CJVT Igre: New Word Games Based on the Digital Dictionary Database of Slovene

Špela Arhar Holdt[1,2], Iztok Kosem[1,2,3]

[1] Faculty of Arts, University of Ljubljana, Aškerčeva 2, 1000, Ljubljana, Slovenia
[2] Faculty of Computer and Information Science, University of Ljubljana, Večna pot 113, 1000, Ljubljana, Slovenia
[3] Institut "Jožef Stefan", Jamova cesta 39, 1000, Ljubljana, Slovenia
E-mail: Spela.ArharHoldt@ff.uni-lj.si, Iztok.Kosem@fri.uni-lj.si

Abstract

CJVT igre (https://igre.cjvt.si/) is a new digital platform offering word games designed to foster lexical awareness and engagement with standard Slovene. Developed by the Centre for Language Resources and Technologies at the University of Ljubljana, the portal currently hosts three games—Cvetka, Besedolov, and Vezalka—with two more in development. Each game utilizes curated lexical data from the Digital Dictionary Database of Slovene, enhanced through targeted lexicographic work to ensure playability, thematic coherence, and age-appropriateness. This includes refining word lists, rating difficulty, and enriching entries with semantic metadata. Cvetka focuses on orthographic guessing tasks with daily thematic prompts, Besedolov on semantic word search challenges within 11x11 grids, and Vezalka on word formation from a constrained letter set. Designed for both educational and general audiences, the games integrate varying levels of difficulty, optional hints, and dynamic scoring. This paper showcases the platform's interface, gameplay mechanics, and the linguistic and technical adaptations required to transform lexicographic resources into effective digital games.

Keywords: language games; Digital Dictionary Database; semantic type; lexicon; Slovene

1. Introduction

CJVT Igre ('CJVT Games' – https://igre.cjvt.si/) is a newly launched digital platform featuring word games designed to promote lexical awareness and playful engagement with the standard Slovene language. The portal, developed at the Centre for Language Resources and Technologies at the University of Ljubljana in cooperation with the Brainylab Creative Studio, is intended for a broad audience, ranging from language enthusiasts to teachers and students at various educational levels. At present, three games are available, with two additional games in development.

Each game draws on different aspects of lexical knowledge. *Cvetka* (approximate translation 'Flower') focuses on orthographic guessing tasks with daily thematic prompts, *Besedolov* ('Word Hunt') on semantic word search challenges within 11x11 grids, and *Vezalka* ('Lace') on word formation from a constrained letter set. The games offer increasing difficulty, optional hints, and rely on a carefully curated lexicon that

excludes offensive, vulgar, or sensitive content, ensuring both age-appropriateness and a positive user experience. The use of lexicographically curated data also supports methodological consistency and transparency. While all games are grounded in data from the Digital Dictionary Database of Slovene (Gantar et al., 2016; Kosem et al., 2021), each required further lexicographic work to ensure playability. This involved refining word lists, rating lexical items by gameplay difficulty, and manually reviewing and enriching word groups based on thematic indicators.

In this paper, we present the portal's interface and underlying architecture, highlight the gameplay mechanics, data integration, and the technical and linguistic strategies employed to adapt dictionary content for gaming purposes. We also outline our ongoing development plans, including new games, expanded lexical content, and features designed to enhance the platform's didactic value.

2. Related Work

Language games have been shown to support vocabulary acquisition, engagement, and cognitive development (Derakhshan & Khatir, 2015; Hidayat, 2016; Bustrillo et al., 2024; Khusaini & Fauziah, 2024). Specific games with features analogous to those presented in our paper have demonstrated notable effects. *Hangman*, which involves letter-guessing and orthographic recognition, led to significant vocabulary gains in secondary students (Munikasari et al., 2021) and an improvement of 48.95% in primary learners' vocabulary test (Dharmayasa, 2022). *Wordle,* with comparable mechanics, has been shown to enhance vocabulary mastery among Indonesian secondary students (Pamungkas, 2021) and was found to reinforce orthographic and semantic patterns (Kwan, 2023). *Semantic Boggle,* a time-constrained letter-grid game that identifies semantically related words, has shown promising results for teaching vocabulary (Toma et al., 2017) and building vocabulary transferability (Halimah & Izzah, 2020). Crossword puzzles also yield significant gains in vocabulary and positive attitudes among learners (Orawiwatnakul, 2017). The findings apply to both first and foreign language learning, and although observed in educational contexts, they also suggest positive effects from leisure-based gameplay.

Central to the value of the games presented in this paper is their grounding in structured linguistic content, derived from lexicographic sources. A similar work has been reported by Mihaljević and Hudeček (2022) who, under the umbrella of the *Mrežnik* project, designed a model for transferring data from dictionary structure into games: quizzes, fill-in-the-blank games, drag-and-drop games, typing games, and crossword puzzles. Mihaljević (2019) also demonstrated the value of gamification in enhancing user engagement and supporting vocabulary learning within e-lexicographic environments. His analysis of 181 online dictionaries and 71 encyclopaedias showed that only a small proportion of these resources incorporated any form of gamified content, with quizzes and scoring systems being the most common elements. The potential of gamification is effectively illustrated by Schoonheim et al. (2012), who report on the

launch of the language game *"Het Verloren Woord,"* which led to a significant increase in user activity and encouraged deeper interaction with the ANW dictionary.

