

DOI 10.57589/srl.v70i4.4088

UDK 811.133.1'374=163.6

Sonia Vaupot

Filozofska fakulteta Univerze v Ljubljani

sonia.vaupot@ff.uni-lj.si

## PRISTOP K IZDELAVI SPLETNEGA SLOVARJA ARTES FRANCOSKO-SLOVENSKEGA STROKOVNEGA JEZIKA<sup>1</sup>

V Sloveniji se je z vstopom v Evropsko unijo pojavila potreba po dvojezičnih in večjezičnih leksikalnih in terminoloških virih, med drugim tudi za francosko-slovenski jezikovni par. Prispevek opisuje pristop k izdelavi francosko-slovenskega terminološkega in frazeološkega spletnega slovarja ARTES. Pojasnjeni bodo zasnova in zgradba slovarja ter pedagoški pristop k oblikovanju slovarja. Z uporabo vnesenih podatkovnih nizov bodo predstavljeni pridobljeni terminološki rezultati s področja krasoslovja ter oblikoskladenjske razlike med generičnimi kolokacijami.

**Ključne besede:** spletni slovar, strokovni jezik, terminologija, frazeologija, Karst

Slovenia's accession to the European Union created a need for bilingual and multilingual lexical and terminological resources, including for the French-Slovenian language pair. This article describes the development of the French-Slovenian online terminological and phraseological dictionary ARTES. The genesis and structure of the dictionary are explained, as well as the pedagogical approach to the creation of the database. The results obtained with regard to the field of Karst studies and generic collocations are presented on the basis of the data sets.

**Keywords:** online dictionary, specialised language, terminology, phraseology, Karst

### 1 Uvod

Francosko-slovenski stiki segajo v obdobje Ilirskih provinc. Valentin Vodnik je leta 1812 sestavil slovensko-nemško-francoski abecednik z dodanim slovarčkom, ki je izšel v glasilu Ilirskih provinc *Télégraphe officiel*. Leta 1924 je Janko Pretnar napisal prvi *Francosko-slovenski slovar*, Janko Kotnik pa je leto dni pozneje izdal prvi *Slovensko-francoski slovar* (1925).

Slovensko-hrvaško-latinsko-nemško-francosko-angleški frazeološki slovar Josipa Pavlice je izšel pozneje, leta 1960. Štejemo ga za prvi tovrstni priročnik na Slovenskem, obravnava pa reke, pregovore in stalne besedne zveze. V slovarju je za vsak jezik zbranih okrog 4000 fraz, ki so razvrščene pod približno 1000 slovenskimi gesli. Kritiki so priročniku sicer očitali »zlasti neenotno in premalo premišljeno ureditev, prirojavanje fraz in pregovorov, prilagajanje slovenskih zgledov tujim vzorcem, neustreznost prevodov, pa tudi naslanjanje na zastarele tuje zglede« (Suhadolnik 1961: 200), vendar je knjiga kljub temu pomembna za področje slovenske leksikografije.

---

1 Prispevek je nastal v okviru bilateralnega projekta 17-18-008, ki ga je financiral ARRS.

V Sloveniji se je z vstopom v Evropsko unijo pojavila potreba po dvojezičnih in večjezičnih leksikalnih in terminoloških virih, med drugim tudi za francosko-slovenski jezikovni par. Na voljo je nekaj splošnih slovarjev v papirni ali elektronski različici. Najbolj razširjeni so *Francosko-slovenski slovar* (Grad 2004, tiskana različica je iz leta 1971), *Slovensko-francoski slovar* (Jesenik, Dembskij 2008; tiskana različica je iz 1990) in različni spletni slovarji (*Pons*, *Glosbe*, *Linguee* itn.).

Med korpusi lahko omenimo *KorToP*, dvojezični dvosmerni vzporedni in primerljivi korpus tolmačenih govorov in prevodov v slovenščini in francoščini, ter prosto dostopni elektronski francosko-slovenski vzporedni korpus *FraSloK* (Mezeg 2020a: 90–95, 2020b).

Kar zadeva terminološke slovarje, se francosko-slovenska jezikovna kombinacija danes pojavlja v *Tolkalnem terminološkem slovarju*, *Gledališkem terminološkem slovarju* in *Planinskem terminološkem slovarju*. Gre za iztočnice v slovenščini z ustreznici v francoščini.<sup>2</sup> Omenimo lahko tudi večjezični *Planinski terminološki slovar: slovensko-angleško-nemško-francosko-italijanski slovar planinskega, alpinističnega, plezalskega izrazja* (2002). Čeprav je na spletu mogoče uporabljati večjezične leksikalne baze *Evrokorpus*, *Evroterm*<sup>3</sup> in *IATE*,<sup>4</sup> ki temeljijo predvsem na korpusih Evropske unije, pa je za francosko-slovenski jezikovni par v splošnem opaziti pomanjkanje specializiranih leksikalnih virov.

Prispevek najprej predstavi rabo korpusov pri izdelavi splošnih in terminoloških slovarjev v frankofonskem prostoru. Opisuje tudi uporabo korpusov v Franciji pri poučevanju prevajanja in pristop, uporabljen pri francosko-slovenskem spletnem slovarju ARTES (Aide à la Rédaction de TExtes Scientifiques), ki je plod bilateralnega projekta med oddelkom UFR EILA francoske univerze Paris Cité in Oddelkom za prevajalstvo Filozofske fakultete Univerze v Ljubljani. Gre za nov, dvojezični strokovni vir, v povezavi s katerim bosta pojasnjeni zasnova in zgradba ter pedagoški pristop k oblikovanju slovarja. Delo obsega več faz, od določitve strokovnega področja in izdelave primerljivih korpusov do vnosa terminoloških oz. frazeoloških sestavkov ter njihovih prevodov oz. ustreznec. Z uporabo vnesenih podatkovnih nizov bodo predstavljeni pridobljeni rezultati s področja krasoslovja in oblikoskladenjske razlike med generičnimi kolokacijami.

