliko od angleškega wordneta, ki se ponaša z odličnim pokrivanjem besedišča iz korpusa, je ujemanje med slovenskim delom korpusa in sloWNetom bistveno slabše, saj je v njem mogoče najti le 54,78 % lem občnih samostalnikov, izluščenih iz korpusa. Pokritost je najboljša in primerljiva z angleško zgolj za najpogostejše besedišče, in sicer 95,14 % za vse besede, ki se v korpusu pojavijo več kot 100-krat in 92,11 % za tiste s frekvenco 50–100, kar glede na način gradnje slovenskega wordneta, s katero smo si prizadevali zajeti predvsem osnovno besedišče, niti ni presenetljivo. Stopnjo prekrivanja med sloWNetom in slovenskim delom korpusa ponazarja Slika 5, iz katere je razvidno, da je ujemanje najslabše za besede s frekvenco 1–5.



Slika 5: Rezultati prekrivanja občnih samostalnikov v slovenskem wordnetu in korpusu glede na njihovo pogostost v korpusu

Med samostalniki, ki se pojavljajo v obeh virih, je glede na sloWNet enopomenskih 32,5 %, kar je v primerjavi z angleščino malenkost manj. Dvo- in tripomenskih je približno še enkrat toliko (32,99 %), podobno kot v angleščini. 4–10 pomenov ima 27,46 % samostalnikov, več kot deset pa jih najdemo za 7,05%, kar je veliko več kot v angleščini in nakazuje na napake, do katerih je prišlo pri avtomatski izdelavi sloWNeta. Da v teh primerih pogosto ne gre za pravo večpomensko, temveč šum v sinsetih, jasno kaže samostalnik *oseba*, ki se v sloWNetu pojavlja v rekordnih 68 pomenih. Te napake je potrebno čim prej odpraviti.

Od besed, ki se v slovenskem delu korpusa pojavljajo več kot 100-krat, jih v sloWNetu manjka le 14. Če izločimo napačno označena lastna imena, jih ostane samo še pet, večinoma ozko specializiranih oz. pogovornih izrazov: *priorstvo*, *pizda*, *gral*, *ent* in *drevje*. Med samostalniki, ki se v korpusu pojavljajo več kot desetkrat, je takšnih, ki jih v sloWNetu ni, 195. Med njimi jih je 17 napačno lematiziranih oz. oblikoskladenjsko označenih (npr. napačno označen glagol *mislita*) ali se razlikujejo v številu (npr. *špaget* v korpusu, *špageti* v sloWNetu). Za preostale besede pa iskanje po največjem slovenskem referenčnem korpusu GigaFida pokaže, da jih ima le 122 frekvenco, normalizirano na milijonski korpus, 1 ali več. Med njimi je 35 takšnih, ki so manjšalnice (npr. *skrinjica*), ženske oblike (npr. *prijateljica*) ali glagolniki (npr. *spreminjanje*) besed, ki jih sloWNet že vsebuje.

Za preostalih 87 manjkajočih samostalnikov smo preverili, ali v sloWNetu manjkajo zgolj zato, ker še ni dosegel zadostne velikosti in bi bila odstopanja odpravljena s preprosto razširitvijo leksikona, ali pa gre za jezikovno- in

kulturno-specifična odstopanja. V prvem primeru je mogoče identificirati sinset, ki v sloWNetu že obstaja, vendar je zaradi pomanjkljivih virov, iz katerih je bil zgrajen, zaenkrat še prazen oz. ne vsebuje tega literala, v nasprotnem primeru pa ustreznega sinseta v sedanji pojmovni mreži ni mogoče najti in bi jo bilo potrebno razširiti s sinseti, specifičnimi za slovenščino. Analiza je pokazala, da za večino manjkajočih samostalnikov obstaja vsaj en primeren sinset, petina (18) pa je takšnih, za katere povsem ustreznega sinseta ni. Večinoma gre za manjšalne oblike samostalnikov (7), ki jih v angleščini ne uporabljajo (npr. *omarica*, obstajata sicer sinseta *omara* in *nočna omarica*, vendar nobeden od njiju ni povsem ustrezen, saj je prvi presplošen, drugi pa preveč specifičen). V 5 primerih težave povzročajo ženske oblike samostalnikov, za katere ni povsem jasno, ali bi jih bilo bolje uvrstiti v isti sinset, ki vsebuje nevtralno moško obliko, ali jih obravnavati kot podpomene (npr. *prijateljica*, obstaja sicer nevtralen sinset *prijatelj*). Preostali problemi pa so jezikovne narave, npr. neštevniki samostalniki, kot sta *drevje* in *kamenje*, za katera sta najbližja sinseta, ki v wordnetu obstajata, *drevo* in *kamen*. Za to besedišče bo v nadaljnjih fazah razvoja slovenskega wordneta potrebno zagotoviti jezikovno specifične koncepte in jih dodati k obstoječi semantični mreži, kot so to že storili pri wordnetih za nekatere druge slovanske jezike, npr. poljščino, češčino in hrvaščino. Med problematičnimi besedami najdemo tudi *stežaj*, ki se v korpusu vselej pojavlja v zvezi *na stežaj* in se v angleščino prevaja z drugo besedno vrsto, kar je v trenutni različici wordneta nemogoče zakodirati. V wordnetu prav tako manjka eden od pomenov samostalnika *dir*, za katerega osnovni pomen, ki označuje način teka pri konjih, obstaja, manjka pa njegov preneseni pomen, ki označuje podobno hitenje pri ljudeh. Tudi sinseta za *kupe* v wordnetu zaenkrat še ni mogoče najti. V angleškem wordnetu sicer obstaja izraz *compartment* (*a partitioned section, chamber, or separate room within a larger enclosed area*), ki pa je nekoliko splošnejši, zato bi mogoče v kateri od prihodnjih različic sloWNeta lahko dodali še podpomenko *kupe*, s specifično definicijo (*zaprt prostor v potniškem vagonu*). Prav tako bo tudi za vse ostale tovrstne pojme v nadaljevanju razvoja slovenskega wordneta potrebno dodati sinsete, specifične za slovenščino.

3.2 Semantična analiza

V drugem delu raziskave analiziramo pomen, ki so zastopani v obeh delih korpusa, preverimo, koliko se prekrivajo, in ugotovimo, v kolikšni meri so usklajeni z wordnetoma v obeh jezikih. S tem želimo po eni strani dobiti vpogled v reprezentativnost pomenov, vsebovanih v wordnetu, glede na dejansko jezikovno rabo, ki je izpričana v korpusu, po drugi strani pa proučiti, kakšne posledice ima na rabo slovenskega wordneta v praksi odločitev za njegovo prevajanje iz angleščine.

3.2.1 Izbor in označevanje besed v korpusu

Pri analizi smo se omejili na vse večpomenske občne samostalnike, ki obstajajo v obeh wordnetih in se vsaj desetkrat pojavljajo v vseh petih knjigah, vključenih v korpus. Na našo odločitev, da v raziskavo zajamemo zgolj besede z razmeroma visoko pogostnostjo, je pomembno vplivalo dejstvo, da je redkejša besedišča v sloWNetu zaenkrat še slabo pokrito. Pogosto besedišče je prav tako tipično bolj zanimivo za opazovanje večpomenskosti. Poleg tega pa je naš osnovni namen raziskave vključeval predvsem medjezikovno proučevanje pomenskega inventarja osnovnega besedišča z dolgoročnejšim ciljem razvoja avtomatskega pomenskega označevalnika in ne identifikacije izjem in redkih pojavov, zato smo s pogoji želeli zajeti le pogosto in splošno besedišče.

V angleščini tem pogojem ustreza 39 besed, v slovenščini pa 35. Med 39 angleškimi besedami se je glede na wordnet kot najbolj večpomenska beseda izkazala angleška beseda *head*, ki ima 33 enobesednih pomenov, najmanj, po štiri, pa jih imajo besede *child*, *hour*, *people* in *year*. Glede na to, da je bil slovenski wordnet zgrajen avtomatsko in vsebuje precej šuma, smo vse pomene izbranih 35 besed pred začetkom označevanja ročno pregledali in po potrebi popravili. Po pregledu ima s 35 enobesednimi pomeni najvišjo stopnjo večpomenskosti beseda *vrsta*, najmanj, po štiri pomene, pa imajo samostalniki *misel*, *oči*, *postelja* in *roka*.

Za semantično označevanje smo za vsako izbrano besedo iz vsake knjige izluščili po pet naključnih stavkov dolžine 5–50 besed, v katerih se pojavlja, se pravi, da bomo za vsako besedo zbrali 25 oznak, kar skupaj pomeni 975 semantično označenih angleških in 875 slovenskih stavkov, ki se med seboj niso prekrivali. Pri tem je potrebno poudariti, da bi v idealnem primeru pri označevanju slovensko-angleškega dela potrebovali vzporedni korpus, v katerem so slovenska besedila izvirniki, angleška pa prevodi. Ker je slovensko-angleški del SPOOK-a zaenkrat še v razvoju, je precej manjši in trenutno vsebuje zgolj znanstvene prispevke s področja jezikoslovja, menimo, da za našo raziskavo ni primeren, zato smo se namesto njega odločili označiti kar del angleško-slovenskega korpusa, vendar v obrnjeni smeri. Čeprav se pri tem se zavedamo, da bomo izgubili nekaj kulturno-specifičnih pomenov, ki se pojavljajo v slovenskem leposlovju, menimo, da izbrani korpus vseeno omogoča zanimiv vpogled v zastopanost, distribucijo in prekrivnost pomenov med jezikoma, ki jo nameravamo v prihodnosti nadgraditi tudi z avtentičnimi slovenskimi literarnimi besedili in njihovimi prevodi v angleščino.

Ker je za to raziskavo ključno, da so oznake zelo zanesljive, smo se označevanja lotili ročno. Delo je potekalo tako, da smo glede na sobesedilo v korpusu in razlago ter semantične relacije v wordnetu najprej pripisali pomen (sinset ID) vsaki izbrani besedi v angleških izluščenih stavkih ter nato v slovenskem

prevodu tega stavka preverili, ali prevodna ustreznica, ki jo je izbral prevajalec, sodi v isti koncept ali ne ter še njej pripisali ustrezen sinset ID. Primer: *The next thing he knew, he was lying on his back on what felt like cushions, with a burning sensation in his ribs and right arm* (eng-30-05563770-n). // *Ko se je spet zavedel, je ležal na hrbtu na nečem mehkeem, morda na blazinah, v prsih in desni roki* (eng-30-05563770-n) *pa ga je žgalo*. V nasprotnem primeru smo zanj poiskali najustreznejši pojem in ji pripisali njegov ID. Primer: »*Remus!*« *Tonks cried as she staggered off the broom into Lupin's arms* (eng-30-05563770-n). // »*Remus!*« *je vzklila, ko se je opotekla z metle v Wulfov objem* (eng-30-00417397-n). Po zaključenem označevanju angleško-slovenskih stavkov smo postopek ponovili tudi v obratni smeri. Ko je bilo označevanje zaključeno, smo za izbrane besede in njihove prevodne ustreznice v obeh jezikovnih smereh skupaj imeli pripisanih 3.700 pomenov iz wordneta, kar je reprezentativno za izbrane besede in primerno za zanesljivo sklepanje o (ne)prekrivanju pojmov med jezikoma in (ne)utemeljenosti izdelave wordneta s prevajanjem.

3.2.2 Rezultati označevanja angleško-slovenskega korpusa

Vse izluščene pojavitve izbranih besed je bilo mogoče označiti z enim od pomenov v angleškem wordnetu, prav tako pa je bilo glede na slovenski wordnet sinset ID mogoče pripisati tudi vsem njihovim prevodnim ustreznicam. Število različnih uporabljenih pomenov iz stavkov, izluščenih iz posamezne knjige, je za vsa dela praktično isto (105–112). Od 372 možnih enobesednih pomenov izbranih besed v wordnetu jih je bilo med označevanjem uporabljenih zgolj 205 oz. 55 %, vendar so med posameznimi besedami velika odstopanja, saj so recimo za besedi *child* in *people* uporabljeni vsi enobesedni pomeni, ki so bili na voljo (4), najmanjši odstotek pomenov pa je bil uporabljen za besedo *head* (7 od 33). Najmanj večpomenska beseda v označenem korpusu je tako *door* (2), najbolj pa *place* in *side* (10).

Najpogostejše izbran pomen je bil pri vseh besedah tisti, ki je bil kot najpogostejši označen tudi v wordnetu. Naj od njih izpostavimo dva, ki sta s kontrastivnega stališča problematična, in sicer samostalnik *hair* v pomenu »*nitast izrastek na koži sesalcev*«, ki ga v slovenščini delimo na *las* in *dlaka*, medtem ko v wordnetu te členitve ni. Drugi pa je prislov *home*, definiran z »*v ali na poti proti prebivališču*«, ki bi ga v slovenščini glede na izbran prislov sklanjali različno, *domov* oz. *doma* in zato najverjetneje tudi ne bi sodila v isti sinset. Poleg enobesednih smo v korpusu pri 14 besedah od skupno 39 izbrali sinset, kjer označena beseda tvori del večbesedne zveze, in tako uporabili 22 različnih večbesednih literalov, ki so bili uporabljeni v 50 stavkih (npr. *arm rest*). Nadaljnjih 11 besed oz. 40 stavkov je bilo označenih z nesamostalniškim literalom (npr. *by heart*).

Ko smo za označene besede v vzporednih stavkih iskali prevodne ustreznice, smo našli 163 različnih. Največ, 8 različic, je bilo uporabljenih za besedo *home*, po ena sama pa za *child*, *foot* in *voice*. Najpogostejši prevod se za vse označevane besede ujema z najpogostejšim pomenom na angleški strani. Najpomembnejši del analize je pregled ujemanja izbranega sinset ID-ja za angleške besede in sinset-ID-ja, ki je bil pripisan njeni prevodni ustreznici v slovenščini. Le-to v povprečju znaša 75 %, odvisno od besede do besede pa niha med 36 % za *thing* in 96 % za *child*, *friend*, *hair* in *wall*. Ker pri raziskavi uporabljamo korpus literarnih besedil, prevodi, ki so pogosto svobodnejši, neujemanje ne pomeni nujno, da pojmi med jezikoma niso prekrivni, saj so že številni avtorji pri proučevanju prevodoslovnih pojavov (glej Baker 1993) ugotovili, da prevajalci tovrstnih besedil radi posegajo po parafrazah in izpustih oz. izvirnik prevajajo s splošnejšim ali bolj specifičnim izrazjem.

Zato smo primere, pri katerih ni ujemanja med pripisanim sinset ID-jem v obeh jezikih, razvrstili v pet kategorij (glej Tabelo 3, med katerimi po pogostosti močno izstopa parafraziranje, ki vključuje 161 primerov, ko je prevajalec iz slogovnih ali individualnih razlogov izvirno besedo nadomestil z drugačnimi jezikovnimi sredstvi, čeprav bi bil neposredni prevod, v katerem bi bil uporabljen literal iz istega sinseta kot v izvirniku, jezikovno povsem ustrezen. Primer: *I'd rather go to bed than get into this.* // *Rajši bi šla malo spat kot pa tole.* (čeprav bi bilo ustrezno tudi *šla v posteljo*).

Tabela 3: Pregled razlogov za neujemanje sinset ID-jev med angleščino in slovenščino.

Razlog za neujemanje	Št. primerov
parafraza	162
spec./generalizacija	44
idiomska raba	21
večbesedna zveza	7
konceptualna razlika	5
skupaj	239

44-krat smo naleteli na prevode, ki so pod- ali nadpomenke oz. mero- ali holonimi izvirnikov, na primer: *The sniper in the rafters transferred the laser dot to Stefan's head.* // *Ostrostrelec je laserski žarek nameril v Štefanovo čelo.* (čeprav bi bilo ustrezno tudi *v Štefanovo glavo*). Pri 21 primerih je bila razlog za neujemanje idiomatska raba nekega izraza v izvirniku, ki je bila prevedena razlagalno oz. nadomeščena s slovenskim idiomom s podobno funkcijo oz. obratno, na primer: *I had some part in that:*

*for I sat in a high place, and I strove with the Dark Tower; and the Shadow passed. // Nekaj **prstov** sem imel pri tem zraven jaz: kajti sedel sem na visokem kraju in se kos-sal s Temnim stolpom; in Senca je prešla. Sedemkrat je do neujemanja prišlo zaradi konceptov, ki so v enem jeziku izraženi z enobesednim leksemom, v drugem pa z večbesedno zvezo, kar je botrovalo izbiri različnih sinset ID-jev v enem in drugem jeziku, bodisi ker večbesedna zveza v enem od jezikov v wordnetu ne obstaja, ali pa se ji v primerjavi z enobesednim literalom spremeni besedna vrsta. Na primer: *Things have changed over the past **year**, explained Ditto, opening a bottle of beer. // Lani se je marsikaj spremenilo, je v pogovor skočil Čvek in si odprl steklenico piva.**

V tej raziskavi so najpomembnejše konceptualne razlike med jezikoma, saj te kažejo na kompleksne težave, do katerih prihaja zaradi prevajanja tujejezičnega leksikalnega vira in ohranjanja tujejezične pojmovne mreže. Zanimivo je, da je tovrstnega razhajanja v označenem korpusu zelo malo, vsa pa izhajajo iz kulturnih razlik, kot so sistemi merskih enot, šolski sistem, politična ureditev in podobno. Primer: *You know it only rises about two **feet** off the ground but he nearly killed the cat and he smashed a horrible vase Petunia sent me for Christmas (no complaints there). // Kot veš, se metla dvigne največ pol **metra** visoko, a skoraj bi ubil mačka in razbil je grozljivo vazo, ki mi jo je Petunija poslala za božič.*

3.2.3 Rezultati označevanja slovensko-angleškega korpusa

Čeprav nam je tudi v slovensko-angleški smeri uspelo označiti vse izluščene besede, je bila kljub predhodnemu pregledu sloWNeta izbira najustrežnejšega pomena za slovenske večpomenske iztočnice težja, ker so se ponavadi pojavljali pomeni, ki jih na prvi pogled ni bilo lahko ločiti. To nakazuje na slabšo razmejitev pomenov v wordnetu, ki jo bo v prihodnje potrebno podrobno proučiti in odpraviti. Tudi v tej smeri je število uporabljenih pomenov po posameznih knjigah podobno (88–101). Število možnih enobesednih pomenov v wordnetu je nekoliko nižje kot za angleški del korpusa (262), vendar je delež uporabljenih za označevanje nekoliko večji (58 %). Večje je tudi nihanje v deležu uporabljenih pomenov po posameznih besedah, saj se je samostalnik *glava* pojavil v samo 2 od 13 pomenov iz sloWNeta, *miza* pa v vseh svojih 4 pomenih. Po popraviljanju napak v sloWNetu smo ohranili samo en pomen za samostalnik *postelja*, zato je monosemna tudi v označenem korpusu, s 24 pomeni pa je najbolj večpomenska beseda *vrsta*.