Apart from games developed for promotional or infotainment purposes, lexically oriented Games with a Purpose (GWAPs) are designed to collect linguistic data through gameplay, enabling users to contribute to lexicographic and NLP resources by completing structured language tasks. Some of the best-known GWAPs for crowdsourcing language data include *Phrase Detectives* (Poesio et al., 2013), *JeuxDeMots* (Lafourcade, 2007), *ZombiLingo* (Guillaume, 2016), and *LikeIt* (Lafourcade et al., 2015). More recent examples include *Fictionary*-style games, where users generate plausible definitions for obscure or rare words (Friðriksdóttir & Einarsson, 2022), and *Word Ladders*, where players construct chains of semantically related words, enabling the collection of associative and hierarchical lexical relations through intuitive interaction (Genovese et al., 2024). For Slovene, *Game of Words*—a gamified mobile application designed to improve and enhance two automatically compiled Slovene dictionaries—was developed and presented by Arhar Holdt et al. (2021). Slovene is also included in the *CrowLL* game developed to flag pedagogically inappropriate corpus examples (Kuhn et al., 2022). Unlike these initiatives, the new CJVT games are primarily intended for leisure-time entertainment. However, we aim to further enhance their educational value and leverage them to guide users toward contemporary digital dictionary resources (see also Chapter 4.4).

3. The Digital Dictionary Database and Data Preparation

The main resource for the data used in the games is the Digital Dictionary Database of Slovene (DDDS; Gantar, 2020; Kosem et al., 2021), which aims to become a one-for-all database for the Slovene language, to be used for both the compilation of language resources and natural language processing tasks. The plans for the database have been described in detail in Klemenc et al. (2017). The database already contains the data for several lexical resources, e.g., Thesaurus of Modern Slovene (Arhar Holdt et al., 2023), Collocations Dictionary of Modern Slovene (Kosem et al., 2023), Slovene Morphological Lexicon (Dobrovoljc et al., 2018), and Comprehensive Slovenian-Hungarian Dictionary (Kosem et al., 2024). The import of larger amounts of other monolingual and bilingual data is currently underway.

The DDDS contains various morphological, semantic, and statistical information about lexical units, which range from single words to compounds and phraseological units (in dictionaries seen as headwords). The information includes semantic indicators, labels, examples, translations, etc., which are found in the aforementioned lexical resources. Additional information found in the DDDS are semantic types, attributed to meanings of lexical units – semantic types are divided into an ontology of 21 top-level categories (e.g. *Act, Artifact, Food, Human, Location*), which are further divided into up to three levels of hierarchical subcategories (Kosem & Pori, 2021).

All this information is essential when preparing data for language games, as it helps to classify, group, and describe words, and exclude unwanted content. Unwanted content comprises sensitive vocabulary, which has no place in pedagogically oriented games (see e.g. Kuhn et al., 2022; Arhar Holdt et al., 2021), but also very rare words that may cause more frustration than motivation in the users. While DDDS already includes dictionary labels for vocabulary expressing discrimination, negative attitudes, and vulgarity (Arhar Holdt et al., 2023), some words (e.g. *rape*) have meanings with negative connotation which are not explicitly labelled as such. On the other hand, there may be polysemous words with one negative meaning (e.g. *donkey* as an insult for a stupid or rude person), which do not warrant exclusion. Consequently, data preparation for the games always requires additional manual inspection.

3.1 Lexical Data Preparation for Cvetka and Besedolov

We exported a list of noun lemmas, including semantic types and indicators related to their meaning(s), and relative frequencies calculated from the Gigafida 2.0 reference corpus of standard Slovene (where available). The data was initially grouped by semantic type. We then manually revised the semantic indicators and the content of each group, adapting the descriptions to be more intuitive and suitable for gameplay. During this process, we merged some smaller groups, split the largest ones, and excluded groups containing overly technical or sensitive content, such as 'Field of psychology' or 'Mode of killing'. An example of the revised lexical data is in Table 1.

Semantic type	**Semantic category**	**Word**	**Relative frequency**
SNOV-naravna ('SUBSTANCE-natural')	material rastlinskega izvora ('material of plant origin')	pepel ('ashes')	7.25
SNOV-naravna ('SUBSTANCE-natural')	material rastlinskega izvora ('material of plant origin')	jantar ('amber')	0.82
SNOV-naravna ('SUBSTANCE-natural')	material živalskega izvora ('material of animal origin')	usnje ('leather')	10.80
SNOV-naravna ('SUBSTANCE-natural')	material živalskega izvora ('material of animal origin')	volna ('wool')	5.63
SNOV-naravna ('SUBSTANCE-natural')	naravna tekočina ('natural liquid')	sirotka ('whey')	0.78
SNOV-naravna ('SUBSTANCE-natural')	naravna tekočina ('natural liquid')	smola ('resin')	12.50

Table 1: A section of lexical data with semantic types, indicators, and relative frequency. Examples are from the same semantic type and three different semantic categories.

In the second step, we identified several additional parameters needed to support gameplay design and difficulty calibration of *Cvetka* and *Besedolov*. For *Cvetka*, we focused on three key features of each word: the number of unique vowels, the number of unique letters, and its relative frequency in the Gigafida 2.0 corpus. Each of these features was assigned a rank on a five-point scale, and the sum of these ranks was used to determine the word's overall complexity for the game. The game progresses from higher rank sums to lower ones, allowing for a gradual increase in difficulty. For *Besedolov* we leveraged frequency data to select only those semantic categories that contained at least five high-frequency nouns. This threshold helped eliminate groups with overly rare or obscure words, ensuring playability across game rounds. Table 2 presents lexical data with the new metrics added.

Word	**Unique vowels**	**Unique letters**	**Relative frequency**	**Rank UV**	**Rank UL**	**Rank Rf**	**Rank sum**
velikan (‘giant’)	3	7	28.27	4	4	5	13
čarovnik (‘mage’)	3	7	5.10	4	4	3	11
vila (‘fairy’)	2	4	19.68	3	2	5	10
kentaver (‘centaur’)	2	7	0.20	3	4	1	8
samorog (‘unicorn’)	2	6	0.36	3	3	1	7
škrat (‘dwarf’)	1	5	3.33	2	2	3	7
jagababa (‘Baba Yaga’)	1	4	0.06	2	2	1	5

Table 2: A section of lexical data with information needed for *Cvetka* and *Besedolov*. All words are from the semantic category ‘Mythological or fairy-tale creature’.