## 2 Korpusni in leksikografski pristop v frankofonskem prostoru

V Franciji se je preučevanje slovarjev začelo v 60. letih 20. stoletja (Quemada 1968). V letu 1960 se je začela tudi digitalizacija primerov iz literarnih del. S to digitalizacijo je nastala besedilna zbirka podatkov *Frantext*,<sup>5</sup> ki predstavlja največji francoski korpus. *Frantext* je bil zasnovan za preučevanje splošne francoščine, pri čemer je spodbudil tudi nastanek referenčnega francoskega korpusnega slovarja *Trésor de*

<sup>2</sup> Vsi slovarji so dostopni na spletnem mestu Terminologišče (na spletu).

<sup>3</sup> Na spletu.

<sup>4</sup> Na spletu.

<sup>5</sup> Na spletu.

*la Langue Française*<sup>6</sup> (TLF), tj. literarnega slovarja besedil 19. in 20. stoletja. Glavni namen besedilnega korpusa *Frantext* je bil omogočiti sestavo zbirke besed in izluščenje primerov, namenjenih urednikom slovarja TLF.

Obstoječe francoske korpuse danes lahko razdelimo na diahrono enojezične korpuse, namenjene proučevanju spreminjanja francoskega jezika skozi posamezna časovna obdobja in znanstveno literaturo, ter sinhrono enojezične specializirane korpuse, ki jih sestavljajo sodobna, žanrsko in/ali področno homogena besedila, zlasti časopisna in znanstvena (Mezeg 2020a: 85–86).

Leta 2005 je Francoski nacionalni center za znanstvene raziskave (CNRS) ustanovil Nacionalni center za besedilne in leksikalne vire (CNRTL), ki združuje niz računalniških jezikovnih virov in orodij za obdelavo francoskega jezika. Vključuje leksikalni portal, francoski govorni korpus, stare in sodobne digitalizirane in računalniško podprte slovarje ter metaleksikografsko bazo podatkov. Danes je terminološko delo na podlagi korpusov zelo priljubljeno področje raziskav, zlasti v Franciji. Potreba po terminoloških podatkih se je med drugim pojavila v podjetjih, ki morajo voditi obsežno dokumentacijo (Condamines 2005).

Na francoskem govornem področju v Quebecu so pred 60. leti 20. stoletja vladale naslednje razmere: francosko govoreče prebivalstvo je sodilo v ekonomsko manjšino, čeprav so bili številčno v večini. Od leta 1969 in v naslednjem desetletju so pod pritiskom francosko govorečega javnega mnenja zaporedne vlade Quebeca razvile dobro zastavljeno jezikovno politiko. *Veliki terminološki slovar* (*Grand Dictionnaire terminologique*)<sup>7</sup> Urada za francoski jezik v Quebecu, javni vmesnik *Quebeške terminološke baze* (*Banque de terminologie du Québec*), ki je nastala na podlagi raziskovalnega korpusa, je postal glavno orodje za širjenje francoskega jezika v 70. in 80. letih prejšnjega stoletja. Hkrati je Univerza v Montrealu leta 1970 razvila bazo podatkov z namenom izdelave obsežnega dvojezičnega angleško-francoskega glosarja, ki bi bil dostopen vsem. Ta glosar je leta 1974 prevzel Prevajalski urad kanadske vlade z namenom standardizacije terminologije v javnih službah in zvezni javni upravi. Nastal je terminološki slovar *TERMIUM. TERMIUM Plus*,<sup>8</sup> ki je bil predstavljen leta 2009 kot brezplačna spletna stran, je danes ena največjih terminoloških in jezikovnih baz podatkov na svetu. Omogoča dostop do več milijonov izrazov v angleščini, francoščini, španščini in portugalsščini s številnih strokovnih področij.

### 3 Korpusi pri poučevanju jezikov in prevajanju v Franciji

V Franciji se je korpusno jezikoslovje začelo razvijati po zaslugi *Frantexta*, prve podatkovne baze francoskih (literarnih, filozofskih, znanstvenih in tehničnih) besedil. To je omogočilo nastanek ne samo različno obsežnih pisnih korpusov, temveč pozneje tudi govornih korpusov. Mnogi formalni jezikoslovci so korpusno jezikoslovje dolgo obravnavali kot neobstoječe področje (Aarts 2001: 5). Boulton (2008: 43) je zaostanek

<sup>6</sup> Na spletu.

<sup>7</sup> Na spletu.

<sup>8</sup> Na spletu.

Francije na tem področju pojasnil s kulturnimi razlogi in tudi z dejstvom, da ni bilo dovolj raziskav o vrednotenju učinkovitosti učenja s pomočjo korpusov.

Uporaba korpusov pri poučevanju jezikov in prevajanju danes ni nič novega. V strokovni literaturi iz devetdesetih let in začetka 21. stoletja najdemo pojem podatkovno usmerjenega učenja (angl. *data-driven learning*) (Johns 1990) in poudarek na uporabi korpusov pri učenju prevajanja (Aston 1999). Vendar kljub razvoju virov in korpusov opažamo, da uporaba slednjih še vedno ni zelo razširjena med predavatelji (Römer 2006) in prevajalci (Kübler 2011). Raziskave o uporabi korpusov med prevajanjem so se začele šele razmeroma nedavno (Pearson 2003).

