
УДК 047.31+811.6

Зоф'я Ева Рудник-Карватова

Інститут славістики Польської Академії наук, м. Варшава

УПОРЯДКУВАННЯ ТЕРМІНОЛОГІЇ ДЛЯ ПОТРЕБ ІНФОРМАЦІЙНО-ПОШУКОВОЇ СИСТЕМИ ISYBISLAW¹

Використання мови ключових слів в інформаційно-пошуковій системі (ІПС) слов'янського мовознавства iSybislaw (див. Bibliograficzna baza danych światowego językoznawstwa slawistycznego, www.isybislaw.ispan.waw.pl) свідчить про надання мовознавчим термінам функції ключових слів, які унаслідок зміст документів, розміщених в інформаційній базі системи. Такий підхід є зручним для користувачів, однак створює низку проблем, не бажаних для ІПС, наприклад, явища синонімії та багатозначності термінів, типові для природної мови.

У багатомовній системі iSybislaw реалізовано оригінальну методологію, що полягає у формуванні рядів еквівалентів термінів, уживаних нині в лінгвістичній літературі певної мови, а також у поєднанні таких рядів у різномовні класи еквівалентів. Така концепція має забезпечити ефективність багатомовної системи для задоволення інформаційних потреб різномовних користувачів. У статті розглянуто як теоретичні проблеми, так і досвід праці над упорядкуванням термінології для системи iSybislaw.

Ключові слова: термінологія, упорядкування термінології, мовознавча термінологія, інформаційно-пошукова система, iSybislaw, ефективність інформаційно-пошукової системи

Інформаційно-пошукова система слов'янського мовознавства становить ініціативу польського славістичного середовища². Реалізація проекту від початку здійснюється у співпраці

¹ Статтю підготовлено в рамках міжнародного наукового проекту «Онлайн-бібліографія світового славістичного мовознавства», який Інститут славістики Польської Академії наук з 2015 р. виконує разом з Інститутом української мови НАН України за Угодою про наукове співробітництво між ПАН і НАНУ. З польської мови текст переклала Є.А. Карпіловська.

² Концепцію системи славістичної інформації було представлено на XII Міжнародному з'їзді славістів у Кракові 1998 р. [7: 41—47].

з мовознавцями як з польських, так і зарубіжних наукових осередків. На сьогодні колектив співпрацівників складають представники Білорусі, Естонії, Македонії, Німеччини, Польщі, Росії, Словаччини, України, Хорватії, Чехії³. Завдяки їх зусиллям і компетенції успішно розвивається інформаційно-пошукова система, представлена в базі даних iSybislaw [1]. На нинішньому етапі роботи ця система славистичної інформації забезпечує користувачам відкритий і безкоштовний доступ до необхідних відомостей, які містяться в понад 21 тис. різномовних наукових праць, а також до кількох сотень повних текстів, опублікованих у період з 1992 р. до сьогодні. Інформаційна база системи постійно поповнюється, розвиваються також функціональні можливості системи (див. рис. 1 і 2).

Терміни у функції ключових слів. У системі славистичної інформації iSybislaw інформаційно-пошуковою мовою є мова ключових слів. Функцію ключових слів, що характеризують зміст документів, які уміщено в інформаційній базі системи і які становлять інструмент пошуку інформації, передусім виконують терміни. Це різномовні терміноодиниці, що відображають теми документів, показових для сучасного слов'янського мовознавства (монографій, збірників, часописів, статей, рецензій і т. ін.). Крім термінів слов'янського мовознавства, які визначають тематичний діапазон системи iSybislaw, у вигляді ключових слів подано й терміни зі сфери теорії мови та загального мовознавства, а також з контрастивного слов'янсько-неслов'янського мовознавства. У значно меншому обсязі в інформаційній системі як ключові слова використано інші одиниці природних мов, що окреслюють тематичні поля, важливі для організації та пошуку інформації у базі репрезентативних документів зі слов'янського мовознавства. Це власні назви: назви осіб (напр., прізвища мовознавців, праці яких є предметом досліджень, описуваних у документах), географічні назви (напр., назви місцевостей та регіонів, де здійснювано дослідження мовних явищ, описуваних у документах, місце і рік конференції, тематику якої реферовано) і збірні назви (напр., назви навчальних закладів і наукових товариств), а також загальні назви: загальномовні одиниці, що становлять предмет досліджень (лексеми, фразеологізми, морфеми і т. ін.), найменування, що мають високу частоту в базі документів (напр., у польській мові *nazwy barw* «назви кольорів», *nazwy części ciała* «назви частин тіла», назви наукових зібрань: *kongres* «з'їзд, конгрес», *konferencja* «конференція») та інші.

