

Developing a comprehensive service for Slovenian language users: the Fran and Franček web portals and language advisory service

Abstract (English)

Since 2014, the ZRC SAZU Fran Ramovš Institute of the Slovenian Language has successfully established a comprehensive service for Slovenian language users. Children and teens can obtain information on nearly 100,000 words via the innovative web portal Franček.si which has been adapted to their needs while general and professional users can simultaneously browse up to forty Slovenian dictionaries on the Fran web portal. In addition, we maintain regular contact with Slovenian language users via two advisory pages: one targeting the general public and one aimed at developing terminology in a multilingual society.

Abstract (Slovenian)

Na Inštitutu za slovenski jezik Frana Ramovša ZRC SAZU smo v času od leta 2014 do danes uspeli vzpostaviti popoln servis za uporabnike slovenskega jezika. Mladi lahko dobijo podatke o skoraj 100.000 besedah na inovativnem, njim prilagojenem portalu Franček.si, splošni in profesionalni uporabniki lahko hkrati brskajo po kar 40 slovarjih slovenskega jezika na portalu Fran, hkrati pa smo v stalnem stiku z uporabniki slovenskega jezika prek dveh svetovalnic: ena je namenjena splošni javnosti, druga pa razvoju terminologije v večjezični družbi.

1. The Fran web portal and language advisory services

In 2014, the Fran Ramovš Institute of the Slovenian Language at ZRC SAZU found itself at an important crossroads. The institute had financial problems and we were unable to meet the new demands of language users at that time. Our online dictionaries could only be accessed one at a time and they did not have a uniform format.¹ Therefore, we decided to completely rework our approach for users of our language manuals.

We combined all major lexicographic sources in one uniform portal system, taking advantage of the fact that practically all basic lexicographic works and sources related to Slovenian had been created at the institute. This allowed us to

¹ Cf. the old website: <http://bos.zrc-sazu.si>.

focus primarily on content-related issues from the very beginning rather than on time-consuming legal procedures to acquire copyright for publishing the works.

The year 2014 saw the launch of the comprehensive Slovenian dictionary portal *Fran.si* (see Perdih 2018, 2020), which brings together all dictionary entries, sources, and materials created by the institute's researchers. The portal, which has recorded an average of over 300,000 searches a day over the past year (bearing in mind that Slovenia's population is just over two million),² provides access to twelve general dictionaries,³ two etymological dictionaries,⁴ five historical dictionaries,⁵ fourteen terminological dictionaries,⁶ five dialect dictionaries,⁷ a dialect atlas,⁸ and two language advisory services, one intended for general users⁹ and one intended for specialists in various fields.¹⁰

The search engine allows users to conduct both simple and complex searches in all dictionaries at once (see Fig. 1); their indexing system allows extremely fast access to as many as 689,941 dictionary entries in various reference works. The order of the displayed search results adapts to users' interests on an ongoing basis (see Fig. 2), making the portal easy to use even for those not familiar with different types of dictionaries, their structures, and their purpose.

² <https://www.fran.si/o-portalu?page=Statistics>.

³ Dictionary of the Slovenian Standard Language, 2nd Edition, Dictionary of the Slovenian Standard Language, 3rd Edition (eSSKJ; 2016–), Synonym Dictionary of Slovenian Language, Slovenian Normative Guide, ePravopis – Slovenian Normative Guide (2016–), Growing Dictionary of the Slovenian Language (2014–), Dictionary of Slovenian Phrasemes, Dictionary of Slovenian Valency, Dictionary of Proverbs and Similar Paremiological Expressions, Dictionary of New Slovenian Words etc.

⁴ Slovenian Etymological Dictionary, NESSJ – New Etymological Dictionary of Slovenian Language (2017–).

⁵ Words of the 16th-Century Slovenian Literary Language, Dictionary of the Slovenian Language in the Works of John Baptist of Sveti Križ, Dictionary of Kastelec and Vorenc (1680–1710), Dictionary of the Language of Marko Pohlin, Dictionary of Old Standard Prekmurje Slovenian, Dictionary of Maks Pleteršnik (1894–1895).