Čeprav v sloWNetu pomeni zaenkrat še niso razvrščeni po pogostosti, so pomeni, ki smo jih najpogosteje uporabili za označevanje korpusa, pričakovani. Pri 12 od 35 besed smo poleg enobesednih uporabili 25 večbesednih literalov (npr. *rojstni dan*), pri petih pa še sedem nesamostalniških (npr. *v redu*), kar je primerljivo z angleškim delom korpusa. Označene besede imajo v korpusu 125 različnih

angleških ustreznice, ena sama je uporabljena za *življenje*, kar devet pa za *vrsto*. Tudi v tej jezikovni kombinaciji se najpogostejše angleške ustreznice vselej ujema-jo z vsebino najpogostejše izbranega sinseta na slovenski strani, pomenov besed, ki v označenem korpusu niso izpričani, pa iz sloWNeta nismo izbrisali, saj ocenjujemo, da je za sprejemanje takšne odločitve korpus premalo obsežen.

Primerjava pomenov, izbranih za iztočnico in njeno prevodno ustreznico, pokaže 78-odstotno povprečno ujemanje, kar je kljub večjim težavam pri označevanju več kot v angleško-slovenski smeri, nihanje med najslabšim in najboljšim ujemanjem pa je enako, od 36 % za besedo *roka* do 96 % za samostalnike *noga*, *obraz* in *vrata*. Analiza primerov, kjer med iztočnico in prevodom prihaja do razhajanja v pripisanem pomenu, v tej smeri pokaže manj parafraz (84, npr. *življenje ali smrt* – *live or die*) in več primerov specializacije oz. generalizacije (84, npr. *prst* – *toe*) kot v obratni smeri, rešitev z večbesednimi zvezami je več (21, npr. *parkirni prostor*), idiomatskih izrazov pa manj (5, npr. *čez glavo*). Zelo smo bili presenečeni nad ugotovitvijo, da v tej smeri med skoraj 900 označenimi stavki nismo naleteli na noben problem, ki bi ga povzročale konceptualne razlike med jezikoma (glej Tabela 4). Poleg drugih razlogov je to nedvomno tudi posledica dejstva, da je bil označen korpus slovenskih prevodov angleških besedil, kjer dejansko ne konceptualiziramo 'slovenskega' sveta, temveč opazujemo prevod konceptualizacije 'angleškega' sveta in se torej težave pričakovano – kot je bilo tudi opaženo zgoraj – pojavljajo v obratni smeri. Zato se v prihodnje nameravamo podrobneje posvetiti tudi proučevanju težav s semantičnim označevanjem slovenskih izvirnih besedil.

Tabela 4: Pregled razlogov za neujezanje sinsetov med slovenščino in angleščino.

Razlog za neujezanje	Št. primerov
parafraza	84
spec./generalizacija	84
večbesedna zveza	21
idiomatska raba	5
konceptualna razlika	0
skupaj	194

4 ZAKLJUČEK

V prispevku smo predstavili uporabo paralelnih wordnetov za semantično označevanje vzporednega korpusa literarnih besedil, ki je bil izluščen iz prevodoslovnega korpusa SPOOK. Primerjava leksikalnega inventarja angleškega in slovenskega

wordneta s korpusom je pokazala, da angleški wordnet dosega odlično pokritost. To niti ni tako zelo presenetljivo, saj ga razvijajo že trideset let in je trenutno najobsežnejši semantični leksikon na svetu. Stanje je bistveno slabše v sloWNetu, ki se je v prvih fazah izgradnje posvečal predvsem najpogostejšemu besedišču, zato je razumljivo, da je manj pogosto besedišče slabše zastopano. Zapolnjevanje identificiranih vrzeli je nujno, saj ima vir, ki slabo pokriva besedišče v korpusu, močno omejeno uporabno vrednost za praktično rabo.

Bolj zanimive izsledke daje semantično označevanje korpusa in primerjava ujemanja pomenov, ki so bili uporabljeni v obeh jezikih. Ob zasnovi eksperimenta smo pričakovali, da bo neujemanja zaradi konceptualnih, jezikovnih in kulturnih razlik med jezikoma bistveno več, kot se je z analizo rezultatov pokazalo, saj je bilo v angleško-slovenski smeri takšnih primerov zgolj pet, pri čemer je skupno število vseh primerov, kjer je prihajalo do razhajanj, 239, število vseh označenih stavkov pa 975. V slovensko-angleški pa nanje, kljub nasprotnim pričakovanjem, nismo naleteli niti v enem od 875 označenih primerov. Stopnjo ujemanja bi bilo v prihodnosti nujno potrebno preveriti tudi na korpusu avtentičnih slovenskih besedil z angleškimi prevodi, kjer pričakujemo večja odstopanja, vendar izkušnje, pridobljene v pričujoči raziskavi, ne kažejo bistvenih konceptualnih razlik, ki bi izdelavo slovenskega wordneta s prevzemanjem semantičnega inventarja iz Princeton WordNeta postavljala pod vprašaj, kar je za nadaljnji razvoj vira zelo spodbudno.

Poleg vpogleda v leksiko-semantični inventar v angleškem in slovenskem wordnetu iz ptičje perspektive in teoretičnih implikacij, ki izhajajo iz analize semantično označenega korpusa, ima opravljena raziskava tudi povsem oprijemljiv rezultat, ki se kaže v obliki prvega vzporednega korpusa za slovenščino, označenega na semantični ravni, ki bo omogočal najrazličnejše leksikološke in komparativne študije, uporaben pa bo tudi za večjezične jezikovnotehnološke aplikacije. Označeni del korpusa je za raziskovalne namene dostopen na <http://nl.ijs.si/slownet/>.

Čeprav se opravljena raziskava ukvarja z eno najpomembnejših posledic prevzemanja tujejezičnega vira, t.j. testiranjem nabora in distribucije pomenov slovenskih besed glede na jezikovno realnost, izpričano v korpusu, v njej nismo preverjali, v kolikšni meri zasnova semantičnega leksikona na obstoječem viru vpliva tudi na strukturo dobljene semantične mreže in na katerih mestih bi bilo zaradi konceptualizacijskih razlik ter jezikovnih posebnosti med angleščino in slovenščino potrebno omogočiti odstopanja od nje. Diagnostični testi, ki jih v sorodnih raziskavah uporabljajo za potrjevanje semantičnih relacij med dvema pojmom, so namreč zanesljivi le na velikih količinah podatkov, bistveno večjih od označenega korpusa, čemur se nameravamo posvetiti v nadaljnjem raziskovalnem delu.

Bibliografija

- Agirre, Eneko in Aitor Soroa, 2009: Personalizing PageRank for Word Sense Disambiguation. *Zbornik 12. mednarodne konference European Chapter of the Association for Computational Linguistics*.
- Artale, Alessandro, Bernardo Magnini in Carlo Strapparava, 1997: WordNet for Italian and Its use for Lexical Discrimination. *Zbornik 5. konference Italian Association for Artificial Intelligence*, Rome, Italy.
- Atkins, Sue, 1991: Building a lexicon: The contribution of lexicography. *International Journal of Lexicography*, 14 (3). 167–191.
- Baker, Mona, 1993: Corpus Linguistics and Translation Studies: Implications and Applications. Mona Baker idr. (ur.): *Text and Technology: In Honour of John Sinclair*, Amsterdam, Philadelphia, John Benjamins. 233–250.
- Bentivogli, Luisa, Pamela Forner in Emanuele Pianta, 2004: Evaluating cross-language annotation transfer in the MultiSemCor corpus. *Zbornik 20. mednarodne konference Computational Linguistics*.
- Burnard, Lou in Guy Aston, 1998: *The BNC Handbook: Exploring the British National Corpus with SARA*. Edinburgh: Edinburgh University Press.
- Erjavec, Tomaž, 2012: Vzporedni korpus SPOOK: označevanje, zapis in iskanje. Vintar, Špela (ur.): *Slovenski prevodi skozi korpusno prizmo*. Ljubljana: Znanstvena založba Filozofske fakultete.
- Fellbaum, Christine, 1998: *WordNet: An Electronic Lexical Database*. Cambridge: MIT Press.
- Ferraresi, Adriano, Eros Zanchetta, Marco Baroni in Silvia Bernardini, 2008: Introducing and evaluating ukWaC, a very large Web-derived corpus of English. *Zbornik 4. delavnice Web as Corpus - Can we beat Google?* Marrakech, Morocco. 47–54.
- Fišer, Darja in Benoît Sagot, 2008: Combining Multiple Resources to Build Reliable Wordnets. *Zbornik 11. mednarodne konference Text, Speech and Dialogue Conference*, Brno, Republika Češka.
- Fišer, Darja in Jernej Novak, 2011: Visualizing sloWNet. *Zbornik 2. mednarodne konference Electronic lexicography in the 21st century: new applications for new user*. Bled, Slovenija.
- Fišer, Darja, 2010: Semantično označevanje korpusov. Vintar, Špela (ur.): *Slovenske korpusne raziskave*. Ljubljana: Znanstvena založba Filozofske fakultete. 110–130.
- Fišer, Darja, 2010: Pristopi za avtomatizirano gradnjo semantičnih zbirk. Ledinek, N., Žagar Karer M. in Marjeta Humar (ur.): *Terminologija in sodobna terminografija*. Ljubljana: Založba ZRC, ZRC SAZU. 357–370.
- Hanks, Patrick, 2000: Do word meanings exist? *Computers and the Humanities*, 34 (1–2).

- Kilgariff, Adam, 1997: I don't believe in word senses. *Computers and the Humanities*, 31 (2), 91–113.
- Lakoff, George, 1987: *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago: University of Chicago Press.
- Miller, George A., Martin Chodorow, Shari Landes, Claudia Leacock in Robert G. Thomas, 1994: Using a semantic concordance for sense identification. *Zbornik delavnice Human Language Technology*.
- Pedersen, Ted in Varada Kolhatkar, 2009: WordNet::SenseRelate::AllWords - A broad coverage word sense tagger that maximizes semantic relatedness. *Zbornik mednarodne konference North American Chapter of the Association for Computational Linguistics - Human Language Technologies*. 17–20.
- Sagot, Benoît in Darja Fišer, 2011: Extending wordnets by learning from multiple resources. *Zbornik 5. mednarodne konference Language Technology Conference 2011*, Poznan, Poland.
- Sagot, Benoît in Darja Fišer, 2012: Cleaning noisy wordnets. *Zbornik 8. mednarodne konference Language Resources and Evaluation*, Istanbul, Turkey.
- Tufiş, D., Cristea, D. in Sofia Stamou, 2004: BalkaNet: Aims, Methods, Results and Perspectives. A General Overview. *Romanian Journal of Information Science and Technology Special Issue*, 7 (1–2), 9–43.
- Vintar, Špela, 2009: Slovenski prevodoslovni korpus. Stabej, Marko (ur.): *Infrastruktura slovenščine in slovenistike*, Ljubljana: Znanstvena založba Filozofske fakultete: 385–391
- Vossen, Piek, 1998: *EuroWordNet: A multilingual database with lexical semantic networks*. Dordrecht: Kluwer Academic Press.

Ptičji pogled na leksikalno ustvarjalnost v slovenskih prevodih

Špela Vintar

Oddelek za prevajalstvo, Filozofska fakulteta, Univerza v Ljubljani
spela.vintar@ff.uni-lj.si

Abstract

The paper addresses the topic of lexical creativity and attempts to implement corpus-based methods to, firstly, identify potentially creative lexemes, and secondly compare translations into Slovene from different source languages with texts originally written in Slovene. The primary resource for our work is the Spook corpus of translated and original contemporary literary texts in Slovene. We attempt to capture lexical innovations via three methods: by looking into the hapax legomena, especially those that do not occur in the large general language corpus Gigafida, by extracting words that occur in only one of the books, and finally by analysing selected productive derivational morphemes in Slovene, specifically *-ovje*, *-anje*, *-išče*, *raz-*, *pre-* and *iz-*. Our quantitative results imply that translators are at least as creative as authors in coining new words or using unexpected word forms, whereby it seems that the English-Slovene language pair contains the largest number of novel lexical items. The analysis of text-specific wordlists reveals the special lexical properties of each single book, including specialized terminology, slang and dialect vocabulary as well as author- or translator-specific neologisms, borrowings and coinings. While these findings cannot be generalized in terms of a prevailing translation strategy, results are encouraging because they show that - at least in our corpus - translators know how to be bold in their lexical choices and do not appear to be inferior to authors in their ability of creating new words.

Keywords: lexical creativity, corpus-based translation studies, hapax legomena, translation corpus

Ključne besede: leksikalna ustvarjalnost, korpusno prevodoslovje, enopojavnice, prevodoslovni korpus

1 UVOD

Ustvarjalnost je inherentna lastnost jezika, saj je vsak govorec jezika sposoben dan za dnem proizvajati še nikoli prej slišane izjave, sprejemniki sporočila pa so jih sposobni razumeti. Kljub temu pod pravo jezikovno ustvarjalnostjo navadno razumemo več kot to; za ustvarjalne označimo pisce ali govorce, ki jezik uporabljajo na izvirni, nepričakovan, neustaljen način. Izvirnost se lahko kaže na vseh ravneh, od besedilne prek skladenjske, semantične, leksikalne vse to fonološke ravni, saj je prav vsak element jezika mogoče uporabiti na nekonvencionalen način. Izvirna raba jezika je osnovni gradnik sloga, skozenj pa pisci in govorniki svojim sporočilom dodajajo osebno noto. Najvišjo stopnjo ustvarjalnosti pričakujemo v leposlovnih besedilih, kjer se ne le s samo propozicijo, temveč s prepletanjem spoštovanja in nespoštovanja jezikovnih konvencij, tke umetniška dimenzija.

Motivacija za pričujočo raziskavo na eni strani izhaja iz samega vira – korpus Spook¹ je bil namreč pred kratkim dokončan in je prvi korpusni vir, ki je zgrajen z namenom opazovanja lastnosti prevodov v primerjavi z izvirniki v slovenskem jeziku, poleg tega korpus vsebuje sodobna leposlovna besedila, ki so za opisovanje prevodnih pojavov še posebej zanimiva. V ozadju želje, da bi s korpusno metodologijo analizirali ustvarjalnost, je cela vrsta tujih prevodoslovnih raziskav, ki so predvsem s preučevanjem tim. prevodnih univerzalij ugotavljale, da se prevajalci pogosto nagibajo k normalizaciji, standardizaciji in eksplicitaciji, kar posredno nakazuje manjšo drznost in ustvarjalnost. Tako smo v naši raziskavi vzeli pod drobnogled besedišče in skušali preveriti, ali se pri izvirnih, svežih ali neustaljenih besedah prevedena besedila bisveno razlikujejo od izvirnih.

Prispevek v drugem razdelku opiše nekaj sorodnih raziskav s tega področja, v tretjem opiše uporabljene vire in metodološki pristop, v četrtem razdelku predstavimo rezultate in o njih razpravljamo, nato pa razpravo sklenemo s končnimi ugotovitvami in zamislimi o nadaljnjih študijah.

2 POGLEDI NA USTVARJALNOST

Z definicijo in analizo jezikovne ustvarjalnosti se je ukvarjala množica avtorjev z najrazličnejših vidikov. Ena vplivnejših študij, ki se ukvarja z ustvarjalnostjo pri literarni produkciji in recepciji, je knjiga *Singularity of Literature* Dereka Attridgea (2004). V njej avtor najprej razpravlja o mejah med ustvarjalnostjo, izvirnostjo in inovativnostjo (*creativity, originality, innovation*), nato pa literarno ustvarjalnost umesti v prostor kulturnih norm, kjer je element presenečenja in

¹ Korpus je nastal v okviru projekta Slovensko prevodoslovje – viri in raziskave, ki je potekal med leti 2009 in 2012; <http://lojze.lugos.si/spook>.

neskladja z ustaljenim – in sicer na strani avtorja in bralca – ključni sestavni del ustvarjalnosti.

Pri nas se ob obsežni komparativistični literaturi o stilističnih značilnostih posameznih avtorjev več jezikoslovcev ukvarja z jezikovno ustvarjalnostjo na leksikalni ravni. Tako Bokal (2003) opisuje jezikovne spremembe z vidika novih besed, Voršič (2009) analizira novotvorjenke v oglasnih sporočilih, Logar Berginc (2006) se podrobneje posveti stilno zaznamovanim novotvorjenkam, cela vrsta študij pa se posveča tvorjenju izrazja na različnih strokovnih področjih, vendar teh tu ne navajamo, saj na strokovnih področjih nove besede tipično nastajajo s poimenovalno motivacijo.

Manj, čeprav še vedno veliko literature najdemo o vprašanju jezikovne ustvarjalnosti pri prevajanju, kjer se razprave v glavnem odvijajo o načinih prenosa avtorjeve, v njegov osebni slog in kulturno ozadje vpete, ustvarjalnosti v ciljno besedilo, pa tudi o vlogi prevajalca, ki lahko zavzame celoten spekter (ne)vidnosti od povsem pritajenega in podrejenega avtorskemu besedilu (Venuti 1995) do vsemogočnega ustvarjalca nove umetnine (Pattison 2006).

Če je že sama definicija ustvarjalnosti izrazito subjektivna, ni presenetljivo, da je vsakršni poskus formalizacije ustvarjalnosti z namenom uporabe korpusnojezikoslovnih pristopov že vnaprej obsojen na očitek neadekvatnosti, v veliki meri upravičen. A čeprav se pri mnogih jezikovnotehnoloških metodah izkaže, da zmorejo zaobjeti le delček kompleksnosti jezikovne komunikacije, to še ne pomeni, da je zato njihova uporaba nesmiselna. Sledeč temu prepričanju se tukajšnja raziskava posveča jezikovni ustvarjalnosti na leksikalni ravni, ki je seveda le ena od številnih ravni, na katerih lahko avtorji in prevajalci odstopajo od uсталjenega. Z uporabo korpusnojezikoslovnih metod želimo samodejno izluščiti tiste leksikalne elemente, ki potencialno odražajo ustvarjalnost, z njihovim kvantitativnim ovrednotenjem ter ročno analizo pa želimo odgovoriti na vprašanje, kaj se z leksikalno ustvarjalnimi enotami zgodi pri prevodu in ali so pri uporabi neustaljenih in nepričakovanih leksemov bolj pogumni avtorji ali prevajalci.

