Information-retrieval System "Base of the World Slavic Linguistics (iSybislaw)" in Language Education

Alla Taran [0000-0001-8091-1477]

Cherkasy Bohdan Khmelnytskyi National University, 81 Shevchenko Boulevard, 18031, Cherkasy, Ukraine
`alla__taran@ukr.net`

Abstract. The article examines the scientific and educational possibilities of the iSybislaw information-retrieval system (infpr ma cyjny System bibliografii slawistycznej). The national module within the joint scientific project framework of the Institute of Ukrainian Language of the NAS in Ukraine and the Polish Academy of Sciences "Online bibliography of modern world Slavic linguistics" is created by Ukrainian linguists. The base is considered as a source of information on language and linguistics; the components of the Ukrainian system module are analyzed: source texts, their bibliographic descriptions and key word dictionary for source and information search. The function of key words is performed by terms of Slavic linguistics and names that have a high frequency of documents in the database. The use of the information-retrieval system for performing linguistic tasks is illustrated. A pilot study was conducted in a group of undergraduate students of the specialty "Applied Linguistics" – according to the notes in iSybislaw selection of terms-paronyms, substantiation of feasibility of their use. Analyzing oppositions of paronyms helps to organize them into equivalence classes in the language of iSybislaw key words, and to create, with their participation, such searchable documents in a database that would meet the user's request.

Keywords: Key words: information-retrieval system, national module, key words, linguistic terminology, equivalence classes.

1   Introduction

Computer science is working to optimize knowledge systems in information retrieval systems. Today, linguists are keenly aware of the problem of systematizing linguistic terminology and the possibility of using information retrieval systems to search the terms in a multilingual environment. [1, 2]. The linguist's help in this search is greatly assisted by the bibliographic system "Base of the world Slavic linguistics (iSybislaw)", which provides an implement for multilingual access and retrieval of such information. It is available at www.isybislaw.ispan.waw.pl. At the institute of PAS in Warsaw. Today, Slavic linguists from different countries, in particular: Poland, Ukraine, Russia, Belarus, Czech Republic, Croatia, Germany, Slovakia, Macedonia, actively participate in the development of the information retrieval system. The

team of Ukrainian scientific participants works in such project which is headed by Professor E. A. Karpilovska, Head of the Department of Lexicology, Lexicography and Structural and Mathematical Linguistics of the Institute of Ukrainian Language of the NAS in Ukraine. The iSybislaw base enables access to scientific works in all Slavic languages in various problems of modern Slavic sciences and opens prospects not only for obtaining background information but also for performing various theoretical and practical research tasks. It covers scientific works and texts published since 1992, and at the same time it is constantly updated interface, improving the functionality of the system. This important and effective search implement for philologists in Slavic studies, experts in the field of Slavic linguistics, provides search in a well-defined linguistic segment, can take the documents with different languages, which certainly promotes scientific dialogue in international Slavic studies. Recently, developers have begun to attach texts to bibliographic descriptions, thus creating the iReteslaw repository.

2   iSybislaw as a Source of Information About Language and Linguistics

iSybislaw has an extensive record information structure (formal and substantive): key words (as of 02/11/2020 – 24601 words), classification, summary and abstract. The searching use for information in various and differ-ent ways of the system provides a dictionary of key words in Slavic linguistics. In iSybislaw, one of the key places is to search for the information you need with the help of a key word language based on terms of linguistics.

The unique key word system of International Committee of Slavic linguists – all Slavonic, also English, German and French, makes it possible to search for documents not in the language but in the context of the language issues being discussed. The capabilities of the special information-retrieval language created by iSybislaw founders of the key word system in the learning process deserve some discussion. There are two types of searches in the base – simple and ad-vanced. Simple is ideal because it involves searching for information on a par-ticular problem by key word. It can also be the author's last name or naming unit. The retrieval system displays results (number of founding sources) that match a given search. The use of logical operators in the advanced search pro-cess can reduce the results obtained and limit, for example, sources only in the Ukrainian language or by time period. For example, if a user needs to find a Ukrainian grammar book published in 2017, he or she will find it easy. An ad-ditional benefit of the iSybislaw system is the ability to search through indexes – alphabetically sorted documents – and view the contents of individual journal numbers, series of publications, and conference materials. This is another search method that makes it easy to access information.

The thematic range of the system covers language theory and general linguistics, Slavic linguistics, as well as contrastive Slavic-non-Slavic linguistics.