3.2 Lexical Data Preparation for Vezalka

The initial lexical dataset for the Vezalka game was exported from the DDDS, focusing on vocabulary also included in the Sloleks 3.0 morphological lexicon, and only on the selected parts of speech (nouns, adjectives, verbs, and adverbs). From this data, proper nouns and adjectives derived from proper names were filtered out to ensure general language applicability, resulting in 131.274 lemmas. The data was then further edited for the game's purpose: we removed entries with digits, non-Slovene characters, and Roman numerals, and translated frequencies into five frequency ranges (5 for most frequent words and 1 for least frequent words in the Gigafida 2.0 corpus).

Based on this data, we generate candidate word sets for the game by selecting a central letter along with six accompanying letters, ensuring that each set allows for at least one pangram—a word that contains all seven letters. Homonyms are treated as a single entry for gameplay purposes (e.g., *lev* as both a noun meaning 'lion' and an adjective meaning 'left'). The resulting word sets are manually reviewed and adjusted to avoid relying too heavily on rare or archaic vocabulary. An example of lexical data is presented in Table 3.

Word	**POS**	**Relative frequency**	**Rank Rf**
igličast ('needle-like')	adjective	0.13	2
čast ('honour')	noun	32.54	5
častiti ('to clean')	verb	10.27	5
tačas ('meanwhile')	adverb	2.12	3

Table 3: Example of lexical data for *Vezalka.* Four words of different POS are presented, all from the same word set (27 words from the letters I, S, T, L, A, G, and the central Č. The word *igličast* is the pangram of the set.

4. The CJVT Igre Portal

4.1 Entry Screens and User Statistics

The entry screen at igre.cjvt.si offers a clean, intuitive interface that allows users to access the daily games as well as the game from the previous day (Figure 1). Players can explore the three available games via clearly marked navigation tiles. Users also have the option to register, in which case their gameplay statistics are stored. The statistics screen provides an overview of performance across all the games, showing both daily and cumulative data such as points earned, relative ranking, total score, number of games played, and the longest streak of consecutive playing days.

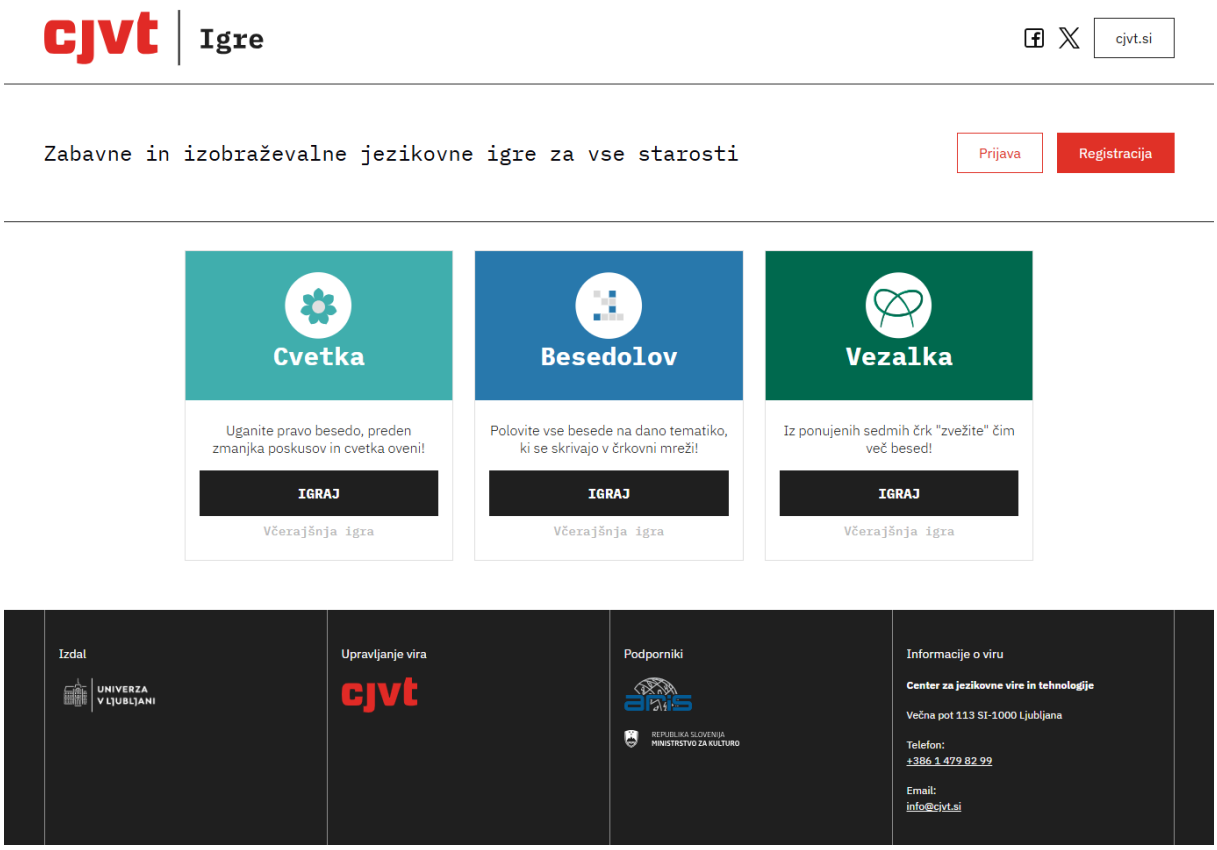


Figure 1: Entry screen of the CJVT language games portal (igre.cjvt.si), offering access to the daily and previous day's versions of the games Cvetka, Besedolov, and Vezalka, with optional user registration.