2.1 Ustvarjalnost in prevajanje

V enem od korpusnomotiviranih poskusov, da bi pojem ustvarjalnosti prijeli za roge, je raziskovalka (Jordanous 2010) zbrala obsežno bibliografijo več kot šestdesetih znanstvenih člankov, ki se na tak ali drugačen način ukvarjajo s pojmom ustvarjalnosti v različnih strokah, nato pa iz njih izluščila ključne besede. Izkaže se, da je daleč najpogostejša beseda, kadar razpravljamo o ustvarjalnosti, »*novo*«. In res, že Orwell v svojem eseju *Politics and the English Languages* leta 1946 (po

Veale in Hao 2009) opozarja na nevarnost jezikovne lenobe, saj ta neizbežno vodi tudi k miselni lenobi. Tako se je po Orwellovem prepričanju pri pisanju treba izogibati predvidljivim frazam in kolokacijam ter si izmišljevati nove.

Carter in McCarthy (2004), ki sta v številnih člankih opisala svoje ugotovitve o jezikovni ustvarjalnosti v govorjenem jeziku, se pri definiciji ustvarjalnosti naslanjata po eni strani na ustaljeni nabor stilističnih figur, kot so metafora, primera, metonimija itd., po drugi pa ugotavljata, da je »ustvarjalnost v svoji učinkovitosti vselej odvisna od namenov in interpretacij udeležencev [komunikacije]«. ² Nekoliko kasneje ugotavljata še, da večina definicij jezikovne ustvarjalnosti implicitno predvideva spremembo, kar lahko razumemo kot premik od obstoječega, znanega, predvidljivega k novemu, inovativnemu.

Kako je mogoče formalno izraziti ustvarjalnost in jo modelirati z računalniškimi pristopi? Kot v svojih mnogih publikacijah dokazuje Veale s soavtorji (2006, 2008), je nekatere stilistične figure mogoče zaobjeti že z regularnimi izrazi, denimo primero z (angleškim) vzorcem *as ADJ as NP*, s čimer je že s pomočjo Googla moč najti ustaljene (*as strong as a bull*) in manj ustaljene primere (*as lost as Paris Hilton in a library*).

Analiza ustvarjalnosti postane še kompleksnejša, če jo opazujemo v prevodni relaciji. Kot že v uvodnem poglavju ugotavlja Venuti (1995), je v anglosaksonski tradiciji prevajanja močno zakoreninjeno prepričanje, da mora biti prevod čim manj opazen oziroma čim bolj prosojen. To med drugim pomeni, da prevajalčev jezik na nobeni ravni ne odstopa od splošne jezikovne norme, še bolj konkretno na ravni besedišča ne vsebuje slengovskih, pidžin ali narečnih izrazov ter je v vseh pogledih tekoč; z drugimi besedami se od prevajalca pričakuje, da besedilo normalizira.

A fluent translation is immediately recognizable and intelligible, "familiarised," domesticated, not "disconcerting[ly]" foreign, capable of giving the reader unobstructed "access to great thoughts," to what is "present in the original." Under the regime of fluent translating, the translator works to make his or her work "invisible," producing the illusory effect of transparency that simultaneously masks its status as an illusion: the translated text seems "natural," i.e., not translated.

(Venuti 1995: 5).

Prevodoslovne teorije so doslej ponudile številne modele klasifikacij prevajalskih strategij (Ivir 1987: 37, Newmark 1988, Schjoldager 2008: 92), s katerimi lahko pojasnujemo ravnanje prevajalcev ob srečanju z ustvarjalnostjo. Mednje sodijo dobesedni prevod, opisni prevod, nadomeščanje (s kulturno ustreznico), izpust, dodajanje (eksplicitacija) in prevzemanje; nekateri avtorji ponujajo podrobnejše taksonomije prevodnih mikrostrategij. Izbor klasifikacije pri našem ukvarjanju z

² Prevedla avtorica prispevka.

ustvarjalnostjo niti ni tako pomemben, saj sama identifikacija mikrostrategije še ne pove prav dosti o prevajalčevi ustvarjalnosti. Kljub temu skuša denimo Holst (2010) v svoji raziskavi prevajalske mikrostrategije, kot jih definira Anna Schjoldager (2008), razdeliti na bolj in manj ustvarjalne. Povsem nobene ustvarjalnosti po njenem ne kažejo dobesedni prevodi, kalki in prevzeti prevodi, medtem ko se pri nadomeščanju, permutaciji, priredbi, parafrazi, eksplicitaciji itd. lahko kaže prevajalčeva ustvarjalnost. S tako postavljenim modelom avtorica kvantitativno analizira izbrana besedila in »meri« stopnjo ustvarjalnosti.

Verjetno najbolj znana na korpusu temelječa raziskava ustvarjalnosti je delo Dorothy Kenny (2001), ki je pogled usmerila predvsem v leksikalne značilnosti besedil in z mnogimi svojimi metodami neposredno navdihnila tudi pričujoči prispevek. Kenny uporablja nemško-angleški korpus literarnih besedil in opazuje predvsem enopojavnice ter redke kolokacije, med katerimi pogosto najdemo ustvarjalnost na ravni zapisa, besedotvorja, nepričakovanih pomenskih asociacij, slikovitih primer in podobno. Vse ustvarjalne leksikalne enote, ki jih identificira v izvirniku, avtorica nato v analizi prevodne rešitve razdeli na dve osnovni kategoriji: tiste, ki jih prevajalec normalizira ali grobo rečeno »povozi« z uporabo ustaljene, ne-nove leksikalne enote, in tiste, ki jih prevajalec v prevodu kompenzira; pri tem Kenny izhaja iz predpostavke, da je za ohranitev ustvarjalnega leksema v prevodu vselej potrebna kompenzacijska strategija, saj neposreden prevod ni možen. Njene kvantitativne ugotovitve večinoma kažejo na močno prevlado normalizacijskih prevodnih rešitev, ki leksikalno bogastvo izvirnika precej oklestijo, čeprav po drugi strani Kenny priznava tudi, da vseh primerov kompenzacije ni enostavno odkriti, saj prevajalci »povoženo« ustvarjalnost morda nadomestijo v drugi povedi ali sploh na drugi besedilni ravni.

3 OPIS VIROV IN PRISTOPA

Tako kot druge raziskave v tej monografiji tudi ta kot glavni vir uporablja korpus Spook, in sicer vse njegove dele. Ker je korpus s tehničnega vidika podrobneje opisan v prispevku Tomaža Erjavca v tej monografiji, tu splošnih števil ne bomo navajali, o njegovi vsebinski sestavi, ki še kako vpliva na rezultate naše raziskave, pa bomo spregovorili v okviru sklepne razprave. Zaenkrat poudarimo le, da vsak od štirih vzporednih delov korpusa vsebuje sodobna leposlovna dela v izvirniku in prevodu v slovenščino, primerljivi slovenski korpus pa zajema sodobna leposlovna dela, izvirno pisana v slovenščini. Jezikovne ustvarjalnosti torej lahko na vseh ravneh pričakujemo veliko, naša raziskava pa je usmerjena v leksikalni inventar besedil.

Izhodiščno vprašanje raziskave je, ali so prevajalci pri izboru besedišča bolj ali manj ustvarjalni kot avtorji besedil, pri čemer ustvarjalnost na besedni ravni razumemo predvsem kot rabo redkih besed ter inovativno kovanje novotvorjenk iz

domačega ali sposojenega morfološkega gradiva. Metodološko je raziskava preprosta, saj v kvantitativnem delu analize opazujemo dve vrsti besed:

- a) Enopojavnice (*hapax legomena*). Iz vsakega dela korpusa izločimo besede, ki se pojavijo le enkrat, ta seznam pa primerjamo še z besednim seznamom splošnega korpusa Gigafida³ in ugotavljamo razmerje med številom neza-beleženih enopojavnic ter celotnim besednim inventarjem korpusa.
- b) Besede, ki se pojavljajo v le enem knjižnem delu. Besedni seznam posamezne knjige primerjamo z besednim seznamom celotnega korpusa in ohranimo le besede, ki se pojavljajo zgolj v enem delu. Število teh besed normaliziramo glede na velikost korpusa in primerjamo razmerja med seboj.

Oba opisana pristopa lahko razumemo tudi kot načina za ožanje preiskovanega prostora, saj tako pridobljeni sezname sicer vsebujejo veliko zanimivega besednega gradiva, vendar lahko šele ročni pregled in analiza sobesedila pokažeta, ali gre res za ustvarjalno rabo, ali je besedo zakrivil tiskarski skrat ali morda napaka med digitalizacijo in oblikoskladenjsko obdelavo. Zato se v drugem delu analize posvetimo pregledu besednih seznamov in raziskavi izbranih produktivnih pripon.

Medtem ko smo za kvantitativni del izdelave in sejanja besednih seznamov uporabili lastne programske skripte v perlu, smo se pri ročnem pregledu izbranih pojavov in sobesedila naslonili na korpusni iskalnik Cuwi⁴, ki s svojimi številnimi možnostmi tudi laičnim uporabnikom omogoča podroben vpogled v drobovje korpusa Spook.

Ker je pogostost posamezne besede v središču naše pozornosti, smo celotno raziskavo izvedli na lemah. To sicer pomeni, da so frekvenčni podatki bolj zanesljivi, vendar se moramo obenem zavedati, da so napake pri lematiziranih oblikah neznanjih, redkih ali prevzetih besed zelo pogoste. Pri ročni analizi smo zato najprej rekonstruirali besedno obliko, ki je dejansko izpričana v korpusu.

4 REZULTATI KVANTITATIVNE IN KVALITATIVNE ANALIZE LEKSIKALNE USTVARJALNOSTI

4.1 Enopojavnice in besede, ki se pojavljajo le v eni knjigi

Enopojavnice ali *hapax legomena* so besede, ki se v posameznem podkorpusu pojavijo le enkrat in so zaradi te svoje edinstvenosti že same po sebi zanimive. V našem primeru enopojavnice opazujemo na ravni lem. Med primeri enopojavnic najdemo tako zatipkane besede (*žužeika*, *utesnjujočc*), povsem pravilne slovenske

³ <http://www.gigafida.net>

⁴ <http://nl.ijs.si/cuwi/spook>

besede, ki so se po naključju pojavile le enkrat (*utopljenka, stožec, pridrveti*), lastna imena in izpeljanke iz njih (*Pittsburgh, potterjevsko, Malfojevi*), in inovativne tvorjenke, ki so plod avtorjeve ali prevajalčeve ustvarjalnosti (*lakotnost, velekača, kvautepazdaj, ksenogamija*).

V skladu z Zipfovimi zakonom razmerje med velikostjo korpusa in številom enopojavnic v njem ni linearno, zato smo vrednosti v Tabeli 1 normalizirali na 10.000 besed, vseeno pa rezultati v Tabeli nakazujejo višji delež enopojavnic v prevodih v primerjavi z izvirnimi besedili. Zadnja dva stolpca tabele navajata še število enopojavnic, ki niso izpričane v korpusu Gigafida, ter njihov normalizirani delež v primerjavi s skupnim številom različnic v podkorpusu.

Tabela 1: Število in delež enopojavnic v posameznih delih korpusa

	Št. različnic	Št. enopojavnic	Norm. delež	Št. enopojavnic, ki niso v Gigafidi	Norm. delež enopojavnic, ki niso v Gigafidi
angleško-slovenski	32.299	13.256	0,4104	4.050	0,125
francosko-slovenski	28.009	10.804	0,3857	2.347	0,083
nemško-slovenski	19.160	8.480	0,4425	1.607	0,084
italijansko-slovenski	19.621	7889	0,4020	1.742	0,088
slovenski izvirni	41.009	16.444	0,4009	3.657	0,089

Če skušamo te številke interpretirati, lahko najprej ugotovimo, da je delež enopojavnic nekoliko višji v prevodnem delu korpusa, saj znaša povprečje vseh štirih delov 0,41. Delež enopojavnic, ki se ne pojavljajo v korpusu Gigafida, je znatno višji v angleško-slovenskem delu korpusa, kar nakazuje na višjo stopnjo ustvarjalnosti. To vsaj na prvi pogled potrjuje tudi pregled seznama teh besed, saj v njem najdemo cele besedne družine novotvorjenk (*bunkljomrzec, bunkljevski, bunkljoliza, bunkljefob, bunkljati, bunkeljstvo; zvarkoslovec, zvarkarstvo, zvarkarjev*), poleg teh pa še številne grozde tvorjenk s produktivnimi predponami (*veleveriga, velevreden, veletrebuh, veleponosen, velekača, veleganten*).

V nemško-slovenskem spisku je novotvorjenk manj, zato pa precej besed, ki so se na seznamu znašle zaradi napačne lematizacije in napak pri skeniranju (*zaprisezi*

-> *zaprizeziti, zaljubljeno -> *zaljubljen, zaktičati -> *zaktičati, zadrževanjern -> *zadrževanjerna), v francosko-slovenskem pa poleg novotvorjenk izstopajo imena in citatne besede v francoščini.

Primerjava povprečnega deleža novih besed v prevodnih podkorpusih (ta znaša 0,095) in v izvirnem delu korpusa se nagiba v korist prevodov; tega podatka vseeno morda ne kaže razumeti kot dokaz, da so izvirniki manj ustvarjalni. Bolj verjetno tiči razlog v večjem številu (tujih) lastnih imen, ki se v prevodih ohranjajo ali le delno slovenijo in s tem povečajo obseg tistih besed, ki jih v korpusu Gigafida ne zasledimo, svoje pa doprinešajo tudi napake, ki so še posebej pogoste v nemško-slovenskem podkorpusu.

4.2 Besede, ki se pojavljajo le v enem delu

Nadalje smo želeli iz korpusa izluščiti besede, ki se pojavljajo zgolj v posamezni knjigi, saj smo predvidevali, da je v teh leksemih skrite največ besedne ustvarjalnosti. Besedni seznam vsake posamezne knjige smo primerjali s seznamom celotnega korpusa in obdržali le tiste, ki se pojavljajo le v posamezni knjigi, pri tem pa pogostost ni bila pomembna.

Tabela 2 kaže povprečno število takih besed za vsak podkorpus, čeprav je v resnici bolj relevanten podatek za vsako knjigo posebej. Ker so bila odstopanja med posameznimi deli precejšnja, podajamo še standardno deviacijo.

Tabela 2: Besede, ki se pojavljajo le v eni knjigi

	Skupni seštevek besed, ki se pojavljajo v samo enem delu	Povprečni delež glede na število različnic	Standardni odklon
angleško-slovenski	18.028	0,21	0,074
nemško-slovenski	13.644	0,24	0,134
francosko-slovenski	13.832	0,15	0,056
italijansko-slovenski	11.398	0,24	0,097
slovenski izvirni	20.473	0,12	0,048

Pregled teh spiskov ogromno pove o žanru in tematiki dela. Pri fantazijskih romanih so novoizumljene besede vezno tkivo, okrog katerega se plete zgodba, tako

najdemo denimo pri Gospodarju prstanov visoko na seznamu besede *vilin*, *kačje-ust*, *orthanec*, *entinja*, *polovnjak*, *mrkolesje*, pri Harryju Potterju pa *hestija*, *biritch*, *Dobrodella*, *frčamrežje*, *bunkljelogija*, *zvarkoslovec*. Pri bolj realističnih žanrih na vrhu spiskov najdemo predvsem osebna in zemljepisna imena ter poimenovanja drugih kulturno specifičnih reči, tako na primer pri Inšpektorju Morseju *Eriksson*, *morsov*, *wythamski*, *seckhamski*, *blenheimski*, *Kidlington*, *Glenfiddich*; pri Leyli *bajram*, *fatma*, *derviš*, *kuruš*, *efendi*; manj pa pravih inovativnih tvorjenk, kot so *plezalček*, *grivar*, *kvazivseznati*. Dela, bogata z dialogi, vsebujejo veliko nižjepogovornih, slengovskih ali narečnih besed, tako denimo pri Belih zobeh *štacuna*, *paki*, *kelnar*, *prefukati*, *sapralot*, *keš*, *lezba*; pri Ni me strah *pička*, *pofočkati*, *jebemomast*, *navaliti*, *štekati*. Sezname se močno razlikujejo tudi po besednovrstni sestavi, tako denimo v Austerlitzu po pogostosti močno izstopajo nekoliko starinski prislovi, npr. *docela*, *malodane*, *spričo*, *navzlic*, *zlagoma*, *prejkone*, *čeravno*.

Da bo slika popolnejša, navedimo še primere besed, ki so značilne za posamezna izvirna slovenska dela. Kot pri prevodih se tudi v izvirnikih skozi specifično besedišče zrcalita tematika in žanr, oboje pa se neločljivo prepleta z avtorskim slogom. Tako se v besednem seznamu dela Umri kaže vojaški besednjak (*desetar*, *intendantski*, *rezervist*, *postolje*, *gušterski*, *piketaš*) s številnimi inovativnimi tvorjenkami (*klobasariti*, *razpljuznen*, *bambusar*), v Potapljanju na dah potapljaško-navtični (*vrša*, *potapljačica*, *neopren*, *valobran*, *kašeta*, *sidrni*, *jeklenka*, *podhlajen*). Nezmotljiv avtorski pečat v divje slikovitih pridevnikih najdemo v Pogledu na Tycho Brahe (*pikljajoč*, *grbinast*, *srkljiv*, *razcufan*, *skrepenel*, *steklinast*, *prilizljiv*), svežo mešanico pogovornih in dialektalnih sestavin v Koledarju Evinih razpoloženj (*mesič*, *horsburger*, *holivudar*, *rugzak*, *fkuhati*), Patosih (*fakinsit*, *brader*, *tehace*, *psihič*, *bajdevej*, *šluk*), Seksu, ljubezni in tem (*eht*, *navlečen*, *skvot*, *fukiš*, *hardkor*) in Omejenem roku trajanja (*košta*, *pronk*, *mrtvica*, *zakraspa*, *kuzlica*, *špringel*).

Sodeč po številkah v Tabelah 1 in 2 je delež potencialno ustvarjalnih besed v prevodih nekoliko višji kot v izvirnikih, in to velja tako za enopojavnice kot za besede, ki se pojavljajo v le enem delu. Vendar je številke, kot vedno, treba jemati z določeno mero previdnosti. Ko namreč sezname pogledamo z budnim očesom, so med avtorsko specifičnimi izrazi v precejšnjem deležu lastnoimenske sestavine, te pa bi zagotovo označili za manj ustvarjalne od pravih novotvorjenk in slikovitih skovank. Za pravi vtis o slogu izbranega avtorja ali prevajalca je torej neizbežno treba pogledati posamezne primere s sobesedilom vred.

4.3 Analiza izbranih produktivnih pon

Če nekako intuitivno privzamemo, da je na leksikalni ravni pravo ustvarjalnost mogoče izživeti predvsem z besedotvorjem, si velja ogledati, ali prevajalci in

avtorji pri kovanju svežih besed uporabljajo enake pristope in v podobni meri. V ta namen smo analizirali pet produktivnih pon, in sicer smo iz spisikov za posamezno delo specifičnih besed izluščili vse tvorjenke z določeno pred- ali pripono in obdržali le tiste, ki jih ne najdemo v SSKJ.