³ Проект офіційно пов'язаний з Інститутом славістики ПАН. Його реалізують співпрацівники з відомих польських і закордонних славистичних осередків: Вроцлавського університету, Університету ім. Адама Міцкевича в Познані, Університету Марії Склодовської-Кюрі в Любліні, Інституту чеської мови АН Чеської Республіки в Празі, Інституту слов'янознавства РАН у Москві, Московського державного університету ім. М.В. Ломоносова, Інституту української мови НАН України в Києві, Центру досліджень білоруської культури, мови та літератури НАН Білорусі у Мінську, Лужицького інституту в Будишині, Центру досліджень з ареальної лінгвістики Академії наук і мистецтв у Скоп'є та інших. Значну частину колективу становлять члени Комісії лінгвістичної бібліографії, яка працює при Міжнародному комітеті славістів.

iSybislaw
Bibliograficzna baza danych światowego językoznawstwa slawistycznego

SZUKAJ W BAZIE
POMOC
O BAZIE
HISTORIA BIBLIOGRAFII
UCZESTNICY
INSTYTUCJE PARTNERSKIE
CENTRUM SLAWISTYCZNEJ
INFORMACJI NAUKOWEJ
PUBLIKACJE
POLECANE SERWISY
MAPA STRONY
KONTAKT

Bibliograficzna baza danych prezentuje **nowoczesny system informacyjno-wyszukiwawczy iSybislaw**, którego koncepcję przedstawiono na Międzynarodowym Kongresie Sławistów w Krakowie w 1998 r. Projekt – instytucjonalnie związany z Instytutem Sławistyki Polskiej Akademii Nauk – jest rozwijany we współpracy z polskimi i zagranicznymi placówkami slawistycznymi przy wsparciu organizacyjnym Fundacji Sławistycznej; we wcześniejszym okresie projekt korzystał przez wiele lat ze wsparcia organizacyjnego Towarzystwa Naukowego Warszawskiego. Jego efekt stanowi ogólnodostępny system informacji bibliograficznej językoznawstwa slawistycznego zawierający obecnie zbiór informacyjny dotyczący ponad 21000 dokumentów. Dostęp do zasobów bazy: opisów bibliograficznych, a także poprzez linki do pełnych tekstów w dedykowanych archiwach, jest otwarty i bezpłatny na licencji CC BY-NC-ND.

Zakres tematyczny
Zakres tematyczny systemu iSybislaw obejmuje nie tylko **językoznawstwo slawistyczne**, **językoznawstwo konfrontatywne** i **językoznawstwo kontrastywne słowiańsko-niesłowiańskie**, lecz także w mniejszym stopniu **językoznawstwo teoretyczne** i **ogólne**, ważne dla slawistyki językoznawczej. System kieruje do różnego typu dokumentów, w których zawarte są informacje relewantne dla użytkowników: do wydawnictw zwartych (monografi, prac zbiorowych, słowników itp.), wydawnictw ciągłych (czasopism) i niesamoistnych (artykułów, recenzji i in.).

Zasięg geograficzny i językowy
W systemie iSybislaw dostępne są **informacje o polskim i światowym językoznawstwie slawistycznym**. **Zasięg językowy systemu jest nieograniczony**, wyznacza go język uwzględnionych w systemie dokumentów. Dzięki zastosowaniu innowacyjnego języka informacyjno-wyszukiwawczego możliwe jest wielojęzyczne wyszukiwanie w różnorodnym zbiorze informacyjno-wyszukiwawczym.

Zasięg chronologiczny
System informacyjno-wyszukiwawczy iSybislaw prezentuje dorobek naukowy językoznawstwa slawistycznego w okresie **od 1992 r. do dnia dzisiejszego**.

