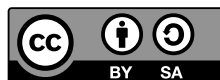


Mojca ŠorliFilozofska fakulteta, Univerza v Ljubljani
mojca.sorli@guest.arnes.si

UDK 811.163.6'374.82:81'322

DOI: 10.4312/vestnik.9.59-81



SPLETNI VIRI IN ORODJA ZA SLOVENŠČINO S POUČENOM NA SLOVENSKO-ANGLEŠKI INFRASTRUKTURI: SLOVARJI V DIGITALNI DOBI

0 UVOD

Kakor je jezikovna situacija slovenskih govorcev in govork posebna zaradi njihovega majhnega števila ter posledično v stiku s tujimi jeziki izrazitega nesorazmerja med številom uporabnikov izhodiščnega in ciljnega jezika, je hkrati tudi univerzalna in določena z mnogimi korenitimi spremembami v pogledih na načrtovanje, uporabo in dostop do jezikovnih virov. Na tem mestu se ne bomo posvečali normativnim priročnikom,¹ saj nas zanimajo predvsem leksikalno-leksikografski viri, v katerih – tipično kot prevajalci – dostopamo do podatkov o pomenu in rabi, torej slovarji, leksikalne podatkovne zbirke in korpusi, ki jih danes izdelujemo s pomočjo jezikovnih tehnologij. Le v manjši meri se v prispevku posvečamo enojezičnim priročnikom za slovenščino, sorazmerno največ pa zaradi aktualnosti angleščine dvojezičnim virom za par slovenščina-angleščina.

Da so digitalni viri relativizirali stare slovarske tipologije, potrjujejo dileme, s katerimi smo se soočali pri organizaciji poglavij, saj se je pokazalo, da teh ni več mogoče oz. smiselno zarisovati v skladu s tradicionalnimi ločnicami med posameznimi kategorijami, npr. eno-, dvo- in večjezični slovarji, temveč predvsem glede na potrebe in morebitne zagate uporabnikov široke palete sodobnih virov na spletu. V skladu z novimi pogoji so vzpostavljena razmerja na eni strani med leksikalnimi in besedilnimi viri ter na drugi med prevajalskimi priročniki oz. viri in orodji.

1 TEMELJNI PREVAJALSKI PRIROČNIKI

Uporabniki klasičnih tiskanih slovarjev smo spoznali prednosti elektronskih oz. prostodostopnih spletnih jezikovnih virov, ki jih je za tuje jezike, zlasti angleščino, na voljo

¹ Takšni priročniki, tipično pravopis, praviloma obravnavajo predvsem tiste segmente jezika, ki so še posebej rezultat konvencij, npr. raba velike in male začetnice ter ločil, zapis datumov ipd., ter nekatera oblikoslovna in besedotvorna vprašanja, medtem ko naj bi pomenske podatke in rabo beležili zlasti slovarji in leksikalne podatkovne zbirke.

že lepo število (npr. Macmillan,² Longman,³ Collins Cobuild,⁴ Cambridge,⁵ Merriam-Webster,⁶ Dictionary.com,⁷ Yahoo's Dictionary⁸ itd.).⁹ Ni naključje, da gre hkrati za najboljše in najuspešnejše angleške pedagoške slovarje oz. slovarje za tujce z dolgo tradicijo – z izjemo zadnjih dveh, ki sta ameriška slovarja za domače govorce –, od katerih so se v zadnjih desetih letih vsi pojavili na trgu v prenovljenih in dopoljenih izdajah. Večina spletnih virov se vsaj v določenem deležu financira z oglaševanjem. Poleg virov, ki jih na spletu ponujajo posamezne založbe, se je vzpostavilo večje število spletnih portalov, ki skušajo uporabnikom ponuditi čim širšo izbiro jezikovnih virov na enem mestu.¹⁰ Poleg številnih prednosti, kot je brezplačnost, imajo spletni priročniki tudi slabosti: pogosto v njih ni mogoče najti osnovnih informacij o npr. avtorskih pravicah, založbi, datumu izdaje, uredniških ekipah, prav tako ne navodil za uporabo itd. (Gabrovšek 2010: 132). Kaže, da je za prevajalca najzanesljivejša pot do ustrezne informacije kombinacija spletnih virov in tiskanih izdaj oz. njihovih izdaj na CD ali DVD-ROMU (ibid.).¹¹

Tudi nekatere sodobne raziskave o rabi slovarjev (npr. Müller-Spitzer 2012, 2014) kažejo, da kljub vedno večjim količinam prostodostopnih jezikovnih podatkov na spletu uporabniki še vedno cenijo strukturiranost podatkov, značilno za klasične slovarje. Poleg tradicionalnih priročnikov in virov zavzemajo v prevajalski praksi vse pomembnejše mesto tudi številni programi in orodja za »avtomatizirano prevajanje« (tukaj v rabi kot skupno poimenovanje za strojno prevajanje in računalniško podprte prevajalske (CAT) sisteme, gl. 1.2.1 in 1.2.2 spodaj), v zadnjem času pa se širi tudi ponudba storitev v oblaku, ki je posledica večanja tehnoloških kapacitet (Vintar 2016: 28).

1.1 Elektronski in spletni slovarji

Poleg programov za avtomatizirano prevajanje (gl. 1.2.1) so druga večja skupina pripomočkov slovarji, pri katerih načeloma vzpostavljamo razliko med elektronskimi različicami obstoječih tiskanih verzij ali izdaj na CD-ROMU oz. DVD-ROMU (tudi na

2 *Macmillan English Dictionary for Advanced Learners* (Rundell 2007), 2. izdaja.

3 *Longman Dictionary of Contemporary English* (Summers 2003), 4. izdaja, 5. izdaja Mayor 2009.

4 *Collins COBUILD Advanced Learner's English Dictionary* (Sinclair 2006), 5. izdaja.

5 *Cambridge Advanced Learner's Dictionary* (Walter 2008), 3. izdaja.

6 *Merriam-Webster's Advanced Learner's English Dictionary* (Perrault 2008).

7 Dictionary.com temelji na *Random House Dictionary*, a navaja tudi druge vire, zlasti ameriške angleščine; slovar je mogoče brezplačno naložiti na Applov iPhone in druge mobilne platforme: <http://www.dictionary.com>.

8 Yahoo's Dictionary ponuja pretežno vsebine *American Heritage Dictionary of the English language*: <http://education.yahoo.com/reference/dictionary>.

9 Za podroben pregled in vrednotenje eno- in dvojezičnih slovarjev, ki vključujejo angleščino, in izčrpno povezano bibliografijo gl. Gabrovšek (2010: 121–143).

10 Za spletno bibliografijo zlasti izbranih angleških virov, gl. Gabrovšek (2010: 143–146).

11 Nekatere založbe se kljub razširjenosti in dobri sprejetosti spletnih slovarjev zaenkrat še odločajo za vzporedne knjižne izdaje, tako na primer vsi navedeni najzgodnejši enojezični angleški slovarji obstajajo tudi v papirni obliki, razen Macmillana, ki od leta 2012 izhaja samo v spletni obliki.

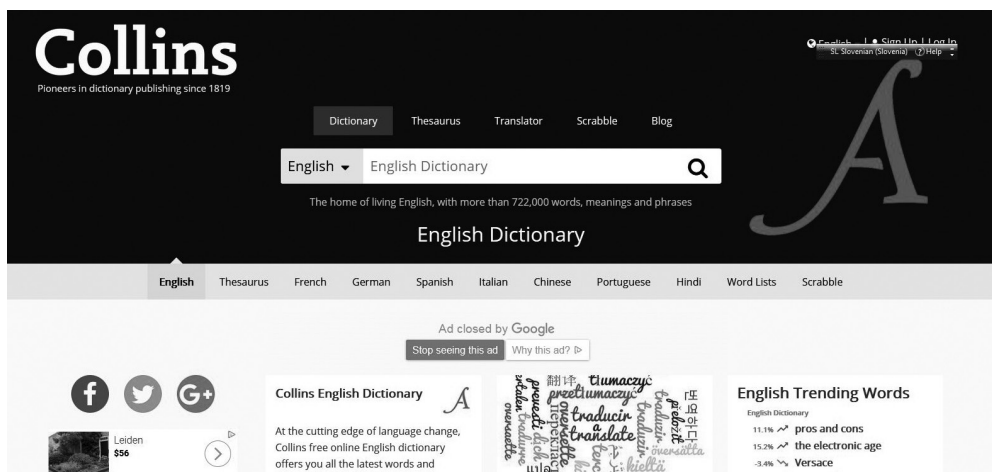
mobilnih telefonih), torej elektronskimi slovarji, ki nastanejo zgolj kot produkt digitalizacije oz. (polnega ali delnega) prenosa knjižnega v digitalni medij, npr. *Slovar slovenskega knjižnega jezika* (brezplačen dostop) ali *Veliki angleško-slovenski slovar Oxford-DZS* (plačljiv dostop), ter t. i. spletnimi slovarji, ki so izvorno izdelani za digitalni medij in ki z ustreznimi uporabniškimi vmesniki omogočajo najhitrejšo in najučinkovitejšo izrabo raznovrstnih podatkov, vse bolj pa so tudi vsebinsko prilagojeni (npr. *Macmillan English Dictionary Online*, *Collins COBUILD Advanced Dictionary*). V praksi so te meje še precej zabrisane in celo spletne različice najuspešnejših angleških spletnih slovarjev naj bi še vedno pretežno izhajale iz tiskanih izdaj.¹²

1.1.1 Angleščina

Zlasti za angleščino je danes na voljo vrsta ne le posameznih spletnih slovarjev, temveč tudi spletnih portalov, ki ponujajo dostop do raznovrstnih jezikovnih virov, na primer splošnih eno- in dvojezičnih slovarjev, pa tudi enciklopedij, delov Wikipedije, glosarjev, terminoloških slovarjev itd. Čeprav so lahko takšni viri zelo koristni, je potrebna dobršna mera kritičnosti pri njihovi uporabi, saj so pogosto dokaj površinski, vključujejo zastarele vire ali zgolj dele posameznih virov, zato niso primerni za sistematično uporabo za prevajalske potrebe (Gabrovšek 2010: 113). Najkvalitetnejše spletne vire lahko izdelujejo seveda založbe same; tak primer je **CollinsDictionary** <<http://www.collinsdictionary.com/>>, ki ponuja dostop do nekaterih svojih najboljših angleških slovarjev (npr. COBUILD Advanced English Dictionary, Collins English Dictionary, oba © HarperCollins Publishers), poleg tega pa še tezaver oz. slovar sinonimov, več dvojezičnih slovarjev, spregatvene paradigme glagolov v angleščini in tujih jezikih, slovnična poglavja, definicije in tudi prevajalni sistem (podpira ga Microsoftov prevajalni sistem) za največje svetovne jezike, vse to z visoko dodelano funkcionalnostjo in vizualno podobo.

