

The Role of Keyword Language in the Database of World Slavic linguistics "iSybislaw"

Alla Taran ¹

Cherkasy Bohdan Khmelnytskyi National University, 81 Shevchenko Boulevard, 18031, Cherkasy, Ukraine

Abstract

The article analyzes the role of key words in the bibliographic database of world Slavic linguistics iSybislaw; it was investigated that key words reflect the content of the document and at the same time allow the user to find relevant information in the whole array of documents (perform meta-information and search functions), and the use of descriptors makes it more efficient to search for information. Particular attention is paid to the semantic relationship between keyword terms. The problems of synonymy, variability and ambiguity of linguistic terms in information retrieval language are considered and the ways of their solution are offered. Experience in organizing terminology for iSybislaw is outlined.

Keywords

Information retrieval system, bibliographic base of world Slavic linguistics, information retrieval language, term, synonym, variant, ambiguity.

1. Introduction

The bibliographic base of world Slavic linguistics iSybislaw is an effective and important tool of scientific research, which allows to search for information in a specific linguistic segment, without dividing it into national schools / languages, which, of course, promotes scientific dialogue in international Slavic studies [3]. iSybislaw is a powerful information retrieval system that allows users to access scientific papers in order to obtain background information and perform a variety of theoretical and practical research tasks.

In addition to information on the content and formal features of the documents, some records have been submitted in full. The quality of information in the system is ensured by the inductive-deductive method of selection and processing of documents, as well as the linguistic competence of indexers. Components of the system bibliography of the world linguistic studies are national modules, formed on a common principle. Since 2015, the author of the article has been working on the development of the Ukrainian module together with colleagues from the Department of Structural and Mathematical Linguistics of the Ukrainian Language Institute of NAS of Ukraine within the international research project "Online Bibliography of World Slavic Linguistics". The coordinator of this project on the Polish side for a long time was Dr. Z. E. Rudnyk-Karvatova together with prof. B. Boyar in 1998 published the concept. Now the project manager from the Institute of Slavic Studies of the Polish Academy of Sciences is the chairman of the Commission on Linguistic Bibliography, Dr. Pavel Kowalski, and the project manager from the Institute of Ukrainian Language of NAS of Ukraine is the head of the Department of Structural and Mathematical Linguistics, Professor Ye. A. Karpilovska.

The article aims to analyze the possibilities of optimizing the information retrieval language of key words iSybislaw. We focus on the problem of ambiguity, variability and synonymy of linguistic terms, their equivalence to terms in other Slavic languages.

COLINS-2022: 6th International Conference on Computational Linguistics and Intelligent Systems, May 12–13, 2022, Gliwice, Poland

EMAIL: email1alla_taran@ukr.net (A.Taran)

ORCID: 0000-0001-8091-1477 (A.Taran)



© 2022 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

2. Key words are the powerful basis for studying Slavic linguistic terminology

The language of information retrieval is the language of key words. The developers of Polish, Ukrainian, Belarusian modules Z. E. Rudnyk-Karvatova [10], P. Kovalskyi and Ya. Banasiak [4], Ye. A. Karpilovska [7], Ye. Volkova [2], Yu. Romaniuk [6] – has repeatedly pointed out that terms are the main tool for creating user queries. The use of terms as key words is effective for users due to their language competence. One of the main tasks of the developers of the Ukrainian module is to create a list of key words to search for information in Ukrainian-language sources, as well as to establish equivalents of such words (especially linguistic terms) in other Slavic languages. As Ye. A. Karpilovska rightly points out, it is important to "fill in the gaps in the index of key words caused by the lack of sources from certain linguistic disciplines and at the same time balance the presentation of such sources, elimination of "distortions" towards certain linguistic disciplines or areas" [8]. To accumulate key words, indexers use different linguistic resources. The user can view documents alternately by title, author or find documents to review the bibliographic description.

The language of key words is close to natural, it is a fragment of cognitive space. The register of key words of Slavic linguistics is an open system, covering elementary language units that coincide with the terms of different languages. In the process of creating a database of key words, translation is not used, the search for functional equivalents in each of the languages is carried out independently, taking into account the content of the document and compatibility in each thematic system. Semantic equivalence is important for establishing the identity of words and phrases that have an identical or close terminological meaning in several languages.

Key words in the iSybislaw system reflect the content of the document (perform a meta-information function) and allow the user to find relevant information throughout the array of documents (search function). An important feature of key word language is pertinence, or compliance of the provided information with the information needs of the user. In 1999, Z. E. Rudnyk-Karvatova and H. Karpinska published a Dictionary of Key words, which facilitates the search for documents in the information retrieval system of Slavic linguistics [11]. Since 2006, an electronic database of key words in various Slavic languages has been created on the basis of the dictionary.

Key words are elementary units of information retrieval language that correspond in form to terms relevant to this field of knowledge and are able to represent the content of a particular document. The following criteria are taken into account when selecting terms for the key word dictionary indexers take into account: frequency (repeatability of word forms in the text, frequency of use exceeds their use in language), information saturation (maximum information in the minimum space), semantic proximity of the word to the topic of the text (synonyms, genus-species relations are considered), and correctness, relevance, conciseness, structural transparency. The hierarchy of these criteria may vary depending on the type of term, its place in the terminological system of a particular language and the role in the text of a particular document.

The terms reflect the content of the documents of modern Slavic linguistics, contrasting Slavic-non-Slavic linguistics, as well as theories of language and general linguistics.