⁶ Concrete Structures, Pharmacy, Law, Automatic Control, Systems and Robotics, Urban Planning, Applied Art, Percussion, Botany, Skiing, Theatre, Beekeeping, Geology, Gemology, Geography, Mountaineering.

⁷ Dictionary of the Črni Vrh Dialect, Dictionary of the Local Dialects of the Dreta Valley, Dictionary of the Bovec Local Dialect, Dictionary of the Kostel Dialect, Dictionary of the Clothing Terminology of the Gail Valley, Local Dialect of the Canal Valley.

⁸ <https://www.fran.si/204/sla-slovenski-lingvisticni-atlas>.

⁹ <https://svetovalnica.zrc-sazu.si>.

¹⁰ <https://isjfr.zrc-sazu.si/sl/terminologisce/svetovanje>.

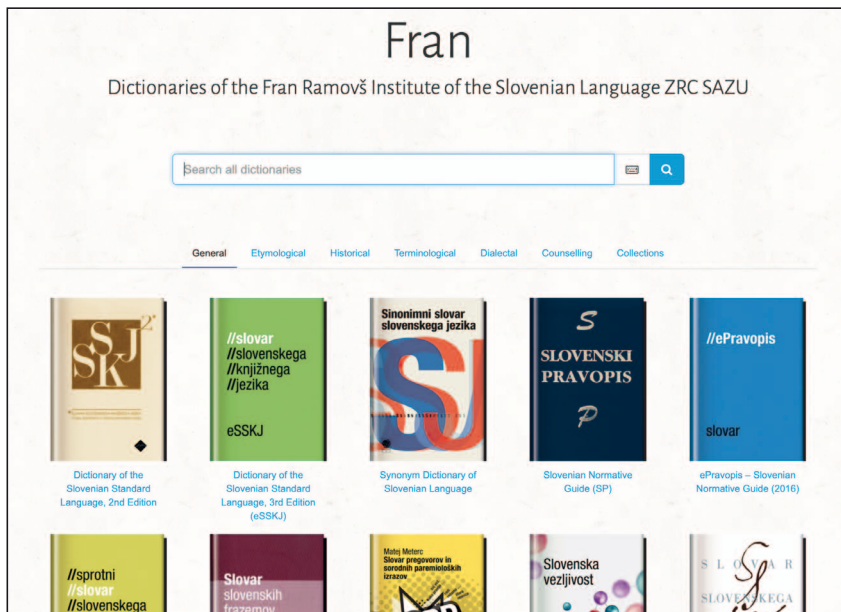


Fig. 1: The Fran web portal: entry page

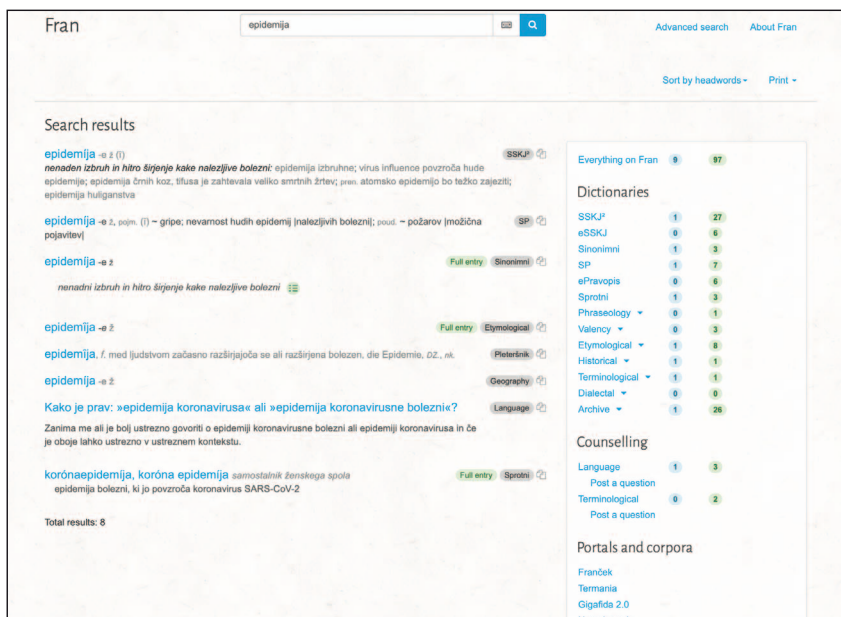


Fig. 2: The Fran web portal: search results

The Fran portal also reacts to its users' queries. Based on search analyses and questions addressed to the general language and terminological advisory service webpages, a dictionary of neologisms¹¹ is being compiled to respond to user needs as they are expressed. For example, during the COVID-19 pandemic we were already able to post explanations of the most frequently used neologisms within a month of the outbreak and we then updated these data on an ongoing basis.

