

Bulgarian-Polish Online Dictionary – Design and Development¹

Ludmila Dimitrova¹, Violetta Koseska², Ralitsa Dutsova³, Rumiana Panova³

¹ IMI-BAS, Bulgaria, ² ISS-PAS, Poland, ³ Veliko Tarnovo University, Bulgaria

Abstract. In the present paper we describe problems of design and development of an experimental Bulgarian–Polish online dictionary. We focus our attention on the ongoing version of the web-based application representing the dictionary. The basis for the first Bulgarian-Polish experimental online dictionary is the ongoing version of the Bulgarian-Polish electronic dictionary, currently developed in MS WORD-format in the framework of the joint research project “Semantics and Contrastive linguistics with a focus on a bilingual electronic dictionary” between IMI-BAS and ISS-PAS. The current version consists of approximately twenty thousand dictionary entries. Some examples of dictionary entries are presented.

Key words: bilingual digital dictionaries, entry classifiers, lexical database, web-based application, online dictionary

1. Introduction

Computer lexicography encompasses computer methods and resources for the automation of lexicographic activity. Such activities include: setting up of basic principles for development, creation and maintenance of dictionaries, recording of linguistic information in databases, creation of electronic indices, etc.

A dictionary (commonly) is a list of words and their meanings arranged in alphabetical order. In it, information is given about the pronunciation, grammar, derivative words, history or etymology of the main word, as well as recommendations for its usage, examples, phraseological expressions, illustrations. Dictionaries are most commonly available as books, but lately ones in electronic form are also gaining recognition.

Dictionary classification is based on multiple criteria. Various classifications exist in different lexicographic and lexicological works.

In the scope of this paper, two classifications are of particular interest:

According to the type of carrier:

- traditional dictionaries - these are developed with a human/computer, but in their final form, they reach the user in paper form;
- electronic dictionaries - these exist in digital format and can be divided into two categories: online (web-based) and local (desktop) dictionaries.

According to the number of languages:

- monolingual dictionaries - dictionaries in which words are defined in the same language;
- bilingual dictionaries - dictionaries which contain a translation of the words in exactly two languages;

¹ The study and preparation of these results have been supported by the EC's Seventh Framework Programme [FP7/2007-2013] under the grant agreement 211938 MONDILEX.

- multilingual dictionaries - dictionaries which contain a translation of the words in more than two languages.

Why choose a Bulgarian-Polish Dictionary? Our reasons are as follows:

- In the past 20-25 years neither Bulgarian-Polish nor Polish-Bulgarian dictionaries have been published.
Existing dictionaries from Bulgarian to Polish or vice versa are a collector's rarity and are outdated. The first Polish-Bulgarian dictionary, prepared by prof. I. Lekov, had been exhausted long before the second one appeared in 1961 – by authors I. Lekov and F. Sławski. The next Polish-Bulgarian dictionary was only published in 1988; its author is S. Radewa. The first Bulgarian-Polish dictionary (PODRĘCZNY SŁOWNYK BULGARSKO-POLSKI), prepared by F. Sławski, was published in 1963 in Warsaw. The second (and last) edition of this dictionary dates from 1987.
- These dictionaries contain about 50-60 thousand words, but they have a significant number of disadvantages (from a contemporary point of view). They contain words that are no longer in use - mostly dialect words and words of Turkish language origin. Moreover, these dictionaries contain a lot of words that were created by the authors themselves (in Bulgarian, it is possible to create new words through the addition of suffixes to certain words, mostly verbs), which are practically unusable.
- There is no existing Bulgarian-Polish online dictionary. So far the communication between languages of the same language group (Slavic) is channelled via other languages such as English or German.

Why an online dictionary? Advantages of online dictionaries over the local (desktop) electronic ones:

- Wide accessibility
- A possibility for a continuous update and editing
- Opportunities for search and hypertext cross-references
- Real-time use by a large number of users
- Real-time editing and update of the dictionary
- No restrictions on volume
- Economizing computer resources – local installation of the dictionary is not required.

2. Problems of the computer realization

The basis for the first Bulgarian-Polish experimental online dictionary is the ongoing version of the Bulgarian-Polish local electronic dictionary [1], [2]. The Bulgarian-Polish electronic dictionary is currently developed in MS WORD-format in the framework of the joint research project “Semantics and Contrastive linguistics with a focus on a bilingual electronic dictionary” between IMI-BAS and ISS-PAS under the supervision of L. Dimitrova and V. Koseska. The current version consists of approximately twenty thousand dictionary entries.