Developers of national modules iSybislaw and Ukrainian in particular, through the description of new scientific works, supplement and reorganize the index of key

words. The function of key words is primarily performed by terms of Slavic linguistics and names that have a high frequency of documents in the database. It is an open set that always provides for the development of the language sys-tem according to the degree of detail and scale of the information required for users.

Title:	Динамічні процеси в сучасному українському лексиконі	
Author/editor:	Н.Ф. Клименко, Є.А. Карпіловська, Л.П. Кислюк	
Publication place:	Київ	
Publisher:	Видавничий Дім Дмитра Бураго	
Year:	2008	
Pages:	336 s.	
Note:	Bibliogr., indeks, streszcz.: rus, eng	
Languages:	ukr	
Abstract:	Розглянуто зміни в лексиконі української мови за період незалежності (від 1991 р.). Проаналізовано типи лексико-словотвірних інновацій: новотвори, новозапозичення, неосемантизми, а також нові поняттєві поля лексики, процеси її термінування й детермінування, взаємодію функціональних стилів мови. До монографії додано індекс нових слів (в тому числі неосемантизмів). (JR)	
Author headings:	ws. Klimenko, Nina Fedorìvna	[Show]
	ws. Karpilovs'ka, Êvgenija Anatoliïvna	[Show]
	ws. Kisljuk, Larisa Pavlìvna	[Show]
Keywords:	1991- r., dynamika językowa, język ukraiński, język współczesny, neologizm, nominacja 1 (procedura), słownictwo, słowotwórstwo 1 (tworzenie wyrazów), tendencja rozwojowa, zmiana językowa	
Tags:	add new tag here add	
Classification:	9.4.6. Ukrainian. Lexicology	[Show]
	9.4.1.3.1. Ukrainian. Nominal formation	[Show]
Inbound references:	Recenzja: WASZAKOWA Krystyna .- "Rocznik Slawistyczny = Revue Slavistique" 2009 [Show]	

Fig. 1. Description of the document in iSybislaw

The concept of iSybislaw is to arrange series of term equivalents in the structure of a key word language that combine the series of terms in Slavic and English (for example, synonymous terms in the same language and equivalent in different languages and different alphabets). To a user who is fluent in such Slavic language and does not speak English, the system shows all the multilingual documents that contain the required information. Conversely, English-speaking users will be able to view the information through an English term selected from the entire equivalence class. The descriptors in the key word-term classes are Polish terms, since the creation of the iSybislaw system under the auspices of the Institute of Slavic Studies, PAS.

Equivalence classes, though simplified, indicate the place of individual terms used in the function of key words in relationships and form the terminology of the separate Slavic languages and in Slavic education at all. It was shown in the Base documents of the world Slavic linguistics. The establishment of inner equivalence classes (Fig. 2) and interlingual (Fig. 3) enables the user to access knowledge through any word belonging to such classes. Therefore, the information value of the iSybislaw system is

increasing and new ways of accessing knowledge about Slavic languages and Slavic lin-guistics are emerging [3].

Descriptor: odmiana regionalna języka
Language: pol (polski)

Keyword versions

Word	Language	Descriptor	Actions
odmiana regionalna języka	polski (pol)	⊙	[Show]
regiolekt	polski (pol)	○	[Show]
regionalna odmiana języka	polski (pol)	○	[Show]
wariant regionalny języka	polski (pol)	○	[Show]
региолект	rosyjski (rus)	○	[Show]
региональный вариант языка	rosyjski (rus)	○	[Show]
regiolekt	chorwacki (scr)	○	[Show]
regionální varianta	czeski (cze)	○	[Show]
варіант літературної мови	ukraiński (ukr)	○	[Show]
регіональна мова 2	ukraiński (ukr)	○	[Show]

[Show documents]
[Show log]
[Check status]

Fig. 2. The class of interlingual equivalence of the semantic field **regiolekt**

Keyword versions

Word	Language	Descriptor	Actions
odmiana regionalna języka	polski (pol)	⊙	[Show]
regiolekt	polski (pol)	○	[Show]
regionalna odmiana języka	polski (pol)	○	[Show]
wariant regionalny języka	polski (pol)	○	[Show]

Fig. 3. The class of intralingual equivalence of Polish terms,synonymous to the **regiolekt**

Each bibliographic description is accompanied by an annotation that will be useful in scientific work. We, the drafters of the national Ukrainian iSybislaw module, formulate an annotation that would cover the most important information about the content of the publication. At the same time, it should not duplicate the abstract of the document itself. First, we are introduced to the document, then begin to form a bibliographic description in the database. After reading the annotation, the user will be able

to understand if this source is useful for him. For example, now we supplement iSybislaw with linguistic monographs and articles in professional publications which were published in Ukraine in 2016–2019, refining the search characteristics of documents, and especially key words. One of the main tasks of the iSybislaw module, is to create a list of key words to search for information in Ukrainian-language documents, and to establish matching words (primarily linguistic terms) in other Slavic languages.