4.3.1 *-ovje, -išče, -anje*

Pripona *-ovje* je srednje produktivna, v vseh spisikih za posamezno delo značilnih besed najdemo skupaj 63 različnic na *-ovje*, od tega 9 v izvirnih besedilih in 54 v prevodih. Semantika te pripone nakazuje določeno količino nečesa, predvsem v prostorskem ali površinskem smislu (*tramovje, mesovje, zidovje*). Med prevodi so le tri tvorjenke take, da jih ne najdemo v SSKJ (*dlakovje, hobitovje, popkovje*), prav tako so tri takšne v izvirnikih (*smetovje, smrklovje, hribčkovje*), kljub temu pa so še številne druge med njimi slikovite in domiselne, čeprav so zabeležene tudi v SSKJ (*meglovje, ritovje, steblovje*). Pomensko od pričakovane semantike nekoliko odstopa *hobitovje*, ki se nanaša na skupnost hobitov, in *smrklovje*, ki je rabljeno v pomenu najstniki.

Pripona *-išče* je nekoliko produktivnejša s 161 različnicami, od tega 34 v izvirnikih. Pomensko označuje kraj, kjer se nekaj nahaja ali dogaja; prav iz tega razloga se veže tako z glagoli kot s samostalniki (*izstrelišče, dirkališče, perišče; grobišče, šotišče, praprotišče*). V naših prevodnih besedilih najdemo pet tvorjenk, ki jih SSKJ ne pozna (*prekrcavališče, pretovarjalsišče, entišče, puščišče, divjišče*), v izvirnikih pa le eno (*alozijevišče*). Morda je zanimivo *divjišče*, predvsem ker je podstava pridevniška, vendar gre tu za tvorjenko pri prevajanju krajinskega imena *Wilderland*.

Od opazovanih pripon je najproduktivnejša *-anje*, saj najdemo kar 2.362 različnic, od tega 634 v izvirnikih. Pripona je namenjena tvorjenju glagolnikov in se veže izključno na nedovršne glagole, ki pa so v našem primeru pogosto novotvorjenke, pogovorne in narečne besede. Pri prevodih jih 27 ne najdemo v SSKJ: *piknikovanje, vagoniranje, sobrenčanje, poročevanje, pogovarjalopoljubljanje, njevanje, parliranje, postrganje, biljardiranje, izdejanjanje, metamorfanje, posmrčevanje, prdnikoliranje, zagrebanje, dozivanje, čvrkotanje, cmerikanje, bobotanje, pogrmevanje, mikroinjiciranje, neravnanje, driblanje, fafanje, čekiranje, samopohabljanje, napikovanje, zoprvanje*; pri izvirnikih pa je takih 38: *kravljanje, prcanje, rvanje, švindlanje, postojevanje, mežikljanje, skimavanje, fafljanje, odpo-tovanje, samoreklamiranje, korigiranje, jajcanje, fajfanje, ricmanje, čmokotanje, mikroskopiranje, dilaanje, skvotanje, avtofajfanje, flancanje, parazitiranje, škljocanje, spregledovanje, granatiranje, preoblačevanje, zdehanje, rentiranje, votkanje, tituliranje, bipanje, bavkanje, aplaudiranje, cigumiganje, direktorovanje, droben-tljanje, klopatanje, šklemfanje, zafrčkavanje*.

4.3.2 *raz-, pre-, iz-*

Glagolske predpone *raz-*, *pre-* in *iz-* so vse precej produktivne, za prvo najdemo 438 različnic, od tega 96 v izvirnikih, za drugo 603 različnic, od tega 153 v izvirnikih, in za tretjo 425 različnic, od tega 78 v izvirnikih. Med bolj izvirnimi v prevodih so *razklobasati*, *razkokodajsati*, *razginiti*, *razznavati*, *razturati*, *razpredalčkati*, *razkrstiti*; *pretleti*, *prelicitirati*, *prežarjati*, *prepetnajstiti*, *presukati*, *preženivati*, *prežuravati*; *izsanjati*, *izstokati*, *iz-mene-dati*, *izklokotati*, v izvirnikih pa *razprhnevati* in *razprsniti*; *premodriti*, *presrati*, *prešlatati*, *prepacati*, *preštrihati*, *prešvercati*, *presestavati*, *prekvasiti*, *premalati*, *prefarbati*, *prešaltati*; *izbehariti*, *izdurati*. Tu je morda edinkrat opazna razlika med prevodi in izvirniki glede pogovornih besed, saj jih je v izvirnikih pri glagolih s predpono *pre-* precej.

Seveda bi bilo zanimivo analizirati še druge pone, prav tako pregled spiskov brez sobesedila ne daje prave slike o ustvarjalni rabi. Kljub temu pa lahko rečemo, da je med zabeleženimi leksemi tako v izvirnikih kot v prevodih veliko resnično izvirnih novotvorjenk.

5 RAZPRAVA IN SKLEP

Med pregledovanjem spiskov enopojavnic, besed, ki se ne pojavijo v Gigafidi in besed, ki se pojavijo le v eni knjigi, se najprej z vso močjo zavemo, kako bogate so besedotvorne zmožnosti slovenščine in kako malo jih uporabljamo v vsakdanji jezikovni rabi. Pisana zmešnjava tujk in lastnih imen, slengizmov, zastarelih in ljudskih izrazov, strokovnih terminov ter pristnih svežih skovank nas naglo prepriča, da je v leposlovju jezik tako v izvirnikih kot v prevodih še kako živ in sočen, prevajalci pa si upajo prav tako drzno zaorati leksikalno ledino.

Pomembnih kvantitativnih razlik med izvirniki in prevodi tudi številke ne kažejo; glede deleža enopojavnic so v rahlem vodstvu prevodi, prav tako glede deleža besed, ki niso v Gigafidi, vendar to zlahka pojasnimo z večjim številom neslovenskih sestavin, predvsem imen in citatno prevzetih besed. Ker prevodna literatura predstavlja okno v tuje kulture, je ta rezultat pričakovan.

V obeh tabelah pa je bilo opaziti pomembno odstopanje angleško-slovenskega dela korpusa, tako glede deleža nenajdenih enopojavnic kot glede deleža za knjigo specifičnih besed. To nakazuje nekoliko večjo besedno ustvarjalnost pri prevajalcih iz angleškega jezika, kar je drzna hipoteza. Dejstvo je, da je med prevajalci iz angleščine tudi Branko Gradišnik, ki slovi po svoji ljubezni do kovanja besed, po drugi strani pa je morda korpus žanrsko nekoliko neuravnotežen in

zato neupravičeno izpostavlja prav ta jezikovni par. Fantazijski roman je namreč gotovo žanr, ki neobhodno zahteva izumljanje besed, in v angleško-slovenskem podkorpusu so kar trije fantazijski romani ter še en znanstvenofantastični roman. V drugih podkorpusih ta dva žanra skoraj nista zastopana.

Druga opazna razlika, ki ne izhaja iz števil, zato pa postane toliko bolj očitna pri ročnem pregledu spiskov, je prisotnost slovenskih slengizmov, pogovornih izrazov in narečnih besed v izvirnikih, medtem ko so prevodi v tem delu besedišča skromnejši. Tudi tu je razlog očiten – izvirniki so kulturno umeščeni v različne socialne in geografske miljeje v Sloveniji, zato tudi protagonisti uporabljajo lokalno besedišče, medtem ko prevodi predstavljajo bodisi tuja bodisi fantazijska kulturna okolja.

Pomemben segment, ki pričujoči raziskavi manjka, je kvantitativno podprta analiza ustvarjalnosti v tujih izvirnikih ter strategij transferja ustvarjalnih leksemov v ciljni jezik. Za to bi morali za vse štiri izvirne jezike pridobiti rojene govorce kot informante, ki bi presojali o ustvarjalni rabi, saj je za razumevanje pomenskih nians in jezikovnih inovacij nujno potrebna najvišja raven jezikovne kompetence.

Čeprav smo za ptičji pogled na ustvarjalnost v prevodih in izvirnikih uporabili razmeroma enostavne korpusne metode, so rezultati – sicer le za ozek segment avtorjev, prevajalcev in žanrov, ki jih zajema Spook – relevantni in spodbudni. Za razliko od sorodnih študij v tujini (prim. Kenny 2001) zaenkrat ni dokazov o poudarjeno normalizacijskih prevodnih strategijah, nasprotno – v prevodih je novega besedišča celo več kot v izvirnikih. A podobno kot pri vsaki korpusni raziskavi je za celovitejši vpogled v jezikovne pojave in prevodne strategije nujno zapustiti spiske in številke in se ukvarjati z besedili, saj je ustvarjalnost preširok in presubtilen pojem, da bi ga lahko opisovali na ravni posameznih leksemov.

Viri

- Attridge, Derek, 2004: *The Singularity of Literature*. London/New York: Routledge.
- Bokal, Ljudmila, 2003: Jezikovne spremembe s stališča novih besed, *Slovenski jezik – Slovene language studies* 4, 13–24.
- Carter, R. A. in McCarthy, M.J., 2004: Talking, creating: interactional language, creativity and context. *Applied Linguistics*, 25 (1): 62–88
- Holst, June L. F., 2010: *Creativity in translation*. BA Thesis, Department of Language and Business Communication, Aarhus University.
- Ivir, Vladimir, 1987: Procedures and strategies for the translation of culture. Tury, Gideon (ur.) *Translation Across Cultures*. New Delhi: Bahri Publications.

- Jordanous, Anna, 2010: *Defining Creativity: Finding Keywords for Creativity Using Corpus Linguistics Techniques. Proceedings of the First International Conference on Computational Creativity (ICCCX)*, Lizbona, Portugalska.
- Kenny, Dorothy: 2001. *Lexis and Creativity in Translation*. Manchester: St. Jerome.
- Logar Berginc, Nataša, 2004: Stilno zaznamovane novotvorjenke. Vidovič Muha, Ada (ur.): *Slovensko jezikoslovje danes*. Ljubljana : Slavistično društvo Slovenije, 2006.
- Newmark, P, 1988: *A Textbook of Translation*. Hemel Hempstead: Prentice Hall.
- Pattison, Ann, 2006: Painting with words. Loffredo, E. in Perteghella, M. (ur.) *Translation and Creativity: Perspectives on Creative Writing and Translation Studies*. London/New York: Continuum International Publishing Group.
- Schjoldager, Anne, 2008: *Understanding Translation*. Aarhus: Academica.
- Veale, Tony in Butnariu, Cristina, 2006: Exploring Linguistic Creativity via Predictive Lexicology. *Zbornik konference ECAI'06: Third Joint Workshop on Computational Creativity*. <<http://afflatus.ucd.ie/Papers/predictiveLex.pdf>>
- Veale, Tony in Hao, Yanfen, 2008: Generating and Understanding Creative Comparisons. *AAAI Spring Symposium: Creative Intelligent Systems 2008*: 123–127
- Venuti, Lawrence, 1995: *The Translator's Invisibility: A History of Translation*. London/New York: Routledge.
- Voršič, Ines, 2010: *Novotvorjenke v tiskanih in spletnih oglasih*. Diplomsko delo. Univerza v Mariboru: Filozofska fakulteta.

Pragmatični pomen in semantična prozodija v medjezični perspektivi: primer slovenščine in angleščine

Mojca Šorli

Trojina, zavod za uporabno slovenistiko
 mojca.sorli@trojina.si

Abstract

The present paper is focused on ways in which the pragmatic (functional) meaning that arises from the elements of the 'extended unit of meaning' and various contextual features comes into play in identifying appropriate translations. The possibility to analyse vast amounts of linguistic data has brought about changes both in methodology as well as in the ways we perceive certain language phenomena. Corpus analysis and new methods in the analysis of pragmatic components of meaning have yielded significant results in areas such as the treatment of semantic prosody. In the introductory part, we present the theoretical background on semantic prosody and the extended unit of meaning as defined by J. Sinclair. In the analysis part, we first examine from a monolingual point the Slovene paradigm for *postati rdeč* and the corresponding English paradigm *to become red*, and conclude by comparing the two and matching up lexical items according to their collocational, colligational and contextual patternings.

Keywords: corpus analysis; translation; unit of meaning; pragmatics; semantic prosody

1 UVOD – EMPIRIČNI PODATKI IN PRAGMATIČNI POMEN

Tudi za pragmatične vidike pomena je pomembno, da so primerno opisani na podlagi empiričnih podatkov o njihovem pojavljanju v realnih besedilih. Ključno za razumevanje pomenotvornega procesa v vsej njegovi celovitosti je opazovanje t. i. semantične prozodije, ki je integralni del razširjene enote pomena (*extended unit of meaning*). Pojem izhaja iz t. i. neofirthovske¹ jezikoslovno-leksikografske tradicije J. Sinclairja (Sinclair 1993, 1996, 2004; Louw 1993; Stubbs 1995; Bublitz 1996; Hunston in Francis 1999; Tognini-Bonelli 2001; Channell 2003; Hunston 2007; Partington 1998, 2004; Philip 2009), temelječe na analizi elektronskih korpusnih besedil, ki frazeološkosti v jeziku ne vidi kot posebnost, temveč kot osnovno načelo njegove organiziranosti, medtem ko je načelo t. i. proste izbire (*open choice*) sicer prisotno, toda ne tipično. Nekateri avtorji opisujejo pojav tudi kot kolokacijska, diskurzna ali pragmatična prozodija (npr. Stubbs 2001). Frazeološkost prinaša s seboj tudi deleksikaliziranost, kar pomeni, da je pomen besed močno odvisen od konteksta. Spoznanje, da je večina konvencionalnega jezika pravzaprav deleksikaliziranega (Sinclair 1991: 113),² spodbuja prevajalca, da dela predvsem s pragmatičnim, funkcijskim pomenom, kar pomeni, da postavlja pomen nad obliko besed in s tem pripomore k natančnejšemu prenosu vsebin (Philip 2003: vii).

1.1 Korpus

V predkorpusnem obdobju so bile možnosti, da bi sistematično raziskovali komponente pomena, ki presegajo raven besed, zelo omejene, saj so te vezane ne le na regularnosti v skladnji oz. skladenjskih vzorcih, tipično na ravni kolokacije, temveč tudi na regularnosti v širšem besedilnem okolju in situacijskem kontekstu, ki jim pravzaprav ni mogoče pripisati enotne kategorije. Poleg splošnih enojezičnih korpusov so za prevodno in leksikografsko prakso, pa tudi prevodoslovno teorijo, nepogrešljivi vzporedni ali primerljivi dvo- ali večjezični korpusi besedil (v slovenskem okolju je tak npr. večjezični korpus SPOOK).³ V korpusih je mogoče v trenutku dobiti vpogled v rabo tako tujega kot materne jezika, kar prevajalcu omogoča, da preveri svojo intuicijo v primerjavi z dejansko rabo.

¹ J. R. Firth je bil med prvimi, ki je poudarjal pomen leksike v opisnem jezikoslovju, po njem pa je prevzeto tudi poimenovanje *prozodija*, ki ga je sam uporabljal za raziskovanje analognih pojavov v fonologiji.

² Kot meni Sinclair: "[...] normal text is largely delexicalized, and appears to be formed by the exercise of the idiom principle, with occasional switching to the open-choice principle".

³ SPOOK: <http://lojze.lugos.si/spook/index.html> (Dostop 30. 9. 2012)

2 SEMANTIČNA PROZODIJA: VREDNOTENJE IN (PRAGMATIČNA) FUNKCIJA

Pomen, ki se udejanja v semantični prozodiji, je pragmatične narave, pomeni pa, nekoliko poenostavljeno, način, na katerega je mogoče navidez nevtralne besede razumeti tudi s pozitivnimi in negativnimi asociacijami, kot jih razvijejo ob pogostem sopojavljanju z drugimi besedami. Namen prispevka je med drugim pokazati, da gre še za mnogo več kot le za preprosto polarizacijo. Izraz *semantična prozodija* (dalje: SP) je doživel precejšen razvoj od časa, ko se je prvič pojavil, kar je povzročilo nekaj nejasnosti glede tega, kaj kateri avtor z njo pravzaprav misli. Kritični pogled na obravnavo oz. rabo pojma v okviru leksikalnega pomena je podal Whitsitt (2005). Medtem ko SP pogosto (vsebinsko) enačimo s semantično preferenco, nekatere novejšje študije (npr. Philip 2009) kažejo, da če pojem obravnavamo v širšem kontekstu, ob podrobni razčlenitvi sestavin razširjene enote pomena, ta vključuje pravzaprav dve ravni pomenske analize, pri čemer je ena besedno, druga pa funkcijsko, frazeološko ali kontekstualno pogojena. Prozodija je rezultat subtilnih pomenotvornih procesov in čeprav gre v osnovi za kolokacijski pojav, jo lahko ugotavljamo šele ob pregledu širšega sobesedja in razbiranju različnih kontekstualnih ključev, predvsem pa z analizo velikega števila zgledov dejanske rabe, saj do nje ne moremo dostopati introspektivno. SP je prisotna v naravnem jeziku tako, da jo dejanska jezikovna raba implicira, vendar le redko neposredno izrazi (Sinclair 1994). Pomensko zakritost SP izpostavlja tudi Hunston (2002: 61, 119, 141–142). Predpostavljamo lahko, da gre pri pojavu SP za težavnost zlasti pri enkodiranju, saj najbrž drži, da načeloma kot rojeni govorci pragmatični pomen obvladujemo nezavedno, vendar je na tem mestu poudarek na ozaveščanju procesov in razumevanju izbir, zlasti pri prevajanju. Kot ugotavlja Jurko (2012: 88), je tudi SP idiosinkratična in ne funkcionira vedno enako v primerjanih oz. kontrastiranih jezikih. Vendar pa kaže, da predstavlja največjo težavo prav možnost, da bomo podcenili njen pomen pri prevajanju. Tudi Gabrovšek (2012: 88) je mnenja, da sodobne izdaje pedagoških enojezičnih slovarjev praviloma kažejo večje zavedanje o tem pojavu – posebej v smislu nakazovanja »pozitivnih« ali »negativnih« vidikov leksikalnega pomena –, vendar je mogoče to področje leksikografskih opisov še nadgraditi, zlasti pa se ob tem zavedati pomena velikih, reprezentativnih korpusov besedil.