The main problems related to the computer realization of dictionaries arise from the fact that they are **simultaneously treated as text and databases**. They obviously look like text and have common points with other types of text. However, users do not normally read dictionaries, from A to Z, as they do with the majority of texts, but rather use them to obtain specific information through a given key (in this case a headword). The information associated with this key can include: pronunciation, grammar information, definitions, etymology, etc. Electronic dictionaries are capable of fulfilling users' requests multiple times

faster than paper dictionaries, as well as providing the possibility to return all entries whose title words contain the user-defined criteria. Despite the fact that dictionary entries resemble a text on the screen, the internal representation of electronic dictionaries is actually done as a database.

Dictionaries are among the **most complex text types** because of the high level of structuring and information content. Every dictionary entry is a structured object which uses different abbreviations and structural units in order to present the whole information succinctly. Furthermore, the structure of dictionary entries varies greatly within the dictionary itself as well as between different dictionaries. In spite of these variations some strict and constant structural rules exist so that the dictionaries can be understood by their readers.

Another kind of problems is related to the part of speech (POS) classifications of the headwords in bilingual digital dictionaries. As we already pointed out, [2], [3], the choice of *POS classifiers* of the headwords in the dictionary entries is very important. The development of a system of multilingual dictionaries on the basis of bilingual ones requires at first a unification of the classifiers in the dictionary entries. The problem turns to the harmonisation of the classifiers for various languages, and its solution has to present a unified selection of classifiers and a standard form of their presentation. The comparison of the Bulgarian and Polish material requires an explanation, which is important for the part of speech classifiers in the dictionary entries of the cited bilingual electronic dictionary. In the current paper we will analyze in short the specifications of verbal forms in both languages.

The POS classification varies across different languages. Often there is more than one possible POS classification for a given language. Next we will briefly review the POS classification of the *participle* (one of the important verbal forms) in the two languages, in comparison to another POS, the *adjective*.

Functions of the participle

The classification of a participle, not only as a verb form, is an important problem: the role of the participle varies significantly across languages, because its properties and functions are different. In the Slavic languages the forms of the participles are inflected, in contrast to English, for instance, where the participle are invariant. In the Slavic languages the forms of the participles contain information about the aspect and tense of the verbal form. The information about the aspect of the verb is important for the Slavic languages, but does not exist, for instance, in English. Bulgarian and Polish distinguish between the following functions of the *participle* form: predicative function, attributive function and adverbial (or semipredicative) function.

Participles and verbs

It is important to emphasize that participles preserve some properties of the main form of the verb, such as voice, tense and aspect. In Bulgarian and in Polish there are active and passive participles.

The properties presented above serve as a proof that participles deserve a separate treatment, different from that of adjectives.

Features of the adjective

Adjectives in Polish can be declined for gender, number and case, while in Bulgarian only for gender and number. The adjective does not express a temporal or aspect relation on its own, unlike the participle.

The main grammatical meaning of an adjective is its attributive meaning. Unlike a participle, which is closely related to the verb tense (with states or events in the past, present and future), an adjective describes properties or qualities of an object, like:

малко дете // małe dziecko // *a little child* //

Adjectives can have not only an attributive but also a predicative function (only as adjective clause). Adjective clauses perform the same function in sentences that adjectives do: they modify nouns. In this function, however, adjective clauses are just a nominal part of the subject, i.e. they do not express independently neither a temporal nor an aspect relation:

Mały dom // Малка къща // *A small house* //

Dom jest mały. // Къщата е малка. // *The house is small.* //

These arguments show that participles must be classified as a separate POS, and not be classified as adjectives. The unification of adjectives and participles is probably possible in languages that do not have verbal forms, or in which the system for description of the aspect and tense of the verbal form is simpler than that of the Slavic languages.

3. Design of Structure and Content of the Entry

Model for dictionary encoding: we choose the dictionary encoding model CONCEDE, developed in the framework of the EC project CONCEDE (Consortium for Central European Dictionary Encoding)². CONCEDE is a model developed in accordance with the TEI standards (Text Encoding Initiative: [7]) and offers a standardized, understandable and intuitive structure and semantics of a dictionary article. The CONCEDE model was used to develop lexical databases in six European languages - Bulgarian, Czech, Estonian, Hungarian, Romanian, and Slovenian. The first lexical database in Bulgarian was created in the framework of this project and contains more than 2700 lexical units from source "Bulgarian Explanatory Dictionary", for more details see ([5].