Uvedba korpusov v jezikovno didaktiko se je v Franciji pojavila kot nova metodologija in del pristopa tako k jezikoslovju kot k didaktiki (Boulton, Tyne 2014). Strokovnjaki za didaktiko na frazeodidaktičnem področju s pomočjo korpusnega jezikoslovja izluščijo reprezentativne frazeološke prvine, nato pa na podlagi korpusa določijo zaporedja leksemov, ki so umeščeni v realni kontekst. To učencem omogoča razumevanje in pravilno uporabo teh jezikovnih prvin. Po drugi strani s tehnologijo podprto pojmovanje jezika, na katerem temeljijo dobro zasnovani korpusi, predstavlja nove vire za didaktiko pa tudi za frazeoterminologijo in strokovno prevajanje. Raziskave, osredotočene na strokovno prevajanje in njegovo poučevanje, preučujejo različne vidike besedil, ki jih ustvarijo študenti (Frankenberg-Garcia 2009).

Od leta 2001 skupina profesorjev, ki se ukvarjajo s strokovnimi jeziki (Kübler 2011), na Oddelku UFR EILA (Izobraževalna in raziskovalna enota za medkulturne študije uporabnih jezikov na pariški univerzi Université Paris Cité) izvaja metodologijo poučevanja strokovnega prevajanja z uporabo orodij korpusnega jezikoslovja, in sicer v okviru magisterija s področja jezikovne industrije in strokovnega prevajanja (ILTS). Ta metodologija združuje pristop korpusnega jezikoslovja, terminološko analizo, ustvarjanje leksikalnih virov pa tudi sodelovanje s strokovnjaki s tega področja, s ciljem izboljšati razumevanje in prevajanje strokovnih besedil. V okviru te metodologije se že od leta 2013 sistematično zbirajo vsa ciljna besedila, pri čemer se izvaja kontrastivna analiza prevodov s pomočjo korpusnih podatkov. Primerjave so narejene z orodji za tekstometrično analizo (Fleury, Zimina 2014). Študenti, vpisani v magistrski študijski program ILTS, smer angleščina, vsako leto za pripravo prevoda s strokovnega področja opravijo terminološke analize in sestavijo po več slovarskih gesel z zbiranjem virov v spletnem slovarju ARTES (Pecman, Kübler 2011; Pecman 2012; Kübler, Mestivier, Pecman 2016; Pecman, Gledhill 2018).

#### **4 Izdelava in opis francosko-slovenske spletne baze podatkov**

Z razvojem digitalne dobe so korpusi postali nepogrešljiv vir na večini področij, ki se ukvarjajo s proučevanjem jezika, med drugim na področju leksikografije, frazeologije, pri poučevanju in v kontrastivnih raziskavah. Velika količina podatkov nam omogoča lažji dostop do gradiva, hkrati pa se pri tem pokaže tudi potreba po poznavanju različnih postopkov in izvornih jezikovnih virov.

Pred začetkom gradnje korpusa strokovnega oz. specializiranega jezika je pomembno določiti področje opazovanja in s tem možen nabor besedil (Gorjanc, Logar 2007: 640). Strokovni korpus je »reprezentativni vzorec jezikovnih zvrsti v okviru določenega strokovnega, socialno ali geografsko opredeljenega specialnega jezika« (Vintar 2008: 79). Namen korpusa strokovnih besedil je lahko tudi pridobitev jezikovnih podatkov za izdelavo terminološkega slovarja določene stroke. Korpusni pristop h gradnji terminološkega slovarja ponuja tehnološko in metodološko naprednejše možnosti vpogleda v inventar terminov na določenem področju (Logar Berginc, Vintar 2008).

V študijskem letu 2017/2018 je med Oddelkom za prevajalstvo Filozofske fakultete Univerze v Ljubljani in Oddelkom UFR EILA Univerze Paris Cité začel potekati bilateralni projekt Proteus. Projekt ARTES se osredotoča na razvoj in izdelavo virov za spletni slovar. Podatkovna baza trenutno vsebuje jezikovne podatke, povezane s strokovno terminologijo in frazeologijo, ter medpodročne leksikološko-slovnicične strukture. To vključuje razvoj terminoloških in frazeoloških virov v francoščini in slovenščini za pomoč pri strokovnem prevajanju z uporabo spletnega slovarja ARTES.<sup>9</sup>

Spletna zbirka podatkov ARTES je leksikografski vir, ki vsebuje veliko uporabnih informacij za strokovno prevajanje in strokovno pisanje, je pa tudi izobraževalno orodje. Študenti, vpisani v magistrski študij na Oddelku za prevajanje, zbrane podatke vsako leto vnašajo v bazo ARTES. Študenti se tako naučijo izdelati slovar za francosko-slovenski jezikovni par, ki bi hkrati lahko služil kot pripomoček prevajalcem in strokovni pripomoček pri pisanju. Dostop do baze podatkov omogoča raziskovalna skupina CLILLAC (Center za medjezikovno posredovanje, leksikologijo in angleško jezikoslovje). Ekipa je poleg informacijskega orodja prispevala tudi različne pristope in svoje izkušnje, saj že več let vodi ta projekt na področju strokovne terminologije in frazeologije za francosko-angleško jezikovno kombinacijo. Kljub temu lahko baza sprejme približno petdeset jezikov.

Viri v bazi podatkov ARTES so prosto dostopni v obliki terminoloških sestavkov ali slovarskih vnosov, do katerih pridemo s pomočjo iskalnega vmesnika. Poizvedba lahko zajema več zavihkov, pri čemer lahko po eni strani uporabimo »Terminološki slovar« (iskanje po terminu, jeziku in področju) in »Slovar izrazov« (iskanje po diskurzivni funkciji), lahko pa uporabimo funkcijo »iskanja po več kriterijih«, ki je prav tako dostopna prek iskalnega vmesnika. Slednja omogoča iskanje po virih z nizom poizvedb, s katerimi opredelimo iskalne kriterije. Iskanje po več kriterijih je mogoče v multidisciplinarnih virih, in sicer po »terminih, strokovnih izrazih, pomenskih razredih in po komentarjih«, obenem pa je mogoče tudi za »transdisciplinarne izraze«. Za funkcije »termini, strokovni izrazi in transdisciplinarni izrazi« je iskanje mogoče po izbranem jeziku, lahko pa se razširi tudi na iskanje po terminu, slovnici kategoriji, vrsti vnosa, kontekstu, izrazu ali področju.