The iSybislaw information retrieval tool meets the deadline. The semantic characteristics of the term also include proper and common names that outline thematic fields important for the organization and search in the database: these are the names of linguists and writers, names of fiction works, dictionaries, monographs, geographical names (regions where research, for example, dialectology), names of conferences (as well as place and year), collective names (scientific societies), common names: tokens, phraseology, which is the object of study, numerical expressions (years and centuries). Morphological characteristics of the term indexing include derived words, complex words, phrases, abbreviations, morphemes.

3. Approaches to solving problems of variance and synonymy

Key words are an important indicator of the quality of indexing. The right choice of key words determines the effective further search for scientifically relevant information. Therefore, their function is to provide an intellectual search for information on the problems of Slavic linguistics. The main criterion for choosing a key word is its potential value for displaying the topic of the document

or for its search. The functioning of the term as a key word in information retrieval language involves the elimination of polysemy, synonymy and variability, which are characteristic for natural languages. However, if descriptors as conditional normative elements are unambiguous, then ascriptors, hierarchically subordinate to descriptors, are capable of logical synonymy, which, in turn, creates conditions for improving search efficiency.

We will illustrate the work with the Ukrainian module on the example of the records of the magazine "Culture of the word: interdepartmental collection" (Kyiv, Institute of Ukrainian Language of the National Academy of Sciences in Ukraine) (see Fig. 1).

Journal issue

Journal: Культура слова : міжвідомчий збірник .- Київ, 1997 - 2019

[\[Show\]](#)

Year 2018

Issue Вип. 88

Articles: MAMIČ M.V.: Вогник чистого серця, або мовотворчість О. Гончара на сторінках журналу "Радянська жінка"	[Show]
SOLOGUB Nadija Mikolaïvna: «Далекі вогнища» мовотворчості Олеса Гончара	[Show]
ĖRMOLENKO Svitlana Jakivna: Думки понад часом	[Show]
GONČAR V.D.: Дух мусить мати свій суверенітет : спомини про мовний світ Олеса Гончара	[Show]
JACENKO Nina Oleksandrivna: «І звідти, з синього неба, білосніжні перла летять...» : кольороназви у новелі Олеса Гончара «Залізний острів»	[Show]
BIBIK Svitlana Pavlivna: Идеалема «пам'ять» у лінгвософії Олеса Гончара : на матеріалі мови роману «Тронка»	[Show]
SINIČJA A.: Інтерпретація прецедентного імені в публіцистиці Олеса Гончара	[Show]
KOC' Tetjana Anatoliïvna: Історія літературної мови в «Письменницьких роздумах» Олеса Гончара	[Show]
SNIŽKO Natalija Volodimirivna: Концептуальні доміанти мововіту Олеса Гончара	[Show]
MAC'KO Ljubov': Лінгвостилістична макроструктура роману Олеса Гончара "Твоя зоря"	[Show]
SAMOJLOVA Ėrina A.: Мова прози Олеса Гончара як об'єкт фразеографії	[Show]
GANŽA A.Ju.: Мовний портрет О. Гончара в українській документалістиці	[Show]
DANYLJUK N.: Мовностилістичні засоби ліризму в новелах О. Гончара	[Show]
MECH Natalija Oleksandrivna: Роман О. Гончара «Собор» як вияв «устремління людини до вічності»	[Show]
PETRENKO L.: Слово Олеса Гончара в Словнику синонімів української мови	[Show]
BOJKO N. I.: Типи лексичних експресивів в ідіолекті Олеса Гончара	[Show]
ĖRMOLENKO Svitlana Jakivna: Філософія мови Олеса Гончара	[Show]
ČEREM'SKA Oľga Stepanivna, MASLO Oľga Volodimirivna: Харківський період мовотворчості Олеса Гончара	[Show]

Figure 1: Description snippet of the digest "Culture of the word" (Issue 88, 2018)

The magazine contains an overview of current issues of the language of works of art, newspaper and magazine journalism and advertising, some problems of grammar, stylistics, dialectology, phraseology, culture of the Ukrainian language and standardization of oral literary language in the public sphere. Usually each issue is devoted to the language world of one of the Ukrainian writers. For example, issue 88 is dedicated to the 100th anniversary of the birth of Oles Terentiyovych Honchar, an iconic figure in the history of Ukrainian literature, language and culture. One of the possible key words for many articles should be *idiostyle*. In the system to the Polish descriptor there are Russian Czech and Ukrainian ascriptors: Pol. *idiostyl* – Rus. *идуостиль* ‘idiostyle’ – Czech *autorský styl* – Czech *individuální styl* – Ukr. *авторський стиль* ‘author's style’– Ukr. *ідиостиль* ‘idiostyle’. In addition, the authors of some articles use the term *idiolect* (Boiko N. I. "Types of lexical expressions in the *idiolect* of Oles Honchar"). In iSybislaw to the Polish descriptor Pol. *idiolekt* there are askriptors Rus. *идуолект* ‘idiolect’ – Belarus. *ідыялект* – Ukr. *ідиолект* ‘idiolect’. However, in linguistics, the terms *idiostyle* (and its variant *individual style*) and *idiolect* are used to denote the set of linguistic features of a writer's work, and researchers often use them as interchangeable. This illustrates the definition given by S. Ya. Yermolenko: individual style, or idiolect, is a set of linguistic and expressive means that perform an aesthetic function and distinguish the language of an individual writer from others [5]. If idiolect as a general linguistic term has long been established in linguistics, and its meaning as a set of individual features inherent in the speech of a particular individual, is more or less stable, the concept of idiostyle is defined insufficiently. Mixing of these concepts is observed in dictionaries. Accordingly, such ambiguity in linguistic terminology affects the structuring of information and search procedures.

