Dejan Gabrovšek – Janoš Ježovnik – Matic Pavlič – Andrej Perdih

Metodologija priprave nabora slik za pedagoški portal Franček: izbor, semantična opredelitev in nadaljnja uporaba

Cobiss: 1.01
https://doi.org/10.3986/JZ.32.1.03

Članek obravnava metodologijo in izzive pri izbiri slik za pedagoški portal Franček, za katerega je bilo ročno izbranih 64.063 licenčno prostih in 120 avtorsko zaščitenih slik za 19.760 iztočnic. Posebna pozornost je namenjena postopku semantične opredelitve slik glede na slovarski pomen. Nato je na primeru psiholingvistične naloge povezovanja besede s sliko predstavljena uporaba slikovne zbirke zunaj slovarskega okvira.
Ključne besede: slika, slovar, pomen, leksikografija, psiholingvistika, semantika

Methodology of Image Set Compilation for the Franček Pedagogical Portal: Selection, Semantic Definition, and Further Use

This article presents the methodology and key challenges involved in selecting images for the Franček educational portal, for which 64,063 free licence and 120 copyrighted images were manually selected for 19,760 headwords. Particular attention is devoted to the process of semantic annotation of images with respect to individual dictionary senses. Finally, the article demonstrates the use of the image collection beyond the dictionary framework through a psycholinguistic word–picture matching task.
Keywords: image, dictionary, sense, lexicography, psycholinguistics, semantics

1 Uvod

Digitalno okolje omogoča enostavnejše vključevanje slik v slovarje v primerjavi s tiskanimi izdajami. Zmanjšale so se prostorske omejitve pri prikazu slik, prav

Dejan Gabrovšek ▪ ZRC SAZU, Inštitut za slovenski jezik Frana Ramovša, Ljubljana ▪ dejan.gabrovsek@zrc-sazu.si ▪ https://orcid.org/0000-0002-4995-1814
Janoš Ježovnik ▪ ZRC SAZU, Inštitut za slovenski jezik Frana Ramovša, Ljubljana ▪ janos.jezovnik@zrc-sazu.si ▪ https://orcid.org/0000-0002-6642-9833
Matic Pavlič ▪ Univerza v Ljubljani, Pedagoška fakulteta ▪ matic.pavlic@pef.uni-lj.si ▪ https://orcid.org/0000-0001-8248-8860
Andrej Perdih ▪ ZRC SAZU, Inštitut za slovenski jezik Frana Ramovša, Ljubljana ▪ andrej.perdih@zrc-sazu.si ▪ https://orcid.org/0000-0002-2248-9666

Prispevek je nastal v okviru projekta Množična raziskava razširjenosti slovenskih besed (J6-50199), ki ga sofinancira ARIS. Temelji na raziskovalnih podatkih, ki se hranijo na Inštitutu za slovenski jezik ZRC SAZU in so dostopni na osnovi utemeljene prošnje.

CC BY SA

tako je na voljo več načinov njihovega pridobivanja. Slikovno gradivo si v tujini postopoma utira pot v enojezične razlagalne slovarje za odrasle rojene govorce,[1] pogosteje pa je uporabljeno v šolskih slovarjih in slovarjih za tujce: za angleščino npr. *Oxford Learners Dictionaries*.[2] V procesu nastajanja enojezičnih razlagalnih slovarjev za rojene govorce je vključevanje slik v slovar sicer zaželeno, vendar ne nujno potrebno (Lišková – Šemelík 2024: 164), z utemeljitvijo, da je že priprava osnovnih vsebin tovrstnega slovarja časovno in kadrovsko izredno zahtevna, zato ni vedno na voljo dovolj časa še za dodatne vsebine. Po drugi strani se je prav zaradi postopnega vključevanja slik v digitalno objavljene slovarje povečalo tudi zanimanje raziskovalcev za obravnavo slikovnega materiala (Biesaga 2016; 2017a; 2017b; Dziemianko 2022; 2024; Kallas idr. 2024; Lew 2010; Lew idr. 2018; Liu 2015). Najbolj odprti za slikovne materiale so slovarji za otroke in tujce, a tudi pri teh obstajajo omejitve glede števila besed oziroma pomenov, ponazorjenih s slikami (Biesaga 2016: 100). Za ta tip slovarjev se praviloma namensko pripravijo ustrezne ilustracije, kar je še vedno zamuden proces. Uporaba že obstoječih slik, npr. v slovarju *Duden online*,[3] lahko proces olajša, pri tem pa se lahko pojavljajo omejitve pravne narave.

V slovenskem prostoru slikovno gradivo vključuje na primer *Slovar za govorce slovenščine kot drugega in tujega jezika* (SLOGOST), ki nastaja pri Centru za jezikovne vire in tehnologije in Centru za slovenščino kot drugi in tuji jezik Univerze v Ljubljani (Klemen idr. 2023). Slike so prikazane pri posameznih pomenih, npr. pri iztočnicah *pošta*, *potovati*, *pouk*.

2 Postopek zbiranja slik za portal Franček

Za portal Franček slovarsko gradivo ni bilo namensko ilustrirano, ampak je bilo s spleta pridobljeno veliko število že obstoječih slik, tako fotografij kot ilustracij. S pomočjo namenske aplikacije,[4] v kateri je bil prek iskalnika Bing[5] zbran širok nabor slik, je bilo za 19.760 od skupno okoli 100.000 iztočnic ročno izbranih 64.063 slik z licencami, ki dovoljujejo odprto uporabo.[6] Nabor je bil razširjen z

---

1 Za nemščino npr. *Duden online* (gl. op. 3), za češčino *Slovník současné češtiny* (*Internetový slovník současné češtiny verze 3.0* založbe Lingea): *meruňka* 'marelica', https://www.nechybujte.cz/slovnik-soucasne-cestiny/merunka?.

2 Npr. *house* 'hiša', https://www.oxfordlearnersdictionaries.com/definition/english/house_1?q=house.

3 Npr. *Stechpalme* 'bodika', https://www.duden.de/rechtschreibung/Stechpalme; *Ananas* 'ananas', https://www.duden.de/rechtschreibung/Ananas – glede na informacijo o avtorskih pravicah slike izvirajo od različnih ponudnikov.

4 Aplikacijo je razvil Klemen Andrejc Kac na Fakulteti za elektrotehniko, računalništvo in informatiko Univerze v Mariboru pod vodstvom Milana Ojsterška. Razvit je bil tako iskalniški del kot uporabniški vmesnik za izbiranje med najdenimi slikami.

5 Za iskalnik Bing smo se odločili, ker je omogočal brezplačno množično pridobivanje slik. Izbira iskalnika je sicer zaradi manjšega nabora slik, ki jih ponuja v primerjavi z drugimi iskalniki, v nadaljnjih korakih lahko vplivala na število in kakovost ustreznih slik.