Rejestrowane dane
Opisy bibliograficzne dokumentów są opracowywane z autopsji przez polskich i zagranicznych językoznawców slawistów, a następnie weryfikowane i redagowane w Centrum Sławistycznej Informacji Naukowej Instytutu Sławistyki PAN. Odpowiednia strukturalizacja zbioru wyszukiwawczego umożliwia użytkownikom szybkie i łatwe dotarcie do potrzebnych informacji. W systemie iSybislaw rozbudowana jest struktura informacyjna rekordów (formalna i rzeczowa), uwzględniająca m.in. **słowa kluczowe** i **klasyfikację**, a także **abstrakty** i **adnotacje** (czasem w dwóch językach: języku obcym i polskim). Korzystanie z zasobów informacyjnych systemu iSybislaw ułatwia **słownik słów kluczowych z zakresu językoznawstwa slawistycznego**.

Рис. 1. Стартова сторінка бази iSybislaw

Serwis iSybislaw korzysta z plików cookie. Pozostając na tej stronie, wyrażasz zgodę na korzystanie z plików cookie. Akceptuję

iSybislaw Instytut Sławistyki Polskiej Akademii Nauk

Użytkownik: Hasło:

Polski English Русский

Start | Osoby | Dokumenty | Serie | Słowa kluczowe | Klasyfikacja | Tagi | Pomoc

Witamy w bazie iSybislaw

Bibliograficzna baza danych światowego językoznawstwa slawistycznego **iSybislaw**

Funkcje wyszukiwawcze w bazie

Indeksy - umożliwiają przeglądanie dokumentów uporządkowanych alfabetycznie w kartotekach: osób (autorów, redaktorów itp.), tytułów dokumentów, serii, słów kluczowych i klasyfikacji. W każdej z kartotek można przeprowadzać wyszukiwanie, wpisując do wyznaczonego pola sekwencję kilku pierwszych liter (znaków) wyrażen wyszukiwawczych.

Szukanie proste - funkcja ta pozwala na wyszukiwanie według zbiorów słów z pól: tytuł, odpowiedzialność, słowa kluczowe, abstrakt, przedmiotowe hasło osobowe oraz ze wszystkich tych pól. W opcji tej można maskować wyrażenia wyszukiwawcze za pomocą znaku * (gwiazdka) oraz łączyć je we frazy za pomocą operatorów logicznych: AND, OR, NOT. W bazie zastosowano również operator ~ (tylda) wyszukiujący wyrażenia podobne.

Szukanie zaawansowane - umożliwia przeszukiwanie bazy za pomocą zbiorów słów z wielu pól równocześnie oraz zawężanie wyszukiwania do typu oraz języka dokumentu. W funkcji tej można również używać operatorów logicznych oraz maskowania przy formułowaniu zapytań.

Użytkownik ma również możliwość zapisania wybranych dokumentów w tzw. koszyku, z którego można wykonać wydruki.

Рис. 2. Головне меню бази iSybislaw з опціями пошуку інформації

Уживання термінів та інших одиниць природних мов (слов'янських мов та англійської мови) як однопорядкових ключових слів в інформаційно-пошуковій мові ефективно для користувачів з огляду на їх мовну компетенцію (однак слід підкреслити, що вправне використання мовознавчої термінології вимагає компетенції не лише мовної, а й мовознавчої). Такий підхід уможливує для користувачів самостійний і зручний пошук чи оцінку важливості знайденої інформації. Застосування в системі iSybislaw мови з одиницями та структурою природної (або подібної до природної) зручне також для індексаторів-розробників системи, оскільки полегшує їм підготовку інформації.

Добір термінів. Для добору термінології використано корпус текстів, показових для сьогочасної літератури предмета, бібліографічних описів таких текстів. Створюваний фонд спеціальної лексики доповнено за авторитетними академічними граматиками, мовознавчими енциклопедіями, тезаурусами та словниками⁴. Лавиноподібне зростання кількості публікацій в галузі слов'янського мовознавства протягом останніх 20 років унеможливує подання в інформаційній базі всієї світової літератури з цієї галузі знання, виданої з 1992 р. Добір документів, інформацію про які вміщено в системі, залежить від того, як ті, хто їх опрацьовує, визначають потреби користувачів, а також від їхнього уявлення про обсяг інформації, необхідної для знання сучасного стану певної мовознавчої дисципліни.

На нинішньому етапі опрацювання термінології реалізовано засаду добору термінологічних одиниць згідно з прийнятими в проєкті критеріями, які окреслюють тематичний діапазон системи. В основному це обсяг знання зі слов'янського мовознавства, до якого входять також контрастивні слов'янсько-неслов'янські дослідження або праці із загальних теоретичних проблем, ілюстровані слов'янським мовним матеріалом.