Other examples. The equivalence classes of the terms *author's language* and the *author's language of the work of art* testify to the partial equivalence (closeness) of the co-referential terms and the types of their variants.

Table 1

The equivalence classes of the terms *author's language* and the *author's language of the work of art*

język autora (pol)	język autora dzieła literackiego (pol)
язык автора (rus)	язык автора литературного произведения (rus)
author's language (eng)	язык автора художественного произведения (rus)
jezik autora (scr)	мова аўтара мастацкага твора (bel)
autorský jazyk (cze)	мова автора художнього твору(ukr)
мова автора (ukr)	

Equivalent linguistic units of different languages that perform the function of denotation, form the equivalence class of the term *language of mass media*. Among the national equivalents of the Polish term *język mediów* is Russian *язык масс-медиа*, Croatian *jezik medija*, Slovenian *jezik medijev*, instead there is no Ukrainian term *мова мас-медіа* ‘language of mass media’ / *мова ЗМІ* ‘language media’. And the reason is not the lack of sources in the database, but the lack of a term in the list of key words in the works` annotations.

Analysis of the key words recorded in iSybislaw provides a basis for constructing a scale of equivalence of terms by the nature and degree of divergence of their content and form, for example, equivalence classes *мова преси* ‘of the language press’ and *мова ЗМІ* ‘the language of the media’.

It is in terminology that the correspondence of form and content must be consistent and non-contradictory, this is required by the specifics of professional and scientific communication. Then it will be clear how the terms *implicature* / *implication*, *repetition* / *repeating*, *illocution* (equivalents in iSybislaw Pol. akt ilokucyjny, Rus. иллокутивный акт ‘illocutionary act’) / *иллокуция* ‘illocution’, Ukr *імплікатура* ‘implicature’ / *імплікація* ‘implication’, Rus. *речевая инерция* ‘speech inertia’ / *речевая интенция* ‘speech intension’, Rus. *мифическое имя* ‘mythical name’ / *мифоним* ‘mythonym’ / *мифологическое имя* ‘mythological name’, Ukr. *антропоцентризм* ‘anthropocentrism’ / Rus. *антропоцентричность языка* ‘anthropocentric language’, Rus. *бранная лексика* ‘swearing’ / *бранные формы* ‘swearing forms’, Rus. *вариантность* ‘variability’ / *вариативность* ‘variability’, Ukr. *варіант* ‘variant’ / *варіант літературної мови* ‘variant of literary language’, Rus. *вводная единица* ‘introductory unit’ / *вводное слово* ‘introductory word’ / *вводный элемент* ‘introductory element’, etc. After all, such complex semantic relationships between terms complicate the work of indexers.

The variability of terms is also a problem when designing a key word language. This is relevant for multilingual users and indexers, as they need access to an effective information retrieval tool. The creation of equivalence classes takes place within the units of the language of formal description, as well as the language of subject description, which reflects the content of the document. For example, in the formal language there are different variants of the author's name and surname: *Shevelov George Y.*, *Ševelev Jurij*, *Ševel'ov Jurij*, *Szerech Jurij*, *Шевелёв Юрий Владимирович*, *Шевельов Юрій*.

In the language of the subject description, variants within one language are given, e.g.: Pol. *językoznawstwo kognitywne* = *kognitywizm* = *lingwistyka kognitywna*; and identical in meaning in many languages: Rus. *когнитивизм* ‘cognitivism’ = *когнитивистика* ‘cognitivistics’ = *когнитивная лингвистика* ‘cognitive linguistics’, Eng. *cognitive linguistics*, Horv. *kognitivna lingvistika*, Czech. *kognitivní lingvistika*, Serbian *когнитивна лингвистика* (Fig. 2).

All terms belong to the same class, and each of them illustrates the same documents in iSybislaw. Therefore, the user receives the same search result regardless of the term specified in the search bar, because the search is performed using key words based on different natural languages.

Word	Language	Descriptor	Actions
językoznawstwo kognitywne	polski (pol)	<input checked="" type="radio"/>	[Show]
kognitywizm	polski (pol)	<input type="radio"/>	[Show]
lingwistyka kognitywna	polski (pol)	<input type="radio"/>	[Show]
КОГНИТИВИЗМ	rosyjski (rus)	<input type="radio"/>	[Show]
КОГНИТИВИСТИКА	rosyjski (rus)	<input type="radio"/>	[Show]
КОГНИТИВНАЯ ЛИНГВИСТИКА	rosyjski (rus)	<input type="radio"/>	[Show]
cognitive linguistics	angielski (eng)	<input type="radio"/>	[Show]
kognitivna lingvistika	chorwacki (scr)	<input type="radio"/>	[Show]
kognitivní lingvistika	czeski (cze)	<input type="radio"/>	[Show]
КОГНИТИВНА ЛИНГВИСТИКА	serbski (scc)	<input type="radio"/>	[Show]

[\[Show documents\]](#)

Figure 2: Class of equivalence of terms with the descriptor *językoznawstwo kognitywne*

Interpretation of the variant as a parallel form of existence of the term allows iSybislaw to take into account all the nominations of the same language concept in all available in the system functional and stylistic varieties of linguistic texts of different times. Of course, the first step in studying the variability of terms should be a careful analysis of the linguistic terminology of individual Slavic languages in a certain period of their functioning with their subsequent comparison in a broader time perspective and in a wider range of areas of life and clarify common and different in their composition, in particular availability, scope and nature of groups of terms-options to denote common concepts. According to Ye. A. Karpilovska, focusing the analysis of variant forms of terms on the purely practical need to search in the database of linguistic sources with certain information encourages the structuring of such a search image of the document by parameters 1) type of text, scope and 2) time of its publication. Taking into account these parameters makes it possible to limit the search for the scope of variant terms to denote the desired user concept [7].