Thanks to Fran's flexibility, the dictionaries produced at the institute can be posted on the portal as works in progress (that is, even before they are completed). The missing dictionary entries are simply covered by those in existing dictionaries.

Users are always told which dictionary they are currently looking at and can quickly access information on that dictionary's characteristics and specific features.¹²

They can send us their questions by e-mail or via a dedicated interface on the language advisory service webpages, and they can also send their suggestions directly via the online form on the Fran portal. In addition to corpus data, the so-called growing dictionaries (that is, dictionaries which are being developed and published in real time) can also rely on data on the most frequently searched words (see Fig. 4), the material on the two advisory pages, and users' direct suggestions and preferences. In this way we can also respond to the language needs of specific groups, such as the deaf, blind, and other vulnerable groups.

Another advantage of Fran is that it displays hits from both advisory pages alongside those from the dictionaries. This means that, in addition to a dictionary entry describing a specific word, users also receive information on special features of its contemporary use.

The problem of the specific role of terminological dictionaries is being solved by incorporating them in the Fran portal. In addition, specialized users can access them on the separate website *Terminologišče*, which provides various kinds of information on terminology, such as a selection of terminological articles and books with active links, various terminological dictionaries, and a terminological advisory service.¹³

¹¹ <https://www.fran.si/132/sprotni-sprotni-slovar-slovenskega-jezika>.

¹² <https://www.fran.si/iskanje?View=1&Query=epidemija>.

¹³ <https://isjfr.zrc-sazu.si/sl/terminologisce>.

The screenshot shows the Fran web portal interface. At the top, there is a search bar with 'epidemija' entered and a search icon. To the right, there are links for 'Advanced search' and 'About the dictionary (SLV)'. Below the search bar, there is a 'Print' button. The main content area is titled 'Entry print-out' and includes several tabs: 'Pomen', 'Stalne zveze', 'Frazologija', 'Izgovor in oblike', 'Normalnost', 'Besedstvoje', and 'Etimologija'. The 'Pomen' tab is selected, showing the entry for 'ébola' (ébola) as a feminine noun. The entry includes a definition: '1. iz medicine nalezljiva virusna bolezen, ki povzroča hude notranje in zunanje krvavitve, razširjena zlasti v Africi'. It also provides information about the word's origin and usage, such as 'Na prizadetem območju je zaradi epidemije ebola pred petimi leti umrlo 66 ljudi.' and 'Virus ebola povzroča močno krvavitev, več kot polovica ljudi, ki pride v stik z njim, pa umre že v nekaj dneh.' Below the definition, there is a section for 'Izgovor in oblike' with a table showing the word's spelling, stress patterns, and IPA transcription.

ZAPIS	JAKOSTNI NAGLAS	TONEMSKI NAGLAS
ebola	[ébola]	[ébola]
	IPA: [ˈeːbola]	IPA: [éːbóla]

Fig. 3: The Fran web portal: a growing dictionary entry

2. The Franček web portal

Our latest achievement is the school web portal Franček (Francek.si) (see Ahačič et al. 2021; Perdih et al. 2021; Petric 2020). The challenge we encountered in designing it was how to introduce primary- and secondary-school students to using dictionaries on a daily basis. Franček is a school portal that gradually develops dictionary skills in young people and complements lexicographic information with grammatical information via links to an online grammar. Franček consists of a list of headwords, which includes nearly 100,000 entries, and is linked to various linguistic databases Perdih (2021).