Sodobni slovarski portali praviloma povezujejo številne možne vidike uporabe jezikovnega priročnika, ki lahko zadovolji najrazličnejše profile uporabnikov, tako zahtevnejše kot tudi laične, npr. z besednimi igrami (Scrabble), poljudnimi članki o jezikih in kulturah, blogi itd. Številni ponujajo uporabnikom možnost aktivnega vključevanja, tipično tako, da ti prispevajo predloge za nova gesla (ang. UGC ali User-generated Content). Eden zdaj že uveljavljenih tovrstnih projektov je ameriški **Wordnik** <<http://www.wordnik.com/>>, ki po pričakovanjih kaže, da uporabniki radi prispevajo predvsem redke besede ali komentirajo nenavadne rabe besed, in tako potrjuje opažanja, da dajejo splošni uporabniki večji pomen partikularnemu kot tipičnemu in pogostemu (npr. Atkins in

12 »Why Publishers are Unable to Develop a Really Good Software Dictionary« je razmislek o prenovljeni 4. izdaji slovarja *Collins COBUILD Advanced Learner's English Dictionary* na <<http://www.antimoon.com/other/goodsoftdic.htm>> (dostop: 10. nov. 2013), ki govori o vprašanih načrtovanja spletnih virov na osnovi njihovih izvornih tiskanih izdaj. Klasičen primer (retrogradnega) prenosa vsebin tiskanih izdaj v spletno okolje je spletišče Fran, vzpostavljeno leta 2014, na katerem je mogoče dostopati do vseh temeljnih del ISJFR ZRC SAZU (SSKJ, Slovar slovenskih frazemov, Slovar sinonimov, Slovenski etimološki slovar, različni terminološki slovarji itd.).



Slika 1: Spletna stran CollinsDictionary <<http://www.collinsdictionary.com/>> (julij 2017)

Rundell 2008: 52). Pozitivna plat podobnih projektov je zlasti v tem, da popularizirajo leksikografske vire in postopke. (V slovenskem prostoru je zgleden primer dobre prakse pri izrabi moči množic **Razvezani jezik** <<http://razvezanijezik.org>>, ki je v devetih letih postal javno prepoznaven projekt s konkretnimi leksikografskimi rezultati – dvema knjižnima izdajama (eni elektronski in eni tiskani) izbora približno 200 geselskih člankov, med katerimi prevladuje sleng oz. »poulični«¹³ jezik.) Hkrati ne gre spregledati tehnoloških pobud, kot je večjezični spletni slovar **Glosbe** <www.glosbe.com>, ki nastaja ob delni podpori množic z ambicijo na enem mestu omogočiti dostop do glosarjev in mednarodnih večjezičnih baz podatkov s pomočjo pomnilnikov prevodov. Dobra plat takšnih tehnoloških projektov je njihova potencialna uporabnost, slaba pa nizka jezikovna ozaveščenost, ki se kaže na primer v uporabi strojno prevedenih nagovorov uporabnikov v različnih jezikih.

1.1.2 Slovenščina

1.1.2.1 Splošni (dvojezični) slovarji

V skupino spletnih slovarjev za slovenščino uvrščamo tudi brezplačne vire na straneh, kot sta **slovarji PONS** <<http://en.pons.eu>> in **Spletni slovar** <<http://www.spletni-slovar.com/>>, ki črpajo podatke iz slovarskih podatkovnih zbirk in katerih večjezične zbirke

13 Treba je izpostaviti sicer morda delno predvidljivo sociolingvistično dejstvo, da gre v velikem deležu geselskih prispevkov za besedišče in govor, ki vsebujeta spolno označene dele telesa, spolne navade, usmerjenosti itd., predvsem skozi frazeologijo, v kateri se pojavljajo spolno označena poimenovanja, pogosto kletvice in slabšalni ali seksistični izrazi.

običajno združujejo obojesmerne slovarje za osrednje evropske jezike (nemščina, francoščina, španščina, italijanščina itd.), v njih pa je mogoče iskati ustreznice posameznih besed, v redkih, npr. PONS, tudi besednih zvez. Takšnim virom so pogosto dodana spremna gradiva ali orodja, npr. črkovalniki, vaje iz slovnice in besedišča, tabele spregatev ali druge posebne priloge. Takorekoč univerzalna pomanjkljivost sicer pogosto jezikovnotehnološko naprednih virov je poleg tega, da ponujajo ozek nabor kontekstualnih omejitev oz. okoliščin rabe navedenih prevodov (te mora prepoznati uporabnik sam), tudi ta, da opisi slovenščine kot izhodiščnega jezika niso nastali na osnovi proučevanja dejanske rabe, temveč so največkrat prevzeti iz starejših, že obstoječih dvojezičnih virov, kar onemogoča ustrezno in izčrpno obravnavo problemov, s katerimi se slovenski govorniki in prevajalci najpogosteje soočajo.

1.1.2.2 Terminološki viri

Kot ugotavlja Gorjanc (2010), je terminoloških slovarjev in glosarjev v slovenskem prostoru precej. Enojezični v veliki meri nastajajo v okviru Terminološke sekcije Inštituta za slovenski jezik Frana Ramovša ZRC SAZU. Težava, ki se postopoma odpravlja, je, da so dokaj nepovezani, poleg tega pa tudi neenotno grajeni in javnosti nedostopni.¹⁴ Eden najbolj izčrpnih, praktičnim potrebam prevajalcev najbližjih slovenskih spletnih portalov je vladni **Spletni slovarji** <<http://evroterm.gov.si/slovar/>>, kjer lahko dostopamo do okrog 900 splošnih in terminoloških virov, vključno z osrednjimi jezikovnimi priročniki za slovenščino, kot sta *Slovar slovenskega knjižnega jezika* (SSKJ) in *Slovenski pravopis*. Med njimi so tudi eno-, dvo- in večjezični korpusi, številni področni slovarji in glosarji, ter povezave na jezikovne tehnologije, kot so (prostodostopni) pomnilniki prevodov, in mnoge druge relevantne vire. Med slovenskimi terminološko-splošnimi spletnimi viri je slovarski portal **Termania** <www.termania.net>, ki je namenjen (hkratnemu) iskanju po več slovarskih zbirkah, omogoča pa tudi urejanje slovarskih zbirk oz. gesel. Vanj je vključenih več Amebisovih brezplačnih terminoloških in večjezičnih slovarjev, pa tudi na primer *SSKJ*, *Slovar sopomenk* in *Slovenski pravopis*. Med brezplačnimi prevajalskimi podatkovnimi zbirkami v slovenskem okolju po kakovosti in obsegu izstopa **Evroterm** <<http://evroterm.gov.si/>>, ki temelji na prevodih evropske zakonodaje. Zbirka, ki podpira 16 jezikov in se nenehno dopolnjuje, ponuja danes spletne povezave tudi do drugih podatkovnih zbirk in virov, npr. IATE,¹⁵ Termanie, SSKJ, korpusa Nova beseda, dvojezičnih slovarjev PONS itd., v njenem okviru pa so dostopni tudi terminološki Evrokorpus <<http://evrokorpus.gov.si/>> (gl. tudi 1.3.2 in Željko 2010), že omenjeni Spletni slovarji in Terminator (»terminološki analizator«, ki avtomatsko označi tiste dele v besedilu, ki so v terminološki zbirki Evroterm). Prav terminologija sodi zagotovo med tista področja, ki

¹⁴ Tu beležimo spremembe na bolje, saj je leta 2014 Inštitut za slovenski jezik FR lansiral slovarsko spletišče Fran, kjer je mogoče dostopati do splošnih, zgodovinskih in nekaterih terminoloških slovarjev, izdelanih v okviru inštituta.

¹⁵ Portal Interactive terminology for Europe.

so največ pridobila z uporabo prevajalskih namizij, tj. spletnih prevajalnikov in še zlasti pomnilnikov prevodov.

1.2 Prevajalske tehnologije

1.2.1 Pomnilniki prevodov

Pomnilniki prevodov so zavzeli trg profesionalnega prevajanja pred dobrim desetletjem in krepko spremenili način dela, a za razliko od strojnih prevajalnikov niso odločilno posegli v vlogo človeka pri nastajanju ciljnega besedila (Vintar 2016: 77). Pomnilnik deluje po načelu vnovične uporabe prevodov (ibid.: 17) in daje najboljše rezultate pri prevajanju novejših različic besedil, ki so bila že prevedena v starejši različici, ki jih je proizvedel človek, pri čemer za optimalno pomoč predpostavlja izgradnjo prevajalčeve lastne podatkovne baze. Prevajanje s pomnilnikom prevodov temelji na predpostavki, da je za določene formulaične oz. standardizirane tipe besedil značilno in tudi zaželeno ponavljanje jezikovnih segmentov. Pomnilniški programi so zato uporabni zlasti pri prevajanju tehničnih oz. visoko terminoloških besedil (ibid.: 19). V tem rangju najdemo vrsto ponudnikov, ki omogočajo iskanje po dvo- in večjezičnih podatkovnih zbirkah, npr. **Eur-lex** <eur-lex.europa.eu> (dostop do celotne zakonodaje EU), **Webitext** <<http://www.webitext.com/bin/webitext.cgi>> (evropska zakonodaja, finance in pravo z omogočenim vzporednim prikazom celotnih dokumentov), **TAUS** <<https://www.taus.net/>> (pravo, farmacevtika in biotehnologija), **MyMemory** <<http://mymemory.translated.net/>> in **Linguee** <www.linguee.com>. Čeprav nekateri integrirajo storitve pomnilnikov in strojnega prevajanja, je za boljše razumevanje sprememb, ki jih tehnološki napredek uvaja v prevajalski proces, nujno razlikovati med računalniško podprtim in strojnim prevajanjem.¹⁶