A separate problem of key word selection and ordering is the ambiguity of terms, which represent different semantic relations, for example: *nomination1* – procedure and *nomination2* result (Figs. 3 and 4).

Descriptor:	nominacja 1 (procedura)
Language:	pol (polski)

Keyword versions

Word	Language	Descriptor	Actions
nominacja 1 (procedura)	polski (pol)	<input checked="" type="radio"/>	[Show]
akt nominacji	polski (pol)	<input type="radio"/>	[Show]
номинация 1	rosyjski (rus)	<input type="radio"/>	[Show]
pojmenování 1	czeski (cze)	<input type="radio"/>	[Show]
номінація 1	ukraiński (ukr)	<input type="radio"/>	[Show]

[\[Show documents\]](#)

Figure 3: Class of equivalence of terms with the descriptor *nominacja 1 (procedura)*

Keyword versions

Word	Language	Descriptor	Actions
nominacja 2 (rezultat)	polski (pol)	<input checked="" type="radio"/>	[Show]
nazwa	polski (pol)	<input type="radio"/>	[Show]
nominat	polski (pol)	<input type="radio"/>	[Show]
название	rosyjski (rus)	<input type="radio"/>	[Show]
наименование	rosyjski (rus)	<input type="radio"/>	[Show]
номинативная единица	rosyjski (rus)	<input type="radio"/>	[Show]
языковая номинация	rosyjski (rus)	<input type="radio"/>	[Show]
номінація 2	ukraiński (ukr)	<input type="radio"/>	[Show]

[\[Show documents\]](#)

Figure 4: Class of equivalence of terms with the descriptor *nominacja 2 (rezultat)*

Without solving the ambiguity problem, the user does not have access to relevant information within a single language.

Equivalence classes begin to construct from the collection of terms in one language, and then organize them, first of all eliminate ambiguity. This is very important, because without solving the problem of ambiguity, the user does not have access to relevant information within a single language.

4. Semantic relations between terms of language dynamics

The use of key words and equivalence classes in iSybislaw can be seen in terms of terms related to language dynamics, as language and time, factors and directions of language evolution, the relationship between language dynamics and social development are problems that scientists are interested in.

In the terminological system we record cases of discrepancy in the scope of concepts that are in the relationship of semantic hierarchy, subordination such as: Pol. *dynamika językowa* – Rus. *языковая динамика* ‘linguistic dynamics’ and descriptors of the Rus. *эволюция языка* ‘evolution of language’ and Pol. *ewolucja języka* – Rus. *развитие языка* ‘language development’ (three descriptors: Ukr. *розвиток мови* ‘language development’, Rus. *развитие языка*, Pol. *rozwój języka* and the Czech equivalent *vyvoj jazyka*). The evolution of language is a natural process of historical changes in the field of language activity. It is the changes in the language structure that encourage the study of dynamic processes in modern Slavic languages, as evidenced by the key word *мовна зміна* ‘language change’ with a Polish descriptor and four equivalents: Pol. *zmiana językowa* – Ukr. *мовна зміна* ‘language change’ – Rus. *языковое изменение* – Croatian *jezična promjena* – Czech. *jazyková změna*.

For the twentieth century characteristic opposition of the *новомови* ‘newspeak’ – the totalitarian *new language*, the weapon of enslavement and obedience and the new language as an expression of the new thinking of society. The primary source of the term *newspeak* is J. Orwell's novel "1984", where the new language is a means of serving Angsots (English socialism). The following equivalents of this Orwellian definition of the language official of the state are used in iSybislaw: Pol. *język totalitarny* – Rus. *тоталитарный язык* ‘totalitarian language’ – Czech. *jazyk totality*; Pol. *język propagandy* – Pol. *propaganda językowa*; Pol. *nowomowa* – Rus. *новояз* ‘novoyaz’ – Croat. *novogovor* – Czech *nový jazyk* and term Rus. *советизм* ‘sovietism’. There is no *renewal of language* as a component of modern lexicon in the species term database, but there is Pol. *postpierestrojka* – Rus. *постперестройка* ‘post-perestroika’, which denotes a social phenomenon.

It is interesting that in the language of iSybislaw key words there are terms *мовна тенденція* ‘language tendency’ with the Polish descriptor *tendencja językowa* and Russian and Ukrainian equivalents *языковая тенденция* – *мовна тенденція* ‘language tendency – tendencja językowa’ and the Ukrainian term *тенденція розвитку* ‘tendency of the development’ with the Polish descriptor *tendencja rozwojowa* and the equivalents Engl. *developmental tendency* – Czech *vyvojová tendence* – Verkhnioluzh. *wuwićowa tendency*. However, a review of documents for the key word *мовна тенденція* ‘language tendency’ shows that in some works it is about development trends, for example: Ohnheiser I. Purpose and concept of the volume // *Słowotwórstwo / Nominacja* (Opole, 2003); Nikolaev G. A. The main trends in the development of word formation in modern Russian language: on the material of dictionaries "New words and meanings" (Opole, 1997); Spiwak O. Something about trends in the development of vocabulary of the modern Ukrainian language (Studia Ukrainica, 2001); Neščimenko G. P. Word formation of colloquial language in the light of the tendency of language economy: on the material of Russian and Czech languages (Innsbruck, 2000) and others.