4.2 Cvetka

Based on the concept of the well-known *Hangman* game, *Cvetka* challenges players to guess a hidden Slovene word or phrase within a limited number of attempts, combining vocabulary knowledge with orthographic intuition. Every day, ten new hidden words

are presented, each accompanied by a thematic category that serves as a hint—for example, *vrsta mesa* ('type of meat') for *perutnina* ('poultry'), or *objekt za določen namen* ('object for a specific purpose') for *počivališče* ('rest stop').

Figure 2 shows a screenshot from the Cvetka game. On the left side of the screen, the category is displayed (in this case, družboslovna veda 'social science field'). The letter grid indicates which letters have already been guessed: blue letters (A, D, E, G, I, K, and P) signify correct guesses, while red letters (F, J, and N) denote incorrect ones. Incorrect guesses cause petals to fall from the animated flower on the right side of the screen; once all seven petals are gone, the guessing concludes. Hints can be accessed via the icon below the letter grid. The interface also provides access to the game rules and displays the player's daily score. After the word is guessed—or once the flower has lost all its petals—the game automatically proceeds to the next daily task. At the end, the total score is presented and added to the daily statistics.

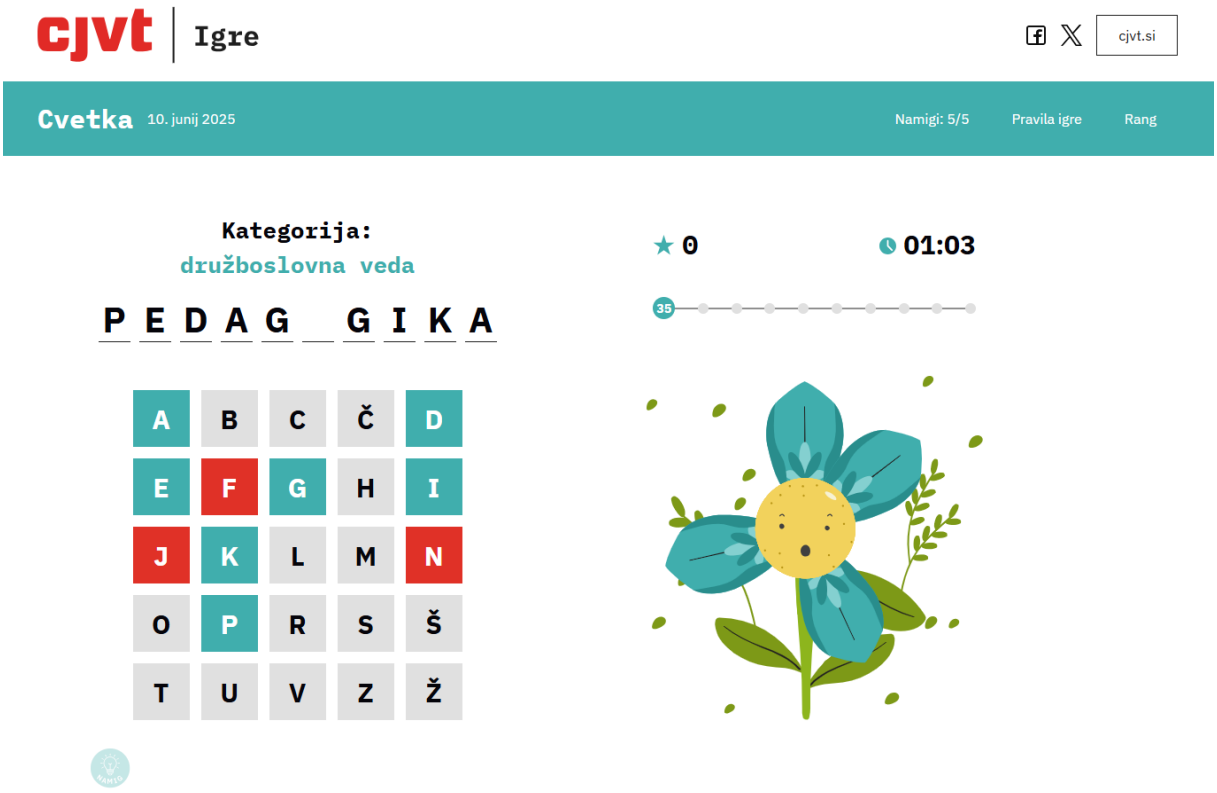


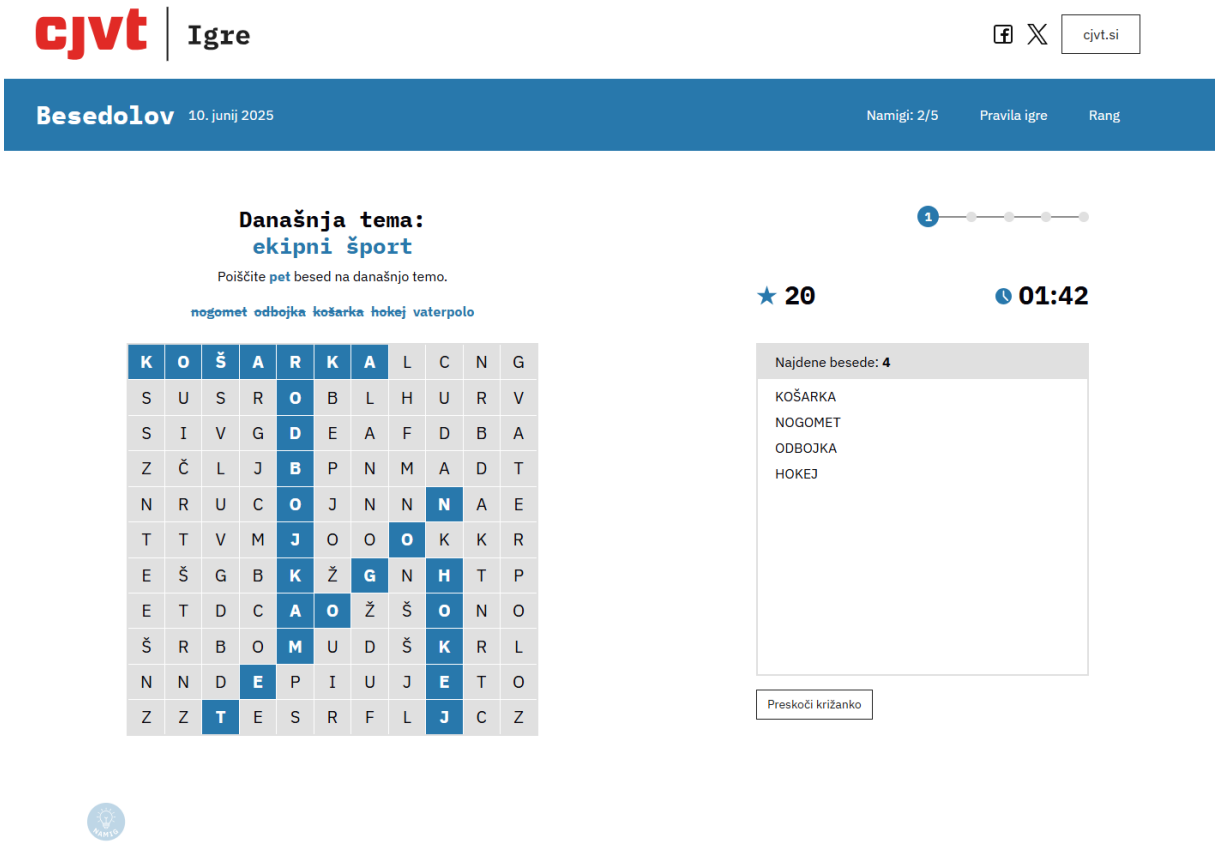
Figure 2: A screenshot from the Cvetka game. The correct answer is *pedagogika* ('pedagogics') from the category *družboslovna veda* ('social science field').

Scoring in the game is as follows: each correct letter earns five points. Guessing the word grants up to fifty additional points, depending on solving speed—three points are deducted for every thirty seconds that pass. If the flower withers, no points are awarded for the word. In Cvetka, using a hint reveals one letter (the first unrevealed letter in the word) and costs 15 points. Only one hint can be used per word.

4.3 Besedolov

Besedolov supports thematic vocabulary development by asking players to locate semantically related Slovene words hidden in a grid of letters. Every day, the game offers five grids, each tied to a different category. For instance, the category *literarno delo* ('literary work') may include hidden words such as *balada* ('ballad'), *pravljica* ('fairy tale'), *ep* ('epic'), *novela* ('novella'), and *satira* ('satire'). Words can appear in any direction, and points are awarded based on word discovery and speed. Advanced players may choose to find all target words without assistance; however, using a hint will reveal all the words for a given grid.