The build-up of electronic dictionaries is a complex and strenuous process, associated with overcoming various difficulties: (1) Lack of a sufficient number of formal models that allow words to be divided into formal language classes and a given word to be automatically included in one or another class. The creation of electronic dictionaries can be done through manual input of the dictionary articles – a process through which paper dictionaries are input into a computer (also possible with a scanner) or new dictionaries are prepared for print (they get printed after being typed up). These dictionaries, known as "machine-readable dictionaries" are different from their paper counterparts mostly in that they exist on magnetic carriers as files and can be processed as files. They follow a certain order and the articles have a concrete structure. As they are meant to be used by a human, their disadvantage from a computer point-of-view is that they are not sufficiently formalized (formal structures are missing from their descriptions) and the extraction of knowledge from them requires the development of special computer modules. (2) A great variety of structures and content, which presupposes a conflict between universality and detail. The conflict between universality and detail is particularly strong for dictionaries due to the large diversity in structures and content, which turns the creation of a standard for dictionary encoding into a major challenge, [8]. With the scope of overcoming this conflict the TEI workgroup created a universal standard for coding different types of dictionaries which encompasses fundamental principles of high degree of structure and diversity of dictionary entries [7].

² <http://www.itri.brighton.ac.uk/projects/concede/>

A dictionary entry – in terms of structure and content – is a complex unit. The structure of dictionary entries varies a lot within the dictionary as well as between separate dictionaries. The external structure (presentation of text) does not completely determine the internal structure (information content in the database).

There is a great diversity of hierarchical structures: in some dictionary entries the hierarchy organization of their structure may be deeply embedded (i.e. it allows many levels), whereas in other cases some structural elements from this hierarchy may be missing.

This makes the database supporting the dictionary logically complex and difficult to create.

4. Lexical Database of the Bulgarian – Polish Dictionary

The structure and content tags of the designed structural unit should fully meet international standards so that the LDB [4] and the electronic dictionaries be compatible with language resources created in other projects and for other languages.

Structure of a dictionary entry:

- Headword
- Formal Features – phonetics, grammar, morphology, syntax, etymology, style
- Semantic information
- Quotations
- Additional information:
 1. Derivatives
 2. Phrases
 3. Examples - phrasal and sentence usages, illustrations

The CONCEDE model [6] is used in the development of LDB for 6 European languages – Bulgarian, Czech, Estonian, Hungarian, Romanian and Slovene. The first LDB for Bulgarian was created in the above framework and contains more than 2700 lexical units from “Български тълковен речник”, [5].

The *structural tags* are **alt, entry, struc**, and the *content tags* are **case, def, domain, eg, etym, gen, geo, gram, hw, itype, lang, m, mood, number, orth, person, pos, q, register, source, subc, time, tns, trans, usg, xr**.

New content tags for **Bulgarian verbs** were added: the **<conjugation>** tag (to represent the conjugation of verbs) and the **<type>** tag (for the type of conjugation).

New information for the **aspect** of verbs in the tag **<gram>** (for perfect aspect and progressive aspect) and for the **transitivity/intransitivity** of verbs in the tag **<subc>** was also added.

Realization of homonyms:

The meanings of homonyms are entered in the dictionary as different database records. On the word-entry page, there is a field where the user must specify a homonym index - a number which shows the order of the meanings. For the representation of the homonym it is necessary to fill in the value of the attribute n (homonym index) in the tag **<entry>**:

<entry n="1">

<entry n="2">

Representation of an entry in the LDB

We choose the following entries from the Bulgarian-Polish electronic dictionary:

завь́ршва|м, -ш *vi.* kończyć, zakańczać; kończyć się
завь́рш|а, -иш *вр. v.* завь́ршвам
завь́ршен *part. adi.* skończoney, zakończoney, wykończoney

Representation of the dictionary entries in the LDB:

```
<entry>
  <hw>завь́ршва|м</hw>
  <pos>v</pos>
  <gram>i</gram>
  <subc>transitive</subc>
    <conjugation>
      <orth>-ш</orth>
      <type>III</type>
    </conjugation>
  <struc type="Sense" n="1">
    <trans>kończyć</trans>
    <alt>
      <trans>zakańczać</trans>
    </alt>
  </struc>
</entry>

<entry>
  <hw>завь́рш|а</hw>
  <pos>v</pos>
  <gram>p</gram>
  <subc>transitive</subc>
  <conjugation>
    <orth>-иш</orth>
    <type>II</type>
  </conjugation>
  <xr>завь́ршвам</xr>
</entry>

<entry>
  <hw>завь́ршен</hw>
  <pos>part</pos>
  <alt>
    <pos>adi</pos>
  </alt>
  <struc type="Sense" n="1">
    <trans>skończoney</trans>
    <alt>
      <trans>zakończoney</trans>
    </alt>
    <alt>
      <trans>wykończoney</trans>
    </alt>
  </struc>
</entry>
```


id	id_bg_word	pl_word	sense_index	alternative_sense_index	latin_translation	id_explanation
1110	663	kończyć	1	1		
1109	663	zakończyć	1	2		
1113	664	skończony	1	1		
1112	664	wykończony	1	2		
1111	664	zakończony	1	3		