1.2.2 Strojni prevajalniki

Ocenjevanje vloge avtomatiziranega prevajanja je lahko zavajajoče, če enačimo različne prevajalske tehnologije¹⁷ in obenem ne definiramo besedilnih zvrsti, ki jih prevajamo z njihovo pomočjo. Dejstvo je, da so ciljno zasnovani prevajalni sistemi (npr. MT@EC) in pomnilniki prevodov nepogrešljivi pri prevajanju standardiziranih tipov besedil; na nekaterih področjih znanosti, na primer v družboslovju in humanistiki, ter vseh oblikah

16 Za podrobnejši opis razlik gl. npr. <http://www.fxm.ch/En/Langues-Traduction/TraductionOrdinateur.en.htm>.

17 "Apart from this function of providing a rough idea of what a given text contains and its utility when deciding whether or not this would be worth translating, MT is only efficient where applied to texts with an appropriate degree of standardisation and coherence. In short, a text that can be translated by a computer must be written in a way that the computer can understand it: there must be no ambiguity, and it must contain only terms contained in the computer's dictionary and which always have the same meaning. This type of controlled language - which imposes major constraints on writers - has few areas of use beyond that of particular types of technical documentation that are sufficiently voluminous to justify the investment" (<http://www.fxm.ch/En/Langues-Traduction/TraductionOrdinateur.en.htm>, dostop 4.7.2017).

prevajanja, kjer igrata vlogo kreativnost in izvirnost, še posebej seveda v literarnem prevajanju, pa so zaenkrat uporabni precej manj. Če drži, da »skoraj 90 % prevajalskega posla predstavljajo tehnična besedila« (Vintar 2016: 17), je bilo v zadnjih desetletjih zaznati drastični porast prav v deležu besedil, ki so posledica oblikovanja informacijske družbe in globalne ekonomije (digitalizacija, deklaracije, lokalizacija ipd.). Ne glede na to osnovne zakonitosti prevajanja, med drugim upoštevanje socialne in funkcijske zvrstnosti besedil, pragmatični ter vsi drugi izzivi, ki so jih raziskovali mnogi prevodoslovni teoretiki (npr. Nida 1984, 1996,¹⁸ Newmark 2000, Baker 2006, 2009,¹⁹ Bassnett 1980²⁰ in mnogi drugi), obstajajo naprej in terjajo temu primerno raznorodnost prevajalskih pristopov. Kljub nespornemu pomenu spletnih jezikovnih virov in orodij je postal za sodobnega uporabnika največji izziv prav dostop do zanesljivih in relevantnih podatkov. Zlasti prostodostopni prevajalni sistemi na spletu temeljijo na t. i. direktnem pristopu (beseda za besedo), npr. **Yahoojev Babel**, kvalitetnejša in plačljiva orodja pa večinoma na t. i. prenosu (tekst na osnovi pravil) (Ooi 2010: 150). Spletni prevajalniki, denimo **Google (Translate)** (več kot 60 jezikov), so nekakšna priučena orodja, vseskozi izboljševana s posebnimi tehnološkimi postopki, ki temeljijo večinoma na statističnem pristopu. Trenutno najboljši brezplačni spletni prevajalniki naj bi bili poleg Googla in Babela naslednji: **Bing Translator** (Microsoft®, 35 jezikov), **Babylon** (30 jezikov) in **World Lingo** (141 jezikov). Od teh ponujajo nekateri tudi plačljive profesionalne, torej človeške prevajalske storitve, npr. Babylon in World Lingo. Med prevajalniki, ki vključujejo slovenščino, je treba omeniti še **ImTranslator** (okrog 50 jezikovnih parov), do katerega imajo dostop registrirani uporabniki, vsebuje pa tudi črkovalnik in večjezični slovar, **Tradukka, Dictionary.com Translator, Intertan**, portal **itranslate4.eu** ter slovenska **Amebis Presis** in **Brezplačno prevajanje** <<http://www.brezplacno-prevajanje.si/>> - slednji zgolj omogoča dostop do nekaterih navedenih prevajalnikov.

Tu so še spletni portali tipa **Lexicoool.com**, ki delujejo kot direktoriji prostodostopnih spletnih dvo- in večjezičnih slovarjev ter glosarjev in, podobno kot omenjeni Glosbe, pogosto vključujejo strojni prevajalnik. Prevajalniki imajo seveda povsem drugačno teoretično podstavo in tehnično izvedbo kot slovarji. Cilj prevajalnikov še vedno ni absolutna točnost, temveč relativna razumljivost, na kar izdelovalci zlasti nekomercialnih prevajalnikov, denimo MT@EC,²¹ običajno tudi opozorijo. Nesporni koristnosti in hitremu napredku storitev strojnega prevajanja navkljub je torej smiselno pristopiti k spletnim prevajalnikom razmeroma kritično; dobro se je zavedati, da gre v pretežni meri za

18 Nida, E. *Translating Meaning*. San Dimas, Calif.: English Language Institute, 1982; *Sociolinguistics of Interlingual Communication*, Brussels: Editions du Hazard, 1996.

19 Baker, M. (ur.) *Critical Readings in Translation Studies*, 2009; *Translation and Conflict: A Narrative Account*, 2006.

20 Bassnett, S. *Translation Studies*, 1980.

21 "Our machine translation service produces raw automatic translations. Use it to grasp the gist of a text or as the starting point for a human-quality translation. If you need a perfectly accurate, high-quality translation, the text still needs to be revised by a skilled professional translator." (https://ec.europa.eu/info/resources-partners/machine-translation-public-administrations-mtec_en/, dostop 27. 11. 2017)

propulzivno komercialno panogo, katere cilj je prodaja storitev, manj pa za področje, ki bi ga še obvladovalo prevodoslovje ali katera druga od jezikoslovnih akademskih disciplin. Obenem je treba priznati, da je razvoj in napredek prevajalskih tehnologij izjemno hiter. Trenutno stanje kaže dokaj visoko stopnjo ustreznosti na ravni leksikalnih izbir (besede in besedne zveze, ang. phrase-based MT), vendar ostaja, kadar frazne ustreznice ni, načelo strojnih mehanizmov prevajanje »beseda za besedo ali besedno zvezo«, brez ustreznega prepoznavanja zakonitosti v skladnji, pragmatiki in besedilnih komponentah obdelovanih besedil ciljnega jezika. Če povzamemo, prevajalniki besedil pomagajo 's pridržkom,' namreč pod pogojem, da uporabniki dokaj dobro poznajo izhodiščni in ciljni jezik: tem uporabnikom lahko pomagajo celo zelo učinkovito. Nikakor pa v tem trenutku ni mogoče strojnih prevajalnikov priporočati tistim, ki se ciljnega jezika šele učijo ali ki imajo le osnovno oz. pasivno znanje ciljnega jezika. Povedano drugače, za uspešno uporabo vseh oblik avtomatiziranega prevajanja je nujno potrebno kvalitetno prevajalsko znanje in spretnosti, ki pa jih ni mogoče pridobiti zgolj ali predvsem s pomočjo omenjenih prevajalskih orodij.

Dvojezičnih besedilnih zbirk je tako več vrst: poleg pomnilnikov prevodov imamo na voljo še vzporedne korpuse (gl. 1.3.2) in sodobne elektronske slovarje, vse od teh zbirk pa imajo svoje specifične, prednosti in omejitve. V nadaljevanju bomo najprej na kratko predstavili splošne slovenske enojezične, nato pa še vzporedne korpuse besedil – ti so večinoma specializirani.

1.3 Prevajalski besedilni viri

V nadaljevanju si bomo na kratko ogledali najbolj razširjene javno dostopne korpuse, za podrobnejši pregled eno- in večjezičnih korpusov za slovenščino gl. tudi Vintar (2008).

1.3.1 Enojezični korpusi izvirnih besedil

Med temi naj izpostavimo prostodostopna enojezična slovenska korpusa, **Nova beseda** <http://bos.zrc-sazu.si/nova_beseda.html> in še zlasti največji, referenčni korpus **Gigafida** <<http://www.gigafida.net/>> (Erjavec 2012; Logar Berginc 2012; Arhar Holdt et al. 2012). Ker mora biti korpus, ki je načrtovan s ciljem prikazovati celovito podobo jezika, primerno taksonomsko uravnotežen, je bil iz korpusa Gigafida izpeljan še reprezentativni 100-milijonski podkorpus **KRES** <<http://www.korpus-kres.net/>> (Logar Berginc 2012; Arhar Holdt et al. 2012).²² Tako Gigafida kot KRES imata tudi manjši različici, namreč cc-Gigafida (100 milijonov) in ccKRES (10 milijonov), ki sta prosto dostopni ne le prek konkordančnika, ampak v celoti, kot podatkovni bazi, zlasti za jezikovnotehnološke namene.

²² Korpusa sta nastala v okviru projekta *Sporazumevanje v slovenskem jeziku*, www.slovenscina.eu. Več na: <http://www.slovenscina.eu/korpusi/gigafida> in <http://www.slovenscina.eu/korpusi/kres>.

◀ prejšnja stran 1 2 3 4 5 6 7 8 9 10 11 naslednja stran ▶

Prikazujem 21-40 od 9.376 konkordanc (2.137 sekund.)