---

<sup>9</sup> Iskalni vmesnik je dostopen na spletu.

## 5 Izdelava spletnega francosko-slovenskega slovarja ARTES

Postopek ustvarjanja virov vključuje gradnjo primerljivega korpusa na strokovnem področju, ki mu sledi ročno dodajanje sestavkov ali transdisciplinarnih izrazov (tj. generične kolokacije) ter njihovega prevoda, pri čemer je treba upoštevati kontekst oziroma diskurz. Študenti se tako seznanijo z iskanjem informacij, korpusnim jezikoslovjem in ustvarjanjem terminoloških in frazeoloških virov. Razvijajo veščine in pridobivajo znanja, ki so pri strokovnem prevajanju pogosto ključna. Poleg strokovnega doprinosu, ki ga takšni projekti nudijo, torej ne gre zanemariti didaktične in pedagoške dodane vrednosti razvoja strokovnega dvojezičnega slovarja.

Gradivo za slovar mora ustrezati zasnovi slovarja. Relevantnost gradiva oz. izbira korpusa glede na koncept je temeljnega pomena (Vidovič Muha 2009: 35). Študenti so od začetka projekta ustvarili korpuse, ki pokrivajo področja medicine, družbene politike, geologije in ekonomije. V študijskem letu 2020–21 so med področja dodali tudi krasoslovje. Cilj projekta je bil razviti leksikografsko bazo podatkov na tem strokovnem področju, zato je bila metodologija, s katero so študenti zagotovili, da so zbrani podatki kakovostni, razdeljena na tri dele: (1) izdelava francosko-slovenskega primerljivega korpusa; (2) izvleček specifičnih terminov in kolokacij ter (3) vnašanje specifičnih terminov in kolokacij v spletni slovar ARTES.

### 5.1 Raziskave in izdelava korpusov

Prednosti uporabe korpusov, zlasti pri strokovnem prevajanju, so bile že večkrat omenjene in dokazane (Kübler 2011). Študenti prejmejo podrobna navodila za izdelavo in uporabo korpusa ter začnejo ustvarjati lastno bazo podatkov z zbiranjem virov iz besedilnih korpusov. Delo na korpusu poteka s pomočjo konkordančne programske opreme *Sketch Engine*, ki omogoča enostavno združevanje številnih dokumentov. Orodje je opremljeno z zmogljivim proizvedovalnim jezikom, ki hitro zagotovi pristne vire in omogoča prikaz podatkov po iskanem terminu oziroma kolokaciji. Učenci lahko tako izbirajo med viri, ki jih predlaga *Sketch Engine*, in odključajo dokumente, ki bi lahko bili vključeni v korpus. Iskanje se izvaja po ključnih besedah in glede na odgovore, vezane na pogostost pojavljanja in/ali težavnost prevoda termina oz. kolokacije.

### 5.2 Uvod v osnove in izdelava strokovnih gesel

Po izdelavi dveh primerljivih korpusov, francoskega in slovenskega, študent izlušči več terminov in kolokacij. Priskrbeti mora vsaj 10 popolnih dvojezičnih sestavkov (imenovanih dolgi sestavki), ki predstavljajo prevajalski problem na določenem strokovnem področju, v tem primeru s področja krasoslovja. Slednje nato umesti v slovar ARTES. Ker imajo študenti malo izkušenj s sestavljanjem strokovnih sestavkov, se pogosto pojavi težava pri izbiri termina oziroma izraza. Opažamo namreč, da študenti radi vnesejo strokovni termin, ki je podoben v obeh jezikih in ne predstavlja prevodne težave: na primer *stalagmite* (fr) in *stalagmit* (sl) oziroma termin, ki je že na voljo v različnih dvojezičnih slovarjih. Dovoljen je samo en takšen vnos, ker študentom

omogoča, da se na začetku naučijo uporabljati bazo podatkov ARTES, ne da bi jih skrbel pretirano zapleten strokovni prevod.

Strokovni sestavki se nato vnesejo v spletni slovar. Primer dolge oblike sestavka je prikazan v preglednici 1. Vključuje naslednje prvine: termin v izvornem jeziku; navedba izvornega jezika; slovnicihna kategorija termina; področje(-a); kontekst(-i); definicija(-e) (vir); kontekst(-i) in vir; kolokacija(-e) in sočasnost(-i); semantične povezave in posredne povezave, če obstajajo; prevod in opombe (jezikovne in/ali tehnične). Enaka shema se uporabi tudi za ciljni jezik. Na koncu vsakega vnesenega termina se prikaže grafični prikaz.

Prav tako je možno vnesti samo glosarski sestavek, ki zajema termine iz določene besedila. Takšen sestavek vsebuje izraz v izvornem jeziku, slovnicihno kategorijo termina, navedbo izvornega jezika, področje, definicijo (njen vir), prevod in obratno za ciljni jezik.

### 5.3 Povzetek rezultatov za krasoslovje

Strokovni sestavki, ki so jih študentje predlagali s področja krasoslovja, nam omogočajo, da ugotovimo in razvrstimo nekatere razlike med obema jezikoma. Kot bi lahko pričakovali, je prevod pogovornih izrazov drugačen. Rezultate smo združili in jih predstavljamo v preglednici 2.