6 Pripisana licenca »Javna last« ali »Free to share and use commercially« (*dovoljena komercialna raba in deljenje*).

dogovorom o uporabi slik, ki jih hrani Slovenski etnografski muzej, v okviru katerega je bilo pridobljenih še 120 slik. Slike so bile ročno izbrane v namenski aplikaciji, v različici portala Franček v letu 2025 pa je po izločitvi nekaterih slovarskih sestavkov v 18.889 slovarskih sestavkih (od skupno 100.393 sestavkov) prikazanih 60.640 slik. Prvotni izbor slik je bil opravljen leta 2019, torej v obdobju pred pojavom enostavno dostopnih umetnointeligenčnih aplikacij za ustvarjanje slik. Slike so bile povezane z ustreznimi slovarskimi sestavki v eSSKJ in drugi izdaji SSKJ, ki predstavljata osnovo geslovnika za portal Franček (Perdih idr. 2021), niso pa še bile pripisane posameznim pomenom znotraj teh sestavkov.

2.1 Postopek iskanja

Aplikacija je samodejno iskala slike, ki bi ustrezale danim besedam, med najdenimi pa so bile nato ročno izbrane najustreznejše. Če aplikacija ustreznih slik ni ponudila, so jih z vpisovanjem pomensko podobnih besed, ustreznih besed v tujih jezikih ter taksonomskih imen za rastline in živali dodatno poiskali ocenjevalci slik. Ti so poskušali slike izbirati tako, da bi odsevale čim več različnih pomenov iztočnice, čeprav slik v tej fazi še niso povezovali s slovarskimi pomeni, torej slike na portalu Franček niso bile pomensko opredeljene. Ocenjevalci so sledili načelu, da se za vsako iztočnico izbere več slik. Izbrane slike so morale ponazarjati vsaj en pomen in nastopati v tipičnih okoliščinah, kjer se ta pomen besede realizira v zunajjezikovnem svetu. Prednost so imele slike brez motečih elementov, ki uporabniku lahko vzbujajo dvome, kaj slika ponazarja. Pomembna je bila tudi njihova ustrezna tehnična kakovost. Zaradi izbire iskalnika in načelne odločitve za slike v odprtem dostopu nekatere kakovostne slike, ki jih sicer najdemo na spletu, niso bile na voljo.

Ker je osrednji namen portala Franček uporaba v pedagoškem kontekstu, so bile izbrane slike, primerne za otroke in mladostnike, zato nekaterim besedam niso bile dodane, in sicer predvsem tistim, ki so kakor koli stereotipno zaznamovane, kot na primer *lepotica*, *lepota* (ker nismo hoteli dati sporočila, kakšna naj bi bila lepa oseba), *cigan* (kjer je iskalnik predlagal le slike umazanih otrok), ter besedju s področja spolnosti ipd.

3 Slike v razmerju do slovarskega pomena

Slika je slikovno razlagalno sredstvo oziroma vizualna upodobitev slovarskega pomena. Posamezni pomen iztočnice lahko ponazarja več slik, pri posamezni iztočnici pa je lahko prikazana tudi le ena slika, čeprav ima leksem več pomenov. Možno je, da več slik ponazarja le en (tj. isti) pomen večpomenske besede.

Slike glede na način prikaza pomena delimo na dve skupini: (1) slovarskopomenske slike in (2) asociativne slike.

3.1 Slovarskopomenske slike

Slovarskopomenske slike pojavnost, ki je opredeljena s slovarsko razlago pri posameznem pomenu, ponazarjajo tako, da se razlaga in slika vzajemno potrjujeta. Slovarsko razlago ‘priprava iz lesene paličice z grafitnim vložkom za pisanje, risanje’ torej ponazarja slika svinčnika. Povezava med sliko in razlago je pri slovarskopomenskih slikah razmeroma neposredna in nedvoumna. Slovarskopomenske slike ponazarjajo isti pomen, kot ga opisuje slovarska razlaga. Navadno prikazujejo en pomen, lahko tudi podpomen, nastal po metonimiji. Ta tip je značilen za prikaz konkretnih pojavnosti, zlasti predmetov, rastlin in živali.[7] Slednje smo večinoma iskali s pomočjo njihovih taksonomskih imen, kar omogoča natančnost prikaza, zlasti če je razlika med posameznimi vrstami organizmov manjša ali je ne poznamo.

S slikami so večinoma ponazorjeni prvi pomeni, neprvi pa le izjemoma. Pomenska členitev na portalu Franček (prim. Perdih idr. 2021; 2024) sledi členitvi v slovarjih SSKJ2 in eSSKJ. Pri samostalniku *ključ* je npr. s sliko ponazorjen prvi pomen (‘kovinska priprava za zaklepanje in odklepanje ključavnice’), (skoraj) nemogoče pa je bilo najti slike, ki bi prikazovale metaforične abstraktne pomene, npr. ‘z dogovorom določen način, sistem, po katerem se kaj rešuje’. Pri besedi *srce* je bilo mogoče najti slike, ki izmed osmih pomenov ponazarjajo prvi (‘votel mišičast organ v prsni votlini, ki poganja kri po žilah’) in zadnji pomen (‘igralna karta z enim ali več znaki v obliki srca’), medtem ko je preostale pomene praktično nemogoče konkretno slikovno ponazoriti, npr. ‘ta organ pri človeku kot središče čustvovanja, zavesti’ ali ‘skupek človekovih značajskih, čustvenih značilnosti’. Beseda *sesalnik* ima zgolj en pomen, ki ga ponazarjajo štiri slike; besedi *sedeti* so dodane slike, ki ponazarjajo le prvega od šestih pomenov.

3.2 Asociativne slike

Asociativne slike ponazarjajo pomen, ki sicer ni povsem prekriven s tistim, ki ga opisuje razlaga, vendar med njima obstaja močna asociativna povezava. Pomen dane besede uporabnik razbere z ugotavljanjem konteksta poimenovanja za pojavnost, upodobljeno na sliki. Pomen torej izhaja iz širšega konteksta, predstavljenega s sliko. Ta način omogoča predstavitev besed, katerih pomena ni mogoče prikazati s slovarskopomenskimi slikami, npr. abstraktnih samostalnikov in glagolnikov, glagolov in pridevnikov.

Tak primer je slika adventnega venca, ki ponazarja besedo *advent*: samostalnika *advent* tako ne predstavlja neposredno, prikazuje pa tipično konkretno pojavnost, povezano z adventom, ki se jo da prikazati s sliko. Pomen uporabe

7 Torej takih, ki so zaznavni z vidom.

asociativnih slik je v preseganju zamejitve prikaza na zlasti konkretne samostalnike (Dziemianko 2022: 219).