The entries of individual words in the portal (see Fig. 5) are adapted for three age groups and cover semantics (“What does this word mean?”), synonymy (“Find words with similar meaning”), phraseology (“Which multi-word unit does this word appear in?”), pronunciation (“How do I pronounce this word?”), inflections (“How is this word inflected?”), dialect use (“How is this word used in dialects?”; users can even record how they pronounce a specific word in their dialect), etymology (“What is the origin of this word?”), and history (“Since when has this word been used?”). The entries differ from the traditional lexicographic presentation in that they approximate how a word is explained by a teacher in class or by parents at home. Every entry on Franček is linked to a dictionary entry in its original format on Fran.

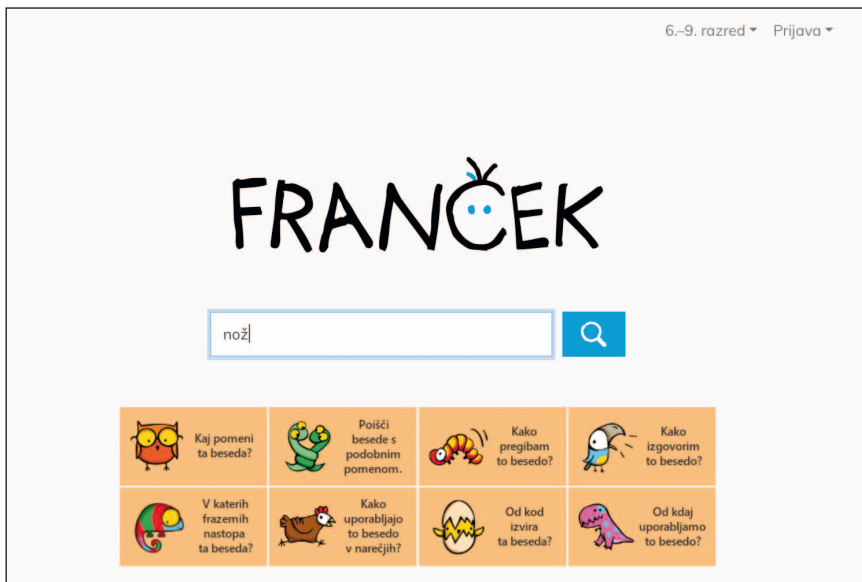
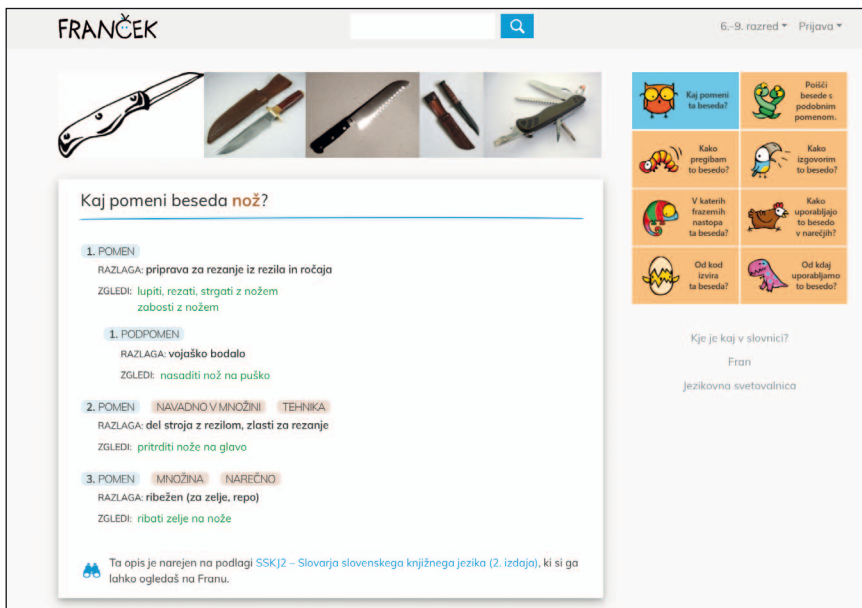


Fig. 4: The Franček web portal: entry page

Fig. 5: The Franček web portal: part of the presentation of the word *nož* ‘knife’

The Franček portal is also connected to a grammar page¹⁴ featuring grammar topics adapted to students of primary and secondary schools (see Fig 6). The presentation of grammar is the same for all three age groups, but the explanation of grammatical topics for primary and secondary schools differs in the level of detail.