Začetne raziskave (Vaupot 2020) na področju medicinske terminologije so pokazale pogostost rabe znanstvenih terminov (90 % v francoščini proti 80 % v slovenščini) in predvsem leksikalnih dvojnic. Številni medicinski termini imajo lahko znanstveno ozadje, ki izvira iz besede latinskega ali grškega izvora, ali pa jih je mogoče prevesti z enakovrednim »ljudskim« izrazom: na primer termin *ictère* (iz latinščine *icterus* ali iz grščine *ikteros*) se v slovenščini lahko prevede bodisi z uporabo izposojenke (*ikterus*) ali njene pogovorne ustreznice (*zlatenica*). Na področju krasoslovja smo prišli do drugačnih zaključkov. Znanstveni termini in leksikalne dvojnice so sicer prisotni (npr. *Protée*, *Proteus*, *človeška ribica*, *Proteus*), vendar ne prevladujejo, zato je utemeljeno izdelati slovar ved o Zemlji in okolju.

## 6 Vključevanje generičnih kolokacij

Vemo, da slovnica in slovar delujeta komplementarno (Žele 2014: 1). Podobno velja za frazeologijo oz. kolokacije, katerih pomen je lahko zanimiv tako za leksikografa kot za opis v slovarju. Po pripravi primerljivih korpusov lahko študent tako preide na drugo vajo, v okviru katere oblikuje diskurzivno frazeološko bazo.

Podatkovna baza ARTES ima tako zavihek »Sestavki«, ki omogoča ustvarjanje zgoraj omenjenih terminoloških sestavkov, pa tudi zavihek »Generične kolokacije« ali »Transdisciplinarni izrazi«. Ker je sočasnost dobra prvina besedilne kohezije, v tem delu prepoznamo kolokacije, ki bi lahko bile pogojene z besedilno zvrstjo ali tipom



**Preglednica 1:** Primer primerljivega strokovnega sestavka.

Sestavek 1 — francoščina	Sestavek 2 — slovenščina
<b>AQUIFÈRE</b> <i>nom, masculin, terme</i>	<b>VODONOSNIK</b> <i>samostalnik, moški spol, termin</i>
<u>Domaine(s)</u> : – Géologie générale. Hydrosphère. Eau en général. Hydrologie [CDU 2004].	<u>Področje(-a)</u> : – Splošna geologija. Hidrosfera. Voda na splošno. Hidrologija [CDU 2004].
<u>Auteur de la fiche</u> : Nom et prénom, Étudiant(e) en Master de traduction, 2020–2021 <u>Définition(s)</u> : – réservoir dans lequel l'eau souterraine circule et s'accumule [Source : définition de l'étudiant] – Formation contenant de l'eau (lit ou strate), constituée de roches perméables, de sable ou de gravier, et capable de céder des quantités importantes d'eau. [Source : <i>Dictionnaire Termium Plus - TERMIUM Plus®</i> : ( <a href="https://www.btb.termiumplus.gc.ca/tpv2alpha/alpha-fra.html?lang=fra">https://www.btb.termiumplus.gc.ca/tpv2alpha/alpha-fra.html?lang=fra</a> )]	<u>Avtor sestavka</u> : Priimek in ime, študent(-ka) magistrskega programa Prevažanje, 2020–2021 <u>Definicija(-e)</u> : – geološka plast, ki zbira in pretaka (podzemno) vodo [Vir: Definicija študenta] – Hranišče talne vode. [Vir: Fran Ramovš idr. – <i>Slovenski pravopis</i> – Slovenska akademija znanosti in umetnosti, Znanstvenoraziskovalni center Slovenske akademije znanosti in umetnosti, Inštitut za slovenski jezik Frana Ramovša ZRC SAZU (2001)]
<u>Contexte(s) et collocation(s)</u> : – Sous les couches profondes de glace chaude, il disparaît (fenêtres thermiques), permettant à l'eau issue du glacier de pénétrer directement dans les <b>aquifères karstiques</b> . [Source : Marian Pulina, Jean-Noël Salomon - <i>Les karsts des régions climatiques extrêmes</i> - Presses Universitaires de Bordeaux - page(s) 11-97 ( <a href="https://books_openedition.org/pub/1043">https://books_openedition.org/pub/1043</a> )]	<u>Kontekst(-i) in kolokacija(-e)</u> : – Prvi, <i>aluvialni vodonosnik z medzrnsko poroznostjo</i> je kvartarne starosti. [Vir: <i>Kakovost podzemne vode v Sloveniji v letu 2009</i> , 3012 – Dravska kotlina – Ocena kemijskega stanja in trendov vodnega telesa podzemne vode ( <a href="http://www.arso.gov.si/vode/podzemne_vode/publikacije_in_poročila/03_Krška_kotlina.pdf">http://www.arso.gov.si/vode/podzemne_vode/publikacije_in_poročila/03_Krška_kotlina.pdf</a> ) (2009) revija.si/dokument.aspx?id=439) (2005)]
<u>Concurrent</u> : modèle de l'aquifère Lien sémantique générique : nappe souterraine	<u>Sočasnost</u> : območje vodonosnika Generična pomenska povezava: podzemna voda
<u>Traduction</u> : – vodonosnik (Slovène) <u>Note linguistique</u> : « Aquifère » peut également s'utiliser comme adjectif (qui porte ou qui contient de l'eau). <u>Note technique</u> : Le terme « aquifère » peut également s'utiliser dans d'autres domaines, notamment dans la géologie, les gisements pétroliers, l'alimentation de l'eau et dans le vocabulaire général.	<u>Prevod</u> : – aquifère (Français) <u>Jezikovna opomba</u> : Sopomenka izraza »vodonosnik« je »akvifer«. <u>Tehnična opomba</u> : Izraz »vodonosnik« se uporablja predvsem na področju hidrologije in hidrogeografije.