Glagol *plavati* se ponazarja s sliko plavalcev, sliko orke in slikama rib v morju: izbiro se da utemeljiti s tem, da izbrane slike plavanja sicer ne prikazujejo neposredno, ampak prikazujejo ljudi in živali, ki plavajo; iz konteksta lahko razberemo, da je prikazano tudi dejanje plavanja. Razmeroma predvidljivo skupino pomenov predstavljajo pari motivirajočih in metonimičnih pomenov, pri katerih vsaj enega od njiju ni mogoče ponazoriti s slovarskopomensko sliko, obenem pa je isto sliko mogoče uporabiti za ponazoritev obeh pomenov. V to skupino se uvršča npr. beseda *akvarel*: izhodiščni pomen je ‘slikarska tehnika, pri kateri se slika na papir s prosojnimi vodenimi barvami’, podpomen pa ‘slika v tej tehniki’. Vse slike na portalu Franček pri geslu *akvarel* prikazujejo primere slik, narisanih v tehniki akvarela, torej prikazujejo le podpomen; po asociativnem pristopu lahko s pomočjo istih slik ponazorimo tudi abstraktni osnovni pomen besede, tj. tehniko, v kateri so te slike nastale. Na podoben način je mogoče prikazovati tudi nekatere vrste medmetov: npr. pri medmetu *bim* je prikazana slika, ki ponazarja zvon (ne pa samega zvoka zvona, ki ga medmet *bim* označuje).

3.3 Kako slike prikazujejo pomen v razmerju do slovarskih razlag

Posamezna slika lahko prikazuje le enega od možnih slikovnih prikazov pomena besede in ne zajame prototipa ali vseh možnosti pojavitve te pojavnosti, npr. vseh možnih oblik, velikosti in barv jabolk: če je na sliki le rdeče jabolko, bi si kdo lahko razlagal, da jabolka drugačnih barv ne obstajajo (Dziemianko 2022: 219). Na portalu Franček se to vprašanje delno rešuje z naborom več slik, s katerimi se skuša prikazati čim več različnih možnosti pojavljanja posamezne pojavnosti. Vsekakor pa to dokazuje, da posamezna slika ne more vedno ustrezno ponazarjati pomena besed v slovarjih.

Slika prikazuje posamezno pojavitev pojavnosti, ki jo na pomenski ravni opredeljuje razlaga. To mora uporabnik posplošiti na vse primere pojavnosti, ki jih ta pomen opredeljuje, medtem ko razlaga predstavlja posplošeno predstavo pomena in ne zajema nujno vseh možnih uresničitev pojavnosti pomena besede – razlago lahko uporabnik nato poveže s posamezno pojavitvijo pojavnosti.

Raziskave (Dziemianko 2022: 220–221, 226; 2024) kažejo, da se pomen prikliče hitreje, če sta na voljo tako zapis besede kot tudi slika, ki predstavlja njen pomen. To je še posebej pomembno ravno pri pedagoškem portalu, kot je portal Franček.

3.3.1 Osnovni statistični podatki o slikah na portalu Franček

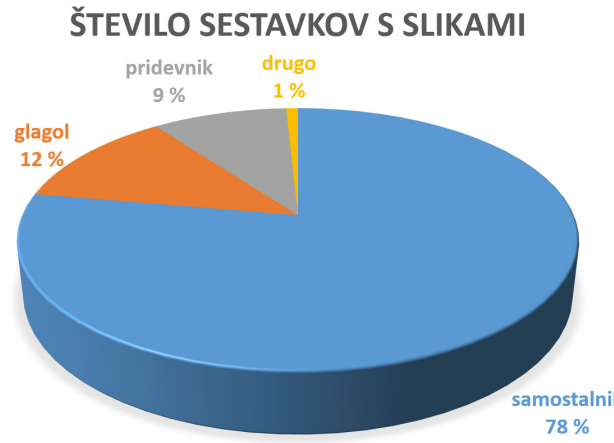
V tem razdelku obravnavamo razporeditev slik glede na besedno vrsto iztočnic, zanima pa nas tudi, koliko slik je pripisanih posameznim besedam. Podatki za najnovejšo različico baze portala Franček iz leta 2025, kot so prikazani v preglednici 1, pričakovano kažejo, da so z vsaj eno sliko najpogosteje ponazorjeni samostalniki (77,75 % vseh iztočnic s slikami), razmeroma dobro zastopani so tudi glagoli (12,14 %) in pridevniki (9,29 %). Iztočnice drugih besednih vrst (in iztočnice, ki jim ni pripisana nobena besedna vrsta) so ponazorjene le redko. Zadnji stolpec preglednice kaže, da je med tistimi slovarskimi sestavki, pri katerih je prikazana vsaj ena slika, povprečno število slik največje pri samostalnikih (3,36), nekoliko nižje pa pri prislovih, pridevnikih in glagolih (med 2,94 in 2,54).

Preglednica 1: Število slovarskih sestavkov na portalu Franček, različica 2025, ki vsebujejo vsaj eno sliko[8]

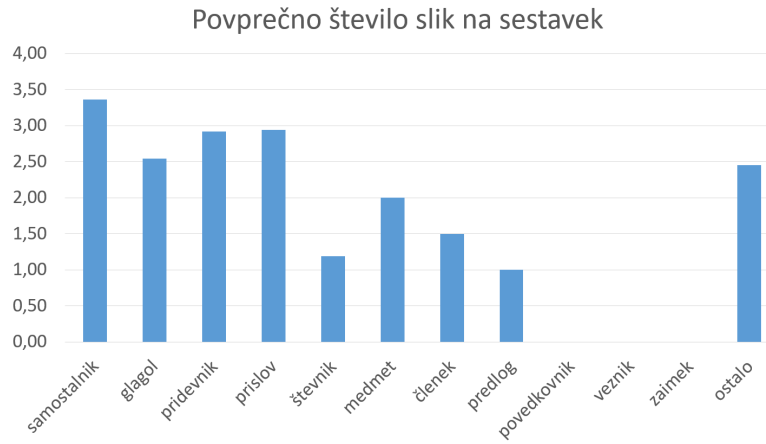
Besedna vrsta	Število sestavkov s slikami	Delež	Število slik	Povprečno število slik na sestavek
samostalnik	14.687	77,75 %	49.344	3,36
glagol	2.294	12,14 %	5.827	2,54
pridevnik	1.755	9,29 %	5.116	2,92
prislov	63	0,33 %	185	2,94
številnik	27	0,14 %	32	1,19
medmet	26	0,14 %	52	2
členek	4	0,02 %	6	1,5
predlog	2	0,01 %	2	1
povedkovnik	0	0,00 %	0	0
veznik	0	0,00 %	0	0
zaimek	0	0,00 %	0	0
ostalo	31	0,16 %	76	2,45
SKUPAJ	18.889	100,00 %	60.640	3,21

Stolpec Delež predstavlja delež sestavkov določene besedne vrste, ki vsebujejo sliko, glede na vse slovarske sestavke, ki vsebujejo slike. Zadnji stolpec prikazuje povprečno število slik na sestavek.