Figure 3 shows a screenshot from the Besedolov game. In this session, the player has already used the hint button (located below the letter grid), revealing five words from the category *ekipni šport* ('team sports'): *nogomet* ('football'), *odbojka* ('volleyball'),



košarka ('basketball'), *hokej* ('hockey'), and *vaterpolo* ('water polo'). The four words already found in the grid are displayed on the right side of the screen, along with the points earned and time spent. The player can continue searching for the last word, and there is also the option to skip the grid (*Preskoči križanko)*. Once all grids have been solved or skipped, the game ends by showing the total daily score.

Figure 3: A screenshot from the Besedolov game. The category is *ekipni šport* ('team sports'). A hint was used to display all five words that need to be found in the grid.

Scoring is as follows: each discovered word earns five points. Solving the entire grid grants up to fifty additional points, depending on solving speed—three points are deducted for every thirty seconds that pass. In Besedolov, a hint reveals the target words for the current grid and can be used only once per grid. Each hint deducts 15 points from the total daily score.

4.4 Vezalka

Vezalka is the most linguistically demanding of the three games. It invites players to form as many valid base-form Slovene words as possible from a given set of seven letters, including a required central letter. For example, with the letters P, K, R, L, E, I and obligatory C, players might produce *kipec* ('statuette'), *rilec* ('trunk'), and *lice* ('cheek'). Points are awarded based on word length, with a bonus for discovering a word that uses all seven letters. One letter set is available per day.

Figure 4 shows the game interface. The set of seven letters is displayed on the right side of the screen. The player can type in a word or select letters by clicking on them. Below the letter circle are buttons for deleting the current input, confirming the word, or reshuffling the letters. On the left, the already guessed words are listed, while above them an indicator shows the player's overall progress. Since many valid words are rare or difficult to guess, players can choose to give up (the *Predam se* button), which ends the game. At that point, the total score is shown, along with a list of missed words. In the example shown, some of the missed words include *klic* ('call'), *cikel* ('cycle'), *preklic* ('revocation'), *cepec* ('flail'), *precep* ('cleft'), *priklic* ('evocation'), *pipec* ('tap, spout'), *celec* ('small target'), *krepelce* ('small clod'), *cepilec* ('splitter'), *lepilec* ('gluer'), and *epicikel* ('epicycle').