**Preglednica 2:** Povzetek rezultatov študentov s področja krasoslovja.

Vrste	Primeri iz študentskih korpusov
<p><b>1. Skoraj enaki znanstveni termini:</b> razlike v črkovanju (npr. naglasno znamenje <i>izgine</i>; francoski <i>c</i> in slovenski <i>k</i>) ali glasoslovju (denazalizacija in izgovarjava končnega <i>n</i>)</p> <p><b>1a.</b> podobna izgovarjava, v slovenščini je na koncu soglasnik</p> <p><b>1b.</b> drugačna predpona, vendar enaka izgovarjava ou = u</p> <p><b>1c</b> v slovenščini se doda končni <i>-a</i></p>	<p>spéléologie/speleologija siphon/sifon catavotron/katavotron</p> <p>dolomi(t)e/dolomit stalagmite/stalagmit stalactite/stalaktit</p> <p>ouvala/uvala</p> <p>diacalse/diaklaza drainage/drenaža estavelle/estavela</p>
<p><b>1č.</b> končnica <i>-gie</i> [i] v slovenščini postane <i>-gija</i> [gija]</p> <p><b>1d.</b> denazalizacija in končnica na <i>-a</i> za samostalnik in na <i>-čen</i> za pridevnik v slovenščini</p> <p><b>1e.</b> različne pripone -tion ali <i>-sion</i> = <i>-ija</i> -ion = <i>-ij</i> -ite = <i>-it</i> -ié = <i>-je</i> ali <i>-is</i></p>	<p>géomorphologie/geomorfologija géologie/geologija</p> <p>concrétion/konkrecija antarctique/antarktičen</p> <p>érosion/erozija alluvion(s)/aluvij pisolite/pizolit polié/polje; lapié/lapis</p>
<p><b>2. Slovníčna inverzija</b> sam.+prid. = prid.+sam.</p> <p><b>3. Različni termini</b></p> <p><b>3a.</b> znanstveni termini z različnimi priponami, raba latinščine v slovenščini</p> <p><b>3b.</b> znanstveni termin, preveden s pogovornim izrazom</p> <p><b>3.c</b> sopomenke francoskih terminov so prevedene z enim in istim slovenskim terminom</p> <p><b>3.č</b> francoski termin, preveden z več slovenskimi sopomenskimi termini</p> <p><b>3.d</b> pogovorni oz. neznanstveni izraz, preveden s svojo ustreznico (najpogosteje)</p>	<p>cône rocheux/skalnati stožec</p> <p>protée (anguillard)/proteus (anguinus)</p> <p>pisolite/jamski biser spéléothème/kapnik</p> <p>exutoire, exurgence/kraški izviri trémie, aven/podor protée, olm/proteus rivière souterraine/ponikalnica poisson humain, salamandre des grottes, salamandre blanche/človeška ribica, močeril</p> <p>lapiaz/lapis, škraplje drainage/drenaža, osuševanje</p> <p>résurgence/ponor; chantoir/ponikva aquifère/vodonosnik fistuleuses/cevčice draperie/jamska zavesa gour/ponvice, itd.</p>

diskurza. V razdelku »Sestavek« študent vključi kolokacije, ki so značilne za področja, zapisana v terminoloških datotekah, medtem ko so generične kolokacije povezane s prevladujočim tipom diskurza in se ne uporabljajo za označevanje informacij ali znanj z določenega področja. Vnesejo se v rubriko »Generične kolokacije«. Študent lahko tako v bazo ARTES doda od pet do deset generičnih kolokacij in njihovih prevodov. Te kolokacije so izbrane glede na pogostost pojavljanja v besedilih. Študent izvede obliko-slovnostno-skladenjsko označevanje, da ugotovi skladske vloge posameznega termina, in filtriranje, da izloči nizkofrekvenčne termine, za katere meni, da so neinformativni ali da ne povzročajo občutnih težav pri prevajanju. Distribucijska analiza kolokacij znotraj besedila omogoča ponazoritev obnašanja in vloge kolokacij na ravni zgradbe besedila. Trenutno je v bazi ARTES navedenih približno 350 specifičnih kolokacij in okoli 250 generičnih slovenskih kolokacij.

Izvedena je bila tudi raziskava, s katero smo želeli poudariti oblikoskladske razlike generičnih kolokacij, vnesenih v podatkovno bazo ARTES in izluščenih iz študentskih prevodov (Vaupot 2021). Ugotovili smo sledeče:<sup>10</sup>

- (1) Slovniciška kategorija **glag.+sam.** ostaja enaka v obeh jezikih, z izjemo dodanega določnega člena v francoščini (*sprejeti ukrepe/adopter les mesures*).
- (2) Slovenska struktura **prid.+sam.** se v francoščini spremeni v **sam.+prid.** (*pravična obravnava/traitement équitable*) oziroma **sam.+pred. (de)+ sam.** (*tiskovna konferenca/conférence de presse*).
- (3) Najpogostejša slovenska kategorija – **sam.+sam.** – se v francoščino prevede s strukturami:
  - **sam.+pred.+sam.** (*varstvo potrošnikov/protection du consommateur*)
  - **sam.+prid.** (*združitev družine/regroupement familial*)
  - **glag.+pred.+sam.** (*odpoved imuniteti/renoncer à l'immunité*)
  - **samostalniška besedna zveza** (*sklenitev sporazuma/mise en place de l'accord*)

Ugotovili smo, da si francoščina in slovenščina delita sledeče slovniciške strukture:

- **glag.+sam.** (pogostost pojavljanja strukture: 86 %)—*imeti pogum za/avoir le courage de*
- **glag.+pred.+sam.** (pogostost: 10 %)—*imeti na zalogi/avoir en stock*
- **glag.+sam.+pred.** (pogostost: 4 %)—*izvajati pritisk na/faire pression sur*

Polekvivalentne (SL/FR) slovniciške strukture so pogostejše:

- **glag.+prid./prid.+glag.** (pogostost: 39 %)—v francoščini pride do zamenjave najpogostejše kategorije: *diplomatsko predstavništvo/mission diplomatique*

<sup>10</sup> Na spletu..