Термінологія з окремих галузей мовознавства не розбудовується в системі однаковою мірою. Вона розростається разом із розвитком певних дисциплін і популярністю порушуваних у них проблем. Найбільша кількість термінів дотична синтаксису, морфології, словотворення, найменша — морфонології і морфотактики. До структури створюваної інформаційно-пошукової мови введено не лише терміни, що мають значне поширення, а й терміни, показові для певних шкіл і методів дослідження, а також терміни з низькою частотою уживання, важливі для певної галузі знання чи уживані з огляду на зацікавлення користувача. Терміни становлять відкриту сукупність, що передбачає її розбудову для потреб мови системи згідно зі ступенем деталізації (глибини) і широти потрібної інформації. Така методика має, з одного боку, запобігти втратам інформації, а з другого, — дати можливість уникнути інформаційного «шуму». Як бачимо, процедури, спрямовані на впорядкування термінології, мають широкий і багатогранний характер. Знайти оптимальні окремі розв'язки — завдання непросте. Узагальнення інформації не пе-

⁴ Бібліографічні відомості про матеріали джерел доступні в базі iSybislaw [1].

редбачене, оскільки користувач має отримати інформацію на певному необхідному рівні глибини її пошуку. Ефективною щодо цього видається, принаймні, часткова розбудова парадигматики мови ключових слів шляхом подавання під час індексації документів як терміна з конкретнішим значенням, підпорядкованого іншому в ієрархічній структурі, так і терміна з загальнішим значенням, вищого в ієрархічній організації мови, напр.: *słownik terminologiczny* «термінологічний словник» — *słownik 1* «словник 1» ‘сукупність слів, упорядкованих за абеткою’, *słownik dwujęzyczny* «двомовний словник» — *słownik 1, zoonim* «зоонім» — *nazwa własna* «власна назва», *chrematonim* «хрематонім» — *nazwa własna, derywacja paradygmaticzna* «парадигматична деривація» — *derywacja słowotwórcza* «словотвірна деривація», *derywacja ujemna* «від’ємна деривація» — *derywacja słowotwórcza, derywacja afiksalna* «афіксальна деривація» — *derywacja słowotwórcza, derywacja sufiksalna* «суфіксальна деривація» — *derywacja afiksalna*. Процедуру впорядкування термінології з огляду на тематичну сферу і релевантність буде застосовано до будови інформаційно-пошукової мови в повному її обсязі.

Багатомовність інформаційно-пошукової бази. Згідно з прийнятою концепцією у плані вираження інформаційно-пошукової мови вживаються всі слов’янські й англійська мови. Немає також обмежень в доступності відомостей про документи зі слов’янського мовознавства, опубліковані іншими мовами. Щодо документів неслов’янськими мовами (крім англійської), формальний опис відтворює мову оригіналу документа, а у власне описі (описі змісту) в інформаційно-пошуковій мові (пошуковій характеристиці документа) використовується одна зі слов’янських мов чи англійська мова, одиниці яких трактовано як ключові слова. Така процедура уможлиблює включення до системи iSybislaw інформації про публікації, дотичні слов’янського мовознавства з неслов’янськомовного обширу.

База документів, унааявлена в системі, різномовна, отже, слід було розв’язати проблему доступу до багатомовного зібрання інформації. Основним завданням було розроблення ефективної інформаційно-пошукової мови, яка б уможлиблювала місткий і багатомовний пошук інформації.

Класи еквівалентності термінів. При проектуванні основного інформаційного знаряддя системи, яким є інформаційно-пошукова мова, проблему становила не лише різномовність сукупності документів, поданих у базі системи, а й різномовність користувачів та індексаторів. У багатомовній системі iSybislaw застосовано оригінальний методологічний підхід, ґрунтований на встановленні класів еквівалентності мовних одиниць як у межах мови формального опису (напр., різні форми імені та прізвища того самого автора в записі різними абетками і в різних варіантах тією ж абеткою), так і у мові власне опису (опису змісту) документів (напр., синонімічні терміни в одній мові і рівнозначні в різних мовах і різних абетках), а також на відповідному рівні структурування засобів пошуку.