The rules and scoring system are as follows: each word must contain at least four letters and include the central (highlighted) letter. Letters can be used more than once within a word. All words must be in their base form—for example, delati ('to work') rather than delam ('I work') or miza ('table') rather than mize ('tables'). The game's word list excludes offensive or hateful language, hyphenated words, and proper nouns. If a word is accepted, a four-letter word earns four points, with one additional point awarded for each extra letter. Each game includes at least one word that uses all seven letters; this word grants a bonus of three points. Players are notified that some valid words may not yet appear in the database, although it is regularly updated and expanded.

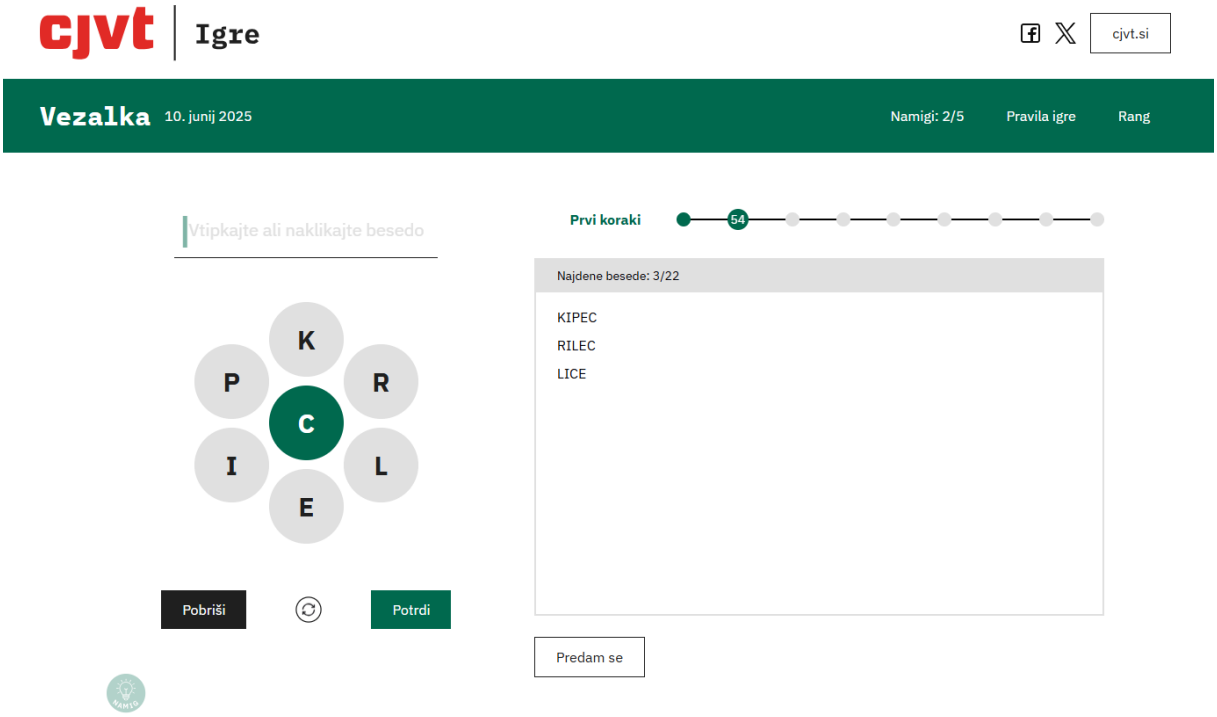


Figure 4: A screenshot from the Vezalka game. Three words have been found so far: *kipec* ('statuette'), *rilec* ('trunk'), and *lice* ('cheek').

4.5 Didactical Value of the Games

As indicated by studies (Chapter 2), *Cvetka* enhances orthographic awareness, spelling accuracy, and recognition of base word forms, while also fostering inferencing strategies through thematic categories. *Besedolov* encourages semantic reasoning and vocabulary expansion by asking players to identify related lexical items within category-based grids, and *Vezalka* promotes morphological sensitivity and lexical recall. Across all games, learning is supported by immediate feedback—whether through revealed solutions or the gradual discovery of missed words—which fosters incidental vocabulary acquisition. The use of thematic prompts, scoring systems, and limited hints introduces a mild form of gamified pressure that boosts focus and engagement, while the emphasis on standard Slovene and exclusion of sensitive language ensures linguistic and pedagogic appropriateness. Together, these features support vocabulary growth, semantic awareness, orthographic skills, and motivation in both educational and casual settings.

An additional pedagogical consideration is the careful calibration of user time and screen exposure. The portal is designed to provide daily language tasks that can be completed within approximately ten minutes, encouraging consistent, focused engagement without contributing to digital fatigue. At the same time, users are free to return to unfinished games later in the day, which is particularly relevant for *Vezalka*, where lexical recall may occur gradually. This balance between time efficiency and

flexible participation supports cognitive processing beyond the immediate session and aligns with current concerns about attention economy.

To build on this foundation, several targeted enhancements could increase the platform's educational impact. Words found or missed in all games could be linked to dictionary entries, allowing motivated users to explore meanings, usage, and word relations in more depth. In *Cvetka* and *Besedolov*, players could be offered the option to revisit semantic categories and be prompted to try similar categories again. *Vezalka* presents a particularly valuable opportunity: currently, legitimate words entered by players but missing from the database are not captured—logging these could inform both lexicographic enrichment and game development. Another possibility is the inclusion of light metalanguage, such as part-of-speech labels or morphological cues, which could help players reflect on word structure and function. (However, overloading the interface with linguistic annotation or excessive instructional elements would risk diminishing user enjoyment.) Finally, a more ambitious but extremely valuable solution could be to develop a specialized teacher dashboard that enables educators to select or import vocabulary by theme or difficulty, monitor student progress, or assign targeted challenges. In such an environment, automated progress tracking and player motivation could further support target vocabulary learning among students.