3 iSybislaw as an Implement for Obtaining Linguistic Information

Digital bibliography is also of practical use, particularly in the training of future philologists and proficients in applied linguistics. In our opinion, students should be aware of the existence of the information-retrieval system iSybislaw base. In the learning process, it can be an additional stimulus that promotes the acquisition of new knowledge from different fields of linguistics. At the initial stage of writing a scholarly work, a student, for example, has the opportunity to find out how well his topic has been researched in the scientific literature, and later – to compile a bibliography using iSybislaw materials. This saves him time compared to searching the library and, at the same time, makes this search a targeted, most up-to-date linguistic source, published in different Slavic coun-tries.

A pilot study conducted in a group of undergraduate students of the specialty "Applied Linguistics" at Bohdan Khmelnytskyi National University of Cher-kasy specializing in the field of Ukrainian studies. It is known that the process of scientific (linguistic) thinking, development and functioning of the linguistic metamova naturally leads to the emergence of ambiguity, paronymy, hyper-hyponymy. The theoretical understanding of the systemic relations between the terms is not only of great practical importance for the ordering and normalization of the iSybislaw key words language, the prospects of developing the Ukrainian module in view of the state and development of the Ukrainian terms system in linguistics, but also of great theoretical importance for studying the ways of developing modern linguistic terminology, common and distinct in some Slavic linguistic terminosystem. After all, the ultimate goal of unifying terminology, according to Y. D. Apresyan, is "to build an optimal system of the most convenient terms", in addition, this system must be built "not on a notion, but on linguistic basis, not on logic, but linguistic reasonably" [1, p. 54].

4 Terms-Paronyms in the "Base of the World Slavic Linguistic (iSybislaw)"

The selection of topical linguistic material, the desire for the represent-ativeness of the key words language, its adequacy of the content contained in the system of iSybislaw documents leads to the appearance of paronyms in it - words that are similar in form and partially different in meaning. For this rea-son, students at iSybislaw were asked to find paronyms, to clarify their meaning and to justify the use of a term.

The indistinguishability of partially similar in the form of lexemes, their confusion in linguistic practice arises because the speaker does not have a clear understanding of the semantics of one or both components of patronym-terms.

Students have captured and analyzed key word terms to the extent that they are presented today in iSybislaw, in its key word language. First of all, the subject of the study was the expression-expressive paronyms.

In linguistics there is no single approach to the linguistic understanding of the terms "expression" and "expressiveness". Linguists try to come up with objective criteria for differentiation of these concepts, and the very problems associated with linguistic mechanisms of creating the effect of expressiveness (expression), need to identify and determine the basis of this phenomenon" [3, p.18].

According to Sh. Balli, the founder of the modern concept and methods of exploring expressive facts of language, expressiveness is "enhancing expres-siveness, imagery, increasing the effective power of what is said" to realize the evaluative-emotional or logical-meaning content of language in the process of verbal communication, the factual learning of reality" [6, p. 107]. Ukrainian linguist V. A. Chabanenko states: "If expression is an" intensifica-tion (increase, amplification) of expressiveness, then expressiveness is already the intensified (increased, enhanced) expressiveness, such a psychologically and socially motivated property of the linguistic sign, which deautomatizes it, maintains keen attention, activates the thinking of the person, provokes the ten-sion of listener`s (reader) feelings. He emphasizes that expression is expressed only at the speech level, and expressiveness is both spoken and linguistic [7, p. 9]. Speech expressiveness prevails over language and is a con-stant source of enrichment for the latter. The linguist distinguishes between ex-pressiveness (expression) of speech and language [7, p. 10]. These are not identical nouns, because expressiveness is a generic concept ("in-tensification of expressiveness"), expression – types (intensified expressive-ness" itself).