Table pl_word

id_bg_word	id_characteristic
662	18
662	57
663	17
663	57
664	5
664	44

Table mm_bg_word_characteristic

id	abbreviation_bg	abbreviation_pl	description_bg	description_pl	description_lat	id_characteristic_type
5	прил	adi	прилагателно име			6
17	мин. нсв.	vi	глагол от несвършен вид			5
18	мин. св.	vp	глагол от свършен вид			5
44	прич	part	причастие			6
57	прех	transitive	преходен глагол			7

Table characteristic

id	name_bg	name_pl
5	Граматически категории за глаголи	
6	Граматически категории (части на речта)	
7	Граматически категории за преходност на глаголите	

Table characteristic_type

5. Web-based Application for the representation of the Bulgarian-Polish online dictionary

The technologies used for the implementation of the web-based application are Apache, MySQL, PHP and JavaScript. We use free technologies originally designed for developing dynamic web pages with a lot of functionalities. With the help of HTML and CSS we created the designs of both administrative and end user modules. The following web based application is experimental, and the structure of the text fields is not permanently determined.

The current version of the Bulgaria-Polish online dictionary works optimally with Internet Explorer 6.0+ (Windows), and with Firefox 2.0.1+ (Windows, Linux). The website resolution is 1024/768 pixels.

Future expansion of the Bulgaria-Polish dictionary is a precondition for any changes in the database and web application.

Main Modules:

The web-based application consists of two modules: an **administrator module** and an **end-user** module.

The administrator module - is intended for the person updating the dictionary, access only for authorized users. The administrator module is used to fill in the database and to offer user- friendly interface to the user who will be responsible for the word management.

The end-user module is bilingual, the user can choose the input language (Bulgarian or Polish) and according to their choice, a virtual Bulgarian or Polish keyboard is displayed.

5.1. Goals and tasks of the end-user module are:

- To present correct and up-to-date information to the user
- To be convenient and easy for searching and finding the meanings of words
- To allow an opportunity for translation from Polish to Bulgarian
- To allow the end-user to report missing words
- To create a user interface in both languages – Bulgarian and Polish (our idea is that both the end-user and administrative parts of the web-based application be bilingual).

The **end-user module** includes three sections.

Section **“Dictionary”**: this section is the home page – the first page of the web-based application. It contains a text field where the users fill in the word whose translation they are searching for. There are a Bulgarian and a Polish virtual keyboards built into this section.

„**About the project**” section: this section represents information concerning the project - authors, ideas, etc. Only the administrator could manage the text in this section. **“Support”** section that visualizes a form containing some text fields. Users can report missing words or inaccuracies in translation.

5.2. Goals and tasks of the Administrator Module

Goals and tasks:

- To allow for enlargement of the number of words in the database (volume of the dictionary)
- The dictionary must be easy to use and must not require a programmer's background from the administrator.
- The administrator must be able to receive, save and store missing words reported by the end-users, to insert new words in the database, to update all pages of the web-based application for the end-user.