- **prid.+sam./sam.+pred.+sam.** (pogostost: 17 %) *izredne razmere/situation d'urgence*
- **glag.+prid.+sam./glag.+sam.+prid.** (pogostost: 12 %) — *imeti ključno vlogo/jouer un rôle clé*
- **sam.+sam./glag.+sam.** (pogostost: 12 %) — ta kategorija vključuje glagolnik na *-anje*, ki se večinoma prevede z glagolom (*izgrajevanje ugleda/promouvoir l'image*).
- **sam.+sam./sam.+pred.+sam.** (pogostost: 12 %) — *varstvo potrošnikov/protection du consommateur*
- **pred.+sam./pred.+sam.+pred.** (pogostost: 8 %) — *pod vodstvom/sous la direction de*

Opažamo, da so glagolske generične kolokacije pogosto prevedene s podobno slovnično strukturo, za razliko od samostalniških kolokacij, ki se prevajajo z bolj raznolikim naborom slovničnih kategorij.

## 7 Zaključek

Izdelava leksikografskega projekta, zbiranje podatkov, prevajanje in kontrastivna analiza so inovativen pristop k preučevanju francosko-slovenskega jezikovnega para. Spletni slovar ARTES je pomemben zaradi svojega univerzalnega pristopa, ki združuje jezikovne, terminološke, frazeološke in diskurzivne prvine, pomembne za prevajanje in sporazumevanje v strokovnem jeziku. Baza podatkov ponuja zanimiv okvir za dvojezični ali večjezični pristop. V tem primeru nam je omogočila razvoj novega pedagoškega pristopa, vezanega na bodoči poklic študentov. S tem leksikografskim podvigom smo podatkovno bazo ARTES obogatili s skoraj 400 popolnimi ali delnimi slovenskimi terminološkimi sestavki, opremljenimi z njihovim francoskim prevodom.

Ta leksikografski pristop se je izkazal za poučnega in koristnega. Študenti lahko v praksi uporabljajo pristop k leksikografski obdelavi, ki jo je razvila francoska univerza, kar omogoča tudi poglobljanje meduniverzitetnega sodelovanja. Naučijo se razvijati metodološko razmišljanje, obenem pa lahko pristop oziroma orodja po potrebi prilagodijo svojim potrebam. Študenti lahko na koncu oddajo poročilo oz. seminarsko nalogo, kjer lahko razberemo tako študentov napredek na področju terminološkega in frazeološkega razmišljanja, kot tudi njegovo obvladovanje francoskega jezika in njegove prevajalske sposobnosti. Predavatelj lahko nato izpelje leksikografske zaključke oz. preveri napredek študenta pri učenju tujega strokovnega jezika.

Podatki, vneseni v podatkovno bazo ARTES, imajo lahko znanstveni namen, saj omogočajo raziskovanje jezikovnih pojavov v povezavi z ustvarjanjem leksikalnih virov, specializiranim prevajanjem in analizo specializiranega diskurza. Podatkovno bazo lahko tako uporabimo kot model za analizo terminologije in diskurzne frazeologije za francosko-slovenski jezikovni par na področju strokovnega prevajanja. Projekt je lahko tudi dobra osnova za pojmovni in kognitivni pristop k terminologiji ali za preučevanje različnih kolokacijskih pojavov.

## VIRI IN LITERATURA

- Bas AARTS, 2001: Corpus linguistics, Chomsky and fuzzy tree structures. *Corpus linguistics and linguistic theory*. Ur. Christian Mair, Marianne Hundt. Amsterdam: Rodopi. 5–13.
- Guy ASTON, 1999: Corpus use and learning to translate. *Textus* 12. 289–313.
- Alex BOULTON, 2008: Esprit de corpus: Promouvoir l'exploitation de corpus en apprentissage des langues. *Texte et Corpus* 3. 37–46.
- Alex BOULTON, Henry TYNE, 2014: *Des Documents authentiques aux corpus: Démarches pour l'apprentissage des langues*. Paris: Didier.
- Neva CEBRON, 2020: Korpusni pristop kot metoda raziskovanja in učenja tujega strokovnega jezika. *Studia Universitatis Hereditati Znanstvena Revija Za Raziskave in Teorijo Kulturne Dediščine* 8/2. 45–64.
- Anne CONDAMINES, 2005: Linguistique de corpus et terminologie. *Langages* 157/1. 36–47.
- Serge FLEURY, Maria ZIMINA, 2014: Trameur: A Framework for Annotated Text Corpora Exploration. *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: System Demonstrations*. Dublin. 57–61.
- Ana FRANKENBERG-GARCIA, 2009: Are translations longer than source texts? A corpus-based study of explicitation. *Corpus Use and Translating*. Ur. Allison Beeby, Patricia Rodríguez-Inés, Pilar Sánchez-Gijón. Amsterdam, Philadelphia: John Benjamins. 47–58.
- Vojko GORJANC, Nataša LOGAR, 2007: Od splošnih do specializiranih korpusov: Načela gradnje glede na njihov namen. *Razvoj slovenskega strokovnega jezika. Obdobja – metode in zvrsti* 24. Ur. Irena Orel, Ljubljana: Center za slovenščino kot drugi/tuji jezik pri Oddelku za slovenistiko Filozofske fakultete Univerze v Ljubljani. 637–50.
- Anton GRAD, 2004 [1971]: *Francosko–slovenski slovar*. Ljubljana: Cankarjeva založba.
- Viktor JESENIK, Narcis DEMBSKIJ, 2008 [1990]: *Slovensko–francoski slovar*. Ljubljana: DZS.
- Tim JOHNS, 1990: From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *CALL Austria* 10. 14–34.
- Natalie KÜBLER, 2011: Working with different corpora in translation teaching. *New Trends in Corpora and Language Learning*. Ur. Ana Frankenberg-Garcia, Lynne Flowerdew, Guy Aston. London: Bloomsbury. 62–80.
- Natalie KÜBLER, Alexandra MESTIVIER, Mojca PECMAN, 2016: Exploitation quantitative de corpus de traductions annotés selon la typologie d'erreurs pour améliorer les méthodes d'enseignement de la traduction spécialisée. *International Conference on Statistical Analysis of Textual Data (JADT)*. Ur. Damon Mayaffre, Céline Poudat, Laurent Vanni. Nice: Université Nice Sophia Antipolis. 731–42.
- Nataša LOGAR BERGINČ, Špela VINTAR, 2008: Korpusni pristop k izdelavi terminoloških slovarjev: od besednih seznamov in konkordanc do samodejnega luščenja izrazja. *Jezik in slovstvo* 53/5. 3–17.
- Adriana MEZEG, 2020a: *Polstavki v francoščini in njihovo prevajanje v slovenščino – korpusni pristop*. Ljubljana: Znanstvena založba Filozofske fakultete Univerze v Ljubljani.