8 Skupina *ostalo* označuje slovarske sestavke brez pripisane besedne vrste in kazalčne sestavke, ki so po večini dejansko samostalniki.



Slika 1: Število slovarskih sestavkov na portalu Franček, različica 2025, ki vsebujejo vsaj eno sliko



Slika 2: Povprečno število slik pri sestavkih, ki imajo vsaj eno sliko

Dodatno smo preverili, kolikšen delež iztočnic, ki pripada isti besedni vrsti, je ponazorjenih s slikami. Čeprav torej večina slik ponazarja samostalnike, je to v resnici pričakovano, saj prav samostalniki predstavljajo največji delež slovarja. Preglednica 2 ob številu slik glede na besedne vrste prikazuje primerjavo s številom vseh iztočnic iste besedne vrste in njune deleže. Pokaže se, da so samostalniki res najlažje upodobljivi, vendar ne tako izrazito: 27,31 % samostalnikov je ponazorjenih z vsaj eno sliko. Slikovno je upodobljenih kar 21,09 % številnikov, in sicer zaradi prikaza s številkami. Številniki tako predstavljajo drugo najbolj s slikami predstavljeno besedno vrsto, sledi kategorija *ostalo*, kjer so zbrani iztočnice brez

besednovrstne opredelitve in kazalčna gesla, pri katerih je besednovrstna opredelitev razvidna šele sekundarno, večina med njimi pa je polnopomenskih. Temu sledijo glagoli (12,32 %) in pridevniki (7,94 %), iztočnice drugih besednih vrst pa so slikovno predstavljene le izjemoma.

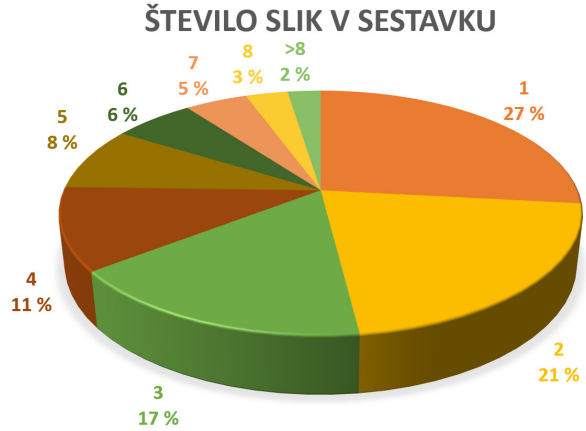
Preglednica 2: Delež slovarskih sestavkov s slikami glede na vse slovarske sestavke iste besedne vrste

Besedna vrsta	Število sestavkov s slikami	Število sestavkov iste besedne vrste	Delež
samostalnik	14.687	53.775	27,31 %
glagol	2.294	18.614	12,32 %
pridevnik	1.755	22.117	7,94 %
prislov	63	4.372	1,44 %
števnik	27	128	21,09 %
medmet	26	676	3,85 %
členek	4	172	2,33 %
predlog	2	113	1,77 %
povedkovnik	0	10	0,00 %
veznik	0	64	0,00 %
zaimek	0	127	0,00 %
ostalo	31	225	13,78 %
SKUPAJ	18.889	100.393	

Največ slovarskih sestavkov, ki vsebujejo slike, je ponazorjenih z eno sliko, kot je prikazano v preglednici 3. Z vsako dodatno sliko se zmanjša število slovarskih sestavkov s takim številom slik.

Preglednica 3: Število slovarskih sestavkov glede na število slik na portalu Franček, različica 2025

Število slik v sestavku	Število sestavkov	Delež glede na vse sestavke s slikami
1	5048	26,72 %
2	4023	21,30 %
3	3117	16,50 %
4	2062	10,92 %
5	1588	8,41 %
6	1128	5,97 %
7	873	4,62 %
8	587	3,11 %
>8	463	2,45 %



Slika 3: **Delež slovarskih sestavkov z določenim številom slik na portalu Franček, različica 2025**

3.3.2 Tehnične lastnosti slik

S tehničnega stališča so slike zelo raznolike glede velikosti, ločljivosti, višine, širine in drugih lastnosti, kar je zaradi njihovega raznolikega izvora pričakovano. Slike s spleta so shranjene v formatu .png, slike Slovenskega etnografskega muzeja pa so lahko shranjene tudi v drugih formatih. Za uporabo na portalu Franček so normalizirane, in sicer so slike, višje od 120 slikovnih točk, zmanjšane na višino 120 slikovnih točk in ob tem dodatno kompresirane. S tem se tako na strežniku kot pri uporabniku zmanjša količina prenesenih podatkov in skrajša čas nalaganja. Ker v okviru portala Franček ni potrebe po tem, da bi se slika ogledovala v izvirni velikosti in kakovosti, se s tem zmanjša tudi potreben prostor za shranjevanje na spletnem strežniku. Dodatno so bile pripravljene tudi slike z višino do 240 slikovnih točk (višje slike so bile zmanjšane in kompresirane), ki so trenutno prav tako shranjene na strežniku portala, vendar niso aktivno uporabljane v okviru Frančkovih vsebin. Kot bomo videli v nadaljevanju, so lahko uporabne za druge namene, ki niso zahtevni z vidika količine prenosa podatkov na strežniku.

3.4 Povezovanje slik s slovarskimi pomeni

V letu 2025 je bilo s pomočjo namenske aplikacije pri tistih samostalniških iztočnicah, katerih pomenski opis na portalu Franček izhaja iz SSKJ2, opravljeno ročno povezovanje slik in ustreznih slovarskih pomenov ter določanje slike, ki dani pomen predstavlja najbolje. Vsaka slika je bila povezana s pomenom oziroma pomeni, ki ga/jih predstavlja. Izmed slik, ki prikazujejo določeni pomen, je bila izbrana najboljša, ki bo primerna za nadaljnjo uporabo. Ročno je do zdaj pomensko označenih 15.555 slik pri 4543 večpomenskih samostalnikih. Pri 9752 enopomenskih samostalnikih je vključenih 32.691 slik, ki prikazujejo edini pomen. Skupno so torej ročno pomensko označene slike pri 97,33 % vseh samostalnikov.