2.1 Udeležba SP glede na vrsto pomena

Philip (2009) ugotavlja, da igra SP vlogo predvsem pri tvorbi t. i. perifernih ali kontekstualno pogojenih pomenov (*non-salient meanings*). T. i. osrednji pomeni (*salient meanings*) so »kodirani pomeni, hitro dostopni zaradi svoje

konvencionalnosti, pogostosti, prepoznavnosti ali prototipičnosti» (Giora 2003: 10, prev. avt.), povezani pa so z besedami kot ortografskimi enotami. Gre torej za pomene, ki so uporabnikom jezika skupni in ki jih včasih imenujemo tudi »slovarski« pomeni. Tudi t. i. sekundarni pomeni, ki niso nujno v enaki meri skupni vsem uporabnikom v jezikovni skupnosti, so vezani na besedo ali izraz, tipično npr. konotacija ali vrednotenjski pomen (na razlike med SP in konotacijo sta na primer skušala opozoriti že Louw (2000) in Partington (1998)). V nasprotju s temi so omenjeni »podoznačeni« (*underspecified*) periferni pomeni, ki so intuiciji tipično nedostopni, določata pa jih kontekst in raba.

2.2 Razširjena enota pomena

V nadaljevanju predstavimo shematsko razgradnjo razširjene enote pomena, kot jo koncipira že Sinclair (1996), njeno sodobno in poglobljeno različico pa izriše Philip (2009) (za prevodoslovni pogled na enoto pomena oz. prevoda gl. npr. tudi Dam-Jensen in Zethsen (2008) in Tognini-Bonelli (2001: 135)). Izvorno štiripartitna razširjena enota pomena dobi v tej še peti element (asociacijo) in vsebuje skupaj: kolokacijo, koligacijo, semantično preferenco, asociacijo, in prozodijo. Da bi lažje razumeli razsežnosti kontekstualnega pomena, je znotraj tako zasnovane sheme pomembno ločiti med dekontekstualiziranim in kontekstualnim pomenom, čeprav gre za shematski prikaz, v katerem lahko potegnemo seveda zgolj shematsko ločnico med kolokacijo, koligacijo in semantično preferenco na eni ter tistimi komponentami, ki opredeljujejo funkcijo pomena, torej SP (po izvorni shemi tudi zunajjezikovno funkcijo), na drugi strani, in za katere menimo, da so prav tako pomembne pri iskanju ustreznega prevoda. Pri realizaciji SP je koligacija pomembna vsaj toliko kot kolokacija – prepozna prozodije v besedilu se pogosto začne prav z identifikacijo koligacijskih preferenc. Ker sta kolokacija in koligacija tako vsebinsko kot terminološko ustaljeni oz. splošno sprejeti, si bomo v nadaljevanju poleg SP od naštetih komponent podrobneje ogledali še semantično preferenco in asociacijo.

2.2.1 Semantična preferenca

Semantična preferenca je abstrahirana pomenska pripadnost kolokatorjev, s katerimi se beseda oz. leksikalna enota v besedilih najraje družijo in ki jim lahko zato pripišemo določen semantični tip. Takšni kolokatorji lahko evocirajo tudi sekundarne pomene, na primer vrednotenjske in asociativne. Semantično preferenco je mogoče še vedno prepoznavati v okviru besednega, dekontekstualiziranega pomena, zato Hoey (2005: 23) poudari potrebo po jasni razmejitvi med semantično preferenco in prozodijo, kot jo je predvidel že Sinclair.

2.2.2 Asociacija

Izraz »asociacija« je uvedel Hoey (2005: 16ff, 22) in sicer kot nekakšen nadomestni izraz za semantično preferenco. Po Philip (2009), ki namesto združevanja kategorij predlaga asociacijo kot dodatno, vmesno kategorijo (med preferenco in prozodijo),⁴ pa se asociacija od semantične preference razlikuje v tem, da je povezana s sekundarnim pomenom, zlasti vrednotenjem in konotacijo, od SP pa v tem, da ne opisuje funkcije pomena.

2.2.3 Semantična prozodija

Medtem ko se semantična preferenca ključno veže na kolokacijo, je SP v skoraj enaki meri odvisna od kolokacije in koligacije ter drugih oprimkov pragmatičnega pomena, ki se nahajajo v širšem situacijskem kontekstu – teh pogosto ni mogoče umestiti v katero od uveljavljenih slovničnih kategorij, saj v besedilu niso eksplicitno izraženi s konvencionalnimi izraznimi sredstvi. V pričujočem prispevku bomo analizirali pomenske enote v njihovi celovitosti, torej z vsemi navedenimi sestavnimi elementi, od katerih vsak na svoji ravni prispeva k izgradnji pragmatičnega pomena. Po teoriji Sinclairja (1996) se vsaka pomenska izbira začne prav na najbolj skrajnem robu razširjene enote pomena, kar vpliva na končno izbiro na ravni kolokacije oz. skladenjskega vzorca, ki naj bi bila tako pogojena z začetno izbiro SP oz. pragmatične funkcije. Obširnejši pregled obravnave in razvoja pojma SP v leksikalnem pomenslovju je podal Stewart (2010), v slovenskem prostoru pa Gabrovšek (2007), ki predlaga obravnavo pojava v sklopu treh pomenskih ravni, tj. posamičnih enobesednih leksemov (konotacija), večbesednih enot (SP) in širšega sobesedila (sintagmatski pomen). Vse tri vrste pomena zajame s krovnim izrazom 'asociativni pomen'. Za opis (enojezičnega) pristopa k pojavu v *Leksikalni bazi za slovenščino*⁵ gl. Šorli (2011). Sintezo različnih interpretacij SP je opravila Philip (2009), ki je dogradila koncept na teoretskih izhodiščih zlasti Sinclairja, Louwa, Stubbsa in Tognini-Bonellijeve. V teh okvirih smo opravili tudi podano dvojezično analizo in sicer z neposredno primerjavo podatkov v korpusu FidaPLUS⁶ s podatki na angleški strani.

2.3 Medjezikovne študije semantične prozodije

Raziskovanje SP na dvo- ali večjezični ravni je še v razmeroma zgodnji fazi. Eno prvih medjezikovnih študij je opravila Lewandowska-Tomaszczyk (1996).

⁴ V izvorni obravnavi leksikalnega niza 'become red' se kategorija asociacije ne pojavlja, vendar jo avtorica uvaja v svojem kasnejšem delu (Philip 2009), namesto tega pa opušča kategorijo t. i. zunajjezikovne funkcije, ki jo na tem mestu za slovenščino obravnavamo v kombinaciji s semantično prozodijo kot tistim vidikom razširjenega pomena, ki – funkcijsko opredeljen – stavi predvsem na širši situacijski kontekst.

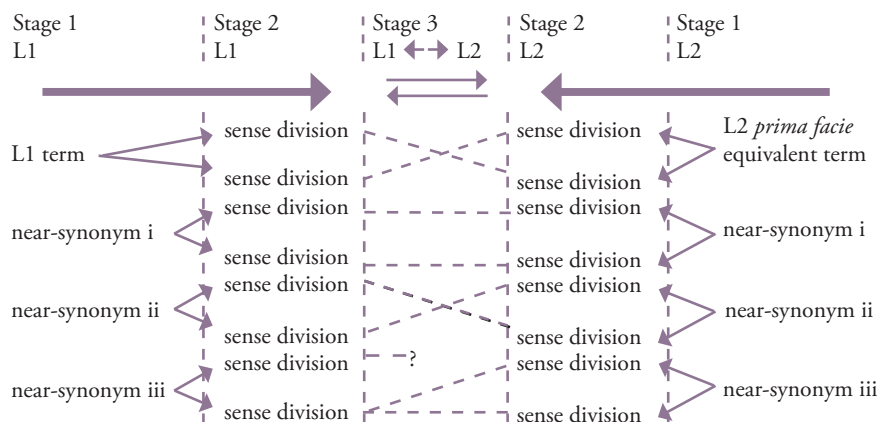
⁵ Leksikalna baza (LBS) (2008–2012) je del projekta *Sporazumevanje v slovenskem jeziku* <www.slovenscina.eu>.

⁶ <www.fidaplus.net>

Med pomembnejšimi prispevki zadnjih desetih let, ki proučujejo druge jezike v kombinaciji z angleščino, naj navedemo Xiao in McEnery (2006) za kitajščino, Berber-Sardinha (2000) za portugalsščino, Partington (1998), Tognini-Bonelli (2001), Philip (2003) in Stewart (2009) za italijanščino, Dam-Jensen in Zethsen (2007) za danščino ter Munday (2011) za španščino; za primerjavo analognih pomenskih enot v paru slovenščina-angleščina gl. tudi Šorli (2012, v pripravi).

2.4 Paradigmatski pristop

Spodnja shema, ki je povzeta po Philip (2003), ponazarja pristop k iskanju prevodno primerljivih enot pomena, pri čemer sta oba raziskovana jezika (tu označena kot L_1 in L_2) obravnavana kot izhodiščna, saj lahko le tako zagotovimo nepristransko analizo, ki ne prihaja do zaključkov o pomenu in rabi enega jezika *glede* na drugega, običajno tistega bolj razširjenega.



Slika 1: Shematski prikaz paradigmatskega pristopa k prevodni ustreznosti (po Philip 2003)

V prvi fazi je treba torej identificirati potencialno prevodno ustreznico in paradigmo delnih sopomenk (*near-synonyms*), ki ji pripada izvorna beseda (L_1), pa tudi tisto, iz katere je mogoče izbrati prevodno ustreznico (L_2). Ko so znani člani obeh paradigem, se identificirajo še njihovi ustrezni pomeni, s čimer dobimo pomenske enote. Pomene ugotavljamo s podrobno analizo vzorcev, tj. kolokacij, koligacij, frazeološke zgradbe, semantične preference, kontekstualnih značilnosti ter SP. Šele po takšni poglobljeni analizi je mogoče iskati primerljive prevode, tako na ravni splošnega kot partikularnega (*ibid.*). Tovrstni prevajalski postopek temelji predvsem na iskanju ustreznih primerljivih vzorcev v obeh jezikih, že Baker (1992: 57) je bila namreč mnenja, da je »*nekoliko izgube, dodajanja ali izkrievljanja pomena pogosto neizogiben del prevajanja*« (prev. avt.).

3 ANALIZA PRIMERA – METODOLOGIJA RAZISKAVE

V neposredni analizi smo primerjali rabo izbranih leksikalnih enot za *postati rdeč* v enojezičnem korpusu FidaPLUS v modulu *besedne skice*.⁷ Pri tem smo ugotavljali, v kolikšni meri se na videz ustrezni posamezni člani paradigem v svojih strukturnih in besedilnih značilnostih tako formalno kot funkcijsko ujemajo s člani paradigem v angleškem jeziku. Za angleško stran smo uporabili obstoječo analizo leksikalnega niza *become red* in sopomenk (Philip 2003),⁸ ki smo jo primerjali s podatki na slovenski strani.⁹ Korpus FidaPLUS je s 621 milijoni besed nekoliko večji od korpusa Bank of English, vendar so rezultati kljub vsemu primerljivi (nabor analiziranih konkordanc na slovenski strani smo omejili na 200). Da je funkcijski pristop k prevajanju in dvojezični leksikografiji nujen, meni Tognini-Bonelli, ki je že pred leti pozvala k izgradnji podatkovnih baz s prevodnimi omrežji (*translation webs*) (2001: 150–154), saj bi te občutno olajšale vrsto nalog v dvojezični praksi.

Paradigma POSTATI RDEČ vs. BECOME RED

Študija paradigem v paru *postati rdeč/ become red* ilustrira Hallidayevo (1992: 15) trditev, da mora vsaka jezikovna teorija temeljiti na teoriji *pomena kot izbire*, če naj bo relevantna za prevod(oslovje). Iz navedenega bo razvidno, da avtor izbere pomen iz vrste različnih možnosti, zato mora prevajalec skušati izbrati ustrezni pomen iz ekvivalentnega niza možnosti. Če skušamo za vsakega člana določiti najustreznejši prevod, moramo identificirati vse tiste njegove značilnosti, ki ga ločijo od drugih, nato pa mu najti najboljšo ustreznico v primerljivi paradigmi ciljnega jezika (Philip 2003: 258).

3.1 Primer: *postati/postajati rdeč* (425/88)

Glagola pripadata paradigmi, ki je vsaj na prvi pogled prevodno neproblematična, vendar podrobnejši pregled sobesedja in situacijskih kontekstov pokaže zanimive razlike. Za izraz *postati rdeč* so značilni kolokatorji, ki označujejo dele telesa, npr. *obraz, koža, oko/oči*, bodisi kot bolezenski simptomi, npr.

⁷ *Besedne skice* (Word Sketches) so povzetki kolokacijskega vedenja besed v programskem okolju Sketch Engine (SkE) <<http://ca.sketchengine.co.uk/>>. Za sekundarno preverjanje relevantnosti podatkov smo uporabili slovenski milijardni korpus Gigafida <<http://demo.gigafida.net>>.

⁸ Za svojo izhodiščno analizo angleške paradigme, izbrane na osnovi ponujenih prevodov v slovarju Ragazzini (1995), je G. Philip uporabila takrat 450-milijonski korpus *Bank of English*. Veljavnost podatkov v leta 2003 opravljeni angleško-italijanski analizi angleškega niza glagolov 'become red' smo preverjali v sodobni angleški leksikalni bazi DANTE (45.000 gesel, 2010, <www.webdante.com>).

⁹ V analizi smo uporabili le najbolj tipične, izstopajoče člene v paradigmatskem nizu, tako npr. v slovensko paradigmo nismo vključili sicer obstoječih, a manj pogostih zagoreti; pordečeti; rdečiti se.

1. *Oko že v nekaj urah po okužbi **postane rdeče**, ščemi in srbi ter kmalu začne izločati rumenkast izcedek oziroma gnoj.*

2. *Koža in meso okoli rane **postaneta rdeča**, vroča in nabrekla.*

ali kot znak odziva telesa na čustveni pretres, tipično jezo ali sram, ali telesni napor, npr.

3. *Čutila sem, kako **postajam rdeča** od sramu, in stopila sem z avtobusa.*

4. *Začutil sem, kako sem **postajal rdeč**, iz nosnic mi je puhala sapa, vroča kot zmajev ogenj.*

Koligacijsko prevladujejo raba v 3. osebi ednine v medicinskih kontekstih ter raba s prislovi stopnje in načina. Med frazeološkimi zvezami, ki opredeljujejo način, najbolj izstopa *postati rdeč kot kuhan rak*. Dopolnila pogosto določajo natančnejši odtенок rdeče barve (škrlatno, temno itd.), tudi ko gre za druge pomene. Preostale konkordance pripadajo predvsem dvema izstopajočima semantičnima skupinama: spreminjanje barve pri morskih živalih, npr.

5. *Brat škarpen je škarpun ali škarpoč, ki je enake oblike, le manjši je in ne **postane rdeč**, ko ga prinesemo na suho, ampak ostane rjav.*

ter rastline in njihovi deli (zlasti listi in plodovi):

6. *Iz belih cvetov nastanejo plodovi, ki **postanejo temno rdeči**, ko dozoriijo.*

7. *Japonski javor ima zelo lepe liste, ki jeseni **postanejo ognjeno rdeči**.*

Tabela 1: Postati/postajati rdeč: kolokacije, vzorci in situacijski kontekst (FidaPLUS, 2000)

izraz	postati(452)/postajati(88) rdeč
kolokacijski vzorci in kolokacije	'kdo postane rdeč kot kuhan rak' (18) obraz (67), koža (46), oko/oči (2), rana (2), beločnica (2), noga, dlesen, lasišče; listi (17), plodovi (3), cvet(ovi) (4), paradižniki (3); številke (4); piki; korale; človek; jeseni (11); čisto (14), še bolj (11), intenzivno, kardinalsko, karminsko, krepko, krvavo, lepo (2), ognjeno (4), opečnato (4), popolnoma (2); od jeze (3), od sramu, od besa
koligacije	prislov stopnje in načina, 3. oseba ednine za medicinske kontekste; končna pozicija oz. pred ločilom
semantična preferenca/asociacija	deli telesa (opisovanje simptomov, čustvena reakcija); morske živali; rastline in njihovi deli;
zunajjezikovne okoliščine	strokovnost, medicinska poročila
semantična prozodija	objektivna-nevtralna

3.2 Primer: *zardeti* (2047)

Na ravni koligacij *zardeti* ne kaže praktično nobene podobnosti s *postati rdeč*. Uporablja se tako v 1. kot 2. in 3. osebi, najpogostejše v tretji, v povednem naklonu in trdilni obliki (nekaj je vprašalnih oblik). V skoraj 100 % se pojavlja v končni poziciji stavka ali pred ločilom, ki uvaja odvisnik. Pojavlja se v nekaterih stalnih, frazeoloških zvezah, zlasti v *zardeti kot kuhan rak*. Primeri, kjer je potrebno dopolnilo, na primer natančnejši odtenek rdeče barve (*škrlatno, temno*), so zelo redki, pogost pa je prislov, ki podrobneje določa način zardevanja (*rahlo, živo, malce, močno, sramežljivo, do ušes* itd.), npr.

8. *Lužin je **zardel do ušes** in začel iskati figure, ki so padle na preprogo.*
9. *Zagledala je vaščane, ki so navdušeno ploskali, in **zardela do lasnih korenin**.*

dokaj pogosto je eksplicitno izražen tudi vzrok (od sramu, zadrege, jeze):

10. *Ko sem vstopila v trgovino, sem takoj opazila, da prodajalka sočutno gleda moj obraz, zato sem v trenutku **zardela od sramu**.*

Iz širšega sobesedila je mogoče razbrati, da gre praviloma za nehotena stanja zunaj posameznikovega nadzora. To nakazujejo tudi posamični primeri v zanikani obliki, npr.