Important for the choice of the term are its formal and semantic features, in particular the aspect of the denoted concept that brings the terms-variants or distances them from each other; suitability of the term for word formation, which provides, as is well known, in Slavic languages structuring of the concept, and hence the establishment of semantic links between terms that affect the formation of linguistic terminological systems and relations between them.

One of the important problems of neology is to determine the criteria for separating neologisms from different types of innovations, as scientists put different meanings into the concept of neologism.

iSybislaw uses *неологізм* 'neologisms' and *інновація* 'innovations' that are similar in meaning to new lexical items. Borrowed in the XIX century in French language, the word *неологізм* 'neologisms' was the first to denote the category of new in the lexical-semantic system of language. In light of this concept, it was considered generic to denote all kinds of new vocabulary and increasingly in linguistics began to be replaced by the term *інновація* 'innovations'. This term is used to denote new phenomena at all levels of the language, as evidenced by the names in the key words of the iSybislaw system "*Innowacje w językach słowiańskich*", "*Innowacje frazeologiczne w powojennej fraszce polskiej*". Under the generic status of the term Pol. *innowacja językowa* and its equivalents Rus. *языковая инновация* 'language innovation' – Croatian *jezična inovacija* – Ukr. *інновація мови* 'language innovation' is subject to a number of species innovations belonging to certain levels of the language system: Pol. *innowacja fonetyczna* – Rus. *фонетическая инновация* 'phonetic innovation', Pol. *innowacja fonologiczna*, Pol. *innowacja frazeologiczna*, Pol. *innowacja leksykalna* – Rus. *лексическая инновация* 'lexical innovation', Pol. *innowacja morfologiczna*, Pol. *innowacja semantyczna*, Pol. *innowacja składniowa*, Pol. *innowacja słowotwórcza*, Pol. *innowacja stylistyczna*.

The Polish descriptor *neologizm* has more equivalents than *інновація* 'innovation' – it Rus. *neologizm* – Croatian *neologizam* – Chezh *neologismus* – Verkhholuzh. *neologism* – Slovenian *neologizmus* – Ukr. *неологізм* 'neologism', but fewer species: Pol. *neologizm frazeologiczny* – *neofrazeologizm*, Pol. *neologizm leksykalny*, *neologizm słowotwórczy*, *neologizm strukturalny*.

The problem of document description and choice of terms *інновація* 'innovation' / *неологізм* 'neologism' is directly related to the problem of terminological nature: it is necessary, firstly, to clarify their meaning; secondly, to define system relations, connections between the concepts that denote these terms, because in one row of key words simultaneously present *innowacja leksykalna* – *neologizm leksykalny*, *innowacja językowa* – *neologizm*, *innowacja leksykalna* – *neologizm leksykalny* (terms *неологізм* 'neologism' and *неолексема* 'neolexem' are absolute synonyms). In our opinion, sometimes the range of features involved in the classification of neologisms is so wide that the term *неологізм* 'neologism' is identified with the term *інновація* 'innovation'.

Add to this the terms from iSybislaw: Pol. descriptor *neosemantyzm* and equivalent *neologizm semantyczny* – Eng. *neosemanticism* Chezh *neosémantismus* – Slov. *neosémantizmus* – Ukr. *неосемантизм* 'neosemanticism'. The doublet is the term *innowacja semantyczna*. Derived meaning of a word or form, which arose on the basis of the main, nominative, in the "Encyclopedic Dictionary of References of Linguistic Terms and Concepts: Russian Language" is interpreted as *а вторичное значение* 'secondary meaning' (Russian descriptor *вторичное значение* 'secondary meaning'). The formation of a new meaning of the token is also called the *вторинною номінацією* 'secondary nomination' – in the key words of the iSybislaw system today there are Rus. *вторичная номинация* 'secondary nomination' – Rus. equivalent *повторная номинация* 're-nomination' – Bel. *другасная намінацыя*. A separate group is formed by semantic borrowings – Pol. *zapożyczenie semantyczne* – Eng. semantic borrowing. The prospect of developing the language of key words in general (classes of equivalence of multilingual terms) and its Ukrainian module in particular may be the addition of terms *чужосемантизм* 'foreign semantics' (Pol. *obcysemantyzm*) or *приховане запозичення* 'hidden borrowing' – new meanings of words that are the result from hidden borrowing and reveal the interaction of found in the secondary nomination on the basis of an already mastered foreign language unit (Pol. *ukryte zapożyczenie*), *англосемантизм* 'anglosemantism' (Pol. *anglosemantyzm*) – *neosemantizm* from the English language in its American version (in the key words of the iSybislaw system the term *neoanglicyzm* is already recorded).

The main processes in the vocabulary with different intensity took place in the language always, in all periods of its functioning, and this is evidenced in the language of key words, in particular: removal from the lexicon of individual words, their departure to periphery (Pol. descriptor *pasywizacja* – Rus. *пассивизация* 'passivization'); return to life of previously irrelevant tokens (Pol. descriptor and equivalent *aktualizacja wypowiedzi* = *aktualizacja* = Czech *aktualizace*); reevaluation of some categories of words, the formation of new connotations (Rus. descriptor *переоценка лексики* 'reevaluation of vocabulary'); dissemination of lexical groups of socially or professionally limited use (Pol. descriptors *żargonizacja*, *wulgaryzacja języka*). The last of these processes is related to the trend of democratization, which, unfortunately, is not recorded, although the database contains sources that indicate the need to introduce this term: Henryk Duda « *Język polski po komunizmie : nowotowa, demokratyzacja języka, zapożyczenia z języka angielskiego, wulgaryzacja* » (Poznań, 2001).