Osnovne oblike	Prevajaj.
▶ prevajati (9.376)	
Vrsta besedila	
▶ Časopisi (4.642)	Kot nosilec signala načeloma lahko uporabljamo praktično vse, kar prevaja električni tok; od radiatorjev in telefonskih vodov do elektroenergetskega
▶ Revije (2.082)	sem prebral, sem resnično užival, besedil nisem nikoli prevajal z namenom, da bi zanj iskal založnike. Tako
▶ Svama besedila (1.217)	danes nadomestili prav priročniki. Nekateri ponudniki priročnike sestavljajo ali prevajajo sami, mnogi pa so se odločili račununalni priložiti raje
▶ Internet (1.083)	za vojne zločine na območju bivše Jugoslavije in ga dvakrat prevajali bosanskim Srbom. Neodvisna organizacija je po teh prevadjanih opravila
▶ Leposlovje (271)	. Dodajmo še, da so številni literati Joycea tudi prevajali, npr: Philippe Sollers, Cesare Pavese, Pablo
▶ Več	35 pesniških zbir, esejev zgodbo in kritik ter prevajal dela Samuela Becketta, Henrika Michauxa in Jacquesa Préverta v
Vir	...da, sitnosti so, kadar ne prevajaš zelo dobrega avtorja ali kadar avtor dela napake. Ali
▶ drugo (2.434)	v sodobnih prevodih starih ruskih kronik izraz Sloveni skoraj vedno prevajajo kot Slovani, pridevnik slovenski pa skoraj vedno kot slovenski
▶ Delo (1.903)	telenovel na televiziji (smeh), sicer pa mi prevaja soigralka Tanja Medved.
▶ Dnevnik (1.509)	To pomeni, da " prevajam " splošne odgovore na splošna vprašanja v teorijo in prakso
▶ DZS (423)	ta obstaja, na tujem uspelo, saj jo izdatno prevajamo tudi Slovenci.
▶ Mladina (423)	Zakonom. Ko sem mu razlagal, da v Sloveniji prevajamo knjige, ne da bi avtorji za to vedeli,
▶ Več	"Nekoliko sem prevajala reškega Italijana Enrica Morovicha, njegove kratke zgodbe in odlomke
Leto	Frequently Asked Questions (FAQ) prevajamo kot pogosta vprašanja. Kratica označuje dokument, kjer so
▶ 2008 (976)	-Depro, za to je poskrbel založnik že prevajajo v angleščino -, potem pa bom napisala le še
▶ 2010 (877)	Magali), 50, 11 snov, ki slabo prevaja električni tok, nepravdnik, 52, 46 ime inašega
▶ 2003 (753)	
▶ 2007 (698)	
▶ 2001 (674)	
▶ Več	

Slika 2: Del konkordančnega niza za »prevajati« v referenčnem korpusu Gigafida.

V letu 2014 smo dobili tudi 2. različico korpusa slovenskih besedil slWac, ki zajema izključno besedila s svetovnega spleta in je izgrajen po vzoru podobnih korpusov za nekatere druge evropske jezike, npr. ukWac (britanska angleščina), deWac (nemščina), itWac (italijanščina), frWac (francoščina) itd. <http://nl.ijs.si/noske/all.cgi/first_form?corpname=slwac;align=> (Erjavec in Ljubešić 2014).²³

user: defaults corpus: slWac 2.0 (Slovene Web, Version 2) Search napaka

Concordance Word List Corpus Info

Query napaka 195,585 (155.3 per million)

Page 1 of 1,956 Go Next Last

www.lisac.si	nalogo porabili trikrat več časa in še kakšno napako bomo naredili. f Današnji nasvet? f Poskrbite
www.lisac.si	postopno dogaja, da znova in znova delamo iste napake . Tudi mi pogosto uporabljamo poslovno strategijo
www.lisac.si	nima nobene knjige o pasji vzgoji. f Velika napaka , ki jo delamo prodajalci, je, da poskušamo
www.lisac.si	poskušamo prodati nekaj, kar nekdo potrebuje. NAPAKA . Prodajati moramo tisto, kar nekdo ŽELI
www.postojna-wireless.si	na določeni postrani pojavi obvestilo o napaki . Ali je to obvestilo povezano z omenjanimi
www.nd-crnuce.si	zaradi nezbranosti v napadu nismo zadeli. Napaka v naši obrambi je obrodila sadove in po
www.lisac.si	učim od njega. f Direktni marketing - osnovna napaka f Tole vidstvo smo dobili večeraj v časopisu
www.lisac.si	najboljšim se večini zgodijo takšne oznome napake . Rešitev! Naslednje naj vsaj izdelek pregledajo
www.magnetit.si	na spletnih straneh pojavijo sporočila o napaki . Ti plikotbi ne zbirajo informacij, ki
www.drustvo-dnk.si	jokom, ko dobi slabo oceno ali ko naredi napako , o kateri bi bilo potrebno obvestiti starše
www.drustvo-dnk.si	neupravičeno sredstvo, ne glede na to, kakšno napako je otrok naredil. Poudariti moramo, da
www.dvignagib.si	ki zagotavljajo delovanje višičarja brez napak f * ker vam omogoča enostavno upravljanje
www.nd-crnuce.si	gostov je dominantnejše pričela dvojab a je po napaki Ekipe A hitro prejela zadetek. Nato je
www.merolovje.si	tip D, različni sistematični in naključni napaki , do tistih katerim se da nastavljati volumen
www.lisac.si	niste res izkušeni, na odru naredili kakšno napako ... f Pa se vrnimo k decembrskemu ljubljanskemu
www.nd-crnuce.si	je v isti minuti igralcem Krima, po grobi napaki ožje obrambe, uspelo znižati izid, vendar
www.nd-crnuce.si	obrodil sadove v sredini drugega polčasa po napaki obrambe, ko domačemu napadalcu ni bilo
www.macka.si	naprodaj. So izjemno lepo ohranjene in brez napak , pack, raztrganin, ... Hranjene so v temi
www.naprejslovenija.si	preobrazbo, da ne moremo samo eni odgovarjati za napake pri svojem delu, ampak da tudi izvoljeni
www.alla.si	preskočila popoldansko spanje. Kapitalna napaka . Prve kilometre sem še slabha sodelovala
www.servotrzi.si	prav v ničemer ne odstopajo od izvirnikov. Napake se pokažejo šele, ko so zamenjani deli
www.great.si	teh spletnih stranih ali za kakršnekoli napake ter pomanjkljivosti v njihovi vsebini. f
www.finiziv.si	dopužča, da bi se podjetja učila na svojih napakah , saj so kazni za napake visoke. Napake

Slika 3: Del konkordančnega niza za »napaka« v referenčnem korpusu slWac (NoSketchEngine).

23 Za tematsko in prenosniško primerjavo med Gigafido in 1. različico korpusa slWac gl. Logar in Ljubešić 2013.

1.3.2 Dvo- in večjezični korpusi: vzporedni in primerljivi korpusi

Glavna prednost vzporednih korpusov je seveda v tem, da ponujajo prevode celotnih delov besedila, ne le posameznih besed ali besednih zvez, s tem pa na širše prevodne enote usmerjajo tudi pozornost prevajalca. Tradicionalni slovarji so (bili) zelo omejeni pri navajanju širšega konteksta in avtentičnih primerov rabe, na primer obstoječih preteklih prevodnih rešitev, ki bi omogočile vpogled v pragmatično ustreznost prevoda (Vintar 2008). Slovenski vzporedni korpusi so večinoma vsaj delno specializirani. Med najstarejšimi večjezičnimi vzporednimi korpusi sta **IJS-ELAN** (Erjavec 2002) in **TRANS** <<http://nl.ijs.si/elan/>> (Vintar 2008), ki sta specializirana prevodoslovna vira. Slovenski **korpus SPOOK** <<http://lojze.lugos.si/spook/index.html>> je najnovejši vzporedni in primerljivi večjezični korpus, ki je prav tako predvsem vir za prevodoslovne raziskave, saj naj bi po svoji zgradbi omogočal medjezikovno primerjavo prevodnih pojavov pri prevajanju v slovenščino (Vintar 2009). Korpus obsega pet jezikov: štiri jezike izvornikov (angleščina, nemščina, francoščina, italijanščina), prevode iz teh jezikov v slovenščino, za primerjavo pa še korpus izvornih besedil v slovenščini. **Evrokorpus** <<http://evrokorpus.gov.si/>> temelji na vzporednih besedilih, zbranih v procesu prevajanja evropske zakonodaje (*acquis communautaire*) kot priprave na vstop Slovenije v EU. Korpus se vseskozi dopolnjuje z novimi prevodi in skupaj s terminološko zbirko **Evroterm** predstavlja dragocen pripomoček za prevajalce in terminologe. V kontekstu splošnih virov naj na koncu omenimo še **Googlov brskalnik**, ki omogoča, da ga uporabimo tudi kot dvojezični primerljivi korpus (<<http://www.2lingual.com/>>).

1.3.3 Specializirani korpusi

Na voljo imamo tudi nekaj ožje specializiranih besedilnih korpusov, na primer **korpus DSI** izrazja informacijskih znanosti (Erjavec in Vintar 2004). Korpus je bil zasnovan kot podpora pri izdelavi **Islovarja** (gl. <http://islovar.org/slovar_oslovarju.asp>), ki zajema temeljno izrazje informatike, informacijske tehnologije in telekomunikacij (gl. Vintar 2008)). Drugi primer, specializirani korpus **KoRP**, ki je prav tako prosto dostopen na spletu, sestavljajo besedila s področja odnosov z javnostmi <<http://www.korp.fdv.uni-lj.si/>>; korpus je bil v letih 2011-2013 nadgrajen v podatkovno zbirko odnosov z javnostmi **TERMIS** <<http://www.termis.fdv.uni-lj.si/>> (Logar 2013, <<http://www.termis.fdv.uni-lj.si/>>). Tu je še **TURK**, korpus turističnih besedil <www.evroterm.govs.si/slovar/>, in nekateri drugi.