11. *Lencii je uspelo, da se je nasmehnila, ne da bi **zardela**.*

Vzrok za zardevanje je pogosto nekaj, kar nekdo izreče, zato je velik del zgledov v obliki spremnih stavkov premega govora (570 od 2047 zadetkov), npr.:

12. *»/.../ Kmalu se bom poročila z Bruceom.« Želel sem ji voščiti srečo, toda nekaj me je stisnilo v grlu. »Zaljubljena si vanj?« Sandy me je pogledala in **zardela**.*

Redki so primeri, kjer gre za vzporednico tipični rabi v angleščini, ko do zardevanja v resnici ne pride, čeprav je to pričakovano in celo zaželeno:

13. *Leta 2003 je bil vržen v brutalni svet politike. Politični razred bi moral **zardeti** in popraviti 18.305 krivic. Namesto sramežljive rdečice vidimo zariple obraze, ki danes na problemu /.../*

Tabela 2: *Zardeti*: kolokacije, vzorci in situacijski kontekst (FidaPLUS, 2000)

izraz	zardeti (2047)
kolokacijski vzorci in kolokacije	'kdo zardi kot kuhan rak (21)', 'kdo zardi do ušes (31)' od sramu (20), od zadrege (9), od jeze (7); do lasnih korenin (7); rahlo (79), živo (22), malce (33), temno (18), škrlatno (9), močno (24), sramežljivo (8), spet (26), opazno (8), nekoliko (19), malo (20), nenadoma (7); obraz (19), lice (8), učiteljica (9), fant (8), dekle (7); v zadregi (14); v obraz (33), v lice (9);
koligacije	1., 2. in 3. oseba prislovi načina (barva) in stopnje končna pozicija ali pozicija za ločilom, ki uvaja odvisnik premi govor (spremni stavki)
semantična preferenca/asociacija	obraz, človek; okolščine sramu, zadrege ali (manj pogosto) osuplosti, telesnega napora
semantična prozodija	neželenost, nezmožnost nadzora oz. nebogljenost pred drugim

3.3 Primer: *pordeti* (588)

Najpogostejše pordi *koža*, *očesna veznica* in deli glave, predvsem *lica*, pa tudi *obraz*, *oči* in *ušesa*; v primeru *lic* in *obraz*a gre v večji meri za leposlovje, če pordi *koža* ali drugi deli telesa, pa za besedila z medicinsko oz. poljudno zdravstveno vsebino (npr. ko pordi mesto ugriza ali vboda itd.). Pretekli deležnik se dokaj pogosto uporablja v privedniški rabi:

14. *Dannyjevo in Brianovo petje se povzdiguje v falzet, Brian se drži za pult in nagiba naprej, obraz ima **pordel** in vlažen.*

Neposredni vzrok je redko izražen, ko gre za rastline, npr.

15. *Najobsežnejša polja te enoletne slanuše so pri nas v Strunjanskih solinah. Jeseni peščene plitvine ob morju zaradi njega popolnoma **pordijo**.*

Za razliko od *zardeti* zahteva *pordeti* v večini rab vsebinsko dopolnilo, najpogostejše načina in vzroka. Kot vzrok prevladujejo zunanji dejavniki, na primer mraz, poseg v telo, bolezen itd. Primeri, kjer pordi obraz ali del telesa zaradi čustvenih pretresov, so redkejši:

16. *Zardevanje je nerodna zadeva, čim bolj se boš zaradi tega vznemirjala, hitreje ti bodo lička **pordela** in zadrega bo še večja. Torej, za zardevanje se ne zmeni.*

17. *Ni gršega, pravijo, kot besna ženska: njen obraz **pordi**, vratne žile se napnejo, oči izbočijo, usta skremžijo, čelo naguba in telo se trese.*

Ponekod je vzrok izražen s predložno zvezo:

18. *Da je med kožo in psiho intimna vez, ve vsakdo iz lastne izkušnje: **pordimo od sramu ali jeze**, pobledimo in potimo se od strahu, zaradi zgroženosti dobimo kurjo polt, tudi rdečica iz zadrege odseva našo dušo.*

Tipični so primeri, kjer se pordelost v običajno priredno postavljeni zvezi povezuje neposredno z oteklino ali drugimi neprijetnimi reakcijami telesa, npr.

19. *Po tednu dni je noga **pordela in otekla**.*
20. *Lasišče običajno navadno **tudi srbi in je pordelo**.*

Nekateri konteksti morda kažejo, da bi šlo lahko za funkcijo poudarjanja čustvenega naboja, vendar zgledov ni dovolj, da bi to lahko z gotovostjo potrdili:

21. *Linda je bila iz sebe. Njen obraz je **pordel** in skoraj se je razpočila od jeze.*
22. *Hitro je dvignil pogled in se nehote zastrmel naravnost v Irisine oči. Ni mu bilo treba dolgo čakati na reakcijo – lica so mu v trenutku **pordela**.*

Tabela 3: Pordeti: kolokacije, vzorci in situacijski kontekst (FidaPLUS, 2000)

izraz	pordeti (588)
kolokacije	koža (72), veznica (10), obraz (23), oči (21), oko (7), lica (3); sklep (6); list (1); oteči (19)/vneti (9) in pordeti; v obraz (3)/ lica (2); rahlo (14), nekoliko (14), močno (11); od mraza (4)/sramu (4)/jeze (3)
koligacije	predložne zveze (vzrok)
semantična preferenca/asociacija	obraz, deli telesa; simptomi bolezni, zdravstvene težave
zunajjezikovna funkcija	dokaj objektivno navajanje različnih vzrokov; humorni učinki? poudarjanje čustvenega naboja?
semantična prozodija	nevtralna-objektivna

3.4 Primer: *pordečiti* (83)

V neprehodni rabi največkrat pordečijo deli telesa ali rastlin:

23. *Tudi pri slednjem ugrizu se je vse srečno končalo, prst je sicer **pordečil**, bil boleč na dotik kakšen teden, hujšega pa ni bilo.*

Pojavi se nekaj zadetkov priredno zložene zveze *pordečiti in oteči*, ki umešča kolokacijo v kontekst bolezenskih oz. poškodbenih simptomov, npr.

24. »Staknil sem nerodno infekcijo kolena, ki mi je naenkrat **pordečilo in močno oteklo**. A se bom zagotovo pozdravil.«

V konkordancah identificiramo – sicer v prehodni obliki – tudi pomen *našminkati* oz. *nanesti rdeče ličilo*, ki se pojavlja večinoma s povratnim zaimkom, zlasti z *ustnice*, *usta*, *lica*, in ki ga ne najdemo pri nobenem drugem glagolu v paradigmi. V tem pomenu se uporablja tudi v 1. osebi, drugod skoraj izključno v 3. osebi. Oblike preteklega deležnika *pordečen* ni zaslediti tako rekoč v nobenem od pomenov.

Tabela 4: *Pordečiti*: kolokacije, vzorci in situacijski kontekst (FidaPLUS, 2000)

izraz	pordečiti (83)
kolokacije	mraz (1), kri (6); lica (9), obraz (3), steza (2), oblak (1); ustnice (7), usta (1); prst (1), koža (4); voda (2); s šminko (3); s krvjo (2)
koligacije	ne izstopajo
semantična preferenca/asociacija	deli telesa, narava in rastline; ličenje
zunajjezikovni kontekst	literarna besedila
semantična prozodija	nevtralna-objektivna

3.5 Primer: *obarvati+(se)+rdeč* (210)

Obarvati se rdeče je relativno pogost vzorec, čeprav je *rdeče* večinoma modificirano z natančnejšimi odtenki ali načinom, npr. *svetlo*, *temno*, *škrlatno*. Prevladujejo konteksti v medicini, kemiji, botaniki in na splošno o pojavih v naravi. Nekaj je primerov, ko gre za barvne približke, npr. z *obarvati se rožnato* (11) (lica, ušesa; ribje meso)/*rdečkasto* (7) (botanika). Rdeče se obarvajo zlasti rastline in njihovi deli, npr. *listi* ali *drevo*, *plodovi* itd., nekaj pa je tudi primerov iz živalskega sveta, npr. *meso*; *mladice*; *plavuti*. Rdeče se obarvajo *lica*, *obraz*, *koža* ali *oči*, slednja dva iz bolezenskih razlogov, prav tako *urin* in *seč*, za obarvanost pri pojavih v naravi pa gre na primer pri kolokatorjih *nebo*, *oblaki*, *vrhovi* in *voda*; *lakmus*, *papir*, *robček*; *ogenj*, *olje*, *raztopina*, *vlakno*; *semafor*.

25. Preden listi odpadejo, **se obarvajo** lepo **rdeče**.
26. Zaljublencu začne srce hitreje biti, lica **se mu obarvajo rdeče** in rdečice ne more prekriti.

27. *Voda se je obarvala rdeče, ker so se na površje dvignile plasti na dnu nako-pičenega železovega oksida.*

Med kolokatorje so vključeni tudi primeri, kjer je rdeča podrobneje opredeljena z odtenkom, npr. kaj se obarva *rahlo* (1)/ *svetlo* (3)/ *temno* (2)/ *temnordeče* (2)/ *škrlatno* (2)/ *škrlatno rdeče* (4)/ *rdečkastol malinasto rdeče* (1):

28. *Oktobra se listi obarvajo škrlatno rdeče, kar daje aroniji okrasno vre-dnost, v nekaj tednih pa odpadejo.*
29. */.../ grlo močno boli, težave pri požiranju, obložen jezik, ki se obarva ma-linasto rdeče. (malinast jezik)*

S predložno zvezo, pogosto v deležniški obliki, je barva lahko izražena kot sred-stvo ali način, zlasti za pojave v naravi: kaj se obarva *z/v rdečo barvo* (in različnimi odtenki), npr.

30. *Pa vendar, bilo je lepo in mrzlo jutro, oblaki so se že obarvali s škrlatno jutranjo barvo in vse je bilo tako spokojno. A vseeno je bilo v ozračju med nama čutiti napetost.*
31. *Najlepši filosilikat iz skupine sljud je ponavadi značilno rožnato obarvan in se pojavlja v obliki ploščic ali prizem psevdoheksagonalnih oblik.*

Tabela 5: Obarvati+(se)+rdeč: kolokacije, vzorci in situacijski kontekst (Fi-daPLUS, 2000)

izraz	obarvati+(se)+rdeč (210)
kolokacije	listi (6), drevo (1), plod (6), lubje, korenine, cvet, grm, poganjki, roža, paprika, trta, brogovita; meso (6); mladice; konice lovk, plavuti; lice (3), obraz, koža, oči; seč (3), urin; nebo (4), oblaki, vrh, vrhovi (2); voda (5); lakmus, papir, robček, ogenj, olje, raztopina, vlakno; semafor (1)
koligacije	pretekli deležnik 3. oseba trpnik* povedni naklon
semantična preferenca/asociacija	rastline in njihovi deli; živalski svet (majhne živali) deli telesa narava; kemijske reakcije
zunajjezikovne okoliščine	botanični in zoološki strokovni in poljudni teksti; navajanje dejstev pri pojavih v naravi, literarnost
semantična prozodija	objektivna-nevtralna (strokovni teksti); v primeru, ko se (škrlatno) rdeče obarva obraz, je mogoče prepoznavati silovitost čustev, včasih posmeh; kontrast barv in hitrost spremembe

4 IDENTIFIKACIJA NAJBOLJŠIH PREVODNIH MOŽNOSTI

Treba je poudariti, da v paradigmi ne gre za izčrpen seznam možnosti, temveč za delno naključni nabor slovarsko izstopajočih ustreznih, izmed katerih na prvi pogled vsi glagoli pomenijo bolj ali manj isto. Prepoznani vzorci kolokacijskega, koligacijskega vedanja ter kontekstualnih značilnosti so do neke mere prekrivni, toda za vsakega člana paradigme smo ugotovili tudi njegove izstopajoče lastnosti, torej tiste, ki ga na katerikoli od pomenskih ravni ločijo od drugih. Znano je pre-pričanje, da popolnih sinonimov ni, saj je treba upoštevati različne ravni enakosti oz. ustreznosti – tako semantične kot slogovne in pragmatične.

Tabela 6: Izbrani paradigmatski niz za *postati rdeč* v slovenščini (621 mil. besed) in angleščini (450 mil. besed)

postati rdeč (494)	to become red (50)
zardeti (2047)	to blush (297)
pordeti (588)	to flush (337)
pordečiti (83)	to go red (108)
obarvati+(se)+rdeč (210)	to redden (239)
	to turn red (225)

Najustreznejša izbira prevoda izhaja iz opazovanja in primerjave izdelanih profilov, in sicer vsake izmed sestavin enote pomena – kolokacije, koligacije, semantične preference in pragmatične funkcije/SP. Kot prvi prevod so spodaj navedeni tisti, ki so neposredno in v celoti gledano najbolj tipični za izhodiščnega člana paradigme, kar seveda ne pomeni, da druge rešitve niso možne. Alternativne ubeseditve za konkretna sobesedila in situacijske kontekste smo navedli tudi na koncu, v povzemalni tabeli (3.6).

4.1 Prevod: *postati rdeč* → *to become red*

Nobeden od glagolov se ne pojavlja v izrazito frazeoloških zvezah, pri obeh gre najpogosteje za bolezenske simptome v povezavi s še drugimi težavami, zlasti zaradi vnetja kože in oči, tudi dlesni, ran, lasišča itd. Vendar pa je *postati rdeč* med člani paradigme povprečno pogost, medtem ko je *become red* najmanj pogost med glagoli svoje paradigme, kar nakazuje manjšo splošno uporabnost. Razvidno je, da se uporablja tako kot *postati rdeč* pretežno za *skin*, *eye(s)* in *face* v medicinskih kontekstih. SP nismo ugotovili ne pri enem ne pri drugem, je pa pri *become red* prepoznavna

zunajjezikovna funkcija, kot jo opredeljuje Philip (2003: 259), namreč nekakšna tehnika distanciranja s predstavljanjem dejstev kot hipotetične možnosti, kar se koligacijsko navezuje na rabo pogojnika. Za rastline in dele rastlin, npr. listi in plodovi, je uporabnejši *turn red*, prav tako v živalskem svetu. Zlasti tam, kjer kaj *postane še bolj/intenzivno/krvavo rdeče*, torej s prislovi stopnje ali načina, sta primernejša *flush* ali *red*, kadar pa gre za predložno dopolnilo *od besa* ali *od jeze*, je najprimernejši *flush*. Opaziti je kontrastivno zanimivo razliko, namreč medtem ko je v angleščini barva natančneje opredeljena v obliki povedkovskega prilastka pri *flush*, npr. *flush bright red/crimson/puce/ mottled maroon*, so realizacije v slovenščini pri *postati/postajati rdeč* z vrinjenim prislovom, ki modificira pridevnik, npr. *postati kardinalske/karminske/ krepke/ krvave/ lepe/ ognjene rdeč*.

Tabela 7: *Become red*: kolokacije, vzorci in situacijski kontekst (po Philip 2003: 260)

izraz	become red (50)
kolokacije	skin (12), eye(s) (9), face (3), scalp (2), ears, feet, fingers, knee red and sore (6), inflamed (5), swollen (4), blotchy (2), hot (2), angry, flaky, irritated, itchy, oversensitive, raw; anger, shame
koligacije	pogojni stavki (13)
semantična preferenca	deli telesa, medicinska poročila ali zdravstveni nasveti
zunajjezikovna funkcija	zdravniški nasvet kot hipoteza; tehnika distanciranja
semantična prozodija	objektivna-nevtralna

4.2 Prevod: *zardevati* → *to flush*

Na prvi pogled najustreznejši prevod *blush* se zaradi posebnosti v prozodiji na angleški strani v naši analizi ne uvršča na prvo mesto. Ugotovljeno je namreč, da *blush* sicer podobno kot *zardevati* izstopa po tem, da ga tipično izzove zadrega ali sram, praktično nikoli pa jeza ali druga negativna čustva, prav tako se oba uporabljata v ožje frazeoloških enotah. Vendar gre večinoma za situacije, ko se zardevanje, čeprav pričakovano, v resnici ne zgodi (Philip 2003: 262). Frazeološke enote *have the grace to blush*, *blush at the thought/idea* ali *something would make (even) somebody blush* kažejo na ustaljene konvencionalizirane rabe, ki bodisi izražajo priznavanje govorčeve zadrege ali ogorčenje nad tem, da nekdo v očitno sporni situaciji zadrege ne občuti. Tipično zardeva človek ali del telesa, pri *blush* pa gre skoraj izključno za človeka. V celoti gledano se

pridobljeni leksikalni profil za *zardeti* najbolj ujema z identificiranimi vzorci pri glagolu *flush*. Sorodni so kolokatorji, ki kažejo, kaj najpogosteje zardi: *obraz*, *lice* in *človek*. Podobno izstopajo prislovi načina, pogostosti in stopnje, predvsem pa se pri obeh izkazujejo tudi okoliščine sramu, zadrege ali (manj pogosto) osuplosti oz. telesnega napora. *Flush* je pogostejše kot v osnovni obliki glagola uporabljen kot deležnik, podobno tudi oblika *zardel*. Gre za glagol, ki včasih potrebuje vsebinsko dopolnilo, zlasti natančnejšo opredelitev barve ali načina. Pri obeh je mogoče ugotoviti neželenost reakcije, nezmožnost nadzora in neobglenost pred drugim, torej podobno SP. *Blush* bo uporaben kot prevod predvsem v kontekstih, kjer ni potrebnih vsebinskih dopolnil oz. tam, kjer stoji *zardeti* v končni poziciji, tj. na koncu stavka ali pred ločilom, pogosto v premem govoru. V nekaterih situacijah se kaže kot najustreznejši prevod *turn red*, zlasti v frazeoloških zvezah in tam, kjer je potrebno izraziti vzrok. Odločitev za bodisi *flush* ali *turn red* je odvisna tudi od tega, ali je izhodiščni glagol pospremljen s prislovom pogostosti ali načina (*rahlo*, *živo*, *malce*, *močno* itd.) – to je izkazani vzorec pri *flush* –, pa tudi od tega, kdo zardi. Če zardi *človek*, bo verjetnejši prevod *flush* (ali v nekaterih primerih *blush*), saj se *turn red* razen v frazeologiji uporablja predvsem za rastline oz. dele telesa.