3.4.1 Možnost uporabe orodij umetne inteligence pri povezovanju slik s slovarskimi pomeni

Na manjšem vzorcu smo preizkusili tudi učinkovitost avtomatskega povezovanja slik s slovarskimi pomeni (Perdih idr. 2025b). Iz celotnega nabora večpomenskih samostalnikov smo naključno izbrali štiri skupine s po sto iztočnicami s pripadajočimi slikami (kot smo že omenili, so bile slike predhodno povezane z iztočnicami, ne pa tudi s specifičnimi pomeni). Uporabili smo jezikovni model GPT-4o mini, ki smo mu kot vhodne podatke podali posamezno sliko, iztočnico in seznam pomenskih opisov za to iztočnico. Analizo smo opravili na vzorcu 1520 procesiranih slik oziroma 398 iztočnic s skupaj 4260 pripadajočimi pomeni (Perdih idr. 2025b: 856). Primerjava z zlatim standardom, ki ga je pripravil človeški ocenjevalec, je pokazala ujemanje v 85,1 % primerov (preciznost 0,807, priklic 0,856); vrednost Cohenove kape $\kappa = 0{,}70$ in vrednost $F1 = 0{,}831$ kažeta na visoko stopnjo ujemanja (Perdih idr. 2025b: 857–858).

Nato sta isti nabor iztočnic, pomenov in slik individualno ocenila dva druga ocenjevalca. Kljub manjšemu neujemanju med človeškimi ocenjevalci je primerjava z GPT-4o mini dala podobne, še marginalno boljše rezultate: 85,7-odstotno ujemanje ($\kappa = 0{,}71$) oziroma 86,9-odstotno ujemanje ($\kappa = 0{,}73$). Skupna primerjava vseh štirih naborov (avtomatsko označenega in trikrat ročno označenega nabora) je pokazala nekoliko nižje, a vseeno močno ujemanje, in sicer v 76,7 % primerov (vrednost Fleissove kape $\kappa_F = 0{,}74$) (Perdih idr. 2025b: 861–862).

Trenutno je zaključena prva faza povezovanja slik s pomeni in izbora najboljše slike za pomen, v nadaljevanju pa bo potrebno dodatno poenotenje meril za povezovanje slik s pomeni na podlagi pridobljenih izkušenj ter na podlagi izpopolnjene metodologije še izvedba povezovanja slik s pomeni, ki so zajeti v eSSKJ.

4 Perspektiva uporabe slik

Na pedagoškem portalu Franček so vse slike trenutno prikazane na ravni celotnega slovarskega sestavka, in sicer v zaporedju, kot so (naključno) zabeležene v bazi podatkov. S pomensko opredelitvijo slik in izborom najboljše slike za posamezni pomen jih bo mogoče razvrščati glede na pomen ter tehnično in vsebinsko kakovost ali pa preiti na prikazovanje slik pri posameznih pomenih z upoštevanjem njihove kakovosti.

Izbrane slike iz baze portala Franček so vključene v *Novi etimološki slovar slovenskega jezika* (Furlan 2017–), pri čemer so ponazorjene tudi iztočnice, ki na portal Franček niso uvrščene, a predstavljajo sinonim (npr. *borovnjak*[9] k *brinovka*[10]).

---

9 URL: https://fran.si/207/nessj-novi-etimoloski-slovar-slovenskega-jezika/4441487/borovnjak.
10 URL: https://www.franček.si/beseda/e001589/brinovka.

Zbrane slike je mogoče uporabiti tudi za druge, ne le slovarske namene. Na uporabnost slik vpliva njihova opremljenost z različnimi podatki – kot je bilo omenjeno, so v tem trenutku na voljo zapis upodobljene besede (iztočnice) ter njena besednovrstna in pomenska opredelitev. Prav pomenska opredelitev glede na slovarski pomen, ki je bila delno opravljena za potrebe psiholingvistične raziskave, omenjene v nadaljevanju, odpira nove možnosti uporabe. Slike v bazi portala Franček so namreč povezane s slovarskimi sestavki eSSKJ in SSKJ2. Poleg tega poteka povezovanje posameznih pomenov med SSKJ2 in *Sinonimnim slovarjem slovenskega jezika* (prim. Treven 2022), saj sta ta dva slovarja na pomenski ravni zaradi svoje sorodnosti do določene mere primerljiva in delno prekrivna. Večina samostalnikov in samostalniških besednih zvez, ki so pomensko opredeljeni v *Sinonimnem slovarju slovenskega jezika* (tj. tisti, ki vstopajo v sinonimno razmerje z vsaj enim svojim pomenom), je uvrščenih v *Pojmovnik SSSJ*, ki je sestavni del tega slovarja in predstavlja umestitev pomenov v hierarhično prikazanem pojmovnem sistemu (Snoj idr. 2018). V *Pojmovnik SSSJ* pa niso uvrščeni vsi samostalniki in samostalniške zveze, saj v nekaterih primerih, npr. *drobnocvetni rogovilček*, njihov »pomen ne vključuje razločevalnih lastnosti, relevantnih za katero od pojmovnih podskupin« (Snoj 2019: 206).

Opredelitev slik (samostalnikov) na pomenski ravni znotraj SSKJ2, delna vzpostavljenost pomenskih povezav s *Sinonimnim slovarjem slovenskega jezika* in njihova opredeljenost v *Pojmovniku SSSJ* v manjši meri že omogoča skupinjenje besed, njihovih pomenov in pripadajočih slik po posameznih pomenskih skupinah.

4.1 Uporaba slik v psiholingvističnih raziskavah

Povezanost slik s semantičnimi podatki predstavlja podatkovni vir za psiholingvistično raziskavo, ki je načrtovana v letu 2026, in sicer za nalogo povezovanja besede s sliko (*word-picture matching task*), ki se uporablja za preučevanje leksikalnega procesiranja pri jezikovnem razumevanju.

Povezovanje besede s sliko je postopek, pri katerem je informantu zaporedno zvočno ali pisno predstavljeno poimenovanje nanosnikov (leksemov), informant pa mora iz nabora upodobitev (navadno med dvema do štirimi slikami) izbrati tisto, ki najbolje ustreza pomenu besede oziroma poimenovanemu nanosniku. V digitalizirani različici naloge to stori s pritiskom na ustrezno tipko, v papirni različici pa pokaže na izbrano upodobitev. Za uspešno opravljanje naloge mora informant primerjati jezikovni in slikovni dražljaj. To je možno le, če si bodisi vizualizira nanosnika, ki je vkodiran v jezikovni dražljaj, ali poimenuje nanosnike, ki so upodobljeni na prikazanih slikah.