The dynamics of the language is also reflected in the time intervals that need to be unified in iSybislaw, because there are key variants: 20/21 w. = 20-21 w. = 21 w. (*początek*) = 21 w. = 20 w. (*druga połowa*) = 20 w. (*koniec*).

Researchers of language dynamics identify the following trends in language development: 1) internationalization (see currently available in the iSybislaw system key words-terms Pol. *internacjonalizacja* – Rus. *интернационализация* ‘internationalization’) – nationalization; 2) intellectualization (Pol. *intelektualizacja języka* – Rus. *интеллектуализация языка* ‘intellectualization of the language’) (“bookkeeping”) – democratization (“talk”, liberalization, substandardization, vulgarization, colloquialization), 3) economy of language means (Pol. *ekonomia językowa* – Czech *ekonomizace řeči*) – transparency and analytical (detailing) of the nomination. This phenomenon, which reflects the linguistic tendency to save means of expression, is known in the linguistic literature under various terms: compression, contraction, semantic condensation, unification or universalization, fusion. The presence of such terms-variants testifies to the complexity and multifaceted nature of the phenomenon of language economy. In iSybislaw from the terminological range to denote the manifestations of the economy of language means recorded universion with the Polish descriptor *uniwersalizacja* and equivalents Rus. *универбація* ‘universalization’ – Czech *univerbace* – Slov. *univerbizácia* – three Slovenian equivalents *poenobesedenje* – *univerbacija* – *univerbizacija* and two Ukrainian doublets *універбація* ‘university’ – *універбізація* ‘universalization’, as well as *усічення* ‘truncation’ with the Pol. descriptor *ucięcie* – Polish equivalent *dezintegracja* – Rus. *усечение* ‘truncation’.

In order to solve the problem of variance of terms, it is necessary to find out the degree of similarity of the content of variants of terms. To determine the dominance in the series of monolingual and multilingual terms, the quantitative and qualitative properties of their functioning in the texts of the iSybislaw database, as well as the degree of closeness of national terms to Polish as the basis of information retrieval language iSybislaw. The formation of groups of term terms allows you to optimize the search in the database without changing the structure of its information retrieval language key words. In addition, the completeness of such groups of variants of the same terms and their synonyms is important as an indicator of the amplitude of the change in the designations of certain concepts of modern Ukrainian linguistics. Coordination of the key word index in a certain language with the texts of the iSybislaw source database makes it possible to arrange the terms in such variant groups according to their decreasing functional weight in the national linguistic tradition of the period under study, taking into account their time perspective.

Terminology of language dynamics is developed on the basis of constant variable empirical material (new types of innovations appear), so it is constantly replenished with new elements of metalanguage and is still in a state of formation and active discussion. Terminological disorder is the result of different understandings of the main categories of neology, so there is an obvious need to organize and unify the terms of neology as an independent linguistic discipline and present it in the language of key words iSybislaw in a form consistent with modern practice. In particular, it is necessary to take into account the semantic relationships between terms that indicate an in-depth semanticization of certain concepts of neology.

5. Paradigmatic relationships between the terms

Other types of paradigm are also established between elements of search language, such as generalization / concretization, logical and / or formal derivation, hypero-hyponymy, and so on. According to hyper-hyponymic connections, i.e. different aspects and the degree of the content's generalization of the same concept, we combine the common root nominations *перфект* ‘perfect’ / *перфектив* ‘perfective’ / *перфектність* ‘perfection’ / *перфективність* ‘perfectiveness’ / *перфективація* ‘perfectivation’.

The terms proposed in linguistic works to denote these phenomena can be schematically depicted as follows:

- *перфект* ‘perfect’ (descriptor *perfectum*) – species-temporal form (V. M. Rusanivskiy);
- *перфектив* ‘perfective’ (descriptor *czasownik perfektywny*) – perfect form (Yu. S. Maslov); correlative in type *перфективи* ‘perfectives’ – *імперфективи* ‘imperfectives’ (M. I. Kalko);

- *перфектність* ‘perfection’ – *перфект* ‘perfect’ (G. O. Zolotova); *перфектний феномен* ‘perfect phenomenon’ (Niels Thelin);
- *перфективність* ‘perfectiveness’ (descriptor *perfektywność*) – the ability of verbs to express the effect;
- *перфективізація* ‘perfectivation’ (descriptor *perfektywizacja*) – the creation of forms of perfect form (Fig. 5).

[Show languages filter]

Keyword	Language	Actions
перфекат	serbski (scc)	[Show]
перфект	rosyjski (rus)	[Show]
перфектив	rosyjski (rus)	[Show]
перфективізація	rosyjski (rus)	[Show]
перфективізація	ukraiński (ukr)	[Show]
перфективізація	serbski (scc)	[Show]
перфективність	rosyjski (rus)	[Show]
перфектное значение	rosyjski (rus)	[Show]

Figure 5: Paradigmatic relationships between the terms

According to hypero-hyponymic relations, we analyze the terms *каузація* ‘causation’ / *каузальність* ‘causality’ / *каузативність* ‘causativity’. The fact that this is an exclusively verbal category is evidenced by the titles of documents in iSybislaw, related to the analysis of verbs in Slavic languages, in particular Belarusian terms *казуативно-маніпулятивні глаголи* ‘casual-manipulative verbs’, *каўзатыўныя дзеясловы са значэннем 'абараняць' у беларускай мове*, ‘causative verbs meaning 'defend' in Belarusian’. At the same time, we note that *каузальність* ‘causality’ (descriptor *przyczynowość*), equivalents Rus. *каузальность* ‘causality’, *причинность* ‘cause’, Czech *kauzalita* – a semantic category that indicates an event or phenomenon that gives rise to another event or phenomenon; *казуативність* ‘casuativity’ (descriptor *kauzatywność*) – realization of causal relations (G. O. Zolotova); *каузація* ‘causation’ – *каузальність* ‘causality’ + *каузативність* ‘causativity’ (G. Glushchuk). However, we record documents where the key word is Pol. *kauzatywność* or Rus. *каузативность* ‘causativity’, and there must be *causal*, eg. : "*Положительная и отрицательная каузальность в русском и польском языках*" ("Positive and negative causality in Russian and Polish").