Tabela 8: *Flush*: kolokacije, vzorci in situacijski kontekst (po Philip 2003: 265)

izraz	flush (337)
kolokacije	face (49), cheeks (27) red (18); bright red (3), deep red, dark red, angry red, scarlet, crimson; pink (2), deep pink, complementing pink, painful pink, blotchy albino pink, gravidlax pink; purple (2), violet, puce, mottled maroon he/she (120); I (14) slightly (8), heavily (2), hotly (2), deeply, brightly, noticeably, visibly, again (6), once more, instantly, quickly, suddenly angrily (5), excitedly, embarrassment (7), anger (6), pleasure (5), rage (3), excitement (2), fury (2), apprehension, delight, envy, mortification, pride, rage
koligacije	premi govor prislovi stopnje, pogostosti feel yourself flush (15)
semantična preferenca	obraz barve čustva
zunajjezikovna funkcija	primeri, ko se zardevanje dejansko zgodi (zlasti literarno)
semantična prozodija	neželenost, nezmožnost nadzora oz. neobglenost pred drugim

Tabela 9: *Blush*: kolokacije, vzorci in situacijski kontekst (po Philip 2003: 263)

izraz	blush (297)
kolokacije	'have the grace to blush' (9) 'blush to admit/recount/tell' (7); 'blush to recall/remember' (4) 'blush at the thought/idea' (4) 'something would make [even] somebody blush'
koligacije	prvoosebni zaimki (51) modalni glagoli (59) kavzativi (81) nikalnost (26)
semantična preferenca	zadrega ali neprijetne situacije
zunajjezikovna funkcija	priznanje zadrege/sramu jeza nad tistimi, ki ob pričakovanih vzorcih vedenja nimajo dovolj sramu
semantična prozodija	pričakovano, toda ne realizirano; težavnost

4.3 Prevod: *pordeti* → *redde*

Najbližji se zdijo kolokacijski vzorci pri *redde* in *become red*. Oba kot ključne subjekte pordevanja navajata *kožo*, *oči* in *obraz*, prav tako se zdijo pogoste okoliščine otekanje in spremljevalne bolečine, kot vzrok prevladujejo zunanji dejavniki, na primer mraz, poseg v telo, bolezen itd. *Pordeti* zahteva torej v večini rab vsebinsko dopolnilo, najpogosteje načina in vzroka, kar je bolj tipično za *redde*, ki se pojavlja pretežno v trpniku, z izraženim agentom, poleg tega so zanj bolj tipični tudi prislovi stopnje, zato je 'redde' boljša splošna ustreznica. Pri *pordeti* in še zlasti pri *redde* je mogoče ugotoviti literarno rabo, v kateri nastopajo kazalci pragmatičnega pomena že v obliki kolokabilnih prislovov, npr. *rahlo*, *nekoliko*, *močno*, *angrily*, *horribly* itd. Koligacije so v obeh primerih zlasti predložne zveze, SP je podobno objektivna-nevtralna, čeprav je na slovenski strani zaznati nekoliko več čustvenega naboja ali humornih učinkov pri rabi glagola. Še zlasti to velja takrat, ko gre za pretekli deležnik, torej *pordela lica* ali *pordel obraz* (za ta primer smo naknadno identificirali potencialni prevod *ruddy cheeks/complexion*, ki bi zahteval nadaljnjo analizo, vendar kaže, da se uporablja tipično v pogovornih kontekstih, ko se rdečica povezuje z zdravjem – ta kontekst je v slovenščini pogosto humoren). Kjer ne gre za simptome zdravstvenih težav, temveč za pordevanje zaradi mraza ali sramu, je mogoče uporabiti tudi prevod *flush*.

4.4 Prevod: *pordečiti* → *redde*

V neprehodni rabi največkrat pordečijo deli telesa ali rastlin, nekaj je zadetkov priredno zložene zveze *pordečiti in oteči*, ki umešča kolokacijo v kontekst bolezenskih oz. poškodbenih simptomov. Glede na to bi bila najustreznejša nevtralna prevoda *redde* in *become red*, se pa nakazuje tudi možnost prevoda z *blush*, na primer, ko se pordeči *oblak* ali *nebo*. Tako kot *flush* je tudi *redde* mogoče pogosto najti v pridevniški rabi. V glagolski rabi ima sicer *pordečiti* zelo podobne kolokatorje in vzorce kot *become red*, na primer izražen vzrok: če gre za zunanega, s predložnim dopolnilom *by ...*, če gre za notranjega pa *with ...*.

Tabela 10: *Redde*: kolokacije, vzorci in situacijski kontekst (po Philip 2003: 269)

izraz	redde (239)
kolokacije	‘feel --self redde’ (6) face (67), eyes (43), skin (37), cheeks (18), hands (9), nose (7), swelling/swollen (10) cold (6), the sun (2), lack of sleep, drink anger (2), embarrassment (2), annoyance, rage, humiliation, effort, strain slightly (5), further(4), again (3), angrily (2), noticeably (2), visibly (2), horribly, instantly, startlingly
koligacije	trpne oblike + agent prislovi stopnje
semantična preferenca	deli obraza
zunajjezikovna funkcija	zadrega, jeza ali napor in njihovi vzroki simptomi bolezni ali njihovi vzroki
semantična prozodija	objektivna-nevtralna

4.5 Prevod: *obarvati*+(*se*)+*rdeč*(/*rožnat*/*škrlaten*) → *to turn red*(/*pink*/*scarlet*)

Najbolj tipično se *rdeče* (tudi npr. *rožnato*, *škrlatno* itd.) obarvajo listi in plodovi, ter cela vrsta drugih rastlin in njihovih delov, tem zato na ravni kolokacij najbolj ustreza *turn red*. Podobno semantično preferenco izkazuje *obarvati se* tudi za živali, zlasti morske, npr. rak, jastog, mladica, plavuti itd., ter mnoge pojave v naravi. Za kolokatorje *nebo*, *oblaki*, *vrh*, *vrhovi* se kot potencialna ustreznica ponuja glagol *blush* (pomen: *become pinkish*). Visoko stopnjo primerljivosti s *turn red* potrjuje

ujemanje na koligacijski ravni: nagnjenost k trpnim oblikam (kaj *se obarva*) oz. pogostejša raba preteklega deležnika (*obarvan rdeče* itd.). Prav tako je mogoče podobnosti identificirati na ravni SP, namreč kontrast oz. poudarjena hitrost pri spremembi barve, tako v naravi kot pri obrazu in licih. Opazna je povezava med hitro spremembo v barvi obraza in silovitostjo čustvene reakcije, še zlasti ko se obraz koga obarva *škrlatno rdeče*, ponekod je zaznati tudi dodani element posmeha. V strokovnih besedilih je SP na obeh straneh pričakovano objektivna-nevtralna.

Pri *turn red* izstopa kolokator *traffic lights* in *light*, torej semantična preferenca *sve-tlobe*, česar pri *obarvati+(se)+rdeč* ni opaziti, zabeležen pa je vendarle en primer kolokatorja *semafor*. Ker ga ni zaslediti pri nobenem drugem glagolu v slovenski paradigmi, lahko to interpretiramo kot potencialno podobnost. *Turn red* se v pomenu obarvanja obraza zaradi jeze ali zadrege dokaj pogosto pojavlja v literarnih besedilih, vendar se – z izjemo primerov s predložnim dopolnilom *from* ali *with* – pretežno ne uporablja za človeka. V smislu prepoznavnosti vzorcev se dokaj dobro ujema z vzorci pri *postati rdeč*. Ko izraža spremembo barve, zlasti v rastlinskem svetu, je bolj kot postopno prehajanje poudarjena nenadnost spremembe. Ta posebnost pojasnjuje tudi kolokabilnost s *traffic lights*, ki je drugod v paradigmi ni najti. Relativna, a povezana posebnost je tudi sopojavljanje z barvami, ki so v močnem kontrastu z rdečimi otenki, npr. *green*, *blue*, *black* itd.

Tabela 11: *Turn red*: kolokacije, vzorci in situacijski kontekst (po Philip 2003: 271)

izraz	turn red (225)
kolokacije	face (22) light (14), traffic lights (7); green (17), blue (14), black (13), orange (5), pink (4), purple (4) rage (7), embarrassment (3), shame (3), anger (2) annoyance, fury leaves (3), rowan, mulberry, cherry, tomatoes, peppercorns, chilli, peppers lobster (2), prawn, crayfish ink
koligacije	predložna zveza: 'from'(3), 'with' (21) + čustva končna pozicija; ločila
semantična preferenca	osnovni izrazi za barve rastlinstvo
zunajjezikovna funkcija	jasni prehodi od ene barve k drugi, kontrast
semantična prozodija	objektivna-nevtralna

Go red opisuje rdečino, ki nastane bodisi zaradi čustev (najpogosteje), bolezenskih simptomov ali namerne spremembe barve las, npr. *to go blonde*, *auburn* itd., vendar v nobenem od teh pomenov ne izkazuje pragmatičnih kazalcev, ki bi označevali osebni odnos govorca do vsebine. Ta glagol je od vseh v paradigmi semantično najbolj izprazen, zato lahko pomeni za avtorja in/ali prevajalca najbolj 'nevtralno' izbiro (Philip 2003: 267).

Tabela 12: *Go red*: kolokacije, vzorci in situacijski kontekst (po Philip 2003: 267)

izraz	go red (108)
kolokacije	'go red in the face' (19) 'feel* her face/-self go red (6) face (15), skin (4), ears (3), neck (2), nose, eyes salon, blonde, brown, auburn, hair
koligacije	--
semantična preferenca	deli obraza
zunajjezikovna funkcija	opis zardevanja zaradi različnih razlogov, vključno z jezo ('go red in the face'), na razmeroma objektivni način
semantična prozodija	objektivna-nevtralna

4.6 Povzetek

Pričakovano imajo vsi posamezni člani paradigem več prevodnih možnosti, pri čemer se ena med njimi običajno kaže kot splošno najuporabnejša. Analiza kaže, da je mogoč bolj ali manj uspešen prenos celotnih (razširjenih) pomenskih enot v ciljni jezik, ki je odvisen od uspešnosti razbiranja implícitnega oz. pragmatičnega pomena v izvornem jeziku. Na ravni enojezične analize lahko skladno z ravnmi razširjene enote pomena določimo štiri korake: identifikacija kolokacijskega profila (leksikalne realizacije), koligacijskih preferenc (leksikogramatikalne realizacije), semantične preference (semantično polje) ter pragmatične funkcije (SP). Na dvojezični ravni naj bi prišlo najprej do primerjave pragmatične funkcije na skrajnih mejah razširjene enote pomena. Na primer, najustreznejši identificirani prevod za *zardevati* v skladu s prepoznano SP ni *to blush*, temveč *to flush*, prvi se namreč uporablja tipično v kontekstih, kjer do zardevanja – kljub pričakovanjem – ne pride. *Blush* se tako tipično uporablja v nikalnih ali pogojnih stavkih. Najbolj idiomatske frazeološke enote, npr. *kdo zardi kot kuhan rak/puran*, imajo pogosto visoko

konvencionalizirane idiomatske ustreznice, v tem primeru *sb turns as red as a beet(root)*, pri čemer gre v prvi vrsti za iskanje ustreznosti na ravni pragmatične funkcije. Seveda tu nakazane prevodne rešitve niso absolutne in izključne. Obstajajo tudi druge možnosti, zlasti opisne oblike prevodov ali takšne, ki so vezane specifično na en sam kolokator ali skladenjsko obliko, npr. *z rdečimi/pordelimi očmi* → *red-eyed*; *močno pordeti* → *to turn a deep shade of red* itd. Do takšnih rešitev na prvi fazi pridemo teže preko neposrednih dvojezičnih virov, lahko pa s poglobljeno analizo korpusa in obrnjenih virov.¹⁰

SLOVENŠČINA → ANGLEŠČINA

postati/postajati rdeč → to become red

čigava koža/obraz postane rdeč/-a → *sb's skin/face becomes red*

čigave oči postanejo rdeče → *sb's eyes become red*

→ to flush

postati karminsko/krvavo rdeč → *to flush crimson*

→ to redden

kdo postaja rdeč → *sb can feel oneself redden?*

čutila sem, kako postajam rdeča → *I could feel myself redden*

→ to go red

kdo postane rdeč v obraz → *sb goes red in the face*

čutila sem, kako postajam rdeča → *I could feel myself go red*

kdo postaja rdeč → *sb can feel oneself go red?*

→ to turn red

kdo postane rdeč kot kuhan rak → *sb turns as red as a beet(root)*

paradižniki/list/rak/korale postanejo rdeči → *tomatoes/leaves/lobster/chorals turn red*

žival postane živo rdeča → *the animal turns bright red*

semafor postane rdeč → *traffic lights turn red*

zardeti → to flush

škrlatno/temno zardeti → *to flush scarlet/dark red*

kdo zardi do ušes → *sb flushes up to their ears*

kdo zardi od sramu/zadrega → *sb flushes with shame/embarrassment*

kdo zardi od jeze → *sb flushes with anger/rage*

→ to blush

nasmehnila se je, ne da bi zardela → *she smiled without blushing*

ob tem bi morali zardeti → *this should make them blush*

¹⁰ Npr. obrnjena baza Velikega angleško-slovenskega slovarja Oxford-DZS (Krek et al., DZS/Amebis 2006).

→ to turn red

kdo zardi kot kuhan rak/puran → *sb turns as red as a beet(root)*

kdo zardi od jeze → *sb turns red with anger/rage*

→ to go red

svetlo/živo zardeti → *to go bright red*

pordeti → to redden

od vetra in mraza pordela lica → *cheeks reddened by the cold*

kdo pordi v obraz (od vročine) → *sb's face/cheeks is/are reddened (by the heat)*

pordel obraz → *a ruddy face?* (He had a ruddy face and purple nose from too much alcohol.)

→ to become red

noga je pordela in otekla → *the foot became red and swollen*

lasišče srbi in je pordelo → *the scalp becomes red and itchy*

pordečiti → redden

mraz/kri pordeči lica komu → *cold/blood reddens sb's cheeks?*

→ become red

koža/prst/koleno pordeči → *the skin/finger/knee becomes red*

→ flush

kaj pordeči komu lica → *sb's cheeks flush with sth*

obarvati (se) rdeče/rožnato* → to turn red/pink*

listi se obarvajo rdeče → *leaves turn red*

Slika 2: Povzetek prevodnih možnosti v smeri slovenščina → angleščina, kot jih narekuje semantično-pragmatična analiza

5 SKLEP

Uspešna prevodna praksa predpostavlja ustrezen prenos vseh vidikov pomena, tudi ali še posebej pragmatičnega, zato je poznavanje jezika v tej njegovi nesporni, pa čeprav morda najbolj izmuzljivi razsežnosti za prevajalca ključnega pomena. Predstavljena analiza celotnih parov paradigem se zdi morda nerealna ali celo nepotrebna v običajnih okoliščinah prevajalskega procesa, in čeprav drži, da je kot metoda pri vsakokratnih prevajalskih nalogah neizvedljiva, je gotovo smiselna, če načrtujemo na primer izdelavo prevajalske podatkovne baze. Ker je izgradnja različnih eno- in večjezičnih baz – tako za namene leksikalnih opisov

kot luščenja terminologije in čim večje avtomatizacije prevajalskih procesov – že uveljavljena jezikovnotehnološka praksa, in ker imamo obenem na voljo že zelo kompleksna statistična orodja za analizo korpusnih podatkov, je treba razmisliti tudi o vsebinski nadgradnji standardnih leksikalnih opisov. Nekateri subtilnejši vidiki pomena, zlasti razširjene enote pomena, kot predstavljeno v prispevku, ostajajo prezrti tudi v nekaterih sodobnih, korpusno zasnovanih leksikografskih virih oz. leksikalnih bazah. Predvidevamo, da je to rezultat časovnega pritiska, saj gre pri ugotavljanju SP za zamudno analizo, pri kateri pravzaprav nikoli ne vemo vnaprej, kaj točno iščemo, poleg tega pa vsaj zaenkrat ni učinkovitih orodij za njeno avtomatsko prepoznavo.

Pričujoči prispevek opisuje večplastni semantični in funkcijski oz. pragmatični pristop, najprej k znotrajjezikovnim, nato pa tudi medjezikovnim razmerjem. Gre za proces iskanja prevodne ustreznosti (primerljivosti) tako, da vidimo jezika L_1 in L_2 kot izvorna jezika, pri čemer razlikovanje med enkodiranjem in dekodiranjem, ali izvirnim in ciljnim jezikom, ne igra več pomembne vloge. Raziskovanje širšega besedilnega okolja, ne le kolokacij, temveč tudi koligacij, semantične preference, asociacij in zlasti funkcijskega pragmatičnega pomena oz. SP je koristno v procesu prevajanja tudi zato, ker ponuja vpogled v širši kontekst jezikovnih izbir avtorja izvirnega besedila, kar posledično vodi v natančnejše in v zavestni izbiri utemeljene prevodne rešitve, ki skušajo reproducirati sporazumevalne učinke v izvornem jeziku.

Bibliografija

Spletni viri

Longman Dictionary of Contemporary English, Longman: <<http://www.ldoce-online.com/>> (LDOCE)

Macmillan Dictionary and Thesaurus: Free English Dictionary Online. <<http://www.macmillandictionary.com/>> (MEDAL)

DANTE: <<http://www.webdante.com/>>

FidaPLUS: <<http://www.fidaplus.net>>

Gigafida: <<http://demo.gigafida.net>>

Sketch Engine: <<http://ca.sketchengine.co.uk/>>

Literatura

Atkins, Sue in Michael Rundell, 2008: *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.

- Atkins, Sue, Charles Fillmore in C. R. Johnson, 2003: Lexicographic Relevance. Selecting information from corpus evidence. *International Journal of Lexicography* 16/3. 251–280.
- Baker, Mona, 1992: *In Other Words: A Coursebook on Translation*. London in New York: Routledge.
- Berber Sardinha, Tony, 2000: Semantic prosodies in English and Portuguese: A contrastive study. *Cuadernos de Filología Inglesa* 9(1). 93–110. Dostopno na: http://www2.lael.pucsp.br/~tony/temp/publications/2000murcia_prosodies.pdf (27. 9. 2012)
- Bublitz, Wolfram, 1996: Semantic prosody and cohesive company: somewhat predictable. *Leuvense Bijdragen: Tijdschrift voor Germanse Filologie*, 85/1-2. 1–32.
- Channell, Joanna, 1999: Corpus-based Analysis of Evaluative Texts. S. Hunston in G. Thompson (ur.): *Evaluation in Text. Authorial Stance and the Construction of Discourse* (2003). Oxford: OUP. 39–55.
- Dam-Jensen, Helle in Karen K. Zethsen, 2007: Pragmatic patterns and the lexical system – a reassessment of evaluation in language. *Journal of Pragmatics* 39(9). 1608–1623.
- Dam-Jensen, Helle in Karen K. Zethsen, 2008: *Translator awareness of semantic prosodies*. *Target* 20/2. 203–221.
- Firth, John R., 1957: *Papers in linguistics, 1934-1951*. London: Oxford University Press.
- Gabrovšek, Dušan, 2007: Konotacija, semantična prozodija in sintagmatski pomen: Tri ravni asociativnega pomena? S. Komar in U. Mozetič (ur.): *English Language Overseas Perspectives and Enquiries. Slovenščina v kontekstu: novi izzivi, novi obeti*. IV/1-2 U. Ljubljana: Društvo za angleške študije.
- Gabrovšek, Dušan, 2012: The Lexicographer, the Bilingual Dictionary and Its User, and the Corpus: Facing Colligational Dilemmas. M. Šorli (ur.): *Dvojezična korpusna leksikografija*. Ljubljana: Trojina, zavod za uporabno slovenistiko.
- Gantar, Polona, Katja Grabnar, Polonca Kocjančič, Simon Krek, Olga Pobirk, Rok Rejc, Mojca Šorli, Simon Šuster in Petra Zaranšek, 2009: *Specifikacije za izdelavo leksikalne baze za slovenščino (Kazalnik 6): Projekt Sporazumevanje v slovenskem jeziku*. Kamnik. Slovenija. http://www.slovenscina.eu/Media/Kazalniki/Kazalnik6/SSJ_Kazalnik_6_Specifikacije-leksikalna-baza_v1.pdf. (Dostop 27. 9. 2012)
- Giora, Rachel, 2003: Literal vs. figurative language: Different or equal? *Journal of Pragmatics* 34. 487–506.
- Halliday, Michael A. K., 1992: Language Theory and Translation Practice. *Rivista internazionale di tecnica della traduzione* 0 (pilotska izdaja). 15–25.
- Hoey, Michael, 2005: *Lexical Priming*. London in New York: Routledge.
- Hunston, Susan in Gill Francis, 1999: *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English*. Amsterdam: Benjamins.