The opposition of the meanings of these terms is outlined: expression is a phenomenon (evaluative, figurative, emotional component of the content of a language sign in the act of speech, use of the sign); expressiveness – the property of a language sign, its ability to express expression.

According to Ukrainian linguists, lexical expressiveness can be not only a phenomenon of linguistic (systemic, visual, inclusive), but also of speech (situational, temporal, contextual, adherent), so it is necessary to distinguish between expressiveness of two types: 1) expressiveness of speech and 2) ex-pressiveness of units, and therefore the content of the terms "expression " and "expressiveness" respectively.

The terms and authors of iSybislaw's scientific works and its develop-ers interpret this term differently. For example, only the key word expressive-ness is given to the bibliographic record of the article by A. A. Zainuldinova, E. F. Kero Hervilli "Methods of semantic categories translation "expressive-ness", "emotionality" and "evaluation" from Russian into Spanish". As the name of the document shows, expressiveness as a semantic-stylistic category reveals a connection with appreciation, emotionality, stylistic meaning, but it cannot be identified with the named concepts and terms for their designation.

The iSybislaw database contains 89 documents with the key word ex-pressiveness, many of which include the term connotation with the Polish de-scriptor konotacja 2 (relacja asocjacyjna) and the equivalents konotacja seman-tyczna (pol.) and connota-tion 2 (in Russian).

Analyzing the bibliographic records, students identified three types of expressive-ness:

1) expressiveness = intensification of expressiveness, as evidenced by key words of Russian terms expressive potential, expressive syntax, expressive vocabulary, ex-pressive language foundation, expressive nominative, expressive word-forming functions by communication. However, the documents them-selves have a much broader content of these concepts than key words.
2) expressiveness = evaluation, emotionality;
3) expressiveness = expression.

There is a term in the iSybislaw key word register for expression (Polish descriptor ekspresywizacja) – the process of enhancing expression, it is recorded in the biblio-graphic description of the article by Z. Novozhenova, A. Pstiga, "Phrasemes in public discourse as a means of expression". The term expression is used only in the Russian module iSybislaw.

The students have not gone unnoticed common root terms expressive – expressive vocabulary. An expressive is a form in which expression is ex-pressed. The system has a Polish descriptor ekspresywizm and the Russian term expressive. The perspec-tives of the Ukrainian module are seen in the addition of a new term to the system of lexical express = express word (N. I. Boyko) and expressive = type of linguistic act by illocutionary force; is a manifestation of greeting, sympathy, attitude, apology, gratitude, wishes (O. O. Selivanova) to-gether with the works of Ukrainian linguists where such terms are used.

Therefore, Ukrainian terms should be divided into equivalence classes and Polish descriptors in the system, that such classes lead:

- expression - Speech Expression (V. A. Chabanenko) – adherent ex-pressiveness (N. I. Boyko) – rus. expression;
- expressiveness – language expression (V. A. Chabanenko) – inherent expres-siveness (N. I. Boyko) – gender. ekspresywność
- expressive 1 - expressive act of speech (I. Melnyk) – expressive act of speech / language act (N. Kuzenna, I. Prykarpatska)
- expressive word – expressive 2 (N. I. Boyko) – rus. expression – gen-der. ek-spresywizm
- expressionism – tendency to expressiveness (V. Bogatko)

One way to express vocabulary is to repeat it. Students recorded the Russian terms paronyms repetition – repeatability.

Keyword

Descriptor: powtórzenie
Language: pol (polski)

Keyword versions

Word	Language	Descriptor	Actions
powtórzenie	polski (pol)	●	[Show]
iteracja	polski (pol)	○	[Show]
повторение	rosyjski (rus)	○	[Show]
паўтор	białoruski (bel)	○	[Show]

Fig. 4. Equivalents of the Russian term repetition in the class with the Polish descriptor powtórzenie

Intra-linguistic equivalents were recorded: the Polish terms – descriptor powtorzenie and ascriptor iteracja. Unfortunately, there are no Ukrainian equivalents yet in this class, but in our opinion, the answer is a deadline term repeat. Its appearance will be provided by works of Ukrainian linguists on this problem.