We will also pay attention to the quality of information retrieval language, which is manifested in the functioning of alternative innovations to the already existing terms of words. Some of them are new or relatively new terms that have arisen together with new phenomena and problems or as a result of a new, different view of phenomena. For example, the increased attention of scientists to language dynamics has become a stimulus for the active development of certain linguistic disciplines, as evidenced by the relevant key words-linguistic terms in the language of key words iSybislaw, namely: neology – sciences of neologisms (Pol. *neologia* – Czech *neologie* – Ukr. *неологія* ‘neology’) *neographies* – the science of the theory and practice of describing neologisms in dictionaries (Pol. *neografia* – Rus. *неография* ‘neography’ – Czech *neografie* – Verkhnluzh. *neografija górnołużycki*) and *neophrasemics (branch)* (Ukr. descriptor). A significant group of neologisms are terms that refer to the sections of linguistics and scientific theories, for example: Pol. *lingwistyka genderowa* and *lingwistyka płci*, Czech *genderová lingvistika*, Rus. *гендерная лингвистика* ‘gender linguistics’; Pol. *lingwistyka korpusowa*, Czech *korpusová lingvistika*, Rus. *корпусная лингвистика* ‘corpus linguistics’; Pol. *lingwistyka kulturowa*, Czech *kulturní lingvistika*, Rus. *лингвокультурология* ‘linguoculturology’. Slavic linguistic terminology responds to the expansion of linguistic boundaries and indexers, in turn, capture new key words in iSybislaw.

6. Conclusions and Further Research

iSybislaw is a relevant intelligent search system, constantly improving, and therefore increasingly satisfying user requests. As rightly noted by O. V. Dobrov, the higher value of the parameter of search

accuracy and rubrication, the less information noise, i.e. irrelevant results; the higher the value of the search completeness parameter, the less inaccessible information for the user [17].

The question of the scope of linguistic terms is related to the principles of building the source base of the iSybislaw information search system, in particular to the criteria for entering texts of certain styles, genres and types or to the chronological boundaries of the publication of such texts. For example, iSybislaw's database of world Slavic linguistics includes texts published from the early 1990s to the present day. Its source is made of scientific works, so outside this database are educational and methodological and popular science publications. However, even clearly defined temporal and functional-stylistic boundaries of the iSybislaw database sources inevitably expand. This is evidenced, for example, by the Ukrainian scientific practice of the early XXI century, journal publications, monographs and collections of terminological conferences. This expansion of the modern linguistic terminology system is due to the fact that Ukrainian scientific publications of the independence period actively discuss the problems of revival of national linguistic terminology, sources, resources and trends in its development, the feasibility of returning specific terms to.

In general, a thorough linguistic analysis of the most common models of terminological units, ways of their semantic differentiation in the iSybislaw system is required. Solving the problems of ambiguity, avoiding synonymy and variability of terms in information retrieval language involves finding out the degree of similarity of terms, the activity of their use in language practice and in iSybislaw texts, which in turn will ensure effective search in the system.

In addition, the search for a key word in iSybislaw not only represents the state of elaboration of the linguistic problem in Slavic studies, but also serves as a scientometric basis on which to establish the citation index of a particular linguist. To date, there have been no such information retrieval databases with powerful Slavic language modules for the humanities. Thus, the evolution of the bibliography has contributed to a more effective implementation of information needs of the user [9].

The prospect for iSybislaw developers is the creation of the Bank of Slavic Linguistic Terminology and on its basis – a new thesaurus of Slavic linguistic terminology in the digital space. Working on such a terminological bank requires from its developers not only linguistic competence, but also mastery of modern technologies for processing language information. Carrying out a joint project with the developers of the iSybislaw system – a team of scientists from the Institute of Slavic Studies of the Polish Academy of Sciences, Ukrainian linguists are delighted to be at the center of the latest concepts and research methods of modern Slavic linguistics.

7. Acknowledgements

We express our sincere gratitude to the academic institutions – the Institute of Ukrainian Language of the National Academy of Sciences of Ukraine, the Institute of Slavic Studies of the Polish Academy of Sciences – for the opportunity to be one of the developers of the Ukrainian module iSybislaw. Special words of gratitude to the head of the Department of Lexicology, Lexicography and Structural and Mathematical Linguistics of the Institute of Ukrainian Language of the National Academy of Sciences of Ukraine, Doctor of Philology, Professor Ye. A. Karpilovska, Dr. P. Kovalskyi, Dr. Z. E. Rudnyk-Karvatova, Z. Mikos for the importance of developing interstate scientific cooperation.