5. Conclusion and Future Work

In this paper, we presented a new digital platform for Slovene word games, developed at the Centre for Language Resources and Technologies. We described three games currently available on the portal—Cvetka, Besedolov, and Vezalka—each based on adapted lexical data from the Digital Dictionary Database for Slovene. We detailed the lexical selection and filtering procedures, including the use of semantic types, frequency thresholds, and measures of word complexity such as letter variety. We described the structure and functionality of the entry and statistics screens, and outlined the gameplay mechanics and scoring systems for each individual game. Finally, we discussed the pedagogical value of the games and proposed several enhancements aimed at increasing their educational impact.

Future development of the platform will focus on enhancing its didactic functionality and expanding its game offering. Several improvements have already been outlined, including richer feedback, integration with dictionary content, and the potential introduction of a teacher dashboard for classroom use. In addition, we aim to increase the variety of games available on the portal. A fourth game, which focuses on identifying collocations and synonyms, is currently under development, and a fifth game is planned, which focuses on clue-based word discovery, placing greater emphasis on semantic inference. Planned user studies will complement these development efforts by gathering systematic feedback on the current games. These studies will also help us refine the profiling of target user groups for each game. In line with our expectations, preliminary user feedback suggests that *Besedolov* and *Cvetka* appeal to younger users, particularly

at the end of primary school, while *Vezalka* tends to engage adult players. Future research will aim to define user segments more precisely, considering factors such as age, language proficiency, and whether Slovene is a first or second/foreign language, to establish appropriate knowledge thresholds and further tailor the user experience.

6. Acknowledgements

The research program Language Resources and Technologies for Slovene (P6-0411), the infrastructure Network of Research and Infrastructural Centers UL (I0-0022) and the projects Large Language Models for Digital Humanities (GC-0002) are funded by the Slovenian Research and Innovation Agency. Ministry of Culture of the Republic of Slovenia is funding the JR-infrastruktura-SJ-2024-2025 project Data completion and gamification of dictionary resources at CJVT UL (PODVIG). We thank the reviewers for their constructive and valuable suggestions.

7. References

Arhar Holdt, Š., Logar, N., Pori, E. & Kosem, I. (2021). Game of Words: Play the Game, Clean the Database. In *Proceedings of the 14th Congress of the European Association for Lexicography (EURALEX 2021)*, pp. 41–49, Alexandroupolis, Greece.

Arhar Holdt, Š., Gantar, P., Kosem, I., Pori, E., Robnik Šikonja, M. & Krek, S. (2023). Thesaurus of Modern Slovene 2.0. In *Proceedings of eLex 2023: Electronic Lexicography in the 21st Century*, Brno, pp. 366–381. https://elex.link/elex2023/wp-content/uploads/82.pdf

Bustrillo, H.T.B., Sumicad, R. & Cuevas, G.C. (2024). Effectiveness of Word Games in Teaching Students Vocabulary. *Journal of World Englishes and Educational Practices*, 6(3), pp. 07–16. https://doi.org/10.32996/jweep.2024.6.3.2

Carr, M. (1997). Internet Dictionaries and Lexicography. *International Journal of Lexicography*, 10(3), pp. 209–230.

Derakhshan, A. & Khatir, E.D. (2015). The effects of using games on English vocabulary learning. *Journal of Applied Linguistics and Language Research*, 2(3), pp. 39–47.

Dharmayasa, I.N.W. (2022). An Implementing the Hangman Game in Teaching English Vocabulary to Elementary School Students. *Jurnal Pendidikan Bahasa Inggris Undiksha*, 10(3), pp. 291–297. https://doi.org/10.23887/jpbi.v10i3.58140

Dobrovoljc, K., Krek, S. & Erjavec, T. (2018). The Sloleks Morphological Lexicon and its Future Development. In *Dictionary of Modern Slovene: Problems and Solutions*, pp. 42–63. Ljubljana: Znanstvena založba Filozofske fakultete.

Friðriksdóttir, S.R. & Einarsson, H. (2022). Fictionary-Based Games for Language Resource Creation. In *Proceedings of the 2nd Workshop on Novel Incentives in Data Collection from People: models, implementations, challenges and results within LREC 2022*, pp. 25–31.

Gantar, P. (2020). Dictionary of Modern Slovene: From Slovene Lexical Database to Digital Dictionary Database. Rasprave Instituta Za Hrvatski Jezik i Jezikoslovlje, 46(2), 589–602. https://doi.org/10.31724/rihjj.46.2.7.

Gantar, P., Kosem, I. & Krek, S. (2016). Discovering automated lexicography: the case of Slovene lexical database. *International Journal of Lexicography*, 29(2), pp. 200–225. https://doi.org/10.1093/ijl/ecw014

Genovese, F., Bolognesi, M.M., Di Iorio, A. & Vitali, F. (2024). The advantages of gamification for collecting linguistic data: A case study using Word Ladders. *Online Journal of Communication and Media Technologies*, 14(2), e202426. https://doi.org/10.30935/ojcmt/14443

Guillaume, B., Fort, K. & Lefebvre, N. (2016). Crowdsourcing Complex Language Resources: Playing to Annotate Dependency Syntax. In *Proceedings of the International Conference on Computational Linguistics (COLING)*. https://hal.inria.fr/hal-01378980/

Halimah, F. & Izzah, L. (2020). Building FL-Vocabulary Transferability through Semantic Boggle. *English Language in Focus (ELIF)*, 2(2), pp. 79–86.