Як уже згадувалося вище, в інформаційних процесах терміни застосовують у функції ключових слів. Інакше кажучи, у процесі опрацювання інформації для системи ключовими словами є терміни, вжиті для опису змісту документів з метою їхнього використання у процесі пошуку [5: 77]. Приймаючи для багатомовної системи концепцію, ґрунтовану на встановленні **класів еквівалентності в структурі створюваної мови ключових слів**, в які об'єднуються пари або ряд рівнопорядкових термінів у слов'янських та англійській мовах, забезпечуємо користувачам доступ через будь-яке позначення з певного класу до всіх різномовних документів, що містять необхідну інформацію. Зокрема, користувачів, які знають одну або дві слов'янські мови і не знають англійської, система спрямує до всіх різномовних документів, у яких міститься потрібна інформація. І навпаки, англomовні користувачі, які знають лише одну слов'янську мову, зможуть доступитися до інформації за допомогою терміна, вибраного з усього класу еквівалентів.

Під час опрацювання інформації для системи термінологія не розбудовується однаковою мірою в окремих мовах. Найбільше просунуто працю над термінологією в польській мові, ініційовану 20 років тому польською славістичною спільнотою. Її результатом став «Словник ключових слів слов'янського мовознавства» («Słownik słów kluczowych językoznawstwa slawistycznego»), оснований на мовознавчій термінології у польській мові, який налічує близько 2500 одиниць [15]. Цей словниковий запас актуалізовано 2006 р. в електронній версії словника. Успішно накопичується та опрацьовується й термінологія з інших слов'янських мов, передусім російської, чеської і хорватської. До класів еквівалентності включено також терміни з праць українською та словацькою мовами, розпочато введення термінів з кількох інших слов'янських мов.

Запропонований інструмент багатомовного доступу та пошуку інформації в інформаційній базі має велику цінність для користувачів. Стратегія пошуку з використанням класів еквівалентності забезпечить повноту одержання необхідної інформації в усіх документах, представлених у системі.

Побудова класів еквівалентності є викликом для розробників системи і всього міжнародного колективу партнерів і співпрацівників. Для реалізації цього завдання необхідна мовна компетенція і належна мовознавча підготовка кожного з членів колективу. Перед цим, на етапі підготовки термінів для включення до класів еквівалентності, вони мають усувати брак точності, неправильність та змістові помилки в термінології, вживаній у працях, представлених у системі. Наприклад, у польській мовознавчій літературі зазвичай не вживають термін *skrót* «скорочення» і відповідне до нього нетермінологічне визначення. В одній зі статей, присвяченій словотворенню назв осіб, автор обговорює три групи формацій: *derywaty sufiksalne* «суфіксальні деривати», *skróty* «скорочення» і *substantywaty* «субстантивати». У розділі праці під заголовком «Skróty» («Скорочення») представлені «скорочені назви осіб. Маємо тут на увазі формації,

утворені від однослівних словотвірних основ» [6: 48] та приклади, що їх ілюструють, серед інших *полит* 'політичний в'язень', *нарком* 'крупний постачальник наркотиків' (пор. *нарком* — *народний комісар*), *юги* 'громадяни Югославії', *зек* ← *зэк* ← *зэка* ← *з/к* ← *заключенный*, *алик* 'алкоголік', *диссик* 'дисидент', *комсюк* 'комсомолец', *летер* 'лейтенант', *милок*, *милютик*, *мильтон* 'міліціонер', *наркоша* 'наркоман' та ін. [Там само: 48-49]. Проблема в тому, що подані приклади не становлять однорідної групи дериватів. Поряд із дериватами з усіченими формантами (від'ємними дериватами) виступають деривати з усіченням основи і додатним суфіксом, деривати з усіченням основи і зміною парадигми, контаміновані деривати, відомі в польській термінології як *derywaty z formantami ujemnymi* «деривати з від'ємними формантами», *derywaty z formantami ujemno-sufiksalnymi* «деривати з від'ємно-суфіксальними формантами» [11], інакше *derywaty z formantami ujemno-dodatnimi* «деривати з від'ємно-додатними формантами» [18], *derywaty z formantami ujemno-paradygmaticznymi* «деривати з від'ємно-парадигматичними формантами» [11], *derywaty paradygmaticzne* «парадигматичні деривати», *kontaminacje 2* «контамінації 2» ('результат процесу *kontaminacja 1* «контамінація 1»') або *skrótowce* «аббревіатури».