8. References

- [1] Bibliograficzna baza danych światowego językoznawstwa slawistycznego iSybislaw, 2022. URL: <http://www.isybislaw.ispan.waw.pl>
- [2] Volkova E. Problemy belorusskoy lingvisticheskoy terminologii i vopros sootvetstviya klyuchevykh slov (lingvisticheskikh terminov) v iSybislaw [Problems of Belarusian linguistic terminology and the question of the matching of key words (linguistic terms) in iSybislaw]. *Slavia Orientalis* 62 (2013) 285–292.
- [3] O. A. Ostapchuk, M. A. Alekseeva. Mnogokomponentnyye terminy v funktsii klyuchevykh slov v informatsionnoy baze slavyanskogo yazykoznaniya iSybislaw [Multicomponent terms in the function of key words in the information base of Slavic linguistics iSybislaw], *Studia z*

- Filologii Polskiej i Slowianskiej 49 (2014) 164–175. doi:10.11649/sfps.2014.016
- [4] P. Kovalsky, I. Banasyak. Klyuchovi slova ta klasy ekvivalentnosti v systemi iSybislaw yak znaryaddya ta predmet doslidzhen' [Key words and equivalence classes in the iSybislaw information-retrieval implement and subject of research]. *Ukrayins'ka mova* [Ukrainian language], Vol. 4 (2017) 14–26.
- [5] S. Ya. Yermolenko, *Narysy z ukrayins'koyi slovesnosti: Stylistyka ta kul'tura movy* [Essays on Ukrainian Literature: stylistics and culture of language], Naukova Dumka, Kyiv, 1999.
- [6] Yu. V. Romaniuk. *Ukrayins'ka avtors'ka terminolohiya v systemii Sybislaw (diyeslovo)* [Ukrainian author's terminology in the Sybislaw (the verb)], *Studia z Filologii Polskiej i Slowianskiej* 49 (2014) 245–258. doi: 10.11649/sfps.2014.022
- [7] Ye. A. Karpilovska, *Varianty terminiv movoznavstva u suchasniy naukovo-osvitniy praktytsi i v chasoviy perspektyvi* [Variants of terms of linguistics in modern scientific and educational practice and in time perspective], in: Zofia Rudnik-Karwatowa (Ed.), *Dynamika wspolczesnego slownictwa slowianskiego w przestrzeni stylowo-funkcjonalnej*, *Prace Slawistyczne – Slavica* 142, Warszawa, 2017, pp. 9–26.
- [8] Ye. A. Karpilovska. *Rol' natsional'nykh moduliv u rozvytku informatsiyno-poshukovoyi systemy* [The role of national modules in the development of information retrieval language of the iSybislaw database], *Ukrayins'ka mova* [Ukrainian language], Vol. 1 (2018) 17–29.
- [9] A. A. Taran. *Information-retrieval system «Base of the world slavic linguistics (iSybislaw)»*, in: *Computational Linguistics and Intelligent Systems. Proc. 4th Int. Conf. COLINS 2020. Vol. I* (2020), pp. 590–599.
- [10] Z. Rudnik-Karwatowa. *Uporyadkuvannya terminolohiyi dlya potreb informatsiyno-poshukovoyi systemy iSybislaw* [Arranging terminology for the needs of information retrieval system iSybislaw]. *Ukrayins'ka mova* [Ukrainian language], Vol. 3 (2016) 105–107.
- [11] Z. Rudnik-Karwatowa, H. Karpińska. *Słownik słów kluczowych językoznawstwa slawistycznego*. *Slawistyczny Ośrodek Wydawniczy*, Warszawa, 2006.
- [12] Jakub Banasiak, Marcin Fastyn and Paweł Kowalski, *Optymalizacja prac naukowych pod kątem wyszukiwarek i pozycjonowania: wprowadzenie do problematyki i wstępne zalecenia dla autorów*, in: in: Paweł Kowalski (Ed.), *Z zagadnień informacji naukowej, terminoznawstwa i językoznawstwa*, Instytut Slawistyki Polskiej Akademii Nauk, Fundacja Slawistyczna, Warszawa, 2018, pp. 37–53.
- [13] Ana Golubović. *Linguistic Bibliographic Database iSybislaw*, *Infotheca*, Vol. 20 №1–2 (2020) 149–159.
- [14] Jadwiga Sadowska, *Wspolczesne modele dziedzinowych bibliograficznych systemow informacyjno- wyszukiwawczych*, in: Paweł Kowalski (Ed.), *Z zagadnień informacji naukowej, terminoznawstwa i językoznawstwa*, Instytut Slawistyki Polskiej Akademii Nauk, Fundacja Slawistyczna, Warszawa, 2018, pp. 275–285.
- [15] V. G. Kulpina, V. A. Tatarinov. *Sinonimiya klyuchevykh slov v sisteme slavisticheskoy bibliografii kak epistemologicheskiye refleksi razvitiya lingvisticheskoy terminologii* [Synonyms of key words in the system of Slavic bibliography as epistemological reflexes of development of linguistic terminology], *Studia z Filologii Polskiej i Slowianskiej* 49 (2014) 188–202. doi:10.11649/sfps.2014.018
- [16] Ана Голубовић. *Библиографска база iSybislaw и Међународна научна конференциј „Језыкознaвство а systemy slawistycznej informacji bibliograficznej – dzis i jutro”*. *Јужнословенск филолог*, Vol. 75 № 1 (2019) 162–171.
- [17] A. V. Dobrov, *Tekhnologii intellektual'nogo poiska i sposoby otsenki ikh yeffektivnosti* [Intelligent Search Technologies and Methods for Evaluating their Effectiveness], in: A. S. Gerd (Ed.), *Структурная и прикладная лингвистика* [Structural and Applied Linguistics], St. Petersburg, 2010, pp. 219–232.