- Adriana MEZEG, 2020b: Definicija polstavka in njegova raba v sodobni prevodni slovenščini (na gradivu prevodov iz francoščine). *Slavistična revija* 68/2. 265–82.
- Bine MLAČ, Marjeta HUMAR, Stanko KLINAR, Bojan LESKOŠEK, Bojan POLLAK, Franci SAVENC, Pavle ŠEGULA, Albin VENGUST, Stanislav Bojan ZUPET, 2002: *Planinski terminološki slovar: slovensko-angleško-nemško-francosko-italijanski slovar planinskega, alpinističnega, plezalskega izrazja*. Ljubljana: Založba ZRC, ZRC SAZU (Slovarji).
- Jennifer PEARSON, 2003: Using parallel texts in the Translator Training Environment. *Corpora in Translator Education*. Ur. Frederico Zanettin, Silvia Bernardini, Dominic Stewart. Manchester: St Jerome Publishing. 15–24.
- Mojca PECMAN, Natalie KÜBLER, 2011: ARTES: an online lexical database for research and teaching in specialised translation and communication. *Proceedings of the First International Workshop on Lexical Resources, WoLeR 2011*. Ur. Benoit Sagot. Ljubljana: European Summer School. 86–93.
- Mojca PECMAN, 2012: Étude lexicographique et discursive des collocations en vue de leur intégration dans une base de données terminologiques. Special issue on Terminology, Phraseology and Translation. *JosTrans* 18. 113–38.
- Mojca PECMAN, Christopher GLEDHILL, 2018: How trainee translators and their teachers deal with phraseological units in the ARTES database. *Equivalences* 45/1–2. 237–59.
- Bernard QUEMADA, 1968: *Les Dictionnaires du français moderne (1539-1863). Etude sur leur histoire, leurs types et leurs méthodes*. Paris: Didier.
- Ute RÖMER, 2006: Pedagogical applications of corpora: some reflections on the current scope and a wish list for future developments. *Zeitschrift für Anglistik und Amerikanistik* 54/2. 121–34.
- Stanislav SUHADOLNIK, 1961: Josip Pavlica: Frazološki slovar v petih jezikih. *Jezik in slovstvo* 6/6. 200–05.
- Sonia VAUPOT, 2020: Création de ressources lexicographiques Français-Slovène d'aide à la traduction spécialisée. *Lexikos* 30. 540–60.
- Sonia VAUPOT, 2021: Creating a bilingual dictionary of collocations: a learner-oriented approach. *Indonesian Journal of Applied Linguistics* 10/3. 762–70.
- Ada VIDOVIČ MUHA, 2009: Poskus določitve meril slovarskega pomena. *Strokovni posvet o novem slovarju slovenskega jezika*. Ur. Vojko Gorjanc. Ljubljana: Založba ZRC, ZRC SAZU. 27–37.
- Špela VINTAR, 2008: *Terminologija: Terminološka veda in računalniško podprta terminologija*. Ljubljana: Znanstvena založba Filozofske fakultete Univerze v Ljubljani.
- Andreja ŽELE, 2014: Primeri glagolske rabe z vidika posebnosti v slovnično-pomenskih razmerjih. *Slavistična revija* 62/1. 1–16.

---

## SUMMARY

Slovenia's accession to the European Union created a need for bilingual and multilingual lexical and terminological resources, including for the French-Slovenian language pair. The article describes the development of the French-Slovenian online terminological and phraseological dictionary ARTES (Aide à la Rédaction de Textes Scientifiques or Dictionary-assisted writing tool for scientific communication). The database is a multilingual, multidomain language resource that offers an approach to lexical resources. Designed in 2010 by the CLILLAC-ARP team of researchers from the University Paris Cité, the database is the outcome of several projects aimed at the creation of resources for LSPs and specialised translation. The article describes the result of a bilateral project between the Department of Etudes Interculturelles de Langues Appliquées (EILA) at the Université Paris Cité and the Department of Translation Studies of the Faculty of Arts at the University of Ljubljana. The genesis and structure of the dictionary are explained, as well as the pedagogical approach to the creation of the database. The work involved several phases, including the determination of the subject area (the field of Karst studies), the creation of comparative corpora, the choice of the terminological and/or phraseological entries, and the search for translation equivalents. The results obtained with regard to the French-Slovenian terminology of the field of Karst studies and generic collocations are presented on the basis of the data entered in the online database.