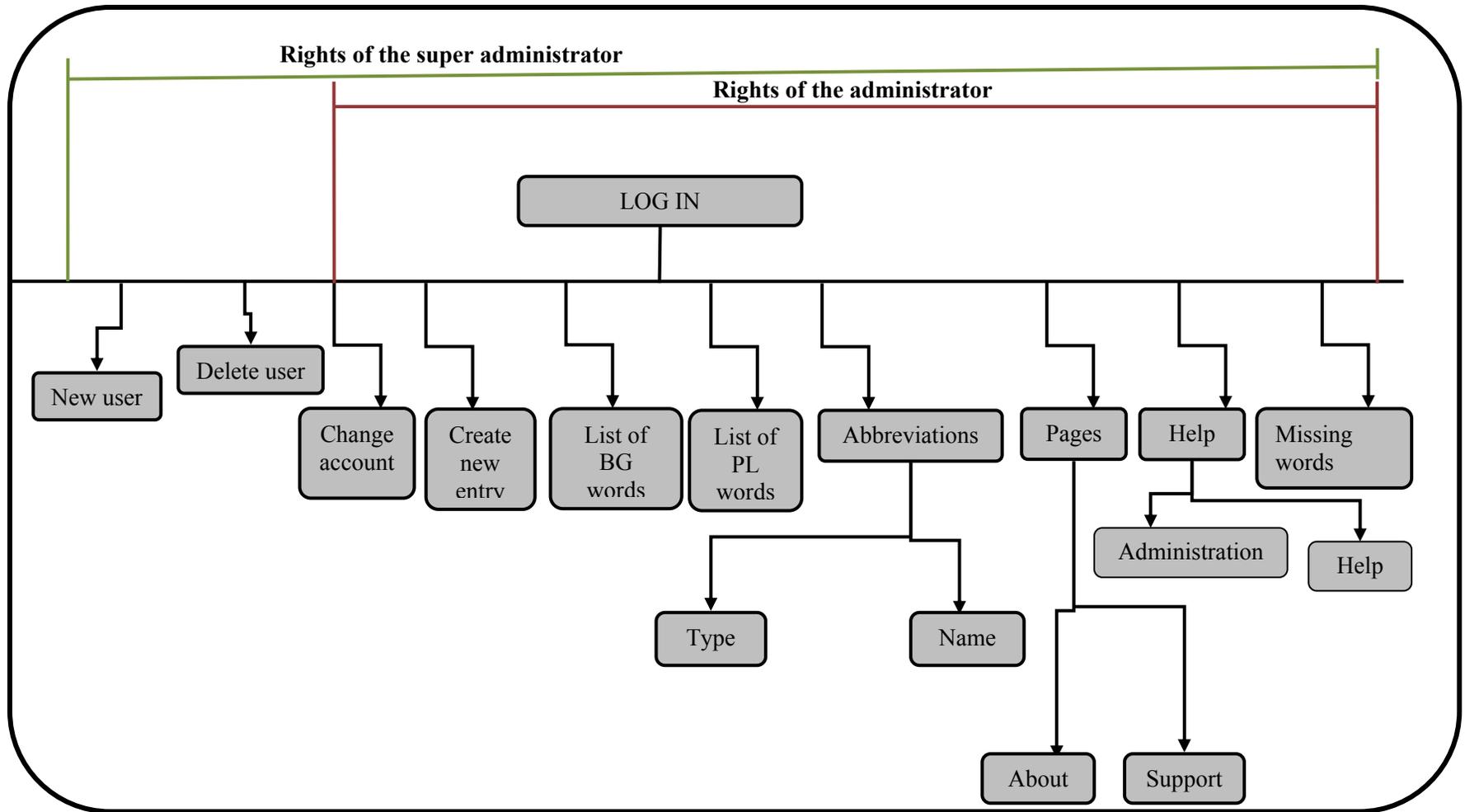
The **administrator module** is intended for the person updating the dictionary. We recognize two types of users here: (1) “**super administrator**”- who has all rights of adding, editing, deleting and searching for words; adding, editing and deleting users and (2) “**administrator**”- who has all rights except creating a new user and deleting an existing one.

Access to the administrative module is permitted only to authorized users. After a user's password and username have been verified, the user is redirected to the administrative module where there are **several sections**: a **section** for entering a new word, **sections** for searching for Bulgarian or Polish words, a **section** where the user can enter new abbreviations, a **section where** end-users report the missing words. The **Help section** serves both the administrators and the end users.

At the beginning the user must choose from a combo box what he/she wants to enter, (what part of speech) and only the corresponding text fields are loaded.



Administrative panel - choosing the type of the word which will be added



The structure of the **Administrator module**

6. How to fill in the LDB?

We present here examples of entering verbal forms: two transitive verbs and a participle. In the Bulgarian-Polish electronic dictionary the corresponding entries are as follows:

завъ'ршва|м, -ш *vi.* kończyć, zakańczać

завъ'рш|а, -иш *vp.* *v.* завъ'ршвам

завъ'ршен *part. adi.* skończony, zakończony, wykończony

(1) Adding the verb *завършвам* / end, finish off, complete, terminate /

The entry in MS WORD format:

завъ'ршва|м, -ш *vi.* kończyć, zakańczać

In the **first step** the administrator or authorised user must fill in the text fields for headword, conjugation of the verb in 2nd person, singular and the conjugation type I, II, III from a drop-down list, to point if the verb is transitive /intransitive, and to choose again from drop-down list the perfect aspect (*vp*) or imperfect aspect (*vi*) of the verb (see the picture below).

There are explanations for the determination of the conjugation type and the definition of transitivity or intransitivity of verbs in the *Help-section* of the administrative module.