According to O. S. Akhmanova`s dictionary, repetition = the same as a repeat: repetition of the root; and for the term repetition there are two meanings:

repeat (doubling, reduplication, repetition) - complete or partial repetition of a root, base or whole word without changing their sound composition (or with minor changes) as a way of creating words, synthetic and descriptive forms of phraseologisms;

repetition - a figure of speech, which consists in repeating in a certain sequence of sounds, words or parts of them, utterances to achieve the corresponding expressive or expressive-pictorial effect" [8, p. 459].

In the article of encyclopedia "Ukrainian language" focuses on two meanings: repetition = figure of language; repeat = reduplication [8, p. 459]. The theoretical material was able to provide the following schematic representation of equivalents:

repetition 1	repetition 2
repeat	repeat
reduplication	phonogram
	lexical repetition
	syntactic repetition

The introduction of the term repeatability is unreasonable: it is used once as a key word in the article by O. K. Grekova "On the actual aspectual, modal-aspectual and modal semantics of species juxtaposition".

In general, we need a thorough linguistic analysis of the most common models of paronyms and their semantic differentiation in iSybislaw. It is in terminology that the correspondence of form and content must be consistent and uncontrovercial, as required by the specificity of professional-scientific communication. How the terms imperfect – imperfective, perfect – perfective, implicature – implication, illocution (equivalents in iSybislaw Pol. aktilokucyjny, Rus. illocutionary act) – illocution, speech inertia – speech intension, syllogism and others.

According to our observation, common in the registry of iSybislaw key words are shared root doublets, which are the result of the development of national term systems due to the internal resources of a particular language, such as Rus. cognitive linguistics – cognitivistics, mythical name – myth, regional language

– regiolect, participle construction – participle turnover, dialogue – dialog speech, borrowing (result) – a borrowed word, a foreign word – a foreign word.

5   Conclusions and Future Work

Therefore, in the iSybislaw information retrieval system, the problem of differentiating similar form terms based on the expression of their most frequent word-forming patterns and cognitive structures in multilingual term systems is an urgent problem. Not only is the ordering of the existing terms of the system promising, but also the development of its Ukrainian module, which illuminates the semantic relations between the existing terms from different languages, as well as the compilation of a dictionary of terminological paronyms, which con-tributes to the process of ordering and systematization of Slavic linguistics.

The prospects of using the information base as a source of scientific search are primarily related to providing access to full-text files of at least the latest new sources in the iReteslaw repository. This is especially true given the low avail-ability of contemporary literature on Slavic linguistics outside the country where the scientific work was published.

In addition, modern students are extremely quick and adept at navigating the information space, so it is important to integrate the information contained in the system into the general flow of scientific information, to relate it to other available external sources of such information (while related to the resources of major national libraries, in particular Vernadsky National Library (www.nbuv.gov.ua), as well as national electronic resources: "Russian-Ukrain-ian dictionaries website (www.r2u.org.ua), the linguistic portal of the labora-tory of Computer Linguistics of Taras Shevchenko National University at Kyiv (www.mova.info), websites of Academic Institute of Ukrainian Language and O.O. Potebnia Institute of Linguistics, Universities of Ukraine.

References

1. Karpilovska Ye.: The role of national modules in the development of the iSybislaw base information-retrievallanguage. Ukrainian language (1), 17–29 (2018).

2. Rudnik-Karwatowa Z., Karpinska H.: Dictionary of key words in Slavic linguistics. Warsaw (1999).
3. Kovalsky P., Banasyak I.: Key words and equivalence classes in the iSybislaw information-retrieval implement and subject of research. Ukrainian language (4), 14–26 (2017).
4. Apresyan Yu.: Lexical semantics: Synonymous means of language. Science, Moscow (1974).
5. Boyko N.: Ukrainian expressive vocabulary: Semantic, Lexicographic, and Functional Aspects. Aspect-Polygraph Publishing House, Nizhyn (2005).
6. Bally Sh.: General linguistics and French language issues. Ed. foreign lit-ry, Moscow (1955).
7. Chabanenko V.: Stylistics of expressive means of the Ukrainian language. ZSU, Zaporozhye (2002).
8. Akhmanova O.: Dictionary of Linguistic Terms. Soviet Encyclopedia, Moscow (1966).
9. Ukrainian language. Encyclopedia. In: Rusanivskyi V. (eds.). Ukrainian Encyclopedia, Kiev (2007).
10. Taran A.: Language dynamics in sources and key words of information-retrieval system iSybislaw. In: Z zagadnień informacji naukowej, terminoznawstwa i językoznawstwa, pp. 321–333. PAN, Warszawa (2018).References