1.3.4 Govorni korpus GOS

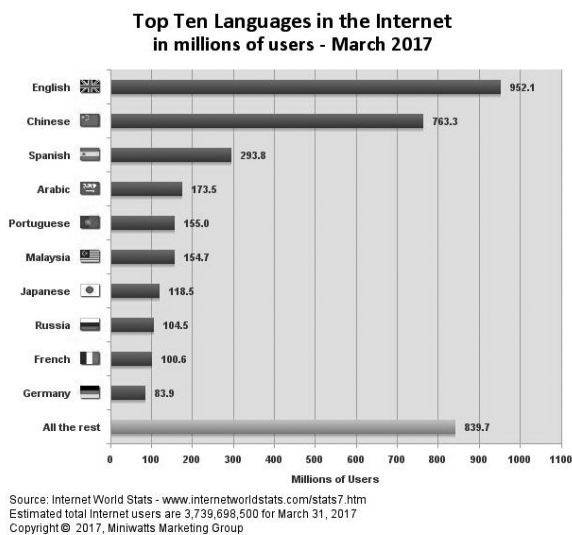
Enomilijonski **GOS**²⁴ je korpus govorjene slovenščine (Verdonik in Zwitter Vitez 2011). Obsega transkripcije okrog 120 ur posnetkov (po)govora v najrazličnejših vsakodnevnih

24 <<http://www.slovenscina.eu/korpusi/gos>>.

situacijah. Govorni korpusi so v prevajalski praksi redko uporabljeni, vendar so lahko koristni na primer pri učenju tolmačenja ali tudi pri prevajanju govornih tekstov ter pri raziskovanju pojavov, ki so tipični za govorni jezik. Ker jih je mnogo težje in dražje sestaviti kot pisne korpus besedil, so običajno tudi dokaj omejeni po obsegu.

1.3.5 Svetovni splet kot korpus

Konkreten zgled orodja, ki omogoča uporabo spleta kot korpusa zlasti jezikoslovcem in jezikovnim specialistom, je **WebCorp** <<http://www.webcorp.org.uk/>>, ki temelji na standardnih spletnih iskalnikih Google in Bing, vendar je opremljen z elementi za bolj prefinjeno uporabo v jezikoslovne namene. Definicija korpusa preprosto kot »zbirke besedil« (Kilgarriff in Grefenstette 2003: 334)²⁵ je sprožila živahne razprave o razmerju med spletom in tradicionalno pojmovanim korpusom, pa tudi postopno preoblikovanje stališč korpusnih jezikoslovcev, ki so bili sprva zelo kritični do ideje »spleta kot korpusa« (gl. tudi zgoraj Googlov dvojezični korpus). Dolgo je bilo mogoče sklepati, da je splet skoraj povsem enojezični angleški jezikovni korpus. Toda raziskava iz leta 2000 (Grefenstette in Nioche) je pokazala, da angleščina sicer zares prevladuje s 66 %, vendar pa udeležba drugih jezikov raste hitreje. Prav večjezičnost naj bi bila ena najzanimivejših lastnosti svetovnega spleta,



Slika 4: Prikaz najbolj razširjenih jezikov na spletu (v milijonih uporabnikov), vir: Internet World Stats.

²⁵ Gre za namerno poenostavljanje, zavračanje mnogih omejitev tega, kar naj bi bil korpus kot »telo«, torej kot stabilna zbirka podatkov, ki naj bi ji bilo mogoče določiti osnovne kriterije, tipično velikost, (taksonomsko) uravnoteženost, reprezentativnost itd.

kar s stališča korpusnega jezikoslovja pomeni, da vsebuje vzporedne in primerljive korpusne za takorekoč vsak pisni jezik na svetu, pri čemer je pomembno to, da pokriva prav vsa področja, teme, zvrsti in žanje (Gatto 2011: 43). Ugotovitve novejših raziskav (marec 2017) distribucije jezikov na spletu so, da sta najbolj razširjena jezika angleščina (z 952.055.837 uporabniki) in kitajščina (s 763.262.224 uporabniki),²⁶ sledijo pa jima jeziki, kot je razvidno iz zgornje tabele. Razlike v relativni (spletni) teži posameznega jezika kažejo na splošen problem t. i. digitalne ločnice med bogatimi in revnimi državami, hkrati pa mnogi poročajo tudi o koristih spleta prav za nekatere manjše in ogrožene jezike (ibid: 43):

2 SLOVAR IN SPLETNO OKOLJE

Prenos klasičnih slovarskih vsebin v spletno okolje je prinesel v leksikografijo številne nove možnosti in odpravil stare, v dobi tiska pomembne omejitve, npr. predstavitvenega prostora (gl. npr. konferenčne prispevke in razprave eLex v Granger in Paquot 2009, Kosem in Kosem 2011, Kosem et al. 2013); iz številnih strokovnih besedil na temo sodobne leksikografije in raziskav uporabniških navad (npr. Müller-Spitzer 2012, 2014) izhaja tudi, da morajo spletni slovarski in širše jezikovni viri bodočnosti uporabnikom predvsem zagotavljati kvalitetne, tj. zanesljive podatke, ti pa morajo temeljiti na sodobnih, tj. korpusnih analizah. Pri tem je zelo pomembna vpetost podatkov v širši kontekst možnosti, ki jih ponujata informacijska in komunikacijska tehnologija. V strokovnih krogih so se tako oblikovale smernice razvoja elektronske leksikografije, ki jih lahko zagotovo povzamemo tudi za slovenski jezik in njegove uporabnike: geselski članki v spletnih slovarjih bodo zasnovani kot vir dinamičnih podatkov, ki bodo ustrezali specifičnim potrebam specifičnih uporabnikov v njihovi specifični situaciji (Granger in Paquot 2009: 307).

2.1 Med slovarjem in podatkovno zbirko

Leksikografi so razvili koncept **leksikalne podatkovne zbirke**, ki nastane na podlagi korpusov besedil – ti so navadno v obsegu do več sto milijonov besed –, v njej pa so zabeleženi vsi jezikovni podatki, zanimivi bodisi za sestavljanje slovarjev bodisi za izgradnjo jezikovnih tehnologij (npr. DANTE, B(ase)L(exicale)F(rancaise)). Pomembna značilnost virov v novem, digitalnem mediju je tudi zabrisanost razlik znotraj klasičnih slovarskih tipologij, saj obstajajo med leksikografskimi orodji, zasnovanimi v skladu z načeli potreb posameznika, zgolj individualizirani dostopi, realizirani v leksikografskih e-orodjih (Granger in Paquot 2009: 309). Leksikalne zbirke so po vsebini in obliki zelo različne, nekatere so zelo blizu tega, kar si predstavljamo kot elektronski slovar (Arhar

26 Od leta 2000 je število uporabnikov obeh, tudi takrat vodilnih jezikov naraslo za skoraj dvakrat.

in Arčan 2011). Iz ene ustrezno zasnovane leksikalne zbirke lahko nastane več različnih slovarjev oz. priročnikov, ki so namenjeni različnim skupinam uporabnikov (npr. splošni slovar, šolski slovar, frazeološki slovar, slovar za tujce, področni slovarji, dvojezični in večjezični slovarji itd.). Gre torej za stopnjo pred izdelavo konkretnih priročnikov ali aplikacij, za katere se lahko – ob predpostavki, da je v leksikalno zbirko zajet celovit opis besedišča – ustrezni podatki na osnovi izbranih kriterijev pridobivajo (pol)avtomatsko. Po že uveljavljenih korpusnih načelih je bila izdelana tudi **Leksikalna baza za slovenščino** (LBS), ki vsebuje v tej fazi 2.500 gesel (gl. 2.1.1), dve sodobni leksikalni zbirki za angleščino, **DANTE** in **WordNet**, po svoji strukturi in namembnosti dokaj različni, pa sta na kratko predstavljeni v poglavjih 2.1.2 in 2.1.3.

2.1.1 Leksikalna baza za slovenščino (LBS)²⁷

LBS vsebuje podatke o pomenu besed, njihovem tipičnem okolju, stilni in pragmatični vrednosti pomenskih enot oz. njihovi rabi, načrtovana pa je bila za izgradnjo slovarjev slovenščine in slovensko-tujejezičnih slovarjev. Drugi pglavitni namen leksikalne baze je zagotoviti jezikovne podatke v obliki, primerni za računalniško obdelavo in razvoj sodobnih jezikovnotehnoloških aplikacij za slovenščino. LBS oz. vizualizacija dela gesel je dostopna na <<http://www.slovenscina.eu/spletni-slovar/>> oz. za prenos celotne baze na <<http://www.slovenscina.eu/spletni-slovar/prenos>>.

2.1.2 Leksikalna podatkovna zbirka DANTE

DANTE (Database of Analysed Texts of English, 2009) je sodobna leksikalna podatkovna zbirka za angleščino, ki podaja podrobno analizo leksiko-gramatičnih vzorcev za približno 45.000 iztočnic jedrnega besedišča na podlagi za ta namen izdelanega korpusa z 1,7 milijarde besed. V njej najdemo sistematični opis pomenov, slovničnega in kolokacijskega vedênja besed ter besedilnih zvrsti in vrst, v katerih se besede pojavljajo, in sicer v kategorijah, prirejenih za strojno obdelavo podatkov. Na osnovi teh podatkov bodo izdelani novi enojezični in angleško-tujejezični slovarji. Zbirka je od leta 2012 v celoti prosto dostopna na <<http://www.webdante.com/>>.

2.1.3 Leksikalna podatkovna zbirka WordNet²⁸

Drugačna je na ontoloških načelih zasnovana podatkovna zbirka **WordNet**, za slovenščino **SloWNet**, namenjena predvsem razvoju aplikacij za računalniško obdelavo naravnega jezika. Zbirka vsebuje podatke za polnopomenske iztočnice, kot so samostalniki, glagoli,

27 Aktivnost *Leksikalna baza za slovenščino* (2008-2012) v okviru projekta *Sporazumevanje v slovenskem jeziku* sta delno financirala ESS ter MŠŠ RS. Spletni slovar, specifikacije ter podatki o izgradnji in vsebini LBS so dostopne na <www.slovenscina.eu>.

28 Gre za projekt, ki se je začel na Univerzi Princeton v ZDA in ga danes poznamo kot Princeton WordNet ter nadaljeval za evropske jezike kot EuroWordNet. Danes obstajajo wordneti za 50 jezikov, krovno združenje pa je the Global WordNet Association <<http://www.globalwordnet.org>>.

pridevniki in prislovi, tako da jih razvršča v pojmovno sorodne leksikalne nize (sinsete), povezana pa je tudi z drugimi leksikalnimi viri, zlasti za namene strojne obdelave podatkov, na primer semantičnega označevanja. SloWNet je dostopen v obliki SloWTool na <<http://nl.ijs.si/slowtool/>>.