Въвеждане на глагол	
Индекс за омоним	<input type="text"/>
Заглавна дума *	<input type="text" value="завъ'ршва м"/> <input type="button" value="търси в списък с думи"/>
2 л. ед.ч. сег. време *	<input type="text" value="аш"/> <input type="button" value="Спрежение на глагола"/> III <input type="button" value="v"/>
Св. / несв. вид на глагола*	<input type="button" value="vi"/> <input type="button" value="v"/>
Преходен / непреходен глагол	<input type="button" value="transitive"/> <input type="button" value="v"/>
добавяне на обяснение към думата >	
<input type="button" value=">>"/>	

Administrative panel – 1st step of adding the verb *завършвам*

It is not necessary to perform the **second step** for this kind of verb, so we move to step 3. In the **third step**, we must fill in the text fields the Polish meanings of the headword. With a button “add” we can enter as many as meanings as necessary. There are also added drop-down fields and extra text fields which can be used to give detailed information for the Polish verbs usage.

вие сте логнат като: admin | нов потребител | изтриване на потребител | изход

създаване на речникова статия | списък- български думи | списък- полски думи | съкращения | страници | помощ | докладвани думи

Данните са успешно запазени

Значение на полски						
№ група на точни значения*	Значение на полски*	(напр./пр.)	Сфера на употреба	Стилистично значение	Латинско значение	
<input type="text" value="1"/>	<input type="text"/>	добави				
1	kończyć					изтрий
1	zakończyć					изтрий

>>

Administrative panel – 3rd step of adding the verb *завършвам*

In the **forth step** we must add examples, derivations and phrases for the current verb. As can be seen from the example above, such information is not given for the verb *завършвам*. That is way one can press the button” finish” and the new word is added to the database.

(2) Adding the verb *завърши* / end, finish off, complete, terminate /

The entry in MS WORD format:

завъ'рш|а, -иш *вр. v. завъ'ршвам*

In the second entry of our example - headword “*завърши*” - there is a reference of type “*see to another verb*”. Due to this fact one must enter at first the verb from the reference - *завършвам*. To add this verb to the database of the web-application, one must perform several steps, some of which are not compulsory while others are obligatory. From the drop-down list the administrator (authorised user) must choose again the option to add a verb. The **first step** is the same as it was described for the previous word.

вие сте логнат като: admin

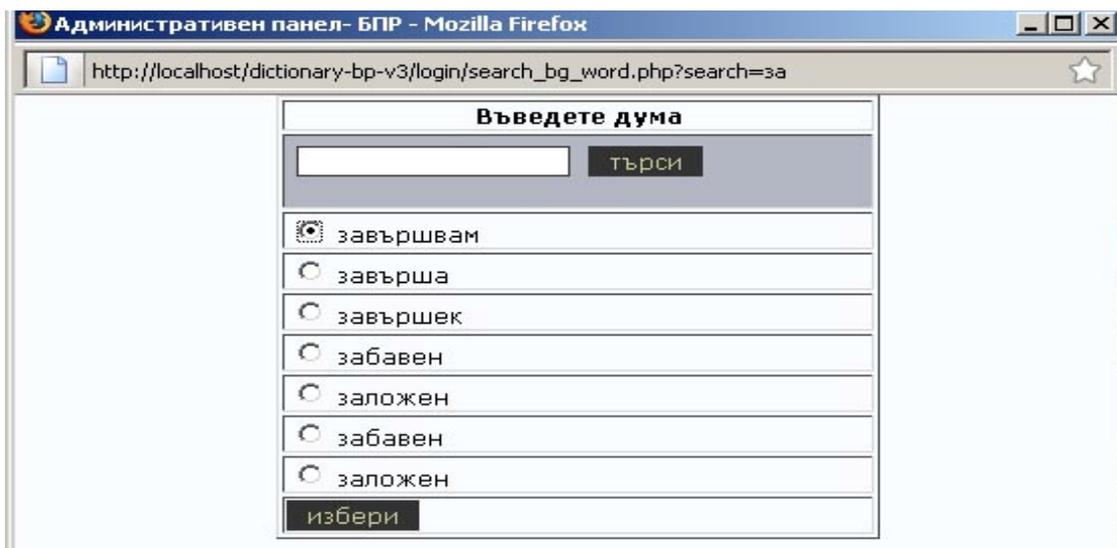
създаване на речникова статия | списък- български думи | списък- полски думи | съкращения

Въвеждане на глагол	
Индекс за ономим	<input type="text"/>
Заглавна дума *	<input type="text" value="завъ'рш а"/> <input type="button" value="Търси в списък с думи"/>
2 л. ед.ч. сег. време *	<input type="text" value="иш"/> <input type="button" value="Спрежение на глагола"/> II <input type="button" value="v"/>
Св. / несв. вид на глагола*	<input type="text" value="вр"/> <input type="button" value="v"/>
Преходен / непреходен глагол	<input type="text" value="transitive"/> <input type="button" value="v"/>
добавяне на обяснение към думата *	