Every piece of lexicographic information is also placed within a grammatical context, which in turn forms the starting point for linking the material to the corpus of the most common errors made by schoolchildren, a special Questions and Answers language page for teachers,¹⁵ and useful exercises.

FRANČEK Kje je kaj v slovnici?

JEZIKOVNE ZVRSTI
Družbene zvrsti
Besedilne vrste

GLAS
Glasovi
Naglas

BESEDA
Zlog
Tvorba besed (besedotvorje)
Odnosi med besedami (besedoslovje)
Odklikoslovje

STAVEK
Besedna zveza
Skladnja stavka
Skladnja povedi

PRAVOPIS
Velika in mala začetnica
Ločila
Deljenje besed
Premi govor

Beseda → **Odnosi med besedami (besedoslovje)**

● Besede, ki imajo isti koren, pripadajo besedni družini.	14	37
● *Besede, ki imajo podoben pomen, pripadajo pomenskemu polju.		38
● Besede so glede na pomen lahko sopomenke, protipomenke, nadpomenke ali podpomenke.	15	39
● *Besede so glede na zapis in izgovor lahko enakozraznice (homonimi).		40
● Beseda ima lahko en pomen (enopomenka) ali več pomenov (večpomenka).	16	41
● Pomen besed je lahko nevtralen, lahko pa je slogovno zaznamovan. Z rabo besed, ki so v knjižnem jeziku slogovno zaznamovane, lahko zelo spremenimo pomen posameznega stavka, povedi ali besedila.	17	42
● Besede so lahko v splošni rabi (*splošna leksika), lahko pa se uporabljajo na različnih specializiranih strokovnih področjih (términi).		

O tej temi poglej še

Pogoste napake

Vsebine v jezikovni svetovalnici

Vaje

Povezave in gradiva

Jezikovna svetovalnica

Črkozmed

Fran

Franček

Fig. 6: The Franček web portal: a grammar page

To sum up, our services currently include: dictionaries on the Fran portal, a user-oriented language advisory service, and an introduction to dictionary use on Franček.

¹⁴ <https://kje-je-kaj-v-slovnici.franček.si/domov>.

¹⁵ <https://svetovalnica.franček.si/domov>.

3. Related activities

We also seek to connect all of this with sociolinguistic research. In 2016 and 2017, we thus conducted the study Slovenia's Language Policy and User Needs, which was the first of its kind to cover, with the help of a nation-wide survey, all basic sociolinguistic problems related to Slovenian (Gliha Komac 2018). Having reviewed legislative and technical documents (Gliha Komac/Kovač 2018), a group of 45 experts and researchers, who possess insight into uses, practices and needs of language users in the Republic of Slovenia and Slovenian language users outside the Republic of Slovenia due to their work and research interests, prepared a sociolinguistic description of Slovenian language community. The central part of the research project was an online survey on language use, knowledge and needs (in media, education, administration and public services, economic and social life etc.) of different groups of language users in the Republic of Slovenia and users of the Slovenian language in neighbouring countries and elsewhere in the world (Ahačič et al. 2017) in which 5,782 language users, i. e. both specialised and general, concerned by Slovenian language policy participated. This was a comprehensive attempt to actively integrate language users in the making of future language policy to the greatest possible extent. Based on this study and a number of other contributions by various researchers, the Slovenian parliament adopted the Resolution on the National Language Policy Program. The program addresses two key spheres: language education and language infrastructure (sources and technologies).

We responded immediately by compiling a new Slovenian normative guide,¹⁶ which is based on material from the advisory pages and which cross-links rules to the dictionary section of the normative guide on the Fran web portal. Users can comment on any part of the normative guide. We will take these comments into account when preparing the final version.

4. Conclusion

Over the past eight years, the ZRC SAZU Fran Ramovš Institute of the Slovenian Language has managed to switch from a mere digital presentation of printed dictionaries to portals that are distinctly user-oriented and also provide instant feedback to lexicographers. This feedback in the form of user questions, comments, and suggestions as well as the possibility of monitoring the search statistics, has reshaped our work in practice. For example, identifying language users' problems facilitates normative work. In addition, user preferences and searches guide our selection of headwords for dictionaries for which the headwords are still incomplete (the ones that have been completed are published on an ongoing basis as part of growing dictionaries).