З мовознавчого погляду *skrót* не мають трактувати як словотвірний термін і ним не можна замінювати терміни *derywat ujemny* «від'ємний дериват», *derywat z ucięciem* «дериват з усіченням», *skrótowiec* та інші згадані вище. Отже, *skrót* означає скорочене графічне представлення часто вживаного слова (а не дериват), напр., *dr* «доктор», *mgr* «магістр», *por.* «порівняй (те)», *prof.* «професор» або сполуки, напр., *itp.* «і таке подібне (пол. *i temu podobne*)», *lp.* «однина (пол. *liczba pojedyncza*)», *m.in.* «серед іншого (пол. *między innym*)», яке використовують у писемних текстах і вимовляють як повне слово. Не лише у польських мовознавчих працях, а й в авторитетних словниках, наприклад в «Іншому словнику польської мови» («*Inny słownik języka polskiego*») [2: 606], знаходимо приклади неправильного, як видно з дефініції, уживання цього терміна стосовно різного типу дериватів, утворених шляхом усічення основи чи за участі (не єдиної) цієї словотвірної техніки⁵. Останнім часом спроби впорядкування цих термінів за допомогою стислих і однозначних дефініцій здійснили польські та чеські автори [10: 125—127; 12: 205; 13: 147—169]. Важливо, що ці терміни обговорюють системно, у взаємовідношеннях.

Неоднорідність і складність процесу усічення спричинює різне його трактування дослідниками. У польських мовознавчих працях функціонує низка термінів на позначення словотворчого процесу, що полягає в усіченні твірної основи деривата: *dezintegracja* (*derywacja dezintegralna*) «дезінтеграція (дезінтеграційна деривація)», *ucięcie* «усічення», *skrócenie* «скорочення», *redukcja* «редукція», *desufiksacja* «десуфіксація», *derywacja*

⁵ Подані в цих працях приклади не містять усіх дефініційних ознак, зокрема щодо графічного характеру скорочення.

ujemna «від'ємна деривація», *derywacja wsteczna* «зворотна деривація», *dekompozycja tematu* «декомпозиція основи», *derywacja alternacyjna* «альтернативна деривація», *derywacja wymienna* «замінна деривація», *mutylacja* «мутиляція», *derywacja redukcyjna* «редукційна деривація» [11: 9]. Серед них можна виділити найкомпактнішу групу термінів, які мають той самий обсяг змісту (денотацію) і нині найширше вживані: *derywacja ujemna*, *dezintegracja*, *derywacja dezintegralna*, *skrócenie*, *ucięcie*. Ці синонімічні терміни становлять частину польськомовного класу термінологічних еквівалентів у створюваній мові, яку буде розбудовано за рахунок іншомовних відповідників таких термінів після їхнього попереднього впорядкування.

Окреме місце в групі дериватів, утворюваних усіченням, до того ж з урахуванням їхнього вимовляння, займають скорочення-аббревіатури (*skrótowce*): літерні, ініціальні (*literowce (skrótowce literowe)*), звукові (*głoskowce (skrótowce głoskowe)*), складові (*grupowce (syłabowce, skrótowce sylabowe)*). Процедура усічення стосується в цьому випадку певних складників кількаслівних назв інституцій, товариств і т.ін. Найвідповіднішим терміном у польській мові, що означає творення скорочень, з огляду на міжнародний діапазон його поширення, є термін *abrewiacja* «аббревіація». Вибір цього терміна як найкращого є проявом процесу інтернаціоналізації термінології й мовного зближення, пор. еквівалентні терміни в усіх слов'янських та англійській мовах: білорус. *абрэвіацыя*, болг. *абрєвиация*, рос. *аббревиация*, укр. *абрєвіація*, чes. *tvoření zkratek, abreviace*, англ. *abbreviation* та інші. Таку тенденцію інтернаціоналізації підтримує спільнокореневий термін (латинського походження), який означає результат процесу аббревіації: білорус. *абрэвіатура*, болг. *абрєвиатура*, рос. *аббревиатура*, укр. *абрєвіатура*, чes. *abreviatura* (поряд з питомим *zkratka, zkratkové slovo*) і т. ін. Можна, проте, зазначити як приклад сильної національної традиції польський термін *skrótowiec*, який взагалі не має альтернативної міжнародної форми (форма *abreviatura* фіксована дуже рідко, переважно в працях русистів). Сталою позицією в системі цей термін завдячує мотивації за допомогою *skrócenie/skracanie* «скорочення/скорочування», а також номінації *skrót* у значенні самого прийому усічення.

Уже на прикладі малого фрагмента термінології видно неоднорідність та складність мовних явищ, які спричинюють інтерпретаційні й номінаційні відмінності. Ці проблеми необхідно вирішити для створення інформаційно-пошукової мови.