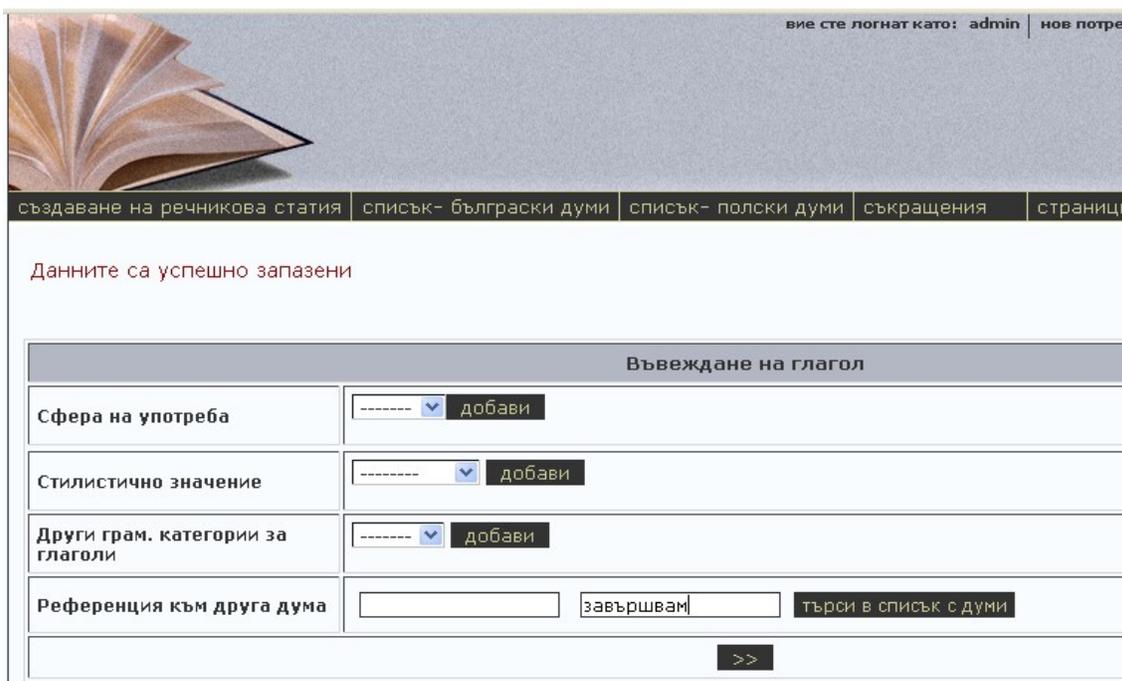
>>

Administrative panel – 1st step of adding the verb *завърши*

In the **second step** one must show that there is reference to another verb. After the administrator clicks on the button “*search from a word list*” a pop up window appears where the administrator (authorised user) can search for a word to create the reference. There is a text field and a “*search*” button. The user can search for a certain criteria. On the image below is the result of a search for all words, which begin with the prefix “*за*”. One can create a reference of type “*see only to words*” that are already stored in the database.



*Administrative panel –2nd step of adding the verb- pop up window for adding a reference to the verb **завърша***



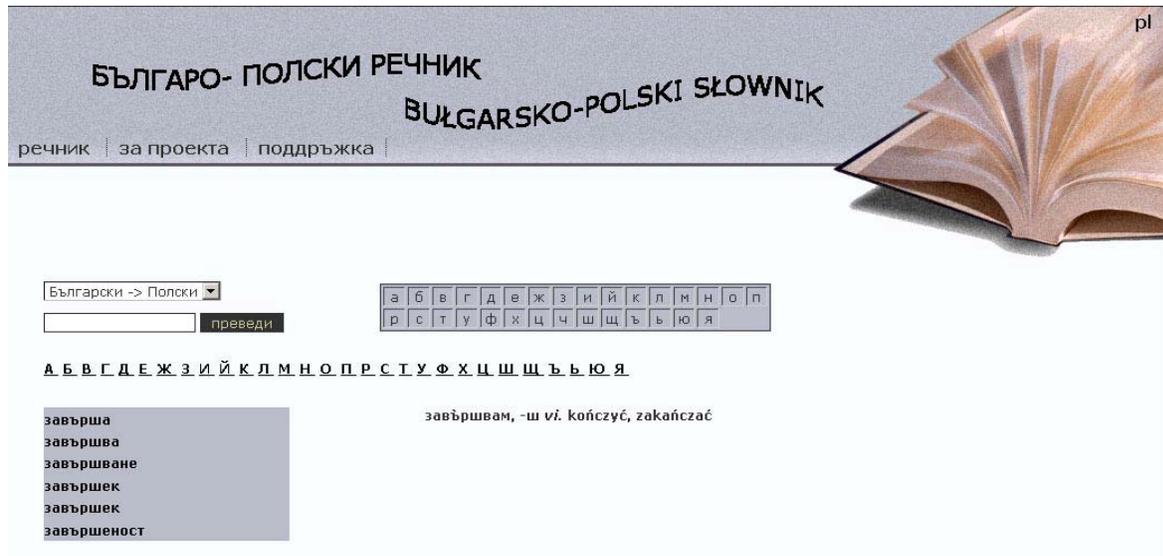
*Administrative panel –2nd step of adding the verb **завърша***