¹⁶ <https://www.fran.si/pravopis8>.

Users have begun to see that we provide a comprehensive service. Just like on Google, where they can find the desired website, they can find all of the information on a word or multi-word unit they are interested in in one place on the Fran and Franček portals. In this way, they grow accustomed to the fact that dictionaries can help us solve various linguistic problems and, at the same time, they increasingly perceive dictionaries, corpora, and linguistic advice as direct linguistic assistance that can help them meet their daily linguistic challenges.

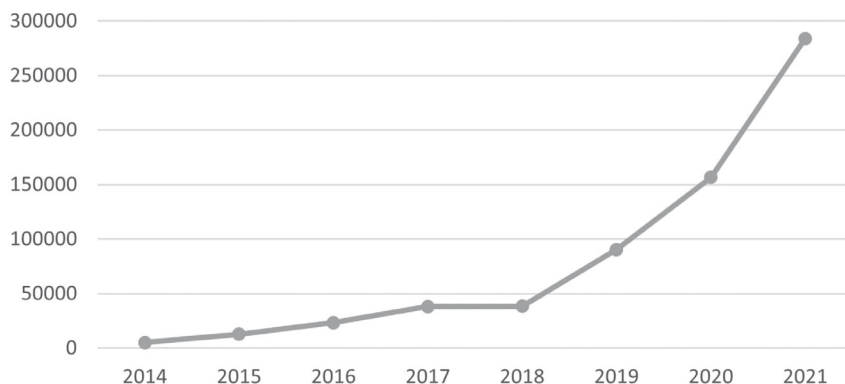


Fig. 7: The Fran web portal: average daily searches by year

References

- Ahačič, K./Ježovnik, J./Ledinek, N./Perdih, A./Petrič, Š./Race, D. (2021): Priprava jezikovnih podatkov za pedagoški portal o slovenščini Franček. In: *Philological Studies* 19/1, 203-224.
- Ahačič, K. et al. (2017): Jeziki v Sloveniji in slovenščina zunaj nje: anketni vprašalnik. Ljubljana: ZRC SAZU, Inštitut za slovenski jezik Frana Ramovša. https://isjfr.zrc-sazu.si/sites/default/files/anketni_vprasanik_1_kod_0.pdf
- Gliha Komac, N. (2018): Jezikovna politika Republike Slovenije in potrebe uporabnikov. In: *Slavia Centralis* 11/2, 7-15.
- Gliha Komac, N./Kovač, P. (eds.) (2018): Pravna ureditev in programski dokumenti o jezikovni rabi in praksah jezikovnih uporabnikov v Republiki Sloveniji in uporabnikov slovenskega jezika v sosednjih državah in po svetu. Ljubljana: ZRC SAZU, Inštitut za slovenski jezik, Založba ZRC.
- Perdih, A. (2018): Dictionary portal Fran: current state and future developments. In: *Slovenská lexikografie počátkem 21. století: sborník příspěvků z mezinárodní konference, Praha 20.-22.4.2016*, 57-65.

- Perdih, A. (2020): Portal Fran: od začetkov do danes. In: *Rasprave Instituta za hrvatski jezik i jezikoslovlje* 46/2, 997-1018.
- Perdih, A. (2021): Indikatorji pri homografih na portalu Franček. In: *Jezikoslovni zapiski* 27/2, 7-21.
- Perdih, A./Ahačič, K./Ježovnik, J./Race, D. (2021): Building an educational language portal using existing dictionary data. In: *Jazykovedný časopis* 72/2, 568-578.
- Petric, Š. (2020): Tipologija razlag v Šolskem slovarju slovenskega jezika. In: *Slavistična revija* 68/3, 391-409.

Bibliographical information

This text is part of the book:

Željko Jozić/Sabine Kirchmeier (eds.) (2022): The role of national language institutions in the digital age. Contributions to the EFNIL Conference 2021 in Cavtat. Budapest: Nyelvtudományi Kutatóközpont/Hungarian Research Centre for Linguistics.

This electronic PDF version of the text is accessible through the EFNIL website at:

<http://www.efnil.org>