- Hunston, Susan, 2002: *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Hunston, Susan, 2007: Semantic prosody revisited. *International Journal of Corpus Linguistics* 12. 247–268.
- Jurko, Primož, 2012: Collocation Errors in Encoding and Possible Cures: Dictionaries Vs. Corpora. M. Šorli (ur.): *Dvojezična korpusna leksikografija. lovenščina v kontekstu: novi izzivi, novi obeti* Ljubljana: Trojina, zavod za uporabno slovenistiko.
- Lewandowska-Tomaszczyk, Barbara, 1996: Cross-linguistic and language-specific aspects of semantic prosody. *Language Sciences*, 18. 153–178.
- Louw, W.E., 1993: Irony in the Text or Insincerity in the Writer?: The Diagnostic Potential of Semantic Prosodies. M. Baker., G. Francis in E. Tognini-Bonelli (ur.): *Text and technology: in honour of John Sinclair*. Amsterdam: Benjamins. 157–176.
- Louw, W.E., 2000: Contextual Prosodic Theory: bringing semantic prosodies to life. C. Heffer in H. Sauntson (ur.): *Words in Context: A tribute to John Sinclair on his retirement*. Birmingham: University of Birmingham. 48–94. Dostopno na: http://www.revue-texto.net/docannexe/file/124/louw_prosodie.pdf (2. 10. 2012)
- Munday, Jeremy, 2011 (ur.): Looming large: A cross-linguistic analysis of semantic prosodies in comparable reference corpora. A. Kruger in K. Wallmach (ur.): *Corpus-Based Translation Studies: Research and Applications*. London in New York: Continuum. 169–86.
- Partington, Alan, 1998: *Patterns and Meanings: Using Corpora for English Language Research and Teaching*. Amsterdam in Philadelphia: John Benjamins.
- Partington, Alan, 2004: Utterly content in each other's company': semantic prosody and semantic preference. *International Journal of Corpus Linguistics* 9. 131–156.
- Philip, Gill, 2003: *Collocation and Connotation: A Corpus-Based Investigation Of Colour Words In English And Italian*. Doktorska disertacija. Velika Britanija: Univerza v Brightonu.
- Philip, Gill, 2009: Why prosodies aren't always there: Insights into the idiom principle. *Corpus Linguistics Conference. Liverpool. Great Britain*. Dostopno na: <http://ucrel.lancs.ac.uk/publications/cl2009/> (17. 10. 2011).
- Sinclair, John M., 1991: *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, John M., 1996: The Search for Units of Meaning. *TEXTUS* 9. 75–106.
- Sinclair, John M., 2004: *Trust the Text*. London in New York: Routledge.
- Stewart, Dominic, 2010: *Semantic prosody: a critical evaluation*. London in New York: Routledge.
- Stubbs, Michael, 1995: Corpus evidence for norms of lexical collocation. G. Cook in B. Seidlhofer (ur.): *Principle and Practice in Applied Linguistics*. Oxford: Oxford University Press. 245–56.

- Stubbs, Michael, 2001: *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.
- Šorli, Mojca, 2011: Pragmatic Components in the Slovene Lexical Database Meaning Descriptions. I. Kosem in K. Kosem (ur.): *Electronic lexicography in the 21st century: new applications for new users: proceedings of eLex 2011*, 10.-12. november 2011, Bled, Slovenija. Ljubljana: Trojina, zavod za uporabno slovenistiko. 251–259. Dostopno na: <http://www.trojina.si/elex2011/Vsebine/proceedings/eLex2011-34.pdf> (2. 10. 2012)
- Šorli, Mojca, 2012 (v pripravi): Corpus-Based Lexicographical Descriptions with a Special Focus on Pragmatics: The Case of the Slovene Lexical Database. S. Komar in U. Mozetič (ur.): *English Language Overseas Perspectives and Enquiries*. Ljubljana: Društvo za angleške študije.
- Tognini-Bonelli, Elena, 2001: *Corpus Linguistics at Work*. Amsterdam in Philadelphia: John Benjamins.
- Tognini-Bonelli, Elena, 2000: Functionally complete units of meaning across English and Italian: Towards a corpus-driven approach. Altenberg in Granger (ur.): *Lexis in Contrast*. Studies in Corpus Linguistics. Benjamins: Amsterdam and Philadelphia. 73–95.
- Vintar, Špela, 2009: Slovenski prevodoslovni korpus. M. Stabej (ur.): *Infrastruktura slovenščine in slovenistike (Simpozij Obdobja)*. Ljubljana: Znanstvena založba Filozofske fakultete. 385–391.
- Whitsitt, Sam, 2005: A critique of the concept of semantic prosody. *International Journal of Corpus Linguistics*, 10/3. 283–305
- Xiao, Richard in Tony McEnergy, 2006: Near synonymy, collocation and semantic prosody: a cross-linguistic perspective. *Applied Linguistics*, 27/1. 103–129.
- Zethsen, Karen K., 2006: Semantic Prosody: Creating Awareness about a Versatile Tool. *Tidsskrift for Sprogforskning*, 4(1). 275–294.

Imensko kazalo

A

Adam, Jean-Michel 123–124, 127, 140
 Agirre, Eneko 146, 163
 Aijmer, Karin 52
 Altenberg, Bengt 57, 62, 68, 205
 Alvstad, Cecilia 52
 Anderman, Gunilla 69
 Anscombe, Jean-Claude 123, 140
 Apothélos, Denis 116
 Arhar, Špela 16, 30
 Armstrong, Nigel 118
 Artale, Alessandro 146, 163
 Aston, Guy 153, 163
 Atkins, Sue 145, 163, 202–203
 Attridge, Derek 167, 177

B

Bajec, Anton 49, 125, 140
 Baker, Mona 35–36, 40, 49, 51–52, 55, 69, 121, 140, 159, 163, 185, 203–204
 Balažic Bulc, Tatjana 122, 140
 Barlow, Michael 101, 116
 Baroni, Marco 163
 Baumgarten, Nicole 36, 49
 Baumgartner Naylor, Sandra 9
 Becher, Viktor 35–37, 48–49, 99, 116
 Beeching, Kate 118
 Bentivogli, Luisa 146, 163
 Berber Sardinha, Tony 203
 Bernardini, Silvia 163
 Biber, Douglas 104, 116
 Bizjak, Kristina 144, 146
 Blum-Kulka, Shoshana 49
 Boch, Françoise 97, 116
 Bokal, Ljudmila 168, 177
 Bond, Francis 57, 68
 Breznik, Anton 71–72, 85, 91–92, 98, 107, 109, 116
 Brown, Dan 65
 Bublitz, Wolfram 181, 203
 Burbeck, Steve 27, 30

Burnard, Lou 153, 163

Butnariu, Cristina 178

C

Cankar, Ivan 72
 Carter, R. A. 169, 177
 Cerar, Vasja 9
 Černelič Kozlevčar, Ivanka 39, 50
 Cestnik, Tatjana 8
 Channell, Joanna 181, 203
 Chesterman, Andrew 41, 50, 52
 Chodorow, Martin 164
 Christ, Oliver 26, 30
 Claudiu, Mihăilă 36, 50
 Colfer, Eoin 151
 Combettes, Bernard 95–97, 113, 116–117
 Conrad, Susan 104, 116
 Cook, G. 204
 Cristea, D. 145, 164
 Crnkovič, Marko 134, 136, 139

D

Dam-Jensen, Helle 185, 203
 Danon-Boileau, Laurent 141
 Darbelnet, Jean 35, 53, 90, 93, 99, 118
 Darrieussecq, Marie 104, 134
 Dexter, Colin 151
 Dubois, Jean 97, 117
 Ducrot, Oswald 122–123, 140

E

Echenoz, Jean 104, 130
 Englund Dimitrova, Brigitta 50
 Erjavec, Tomaž 7, 9, 14–17, 20, 23, 30, 32, 71, 92–93, 101, 117, 151, 163, 170

F

Fabricius-Hansen, Cathrine 36, 50
 Fellbaum, Christine 146–147, 163
 Ferraresi, Adriano 153, 163
 Fillmore, Charles 203

Firth, John R. 181, 203
 Fišer, Darja 30, 33, 53, 144–146,
 148–149, 163–164
 Fligelstone, Steven 93
 Forner, Pamela 146, 163
 Francis, Gill 49, 68, 140, 181,
 203–204
 Frankenberg-Garcia, Ana 50
 Fuertes-Olivera, Pedro A. 63, 66, 68

G

Gabrovšek, Dušan 182, 184, 203
 Gadet, Françoise 118
 Gantar, Polona 203
 Gašparič, Špela 9
 Gaudé, Laurent 136
 Gerzymisch-Arbogast, Heidrun 52
 Giacalone Ramat, Anna 55–57, 61,
 68
 Giacomo, Mathée 117
 Gile, Daniel 50
 Giora, Rachel 183, 203
 Gorjanc, Vojko 16, 30, 33, 39, 47,
 50, 122, 140
 Grabnar, Katja 203
 Grad, Anton 42, 50, 125, 141
 Gradišnik, Branko 9, 45, 176
 Grad, Martin 9, 33, 50
 Grafström, Åke 57, 68
 Grahek Križnar, Nina 9
 Greenbaum, Sidney 69
 Greer, Derek 27, 31
 Grimbert, Philippe 104, 129
 Grum, Jožica 129
 Guespin, Louis 117
 Guibert, Hervé 128
 Gumul, Eva 50

H

Halliday, Michael A. K. 186, 203
 Hamel, Marie-Josée 53, 118
 Hanks, Patrick 145, 163

Hansen, Gyde 50, 52
 Hansen-Schirra, Silvia 37, 50
 Hao, Yanfen 169, 178
 Havu, Eva 97, 117
 Heffer, C. 204
 Hirci, Nataša 32–33, 50–51
 Hoey, Michael 183–184, 203
 Holst, June L. F. 170, 177
 House, Julianne 36, 49, 51
 Humar, Marjeta 163
 Hunston, Susan 99, 117, 181–182,
 203–204

I

Ide, Nancy 17, 31
 Ignat, Camelia 30, 117
 Ivir, Vladimir 169, 177

J

Jakop, Nataša 39, 51
 Javoršek, Jan Jona 30, 138–139
 Jesenšek, Marko 72–73, 92, 98, 117
 Johansson, Stig 50, 63, 68
 Johnson, C. R. 203
 Jordanous, Anna 168, 178
 Jurko, Primož 34, 43, 51, 182, 204

K

Kawaoka, Tsukasa 68
 Kenda, Jakob J. 9
 Kenny, Dorothy 170, 177–178
 Kilgariff, Adam 164
 Klaudy, Kinga 36, 48–49, 51–52
 Klinar, Stanko 34, 51
 Knowles, Gerry 93
 Kocjančič, Polonca 203
 Kolhatkar, Varada 146, 164
 Komar, Smiljana 51, 203, 205
 Koncut, Suzana 130, 132, 139
 Konšalová, Petra 36, 51
 Kosem, I. 205
 Kosem, K. 205

Kovač, Miha 8
 Kranjc, Simona 51–53
 Krek, Simon 15–17, 30, 200, 203
 Krisztina Karoly 48–49, 51–52
 Križaj Ortar, Martina 126–127, 142
 Kruger, A. 204
 Kujamäki, Pekka 52, 69

L

Lagane, René 97, 117
 Lakoff, George 145, 164
 Lambert, José 51
 Landes, Shari 164
 Laurent, Damien 97, 116
 Laviosa, Sara 63, 69
 Leacock, Claudia 164
 Ledinek, N 163
 Leech, Geoffrey 69
 Leeming, Henry 42, 50
 Lehmann, Christian 55, 69
 Lenassi, Nives 74, 93
 Lewandowska-Tomaszczyk, Barbara 184, 204
 Loffredo, E. 178
 Logar Berginc, Nataša 168, 178
 Louw, W.E. 181, 183–184, 204
 Lujić, Andra 98, 117

M

Magnini, Bernardo 146, 163
 Mahkota, Tina 9
 Makine, Andreï 104
 Malmkjær, Kirsten 50, 121, 141
 Marc Bratina, Karin 93
 Marcellesi, Christiane 117
 Marcelles, Jean-Baptiste 117
 Mauranen, Anna 51–52, 55, 69
 McCarthy, M.J. 169, 177
 McDaniel, Adam 27, 31
 McEnery, Tony 93
 Mele, Miha 9
 Mérot, Pierre 104, 138

Mével, Jean-Pierre 117
 Meyer, Bernd 36, 49
 Mezeg, Adriana 8, 9, 72–73, 93–94, 99, 100, 102, 117, 121, 127, 141
 Mikolič Južnič, Tamara 8, 9, 33, 49, 51, 70, 74, 91, 93
 Mikolič, Vesna 51, 52, 93
 Miller, George A 146, 164
 Milojević Sheppard, Milena 72, 93, 98, 116–117
 Moder Saje, Alenka 133–134, 139
 Mohar, Tjaša 136, 139
 Morel, Mary-Annick 124, 141
 Mozetič, Brane 128–129, 132, 139
 Mozetič, Uroš 51, 205
 Munday, Jeremy 185, 204

N

Neumann, Stella 50
 Neveu, Franck 97, 116–117
 Newmark, P 169, 178
 Noailly, Michèle 97, 117
 Novak, Jernej 149, 163

O

Ogura, Kentara 68
 Oksefjell, Signe 50
 Olohan, Maeve 51–52, 69, 141
 Øverås, Linn 36, 52
 Özçetin, Demet 36, 49

P

Pápai, Vilma 36, 52
 Partington, Alan 181, 183, 185, 204
 Pattison, Ann 168, 178
 Pavelin Lešić, Bogdanka 142
 Pedersen, Ted 146, 164
 Pellat, Jean-Christophe 97, 117
 Perteghella, M. 178
 Petit, Robert 122
 Philip, Gill 181–186, 194–199, 204

Pianta, Emanuele 146, 163
 Pinchon, Jacqueline 98, 118
 Pisanski Peterlin, Agnes 33, 52, 54,
 62, 69, 121, 141
 Pobirk, Olga 203
 Pokorn, K. Nike 33, 52–53, 62, 68,
 72, 92–93, 141
 Pollak, Senja 9
 Pouliquen, Bruno 30, 117
 Prunč, Erich 93
 Puurtinen, Tinna 37, 52
 Pym, Anthony 36–37, 48, 52, 55,
 69

Q

Quirk, Randolph 57, 69

R

Ragnarsdóttir, Hrafnhildur 56, 69
 Ramat, Paolo 68
 Rejc, Rok 203
 Reppen, Randi 104, 116
 Riccardi, Alessandra 93
 Riegel, Martin 97, 117
 Rioul, René 97, 117
 Rogers, Margaret 69
 Roma, Elisa 68
 Rossi-Gensane, Nathalie 95, 118
 Rowling, J. K. 65, 151
 Rundell, Michael 202

S

Sager, Juan C. 49, 53, 118
 Sagot, Benoît 148–149, 163–164
 Saldanha, Gabriela 36, 52
 Sansò, Andrea 55–57, 61, 68
 Sa, Shan 104
 Sauntson, H. 204
 Schjoldager, Anna 169–170, 178
 Schlamberger Brezar, Mojca 33, 52,
 72–74, 82, 91, 93, 120–124,
 127, 140–142

Schmid, Helmut 16, 31, 101, 118
 Schmidt, Douglas C. 27, 31
 Schubert, Klaus 51
 Seidlhofer, B. 204
 Shlesinger, Miriam 69
 Siemund, Peter 53
 Sijie, Dai 133
 Simeoni, Daniel 69
 Sinclair, John M. 7, 49, 71, 93, 99,
 118, 140, 163, 180–184, 204
 Škerlj, Ružena 50
 Skubic, Andrej 39, 53
 Sladanha, Gabriela 51
 Smith, Zadie 151
 Smolej, Mojca 38–39, 53
 Sohár, Anikó 51
 Šolar, Jakob 116
 Somers, Harold 49
 Šorli, Mojca 180, 203–205
 Soroa, Aitor 146, 163
 Sperber, Dan 112, 118
 Stabej, Marko 30, 53, 93, 118, 164,
 205
 Stamou, Sofia 145, 164
 Steinberger, Ralf 30, 117
 Steiner, Erich 37, 50, 53
 Stewart, Dominic 184–185, 204
 Strapparava, Carlo 146, 163
 Strömqvist, Sven 56, 69
 Štrukelj, Inka 141
 Stubbs, Michael 181, 184, 204
 Šuštaršič, Rastislav 72, 92–93
 Šuster, Simon 203
 Svartvik, Jan 69

T

Thomas, Robert G. 164
 Tirkkonen-Condit, Sonja 55, 63, 69
 Tognini-Bonelli, Elena 49, 140, 181,
 183–186, 204–205
 Tolkien, J. R. R. 45, 65, 151
 Tony McEnergy 185, 205

Toporišič, Jože 37–40, 53, 63, 69, 72,
90–91, 93, 96, 118, 124–125,
142
Toury, Gideon 53, 55, 69, 72, 93,
177
Trubar, Primož 33
Tufiş, Dan 145, 164
Tutin, Agnès 97, 116

V

van der Auwera, Johan 62, 69
Vanderauwera, Ria 35, 53
Veale, Tony 169, 178
Venuti, Lawrence 168–169, 178
Veronis, Jean 17, 31
Vear, Štefan 9
Viberg, Åke 56, 69
Vidovič Muha, Ada 178
Vilkou-Poustovaia, Irina 141
Vinay, Jean-Paul 35, 53, 90, 93, 99,
118
Vincenot, Caude 125, 142
Vincenot, Claude 125
Vintar, Špela 33–35, 50–53, 71,
74, 93, 101, 118, 121, 150,
163–164, 166, 205
Vogrinc Javoršek, Urša 9
Voršič, Ines 168, 178
Vossen, Piek 145, 164

W

Wagner, Robert-Léon 98, 118
Wallmach, K. 204
Werner, Koller 51
Whitsitt, Sam 182, 205
Wichmann, Anne 93
Wilson, Deirdre 112, 118

X

Xiao, Richard 141, 205

Z

Žagar, France 72, 93
Žagar, Igor Ž. 127, 142
Zanchetta, Eros 163
Zaranšek, Petra 203
Zethsen, Karen K. 183, 185, 203,
205
Žganec Gros, Jerneja 30, 93



Univerza v Ljubljani
FILOZOFSKA
FAKULTETA