3 POTREBE UPORABNIKOV

Prevajalci pri svojem delu (poleg spletnih besedil in področnih virov) tipično uporabljajo tudi splošne enojezične slovarje, od katerih upravičeno pričakujejo tudi podatke o pragmatičnem pomenu in rabi besed. Slovarjev, ki bi nastali na osnovi podatkov iz leksikalne zbirke, še ni prav veliko,²⁹ obstaja pa več enojezičnih korpusnih slovarjev, zlasti za tujce, ki so nastali predvsem v angleškem jezikovnem okolju (gl. 1).

Če je zaradi »glosarske« usmerjenosti dolgo veljalo, da je na primer posvetovanje z dvojezičnim slovarjem šele začetek procesa iskanja prevodne ustreznosti, in če je danes samoumevno, da so jezikovni viri zasnovani v elektronski obliki, pa vendar še ni dokončnega odgovora na vprašanje, kako in v kolikšni meri tradicionalne slovarje nadomestiti z novimi aplikacijami oziroma kako njihov osnovni koncept (in poimenovanje)³⁰ ustrezno integrirati v večfunkcionalne aplikacije, v katerih je slovar zgolj ena od komponent, ki uporabniku pomagajo pri tvorbi (določenih tipov) besedil, preverjanju jezikovne ustreznosti itd., torej v okviru orodij, ki jih v najširšem pomenu razumemo kot vsestransko »podporo« pri pisanju in uporabi (maternega ali tujega) jezika. Kot razmišlja angleški leksikograf, nekdanji urednik Longmanovih, danes pa Macmillanovih angleških slovarjev za tujce, M. Rundell (2011), bi lahko takšna orodja dolgoročno celo povsem nadomestila slovar, vsaj za enkodirne namene, saj bi uporabniku precej bolje kot klasični slovar pomagala pri istih nalogah. Čeprav so napovedi prihodnosti slovarjev v tem trenutku še tvegane, pa Gabrovšek (2010) po drugi plati meni, da potrebe in spretnost uporabnikov slovarjev precej zaostajajo za napredkom in možnostmi, ki jih omogočajo nove tehnologije.

3.1 Kaj in kako uporabnik pravzaprav išče?

Nekatere raziskave (Granger in Paquot 2009) nakazujejo, da je treba zasnovano spletnih slovarjev in priročnikov načrtovati s pomočjo abstrahiranja tipa uporabnika, tipa iskalnih

29 Leksikalna baza DANTE je nastala prav z namenom izdelave novega angleško-irskega slovarja (NEID).

30 Poimenovanje »slovar« se zdi še vedno dovolj uporabno, danes zlasti kot splošnejši izraz za spletna mesta, saj ti ostajajo najbolj aktualni viri za leksikalne poizvedbe tudi v spletnem okolju. Kot nakazujejo rezultati iskanj zlasti posameznih (dekontekstualiziranih) leksemov, npr. Googlov brskalnik povečini med prvimi zadetki - vsaj za angleščino - ponuja prav povezave do bodisi uveljavljenih slovarskih virov z znanimi uredniškimi zasedbami in dolgotrajno tradicijo (npr. Webster, Cambridge itd.) ali novih (participativnih) slovarskih projektov tipa Urban Dictionary <www.urbandictionary.com>.

situacij, posameznikovih potreb in tipa podatkov, ki zadovoljuje takšne potrebe. Analiza konkretnih iskanj namreč pokaže, da ni neke sistematične iskalne poti, na osnovi katere bi lahko ugotavljali, kako uporabniki pravzaprav iščejo. Izdelani so že nekateri modeli, ki uspešno izrabljajo jezikovne tehnologije in nove tehnike za optimizacijo uporabniških orodij pri uporabi spletnih virov.³¹ Novejše razprave na temo dobrega slovarja sicer izpostavljajo predvsem dejstvo, da – v nasprotju s tradicionalnim prepričanjem – velika količina gradiva ne pomeni tudi nujno velike prednosti. Tisto, kar naj bi odlikovalo sodobni slovar, je predvsem omogočiti uporabniku, da pride do relevantne informacije v najkrajšem možnem času. Kot pomembna se kaže tudi možnost povezav z drugimi tipi relevantnih podatkov, predvsem besedilnimi korpusi – to pomeni, da ima uporabnik možnost s klikom na ustrezno besedo v takšnem slovarju priti do podatkov o njeni dejanski rabi v korpusu. Za slovenščino velja v tem kontekstu omeniti povezano dostopanje preko spletnega vmesnika projekta *Sporazumevanje v slovenskem jeziku* do projektnih vsebin, tj. omenjenih korpusov Kres, GOS in Gigafida, Slogovnega priročnika, Leksikona besednih oblik in Spletnega slovarja (več na <www.slovenscina.eu>).

3.2 Stereotipi: slovar kot glosar ali »slovarji se ukvarjajo z besedami«

Splošno razširjene predstave o slovarjih kot seznamih besed in njihovih pomenov so morda točne za del slovarske produkcije, zlasti predkorpusne. Toda leksikografska teorija se razvija in z njo praktične metode izdelave samih slovarjev. Tudi v slovenskem prostoru je že dobila domicil praksa korpusnega pristopa k izdelavi slovarjev, toda ta je zaenkrat omejena na posamezne (pretekle) projekte, ki iz različnih razlogov niso dovolj vpeti v vsakdanjost slovenskih slovarskih uporabnikov, celo profesionalnih ne. Glede na razširjenost angleščine je v danem trenutku zagotovo najbolj nujna kakovostna obravnava na relaciji **slovenščina-angleščina**, pri čemer smo že zagotovili kvalitetne splošne vire za potrebe prevajanja v slovenščino, ne pa tudi v angleščino.

3.2.1 Slovenski (dvojezični) korpusni slovarji

Že omenjeni *Veliki angleško-slovenski slovar Oxford-DZS*³² (VASS), pa tudi *Priročni angleško-slovenski in slovensko-angleški slovar* (2010, DZS) ter *Mali angleško-slovenski in slovensko-angleški slovar DZS* (2006) so želeli preseči ozko pojmovanje »slovarskega« pomena, ki se je v jezikoslovnih razpravah uveljavil za označevanje tistega, čemur protipostavljamo kontekstualni in pragmatični pomen. Z vidika tradicionalnejšega glosarskega pojmovanja slovarja naj bi bile »besede« v dvojezičnem slovarju prevedene v ustrezne »besede« na desni strani slovarja, torej v ciljnem jeziku, toda z razvojem

31 Tak primer je npr. vmesnik *B(ase)L(exicale)F(rancaise)* S. Verlindeja z belgijske Univerze v Louvainu <<http://ilt.kuleuven.be/blf/>> (Dostop: 6.12.2012).

32 VASS je bil s strani avtorice obširneje predstavljen v *Mostovi* 42, 2009/2010.

korpusnega jezikoslovja se je okrepilo spoznanje, da morajo biti osnova slovarskega opisa predvsem širše pomenske enote. Te so sestavljene iz zgradbeno raznovrstnih besednih sopojavljanj, od dvo- in trobesednih kolokacij do širših frazeoloških enot, mnoge od pomenskih enot pa je mogoče opisovati pravzaprav šele na ravni besedila. S pomenom imamo torej opraviti tako na ravni leksike kot slovnice in seveda besedila. Sodobni slovarski viri so zato usmerjeni vse manj glosarsko, tj. besedno, in vse bolj besedilno.

3.2.2 Obrnjeni Veliki angleško-slovenski slovar Oxford-DZS: baza podatkov o kontrastivnih razmerjih med jezikoma

Vsaka dvojezična slovarska zbirka je do neke mere – odvisno od njene velikosti in dometa –, kontrastivna analiza pomenskih in drugih razmerij med jezikoma, zato je naravno in logično izhodišče za vsak nadaljnji sistemski opis specifik na ravni dveh sopostavljenih jezikov. Ko je v letih 2005–6 izšel pri nas do sedaj največji, korpusni *Veliki angleško-slovenski slovar Oxford-DZS* s 120.000 iztočnicami, je bila sprva predvidena tudi izdelava slovarja v obratni smeri, ki pa se ji je založba zaradi visokih stroškov prvega slovarja odpovedala. Obsežna dvojezična podatkovna zbirka je bila avtomatsko “obrnjena”,³³ s čimer smo dobili ne le idealno orodje za pripravo kvalitetnega slovensko-angleškega besedila, temveč tudi neprecenljivo podatkovno zbirko, ki bi jo lahko uporabili za poglobljene kontrastivne in prevodoslovne raziskave (gl. tudi Krek et al. 2008; Šorli 2009; Srebnik 2012). Žal ostaja obrnjena baza nedostopna tako za splošnega uporabnika kot tudi za profesionalne (prevajalske in pedagoške) ter raziskovalne namene. Navajamo primer geselskih člankov, ki kaže tip in naravo tako pridobljenih podatkov:

ročaj 25 / 5 / 9 / 0 /

arm

(o očalih) **ročaj**

bail

(npr. pri čajniku, vedru) **ročaj**; (pri baldahinu) polkrožni opornik; (pri pisalnem stroju) držek (za papir)

bend

optika (pri očalih) **ročaj**

earpiece

optik (pri očalih) **ročaj**

grab bar

(pri prhi) **ročaj**, držaj

grab handle

(npr. v avtobusu) **ročaj**, držaj

grab rail

(npr. v avtobusu) **ročaj**, držaj

grip

ročaj, držaj

Slika 5: Del gesla ROČAJ, 1. skupina (obrnjena baza VASS)

nosečnost 1 / 0 / 2 / 3 /

antenatal

predporoden; **nosečnost**, nosečniški

1

pre-eclampsia

med. preeklampsija, **nosečnostna** toksemija

stretch marks

med. strije; **nosečnostne** proge

1

negative

nosečnostni test je bil negativen *the pregnancy test was negative*

positive

med. nosečnostni test je bil pozitiven [*reaction, result, test*] *the pregnancy test was positive*

pregnancy

med. nosečnostni test; test nosečnosti *a pregnancy test*

Slika 6: Del gesla NOSEČNOSTEN, 1., 3. in 4. skupina (obrnjena baza VASS)

4 SKLEP

Pričujoči prispevek je nastal med drugim kot podlaga za razmislek, kateri leksikalni viri in orodja za slovensko-angleški jezikovni par naj bi bili prioritetni v novem in nadaljnjih slovenskih akcijskih načrtih za večjezično opremljenost. Na tem mestu so predstavljeni zlasti novejši enojezični in dvojezični leksikografski in besedilni viri, v katerih lahko poiščemo pomen besed in širših jezikovnih enot, tj. slovarji, leksikalne podatkovne zbirke, korpusi besedil in spletna prevajalska orodja. Spregovorili smo tudi o prevajalskih tehnologijah in na kratko predstavili spremembe pri načrtovanju, izdelavi in uporabi klasičnega slovarja v novem, digitalnem okolju. Spomnili smo na nekatere stereotipe o uporabi slovarjev ter izpostavili nekatere pomisleke ob vrednotenju vloge slovarjev v navezavi na digitalno okolje in jezikovne tehnologije. Vsak prihodnji akcijski načrt bo moral upoštevati omejena sredstva, namenjena razvoju jezikovne infrastrukture, in dejstvo, da slovenska govorna skupnost še vedno čaka na nekatere osnovne jezikovne priročnike: sodoben slovar standardne slovenščine, korpusni slovensko-angleški slovar, pedagoški slovar slovenščine itd. Vsi ti viri so, med drugim, nujni za uspešen razvoj jezikovnih tehnologij za slovenščino.