Najpomembnejši vir standardiziranih slikovnih dražljajev, uporabljen v več tisoč raziskavah in prilagojen desetinam jezikov, je nabor črno-belih črtnih risb vsakdanjih predmetov (Snodgrass – Vanderwart 1980). Ta nabor je bil kasneje razširjen v zbirko 520 predmetov in 275 dejanj v okviru projekta International

Picture-Naming Project (Bates idr. 2003; Székely idr. 2004), v nadaljevanju pa so nastale tudi druge pomembne slikovne baze (Bonin idr. 2003; Cycowicz idr. 1997; Moreno-Martínez – Montoro 2012). Med njimi velja posebej omeniti dve. The Multilingual Picture Databank (MultiPic) obsega 750 standardiziranih črtnih risb z barvnim polnilom ter normami za usklajenost poimenovanja in kompleksnost upodobitve v španščini, nemščini, angleščini, nizozemščini, italijanščini in francoščini; namenjena je večjezičnim raziskavam (Duñabeitia idr. 2018). Bank of Standardized Stimuli (BOSS) pa obsega 930 barvnih fotografij brez ozadja skupaj z angleškimi normami za usklajenost poimenovanja, gledišča, nanosnika in njegove kategorije ter normami za poznavanje kompleksnosti upodobitve in uprizorljivost rokovanja z nanosnikom (Brodeur idr. 2010; 2014). Standardizirane slikovne baze so pomemben vir dražljajev, saj so slike poenotene glede ključnih lastnosti, med katerimi izstopata kompleksnost upodobitve nanosnika in njegova oprijemljivost.

Raziskave učinkov oprijemljivosti (*concreteness*) nanosnikov v različnih kognitivnih nalogah kažejo, da testiranci lekseme z oprijemljivimi nanosniki prikličejo hitreje in natančneje kot lekseme z neoprijemljivimi nanosniki (Paivio 1991). Posledično se v nalogah poimenovanja in povezovanja praviloma uporabljajo le leksemi z oprijemljivimi nanosniki, ki jih je mogoče neposredno upodobiti (medtem ko je neoprijemljive nanosnike mogoče prikazati le posredno, na primer s simboli). Iz težnje po čim večji oprijemljivosti nekateri raziskovalci namesto risb priporočajo uporabo fotografij nanosnikov (Stern – White 2003), kar so upoštevali tudi pri zbirki BOSS (Brodeur idr. 2010; 2014). Res pa je, da fotografije lahko vključujejo tudi (pre)več nepomembnih podrobnosti, ki jih je v risbah mogoče izpustiti.

Na podlagi obstoječih podatkov slik v okviru podatkovne baze Franček in njihove pomenske opredelitve z navezavo na SSKJ2 in *Pojmovnik SSSJ* smo pregledali večpomenske besede in odstranili slike, ki so se nanašale na neosnovne pomene, tako da so ostale le slike, ki predstavljajo osnovni pomen posamezne besede. Pri tem smo izhajali iz seznama besed, ki je bil narejen za množično raziskavo razširjenosti slovenskih besed (Perdih idr. 2025a). Poleg tega smo, kot je predstavljeno v razdelku 3.4, za vsak osnovni pomen besede z več prirejenimi slikami izbrali eno, najbolj reprezentativno sliko. Tako je ostalo 1333 besed, ki so razvrščene v hierarhične pomenske skupine na treh ravneh (npr. za besedo *gajba*: predmet – posoda – shranjevanje/tovorjenje). Tem besedam smo nato dodali še podatek o pogostosti rabe na podlagi deduplicirane različice besedilnega korpusa Gigafida 2.0 (Krek idr. 2020).

Iz nabora 1333 besed smo ročno izbrali 270 ciljnih besed in pripadajočih ciljnih slik z upodobitvami nanosnikov, ki jih te besede vkodirajo. Za vsako ciljno sliko smo dodali dve motilni sliki, ki prikazujeta nanosnik iz iste pomenske skupine kot ciljna slika. Če bi motilne slike prikazovale nanosnike iz povsem drugih skupin, bi bilo nalogo mogoče rešiti že na osnovi osnovnega prepozna-

vanja pomena, brez polnega dostopa do ciljnega leksema v mentalnem slovarju. Udeleženec bi lahko na primer tudi brez celostnega poznavanja pomena besede *gajba* izločil slike živali ali ljudi in pravilno izbral sliko predmeta. Motilci iz iste pomenske skupine zato povečajo kognitivno zahtevnost naloge in zagotovijo, da udeleženec razlikuje med semantično sorodnimi koncepti, kar omogoča občutljivejše merjenje leksikalnega dostopa in semantične diskriminacije.

V nadaljevanju smo pri izbiri motilnih slik upoštevali tudi frekvenco besed, ki vkodirajo nanosnike na motilnih slikah: izbrali smo tiste, ki imajo glede na ciljno besedo čim bolj podobno frekvenco v korpusu Gigafida 2.0. Pomembno je namreč, da imajo ciljna beseda in besede, ki vkodirajo nanosnike na motilnih slikah, približno enako frekvenco, saj ta močno vpliva na hitrost in natančnost leksikalnega procesiranja. Visokofrekvenčne besede so praviloma procesirane hitreje in z manj napakami kot nizkofrekvenčne, ne glede na vrsto naloge. Če bi se ciljna beseda po frekvenci bistveno razlikovala od motilnih, bi razlike v odzivnem času ali pravilnosti odgovorov lahko odražale učinke pogostosti, ne pa procesov, ki jih želimo preučevati.

Specifične grafične zahteve uporabniškega vmesnika naloge povezovanja besede s sliko nato zahtevajo še strojno ali ročno obrezovanje nekaterih slik, da ustrezajo zahtevanemu razmerju med višino in širino.

Z upoštevanjem pomenske skupine nanosnikov in pogostosti rabe besed med ciljnimi in motilnimi postavkami smo torej zagotovili, da noben odziv nima sistematične prednosti pri procesiranju. S tem smo povečali notranjo veljavnost naloge in primerljivost odzivnih časov ter zmanjšali možnost strategijskega odgovarjanja, pri katerem bi udeleženci do pravilnega odgovora prišli z izločanjem postavk, ki izstopajo po vizualnih, konceptualnih ali leksikalnih značilnostih.

5 Zaključek

Uporaba slikovnega gradiva je v leksikografiji uveljavljena zlasti pri pedagoških slovarjih, obenem pa predstavlja velik potencial tudi za druge vrste slovarjev. V zadnjem obdobju se je z razmahom lažje in hitreje dostopnih slik ter z razvojem orodij za njihovo strojno generiranje proces vključevanja slik v slovarje poenostavil. Pri tem je pomembno ustrezno razumevanje njihove vloge, saj besedilno izražene razlage pomenov praviloma le dopolnjujejo. Tudi pri izbiri slik za portal Franček se je potrdilo, da je slikovna upodobitev najbolj uspešna pri samostalnikih, v manjši meri pa tudi pri glagolih in pridevnikih. Števniki predstavljajo zanimivo skupino, saj je nekatere mogoče ponazoriti s slikovno upodobitvijo številk. Pristop k izbiri slik za portal Franček je bil širok: poleg slik, ki neposredno ustrezajo slovarskemu pomenu, so vključene tudi slike, ki so s pomenom povezane asociativno.