7. How the online dictionary works?

If an end-user starts to search for a word, on the left side of the screen a list of words, starting with the given entry, is displayed. When clicking on any of these words in the list the translation is visualized in the right frame.

If one translates from Bulgarian to Polish, the whole information saved in the RDB is displayed.

When translating from Polish to Bulgarian, only the Bulgarian headwords are visualized.



The screenshot shows a web interface for a Bulgarian-Polish dictionary. At the top, the title "БЪЛГАРО- ПОЛСКИ РЕЧНИК" (BULGARO-POLSKI SŁOWNIK) is displayed in both Cyrillic and Latin scripts. Below the title, there are navigation links: "речник", "за проекта", and "поддръжка". A search bar is present with a dropdown menu set to "Български -> Полски" and a "преведи" button. To the right of the search bar is a grid of letters for navigation: a b v г д е ж з и й к л м н о п and р с т у ф х ц ч ш щ ъ њ ю я. Below the search bar, a list of Bulgarian words is shown: "завърша", "завършва", "завършване", "завършек", "завършек", "завършеност". To the right of this list, the Polish translation "zawършвам, -ш vi. kończyć, zakańczać" is displayed. An image of an open book is visible in the top right corner of the interface.

*Web page for end user - translation of a Bulgarian word **завършвам***

8. Possibilities for the Future Improvements

- Improvement and extension of the databases (lexical and relational).
- Increase in speed of the database search.
- User feedback via a web-based application, which would allow a continuous update and extension of the dictionary resources.

References

- [1] Dimitrova, L., V. Koseska–Toszewa. Digital Dictionaries – Problems and Features. In: *Proceedings of the Jubilee International Conference Mathematical and Computational Linguistics, 6 July 2007, Sofia, Bulgaria*. 25-34. ISBN 978-954-8986-28-1.
- [2] Dimitrova, L., V. Koseska–Toszewa. Some Problems in Multilingual Digital Dictionaries. In: *International Journal Études Cognitives*. SOW, 8, 237–254. ISSN 1974-0818 ; 1
- [3] Dimitrova, L., V. Koseska–Toszewa. The significance of entry classifiers in digital dictionaries. In: *Lexicographic Tools and Techniques. Proceedings of the MONDILEX Open Workshop “Lexicographic Tools and Techniques”, Moscow, 3–4 October 2008*. 89-97. ISBN 978-5-9900813-6-9.
- [4] Dimitrova, L., R. Panova, R. Dutsova. Lexical Database of the Experimental Bulgarian–Polish Online Dictionary. In: *Metalanguage and Encoding scheme Design for Digital Lexicography. Proceedings of the MONDILEX Third Open International Workshop, Bratislava, Slovak Republic, 15–16 April 2009*. 36-47. ISBN 978-5-9900813-6-9.
- [5] Dimitrova, L., Pavlov, R., Simov, K. (2002). The Bulgarian Dictionary in Multilingual Data Bases. *Cybernetics and Information Technologies*, 2(2), 33–42.
- [6] Erjavec, T., R. Evans, N. Ide, A. Kilgarriff. The Concede Model for Lexical Databases. In: *Proceedings of the Second International Conference on Language Resources and Evaluation, LREC'2000*. 355-362, ELRA, Paris.
- [7] Ide N. M., C. M. Sperberg-McQueen. The TEI: History, Goals, and Feature. *In Computers and the Humanities*, 29, 5-15, 1995.
- [8] Ide, N., J. Véronis. Encoding dictionaries. In: Ide, N., Veronis, J. (Eds.) *The Text Encoding Initiative: Background and Context*. Dordrecht: Kluwer Academic Publishers, 167-179, 1995.
- [9] CONCEDE: <http://www.itri.brighton.ac.uk/projects/concede/>
- [10] TEI: <http://www.tei-c.org/index.xml>