Na tem mestu velja omeniti raziskavo *Jezikovna politika RS in potrebe uporabnikov* pod vodstvom ZRC SAZU (<http://isjfr.zrc-sazu.si/sl/programi-in-projekti/>)

jezikovna-politika-republike-slovenije-in-potrebe-uporabnikov#v), v kateri so sodelovale številne izobraževalne in raziskovalne institucije. Gre za obsežno (sociolingvistično) raziskavo o navadah, potrebah in prepričanjih uporabnikov slovenščine (govorcev slovenščine kot maternega ali drugega/tujega jezika) in tujih jezikov, prav tako govorcev s posebnimi potrebami, ki se je začela oktobra 2016 in zaključila septembra 2017. V delu raziskave, ki se osredotoča na opremljenost skupnosti za večjezičnost, smo posebej naslovili skupine jezikovnih delavcev – prevajalcev, tolmačev in drugih strokovnjakov, ki uporabljajo vsaj en tuji jezik, pri čemer je bil cilj raziskati tudi ustaljene delovne navade in strategije pri reševanju poklicnih izzivov, predvsem pa odnos do obstoječe jezikovne infrastrukture in potrebe na področju opremljenosti z viri in orodji. Raziskava v obliki spletne ankete sloni na štirih vsebinskih sklopih: odnos do jezikovnih vprašanj, sporazumevalne prakse, potrebe uporabnikov, jezikovni opis in jezikovna opremljenost. Na podlagi končnih rezultatov projekta bo mogoče ustrezneje načrtovati razvoj jezikovnih virov, orodij in tehnologij, zato bodo ti aktualni tudi za načrtovalce jezikovne politike in razvijalce jezikovnih tehnologij.

BIBLIOGRAFIJA

Slovarji:

- Collins COBUILD English Language Dictionary. Sinclair, John McH. (ur.), 1995, 2. izdaja. London: HarperCollins. (COBUILD)
- Collins COBUILD Advanced Learner's English Dictionary. Sinclair, John McH. (ur.), 2003, 4. izdaja. Glasgow: HarperCollins Publishers.
- Macmillan English Dictionary for Advanced Learners. Rundell, Michael, 2007, 2. izdaja. Oxford: Macmillan Education.
- Mali angleško-slovenski in slovensko-angleški slovar DZS (Zaranšek 2006).
- Priročni angleško-slovenski in slovensko-angleški slovar DZS (Drinovec Sever et al. 2010).
- Veliki angleško-slovenski slovar Oxford-DZS (Krek 2005-6).

Literatura:

- ARHAR HOLDT, Špela/Iztok KOSEM/Nataša LOGAR BERGINC (2012) Izdelava korpusa Gigafida in njegovega spletnega vmesnika. T. Erjavec, J. Žganec Gros (ur.), *Zbornik Osmе konference Jezikovne tehnologije*. Ljubljana: Institut Jožef Stefan.
- ARHAR HOLDT, Špela/Mihael ARČAN (2011) Avtomatsko pridobivanje besednih zvez iz korpusa z uporabo leksikona SSJ. S. Kranjc (ur.), *Meddisciplinarnost v slovenistiki*, (Obdobja, Simpozij, = Symposium, 30). Ljubljana: Znanstvena založba Filozofske fakultete.

- ATKINS, Sue/Michael RUNDELL (2008) *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- BAKER, Mona (ur.) (2009) *Critical Readings in Translation Studies*. London & New York: Routledge.
- BAKER, Mona (2006) *Translation and Conflict: A Narrative Account*. New York & London: Routledge.
- BASSNETT, Susan (1980) *Translation Studies*. Methuen & Co. Ltd.
- CRYSTAL, David (2012) *English as a Global Language*. 2. izdaja. Cambridge University Press.
- ERJAVEC, Tomaž/Nikola LJUBEŠIČ (2014) The slWaC 2.0 corpus of the Slovene web. T. Erjavec, J. Žganec Gros (ur.) *Jezikovne tehnologije : zbornik 17. mednarodne multikonference Informacijska družba - IS 2014, 9. - 10. oktober 2014, [Ljubljana, Slovenia] : zvezek G*, 50—55. Ljubljana: Institut Jožef Stefan, 2014.
- GABROVŠEK, Dušan (2010) About Dictionaries – English, English and Slovene, and a Handful of Others: the Good, the Better, and the Useful. *Mostovi* 42, št. 1-2, letnik 2009/10, 110—146. Ljubljana: Društvo znanstvenih in tehniških prevajalcev Slovenije.
- GANTAR, Polona, et al. (2009) Specifikacije za izdelavo leksikalne baze za slovenščino: standard za izdelavo posamezne leksikalne enote v leksikalni bazi. Projekt *Sporazumevanje v slovenskem jeziku ESS in MŠŠ* (2008–2013).
- GATTO, Maristella (2011) “The ‚body‘ and the ‚web‘: The Web as Corpus Ten Years On.” *ICAME Journal* no. 35 (April), 35—58. Dostopno tudi na: http://icame.uib.no/ij35/Maristella_Gatto.pdf/ (8. november 2017).
- GRANGER, Sylviane/Magali PAQUOT (ur.) (2009) *eLexicography in the 21st century: New challenges, new applications. Proceedings of eLex 2009*, Louvain-la-Neuve, 22–24 October 2009.
- GREFENSTETTE, Gregory/Julien NIOCHE (2000) Estimation of English and non-English language use on the WWW. In *Proceedings of the RIAO (Recherche d'Informations Assistée par Ordinateur)*, Paris, 12–14 April 2000, 237—246. Dostopno na: <http://arxiv.org/ftp/cs/papers/0006/0006032.pdf/> (8. november 2017).
- GORJANC, Vojko (2010) Zaključno poročilo o rezultatih raziskovalnega projekta: ARRS-RPROJ-ZP-2010-1/46.
- KILGARRIFF, Adam/Gregory GREFENSTETTE (2003) Introduction to the special issue on the web as corpus. *Computational Linguistics* 29 (3), 333—347.
- KOSEM, Iztok/Karmen KOSEM (ur.) (2011) *eLexicography in the 21st century: Proceedings of eLex 2011*, 10-12 November, Bled, Slovenija.
- KOSEM, Iztok, et al. (ur.) (2013) *eLexicography in the 21st century: Proceedings of eLex 2013*, 17-19 October 2013, Tallinn, Estonia.
- KREK, Simon/Mojca ŠORLI/Polonca KOCJANČIČ (2008) The Funny Mirror of Language: The Process of Reversing the English-Slovenian Dictionary to Build the

- Framework for Compiling the New Slovenian-English Dictionary. E. Bernal, J. DeCesaris (ur.) *Proceedings of the xii Euralex International Congress*. Barcelona: Universitat Pompeu Fabra, 535—542.
- LOGAR, Nataša (2013) Aktualni terminološki opisi in njihova dostopnost. A. Žele (ur.) *Družbena funkcijskost jezika (vidiki, merila, opredelitve): Obdobja 32*. Ljubljana: Znanstvena založba Filozofske fakultete, 247—253.
- LOGAR BERGINČ, Nataša/Špela ARHAR HOLDT/Miha GRČAR/Marko BRAKUS/Simon KREK (2012) *Korpusi slovenskega jezika Gigafida, KRES, ccGigafida in ccKRES: gradnja, vsebina, uporaba*. Ljubljana: Trojina, zavod za uporabno slovenistiko; Fakulteta za družbene vede.
- LOGAR, Nataša/Nikola LJUBEŠIČ (2013) Gigafida in slWaC : tematska primerjava. *Slovenščina 2.0*, 2013, št. 1, 78—110.
- MAYOR, Michael (ur.) (2009) *Longman Dictionary of Contemporary English* (5. izdaja). Harlow, Essex: Pearson Education / Longman.
- MÜLLER-SPITZER, Caroline/Alexander KOPLENIG/ Antje TÖPEL (2012) Online dictionary use: Key findings from an empirical research project. S. Granger, M. Paquot (ur.) *Electronic Lexicography*. Oxford: Oxford University Press, 425—457.
- MÜLLER-SPITZER, Caroline (ur.) (2014) *Using online dictionaries. Lexicographica. Series maior*. Berlin: de Gruyter.
- NEWMARK, Peter (2000) *Učbenik prevajanja*. Ljubljana: Krtina.
- NIDA, Eugene A. (1982) *Translating Meaning*. San Dimas, Calif.: English Language Institute.
- NIDA, Eugene A. (1996) *The Sociolinguistics of Interlingual Communication*. Brussels: Editions du Hazard.
- OOI, Vincent B. Y. (2010) English Internet Lexicography and Online Dictionaries. *Lexicographica 26*. Berlin: de Gruyter, 143—154.
- PERRAULT, Stephen J. (ur.) (2008) *Merriam-Webster's Advanced Learner's English Dictionary*. Springfield, MA: Merriam Webster.
- RUNDELL, Michael (2011) The future of dictionaries. Too soon to tell. Blog o konferenci *eLexicography in the 21st Century: New Applications, New users* /<http://www.trojina.si/elex2011/index.html/>. Dostopno na: <http://www.macmillandictionaryblog.com/the-future-of-dictionaries-too-soon-to-tell/>. (8. november 2017)
- RUNDELL, Michael (ur.) (2007) *Macmillan English Dictionary for Advanced Learners* (2. izdaja). Oxford: Macmillan Education.
- SINCLAIR, John (ur.) *Collins COBUILD Advanced Learner's English Dictionary* (5. izdaja). Glasgow: HarperCollins Publishers.
- SREBNIK, Anita (2012) Obrnjena nizozemsko-slovenska slovarska baza kot korak do novega slovensko-nizozemskega slovarja. M. Šorli (ur.) *Dvojezična korpusna leksikografija. Slovenščina v kontrastu: novi obeti, novi izzivi*. Ljubljana: Trojina, zavod za uporabno slovenistiko.