Використання термінів у функції ключових слів пов'язане з необхідністю розв'язання проблеми синонімії та багатозначності, явищ, не бажаних в інформаційно-пошуковій мові. В окремих мовах наявне відношення синонімії між двома і більше термінами, традиційно окреслюване як абсолютна, або класична, синонімія, а одиниці, що вступають у такі відношення, визначають як дублети, або абсолютні, класичні синоніми, пор. пол. *językoznawstwo — lingwistyka*, укр. *мовознавство — лінгвістика*, рос. *языкознание — лингвистика*. У цьому випадку процедура творення класів еквівалентності термінів проста: терміни, що належать певному

поняттю (це можуть бути окремі терміноодиниці чи абсолютні синоніми, дублети, що виступають у межах однієї мови та однієї концепції), об'єднуються з іншомовними дублетами або окремими одиницями в клас еквівалентності, пор. пол. *językoznawstwo*, *lingwistyka*, чес. *lingvistika*, *jazykověda*, словац. *lingvistika*, *jazykoveda*, хорват. *jezikoslovlje*, влуж. *rečespyt*, білорус. *мовазнаўства*, *лінгвістыка*, рос. *лингвистика*, *языкознание*, укр. *мовознавство*, *лінгвістика*, англ. *linguistics* та ін. У мовознавчій термінології наявні чимало пар або навіть тріад абсолютних синонімів, які мають питомий та запозичений (міжнародний) характер, пор. чес. *odvozenina* — *odvozené slovo* — *derivát* та їхні еквіваленти пол. *wyraz pochodny* — *derywat*, укр. *похідне слово* — *дериват*, рос. *производное слово* — *дериват* і т. ін.

Вищої мовознавчої компетенції вимагає від індексаторів об'єднання в класи еквівалентності термінів, між якими встановлюються відношення аспектною синонімії (синонімії концептуальної). Такий клас еквівалентності представляють, наприклад, словотвірні терміни: пол. *derywacja paradigmatyczna* «парадигматична деривація», *derywacja fleksyjna* «флексійна деривація», *derywacja zerowa* «нульова деривація», *derywacja bezsufigksalna* «безсуфіксна деривація», рос. *нулевая деривация*, *нулевая аффиксация*, *нулевая суффиксация*, укр. *безафіксне словотворення*, *нульсуфіксація*, *нульова суфіксація*, *нульсуфіксальний спосіб словотворення*⁶ та ін. Маємо тут ілюстрацію багатоаспектного осмислення того самого процесу деривації. Залежно від того, що прийнято за головні (вищі в ієрархії) — зміну флексії чи суфіксальний формант (а власне, його відсутність), застосовують і різні терміни. Справа ускладнюється, коли у дослідників наявні розбіжності в інтерпретації частини цієї сукупності дериватів. Зокрема, формації на зразок *brąz* «колір бронзи», *tachóń* «колір махагон, каштановий з червонуватим відтінком», які більшість мовознавців визначає як парадигматичні деривати, І. Капронь-Хажинська трактує як наслідок від'ємної деривації з ієрархічно нижчою, підпорядкованою їй зміною парадигми [11: 30]. Під час індексації вказівки на вищий ступінь в ієрархії стосуються змісту документа, тому мають бути відображені терміни, вживані в описуваній праці, особливо якщо вони функціонують у сучасному мовознавстві. Натомість інтерпретаційні розбіжності можна представити в бібліографічному описі в полі реферату (анотації).

Багатозначність термінів. Проект передбачає, що в створюваній мові ключових слів багатозначні терміни будуть подані як однозначні лексичні одиниці. Виділені ключові слова будуть споряджені текстовими поясненнями і номерами 1, 2, 3 тощо. Наприклад, багатозначний у польській мові термін *gramatyka* представлений у слов'янських та англійській мовах трьома класами еквівалентних одиниць з окремими значеннями у формі текстових пояснень і номерів 1, 2, 3: *gramatyka 1* 'зведення правил функціонування одиниць, які складають систему мови', *gramatyka 2*

⁶ Сердечно дякую проф. Євгенії Карпіловській за консультацію стосовно термінології в українській мові.