Ker leksikografija sloni na urejenih in strukturiranih podatkih, je nabor slik v povezavi s slovarskimi podatki prenosljiv na slovarje različnih tipov, med njimi tudi take, kjer jih donedavna nismo bili vajeni, npr. na etimološki slovar, uporaben pa je tudi zunaj meja leksikografije. Na primeru priprave slikovnega in slovarskega gradiva za psiholingvistično nalogo povezovanja besede s sliko smo dodatno potrdili smiselnost povezovanja slik s posameznimi slovarskimi pomeni, medsebojno povezovanje slovarjev na pomenski ravni in pomembnost gradnje hierarhično urejenega sistema pomenskih skupin. Slikovno gradivo se tako vklaplja v širši nabor semantičnih in s semantiko povezanih podatkov.

Viri in literatura

Bates idr. 2003 = Elizabeth Bates idr., Timed picture naming in seven languages, *Psychonomic bulletin & review* 10.2 (2003), 344–380, DOI: https://doi.org/10.3758/BF03196494.

Biesaga 2016 = Monika Biesaga, Pictorial Illustration in Dictionaries: the State of Theoretical Art, v: *Proceedings of the 17th EURALEX International Congress, EURALEX 2016*, ur. Tinatin Margalitadze – George Meladze, Tbilisi: Lexicographic Centre, Ivane Javakhishvili Tbilisi State University, 2016, 99–108, http://euralex.org/wp-content/themes/euralex/proceedings/Euralex%20 2016/euralex_2016_007_p99.pdf.

Biesaga 2017a = Monika Biesaga, Pictorial Illustrations in Encyclopaedias and in Dictionaries – a Comparison, v: *Electronic lexicography in the 21st century: proceedings of eLex 2017 conference*, ur. Iztok Kosem idr., Brno: Lexical Computing CZ, 2017, 221–236, https://elex.link/elex2017/wp-content/uploads/2017/09/paper13.pdf.

Biesaga 2017b = Monika Biesaga, Dictionary Tradition vs. Pictorial Corpora: Which Vocabulary Thematic Fields Should Be Illustrated?, *Lexikos* 27 (2017), 132–151, DOI: https://doi.org/10.5788/27-1-1397.

Bonin idr. 2003 = Patrick Bonin – Nathalie Malardier – Alain Méot – Marylène Chalard, A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies, *Behavior Research Methods, Instruments, & Computers* 35.1 (2003) 158–167, DOI: https://doi.org/10.3758/BF03195507.

Brodeur idr. 2010 = Mathieu B. Brodeur – Emmanuelle Dionne-Dostie – Tina Montreuil – Martin Lepage, The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research, *PloS one* 5.5 (2010), e10773.

Brodeur idr. 2014 = Mathieu B. Brodeur – Katherine Guérard – Maria Bouras, Bank of Standardized Stimuli (BOSS) phase II: 930 new normative photos, *PloS one* 9.9 (2014), e106953.

Cycowicz idr. 1997 = Yael M. Cycowicz – David Friedman – Mairav Rothstein – Joan Gay Snodgrass, Picture naming by young children: norms for name agreement, familiarity, and visual complexity, *Journal of Experimental Child Psychology* 65.2 (1997), 171–237, DOI: https://doi.org/10.1006/jecp.1996.2356.

Duden = *Duden online*, https://www.duden.de/.

Duñabeitia idr. 2018 = Jon Andoni Duñabeitia – Davide Crepaldi – Antje S. Meyer – Boris New – Christos Pliatsikas – Eva Smolka – Marc Brysbaert, MultiPic: a standardized set of 750 drawings with norms for six European languages, *Quarterly Journal of Experimental Psychology* 71.4 (2018), 808–816, DOI: https://doi.org/10.1080/17470218.2017.1310261.

Dziemianko 2022 = Anna Dziemianko, The usefulness of graphic illustrations in online dictionaries, *ReCALL* 34.2 (2022), 218–234, DOI: https://doi.org/10.1017/S0958344021000264.

Dziemianko 2024 = Anna Dziemianko, Pictures in Online Dictionaries: Shall We See Them?, *GEMA Online Journal of Language Studies* 24.2 (2024), 31–53, DOI: https://doi.org/10.17576/gema-2024-2402-03.

eSSKJ 2016– = *eSSKJ: Slovar slovenskega knjižnega jezika*, 2016–, https://www.fran.si/iskanje?FilteredDictionaryIds=201&View=1&Query=%2A.

Fran = Portal Fran, www.fran.si.

Franček = Portal Franček, www.franček.si.

Furlan 2017– = Metka Furlan, *Novi etimološki slovar slovenskega jezika*, 2017–, https://www.fran.si/iskanje?FilteredDictionaryIds=207&View=1&Query=%2A.

Kallas idr. 2024 = Jelena Kallas – Kristina Koppel – Katrin Tsepelina, The EKI Picture Dictionary. A Multilingual Tool for A1–B1 Learners, v: *Lexicography and Semantics. Book of Abstracts of the XXXI EURALEX International Congress*, ur. Kristina Despot idr., Zagreb: Institut za hrvatski jezik, 2024, 108–112, https://euralex.jezik.hr/wp-content/uploads/2021/09/Euralex_boa_20.pdf.

Klemen idr. 2023 = Matej Klemen – Špela Arhar Holdt – Senja Pollak – Iztok Kosem – Eva Pori – Polona Gantar – Mihaela Knez, Building a CEFR-Labeled Core Vocabulary and Developing a Lexical Resource for Slovenian as a Second and Foreign Language, v: *Electronic lexicography in the 21st century: proceedings of the eLex 2023 conference*, ur. Marek Medveď idr., Brno: Lexical Computing CZ, 2023, 664–678, https://elex.link/elex2023/wp-content/uploads/118.pdf.

Krek idr. 2020 = Simon Krek – Špela Arhar Holdt – Tomaž Erjavec – Jaka Čibej – Andraž Repar – Polona Gantar – Nikola Ljubešić – Iztok Kosem – Kaja Dobrovoljc, Gigafida 2.0: the reference corpus of written standard Slovene, v: *LREC 2020: Twelfth International Conference on Language Resources and Evaluation*, ur. Nicoletta Calzolari, Paris: ELRA - European Language Resources Association, 2020, 3340–3345, http://www.lrec-conf.org/proceedings/lrec2020/LREC-2020.pdf.