- SUMMERS, Della (ur.) (2003) *Longman Dictionary of Contemporary English* (4. izdaja). Harlow, Essex: Pearson Education / Longman.
- ŠORLI, Mojca (2009) Pridobivanje podatkov o slovenščini za izdelavo slovensko-tujejezičnih slovarjev. *Infrastruktura slovenščine in slovenistike*. M. Stabej (ur.) *Simpozij Obdobja 28*. Ljubljana: Znanstvena založba Filozofske fakultete, 359—369.
- VERDONIK, Darinka/Ana ZWITTER VITEZ (2011) *Slovenski govorni korpus Gos*. Ljubljana: Trojina, zavod za uporabno slovenistiko.
- VINTAR, Špela (1999) Računalniška orodja za prevajanje. *Mostovi 1999/XXXIII*, Ljubljana: Društvo znanstvenih in tehniških prevajalcev Slovenije.
- VINTAR, Špela (2001) Računalniška orodja za jezikoslovce in prevajalce. *Zbornik 37. seminarja slovenskega jezika, literature in kulture*. Ljubljana, Filozofska fakulteta.
- VINTAR, Špela (2003) Kaj izvira iz jezikovnih virov. *Jezik in slovstvo XLVIII(2-3)*, 77—88.
- VINTAR, Špela (2008) Corpora in Translation: A Slovene Perspective. *Journal of Specialized Translation*, Issue 10, http://www.jostrans.org/issue10/art_vintar.php/. (8. november 2017)
- VINTAR, Špela/Tomaž Erjavec (2008) iKorpus in luščenje izrazja za Islovar. T. Erjavec, J. Žganec Gros (ur.) *Zbornik Šeste konference Jezikovne tehnologije, 16. do 17. oktober 2008, Ljubljana, Slovenia : zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, zvezek C : proceedings of the 11th International Multiconference Information Society - IS 2008, volume C*, (Informacijska družba). Ljubljana: Institut „Jožef Stefan“, 2008, 65—69.
- ŽELJKO, Miran (2010) Pomnilnik prevodov EU. *Mostovi 42*, št. 1-2, letnik 2009/10, 72—75. Ljubljana: Društvo znanstvenih in tehniških prevajalcev Slovenije.
- WALTER, Elisabeth (ur.) (2008) *Cambridge Advanced Learner's Dictionary* (3. izdaja). Cambridge: Cambridge University Press.

Spletni viri:

- Amebis d.o.o. 9. julij 2017. <http://www.amebis.si/>.
- Babelfish. 9. julij 2017. <http://www.babelfish.com/>.
- Babylon. 9. julij 2017. <http://translation.babylon.com/>
- Base Lexicale Française (BLF). 6. december 2016. <http://ilt.kuleuven.be/blf/stop/>.
- Bing translate. 9. julij 2017. <https://www.bing.com/translator/>.
- Brezplačno prevajanje. 9. julij 2017. <http://www.brezplacno-prevajanje.si/>.
- Cambridge Advanced Learner's Dictionary. 9. julij 2017. <http://dictionary.cambridge.org/>.
- Collins COBUILD English Dictionary. 9. julij 2017. <http://www.collinsdictionary.com/>.
- Collins Dictionary. 9. julij 2017. <http://www.collinsdictionary.com/>.
- DANTE Database. 9. julij 2017. <http://www.webdante.com/>.
- Dictionary.com Translator. 9. julij 2017. <http://translate.reference.com/translate/>.
- eLex2011/2013. 9. julij 2017. <http://www.trojina.si/elex2011/Vsebina/proceedings.html/>.

- Eur-lex. 9. julij 2017. <http://eur-lex.europa.eu/>.
- Evrokopus. 9. julij 2017. <http://evrokopus.gov.si/>.
- Evroterm. 9. julij 2017. <http://evroterm.gov.si/>.
- Fran. 9. julij 2017. <http://www.fran.si/>.
- Free dictionary. 9. julij 2017. <http://www.thefreedictionary.com/>.
- Gigafida. 9. julij 2017. <http://www.gigafida.net/>.
- Glosbe.com. 9. julij 2017. <http://glosbe.com/>.
- Google Translate. 9. julij 2017. <http://translate.google.com/>.
- ImTranslator. 9. julij 2017. <http://translation.imtranslator.net/translate/default.asp?loc=en/>.
- Islovar. 9. julij 2017. <http://www.islovar.org/>.
- Kres. 9. julij 2017. <http://www.kopus-kres.net/>.
- Lexicool.com. 9. julij 2017. <http://www.lexicool.com/>.
- Linguee. 9. julij 2017. <http://www.linguee.com/>.
- Longman Dictionary of Contemporary English, Longman. 9. julij 2017. <http://www.ldo-ceonline.com/>.
- Macmillan Dictionary and Thesaurus: Free English Dictionary Online. Oxford: Macmillan Education. 9. julij 2017. <http://www.macmillandictionary.com/>.
- Merriam Webster's Learners' Dictionary. 9. julij 2017. <http://www.learnersdictionary.com/>.
- META-NET, Zbirka Bela knjiga (Krek 2012). 9. julij 2017. <http://www.meta-net.eu/whitepapers/e-book/slovene.pdf/>.
- Microsoft Translator. 9. julij 2017. <http://www.microsofttranslator.com/>.
- MyMemory. 9. julij 2017. <http://mymemory.translated.net/>.
- Nova beseda. 9. julij 2017. http://bos.zrc-sazu.si/a_beseda.html.
- 'OneLook Dictionary'. 9. julij 2017. <http://www.onelook.com/>.
- Razvezani jezik. 10. december 2016. <http://razvezanijezik.org/>.
- Slovar slovenskega knjižnega jezika. 9. julij 2017. <http://bos.zrc-sazu.si/sskj.html/>.
- Slovarji PONS. 9. julij 2017. <http://sl.pons.com/prevod/>.
- Spletni slovar. 9. julij 2017. <http://www.spletni-slovar.com/>.
- Spletni slovarji. 9. julij 2017. <http://www.evroterm.gov.si/slovar/>.
- SloWTool. 9. julij 2017. <http://nl.ijs.si/slowtool/>.
- Sporazumevanje v slovenskem jeziku. 9. julij 2017. <http://www.slovenscina.eu/>.
- TAUS. 9. julij 2017. <https://www.taus.net/>.
- Termania. 9. julij 2017. <http://www.termania.net/>.
- 2lingual Google Search. 9. julij 2017. <http://www.2lingual.com/>.
- Vir definicij. 9. julij 2017. <http://www.definitions.net/definition/>.
- WebCorp. 9. julij 2017. <http://www.webcorp.org.uk/>.
- Webitext. 9. julij 2017. <http://www.webitext.com/bin/webitext.cgi/>.
- Wikipedia. 9. julij 2017. <http://en.wikipedia.org/wiki/>.
- Wordnik. 9. julij 2017. <http://www.wordnik.com/>.
- Worldlingo. 9. julij 2017. <http://www.worldlingo.com/>.

IZVLEČEK

V prispevku najprej predstavimo nabor enojezičnih in dvojezičnih leksikografskih in besedilnih virov za slovenščino – med dvojezičnimi prednostno prevajalske priročnike v kombinaciji z angleščino –, v katerih lahko poiščemo pomen besed in širših jezikovnih enot, tj. slovarje, leksikalne podatkovne zbirke, spletna prevajalska orodja in različne korpuse besedil. V ločenem podpoglavju spregovorimo o prevajalskih tehnologijah in ovrednotimo njihovo vlogo pri delu sodobnega prevajalca. V drugem in tretjem poglavju na kratko predstavimo spremembe pri načrtovanju, izdelavi in uporabi klasičnega slovarja v novem, digitalnem okolju ter razmerja med slovarjem in sorodnimi viri, kot so leksikalne podatkovne zbirke. Dotaknemo se uporabnikov in nekaterih stereotipov o uporabi slovarjev ter ob koncu povzamemo stanje leksikalnih, na korpusu temelječih podatkovnih zbirk za slovensko-angleški par, z mislijo na prioritete bodočega akcijskega načrta za večjezično opremljenost slovenske govorne skupnosti.

Ključne besede: spletni viri, slovenščina, angleščina, leksikografija, korpus, slovar, leksikalna zbirka, prevajanje, avtomatizirano prevajanje, strojno prevajanje

ABSTRACT

Web Resources and Tools for Slovenian with a Focus on the Slovenian-English Language Infrastructure: Dictionaries in the Digital Age

The article begins with a presentation of a selection of electronic monolingual and bi/multilingual lexicographic resources and corpora available today to contemporary users of Slovene. The focus is on works combined with English and designed for translation purposes which provide information on the meaning of words and wider lexical units, i.e., e-dictionaries, lexical databases, web translation tools and various corpora. In a separate sub-section the most common translation technologies are presented, together with an evaluation of their role in the modern translation process. Sections 2 and 3 provide a brief outline of the changes that have affected classical dictionary planning, compilation and use in the new digital environment, as well as of the relationship between dictionaries and related resources, such as lexical databases. Some stereotypes regarding dictionary use are identified and, in conclusion, the existing corpus-based databases for the Slovenian-English pair are presented, with a view to determining priorities for the future interlingual infrastructure action plans in Slovenia.

Key words: web resources, Slovene, English, lexicography, corpus, dictionary, lexical database, translation, automated translation, machine translation