Hidayat, N. (2016). Improving Students' Vocabulary Achievement through Word Game. *JEES (Journal of English Educators Society)*, 1(2), pp. 95–104. https://doi.org/10.21070/jees.v1i2.446

Khusaini, F. & Fauziah, N. (2024). The Implementation of Word Games to Improve Students' English Vocabulary Proficiency. *Journal of Classroom Action Research*, 6(3), pp. 592–595. https://doi.org/10.29303/jcar.v6i3.8746

Klemenc, B., Robnik-Šikonja, M., Fürst, L., Bohak, C. & Krek, S. (2017). Technological Design of a State-of-the-art Digital Dictionary. In V. Gorjanc, P. Gantar, I. Kosem & S. Krek (eds). *Dictionary of modern Slovene: problems and solutions.* Ljubljana: Ljubljana University Press, Faculty of Arts, pp. 10-22.

Kosem, I. & Pori, E. (2021). Slovenske ontologije semantičnih tipov: samostalniki. In I. Kosem (ed.) *Kolokacije v slovenščini*. Ljubljana: Znanstvena založba Filozofske fakultete, pp. 159–202.

Kosem, I., Krek, S. & Gantar, P. (2021). Semantic data should no longer exist in isolation: the Digital Dictionary Database of Slovenian. In Z. Gavriilidou, L. Mitits & S. Kiosses (eds.) *EURALEX XIX, Congress of the European Association for Lexicography, Lexicography for inclusion: 7–9 September 2021, virtual: abstracts book*. Komotini: Democritus University of Thrace, pp. 81–83. https://euralex2020.gr/wp-content/uploads/2021/09/EURALEX2020_BookOfAbstracts-Preview-1.pdf

Kosem, I., Arhar Holdt, Š., Gantar, P. & Krek, S. (2023). Collocations Dictionary of Modern Slovene 2.0. In *Proceedings of eLex 2023: Electronic Lexicography in the 21st Century*, Brno, pp. 491–507. https://elex.link/elex2023/wp-content/uploads/100.pdf

Kosem, I., Bálint Čeh, J., Ponikvar, P., Zaranšek, P., Kamenšek, U., Koša, P., Gróf, A., Gantar, P., Krek, S., Arhar Holdt, Š. & Gorjanc, V. (eds.) (2024). *Veliki slovensko-madžarski slovar = Szlovén-magyar nagyszótár*. Ljubljana: Založba

Univerze. https://doi.org/10.4312/9789612973049

Zingano Kuhn, T., Arhar Holdt, Š., Kosem, I., Tiberius, C., Koppel, K. & Zviel-Girshin, R. (2022). Data preparation in crowdsourcing for pedagogical purposes: the case of the CrowLL game. *Slovenščina 2.0: Empirične, aplikativne in interdisciplinarne raziskave*, 10(2), pp. 62–100. https://doi.org/10.4312/slo2.0.2022.2.62-100

Kwan, Y.H. (2023). Guess the Right Word: A Review of the Language Game Wordle. *RELC Journal*, 55(3), pp. 855–859. https://doi.org/10.1177/00336882231192121

Lafourcade, M. (2007). Making people play for Lexical Acquisition with the JeuxDeMots prototype. In *Proceedings of the 7th International Symposium on Natural Language Processing (SNLP'07)*, Pattaya, Thailand.

Lafourcade, M., Joubert, A. & Le Brun, N. (2015). Collecting and evaluating lexical polarity with a game with a purpose. In *Proceedings of the International Conference Recent Advances in Natural Language Processing*, pp. 329–337.

Mihaljević, J. (2019). Gamification in E-Lexicography. *INFuture*, pp. 155–164.

Mihaljević, J. & Hudeček, L. (2022). Model for developing educational games based on data from dictionary structure. *Studia Lexicographica*, 16(30), pp. 111–133. https://doi.org/10.33604/sl.16.30.6

Munikasari, M., Sudarsono, S. & Riyanti, D. (2021). The Effectiveness of Using Hangman Game to Strengthen Young Learners' Vocabulary. *Journal of English Education Program*, 2(1), pp. 57–65. https://jurnal.untan.ac.id/index.php/JEEP/article/view/57-65/0

Orawiwatnakul, W. (2017). El Uso de Crucigramas como Herramienta en el Desarrollo del Vocabulario. *Electronic Journal of Research in Educational Psychology*, 11(30), pp. 413–428. https://doi.org/10.14204/ejrep.30.12186

Pamungkas, N.A.R. (2021). The Effects of Wordle Media on Students' Vocabulary Mastery. *JETAL: Journal of English Teaching & Applied Linguistic*, 2(2), pp. 56–61. https://doi.org/10.36655/jetal.v2i2.531

Poesio, M., Chamberlain, J., Kruschwitz, U., Robaldo, L. & Ducceschi, L. (2013). Phrase Detectives: Utilizing Collective Intelligence for Internet-scale Language Resource Creation. *ACM Transactions on Interactive Intelligent Systems*, 3(3), pp. 1–44.

Schoonheim, T., Tiberius, C., Niestadt, J. & Tempelaars, R. (2012). Dictionary use and language games: Getting to know the dictionary as part of the game. In *Proceedings of the 15th EURALEX International Congress*. Oslo: Department of Linguistics and Scandinavian Studies, University of Oslo, pp. 974–979.

Toma, I., Alexandru, C.E., Dascalu, M., Dessus, P. & Trausan-Matu, S. (2017). Semantic Boggle: A Game for Vocabulary Acquisition. In É. Lavoué, H. Drachsler, K. Verbert, J. Broisin & M. Pérez-Sanagustín (eds.) *Data Driven Approaches in Digital Education. EC-TEL 2017*. Lecture Notes in Computer Science, vol. 10474. Cham: Springer, pp. 603–607. https://doi.org/10.1007/978-3-319-66610-5_73