Lew 2010= Robert Lew, Multimodal Lexicography: the Representation of Meaning in Electronic Dictionaries, *Lexikos* 20 (2010), 290–306, DOI: https://doi.org/10.5788/20-0-144.

Lew idr. 2018 = Robert Lew – Rafał Kaźmierczak – Ewa Tomczak – Mateusz Leszkowicz, Competition of Definition and Pictorial Illustration for Dictionary Users' Attention: an Eye-Tracking Study, *International Journal of Lexicography* 31.1 (2018), 53–77, DOI: https://doi.org/10.1093/ijl/ecx002.

Lišková – Šemelík 2024 = Michaela Lišková – Martin Šemelík, Show me the meaning of being lonely... Graphic Illustrations in The Academic Dictionary of Contemporary Czech, v: *Lexicography and Semantics. Book of Abstracts of the XXXI EURALEX International Congress*, ur. Kristina Despot idr., Zagreb: Institut za hrvatski jezik, 2024, 163–165, https://euralex.jezik.hr/wp-content/uploads/2021/09/Euralex_boa_20.pdf.

Liu 2015 = Xiqin Liu, Multimodal Definition: the Multiplication of Meaning in Electronic Dictionaries, *Lexikos* 25 (2015), 210–232, DOI: https://doi.org/10.5788/25-1-1296.

Moreno-Martínez – Montoro 2012 = Francisco Javier Moreno-Martínez – Pedro R. Montoro, An ecological alternative to Snodgrass & Vanderwart: 360 high quality colour images with norms for seven psycholinguistic variables, *PloS one* 7.5 (2012), e37527, DOI: https://doi.org/10.1371/journal.pone.0037527.

Oxford Learners Dictionaries, https://www.oxfordlearnersdictionaries.com/.

Paivio 1991 = Allan Paivio, Dual coding theory: retrospect and current status, *Canadian Journal of Psychology* 45.3 (1991), 255–287, DOI: https://doi.org/10.1037/h0084295.

Perdih idr. 2024 = Andrej Perdih – Kozma Ahačič – Nataša Jakop – Nina Ledinek – Špela Petric Žižić, Semantic Information on the Franček Educational Language Portal for Slovenian, v: *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress*, ur. Kristina Despot idr., Zagreb: Institut za hrvatski jezik, 2024, 155–168, https://euralex.jezik.hr/wp-content/uploads/2021/09/Euralax-XXI-final-web.pdf.

Perdih idr. 2021 = Andrej Perdih – Kozma Ahačič – Janoš Ježovnik – Duša Race, Building an Educational Language Portal Using Existing Dictionary Data, *Journal of Linguistics / Jazykovedný časopis* 72.2 (2021), 568–578, DOI: https://doi.org/10.2478/jazcas-2021-0052.

Perdih idr. 2025a = Andrej Perdih – Dejan Gabrovšek – Matic Pavlič, Izdelava seznama besed za množično raziskavo razširjenosti slovenskih besed, *Slavistična revija* 73.1 (2025), 121–138, DOI: https://doi.org/10.57589/srl.v73i1.4231.

Perdih idr. 2025b = Andrej Perdih – Dejan Gabrovšek – Janoš Ježovnik, Image-to-sense alignment using AI tools, v: *Electronic lexicography in the 21st century (eLex 2025): intelligent*

lexicography. Proceedings of the eLex 2025 conference, ur. Iztok Kosem idr., Bled: Lexical Computing CZ s.r.o., 2025, 852–865, https://elex.link/elex2025/proceedings/.

SLOGOST = *Slovar za govorce slovenščine kot drugega in tujega jezika*, https://lexonomy.cjvt.si/slovar-za-govorce-slovenscine-kot-drugega-in-tujega-jezika/.

Snodgrass – Vanderwart 1980 = Grace Snodgrass – Mary Vanderwart, A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity, *Journal of Experimental Psychology: Human Learning and Memory* 6.2 (1980), 174–215, DOI: https://doi.org/10.1037/0278-7393.6.2.174.

Snoj 2019 = Jerica Snoj, *Leksikalna sinonimija v Sinonimnem slovarju slovenskega jezika*, 2019, DOI: https://doi.org/10.3986/9789610501510.

Snoj idr. 2018 = Jerica Snoj – Martin Ahlin – Branka Lazar – Zvonka Praznik, *Sinonimni slovar slovenskega jezika*, Ljubljana: Založba ZRC, 2018, www.fran.si.

SSKJ2 = *Slovar slovenskega knjižnega jezika*, druga, dopolnjena in deloma prenovljena izdaja, 2014 ([1]1970–1991), www.fran.si.

Stern – White 2003 = Robert A. Stern – Travis White, *NAB, Neuropsychological Assessment Battery: Administration, Scoring, and Interpretation Manual*, Lutz: Psychological Assessment Resources, 2003.

Székely idr. 2004 = Anna Székely idr., A new on-line resource for psycholinguistic studies, *Journal of Memory and Language* 51.2 (2004), 247–250.

Treven 2022 = Matevž Treven, *Medslovarsko pomensko povezovanje na primeru SSSJ in SSKJ 2: nabor iztočnic s sklicevalno razlago glagolnik od X*, diplomska naloga, Univerza v Ljubljani, Filozofska fakulteta, 2022, https://repozitorij.uni-lj.si/IzpisGradiva.php?id=141545.

Summary

Methodology of Image Set Compilation for the Franček Pedagogical Portal: Selection, Semantic Definition, and Further Use

The use of images in lexicography is mostly aimed towards younger users, while at the same time it also represents considerable potential for other types of dictionaries. In recent years, the process of incorporating images into dictionaries has been simplified by the increasing availability of easily and quickly accessible images and by the development of image generation AI tools. In this context, an appropriate understanding of their role is important, as images generally only complement textually explained meanings. The selection of images for the pedagogical Franček portal for Slovenian likewise confirmed that visual representation is most effective for nouns and, to a lesser extent, for verbs and adjectives. Numerals constitute an interesting group, as some of them can be illustrated by visual representations of numbers. The approach to selecting images for the Franček portal was broad: in addition to images that directly correspond to a dictionary sense (strictly lexical images), images that are only associatively related to the meaning were also included (associative images).

Since lexicography relies on organized and structured data, the set of images linked to dictionary data is transferable to dictionaries of various types, even in those where such material has not traditionally been used, for example etymological dictionaries, and it is also applicable beyond the field of lexicography. Using the example of preparing visual and lexicographic material for a psycholinguistic word-picture matching task, we further confirmed the usefulness of linking images to individual dictionary senses, establishing links between senses of different dictionaries, and the importance of building a hierarchically structured system of semantic groups. Visual material thus integrates into a broader set of semantic or semantically related data.