‘мовознавча дисципліна (піддисципліни), яка описує зведення правил функціонування одиниць, що складають систему мови’⁷ і *gramatyka 3* ‘документ, який містить опис правил функціонування одиниць, які складають систему мови’; багатозначний польський термін *morfologia* представлений двома класами еквівалентних лексичних одиниць у польській та інших слов’янських мовах, пор. у польській мові: *morfologia 1* ‘будова слів і форм слів або відмінювання слів’ і *morfologia 2* ‘мовознавча дисципліна (піддисципліна), яка вивчає будову слів і форм слів або відмінювання слів’. Подібною є семантична структура термінів: пол. *słowotwórstwo 1 і 2, etymologia 1 і 2, fonetyka 1 і 2, fonologia 1 і 2, fleksja 1 і 2, morfonologia 1 і 2, składnia* «синтаксис» *1 і 2* та ін. (див. рис. 3). Як бачимо, розкладення (декомпозиція) значень таких термінів дозволяє уявити типові вужчі значення, а також семантичні відношення між ними. Слід підкреслити, що ці процедури становлять лише початок конструювання мови ключових слів. Наступним завданням є порівняння та об’єднання в класи еквівалентів однопорядкових термінів у слов’янських та англійській мовах.

Відношення між виділеними значеннями вищезгаданих пар і тріад термінів мають регулярний характер, реалізований тут за допомогою певної семантичної моделі⁸. У семантичній структурі описуваної групи термінів та їхніх іншослов’янських еквівалентів можна виокремити типові відношення: ‘явище’ — ‘дисципліна (піддисципліна), що досліджує таке явище’ [4: 203].

Інший тип відношень можна виокремити в семантичній структурі багатозначних термінів: *nominacja, kompozycja, kontaminacja* та ін. Регулярний характер тут має семантичне відношення ‘дія/процес’ — ‘результат’, напр., *nominacja 1* ‘творення мовних одиниць, які мають номінативну функцію’, *nominacja 2* ‘мовна одиниця, яка має номінативну функцію’, *kontaminacja 1* ‘об’єднання різних мовних елементів в слово або словосполучення’, *kontaminacja 2* ‘слово або словосполучення, які виникли внаслідок об’єднання різних мовних елементів’.

В усуненні багатозначності термінів допомагають словотвірні механізми. Це можна продемонструвати на прикладі багатозначних термінів з ономастики. Ще не так давно в польській мовознавчій літературі терміни: *antroponimia, toponimia* уживали в двох значеннях: 1. ‘сукупність антропонімів, антропонімічна система’, 2. ‘дисципліна (піддисципліна), яка вивчає антропоніми’ і 1. ‘сукупність топонімів, топонімічна система’, 2. ‘дисципліна (піддисципліна), яка вивчає топоніми’. Ці значення і відношення між ними можна узагальнити в типовому семантичному відношенні ‘явище, сукупність/система об’єктів’ — ‘дисципліна’, яке було представлено раніше. З розвитком ономастичних досліджень та їхньою спеціалізацією змінюється термінологія. Завдяки міжнародним зразкам виник новий тип структур, що дав початок серії: *antroponomastyka, toponomastyka,*

⁷ Зузанна Тополінська пропонує введення до польської мовознавчої термінології терміна *morfosyntaksa* «морфосинтаксис», сумірного з *gramatyka 2* «граматика 2» [17: 210].

⁸ Докладно це явище описує Ю.Д. Апресян [4: 187—193].

Słowo kluczowe			
Deskryptor: slowotwórstwo 1 (tworzenie wyrazów)			
Język: pol (polski)			
Wersje słowa kluczowego			
Słowo	Język	Deskryptor	Akcje
slowotwórstwo 1 (tworzenie wyrazów)	polski (pol)	⊙	[Wyświetl]
словообразование 1	rosyjski (rus)	⊙	[Wyświetl]
словопроизводство	rosyjski (rus)	⊙	[Wyświetl]
словотворчество	rosyjski (rus)	⊙	[Wyświetl]
tvorba riječi 1	chorwacki (scr)	⊙	[Wyświetl]
tvoření slov 1	czeski (cze)	⊙	[Wyświetl]
slowotwórba 1 (tworzenie wyrazow)	górnoluzycycki (hsb)	⊙	[Wyświetl]
творба речи 1	serbski (scc)	⊙	[Wyświetl]
slovotvorba 1	słowacki (slo)	⊙	[Wyświetl]
besedotvorje 1	słoweński (slv)	⊙	[Wyświetl]
словотвір 1	ukraiński (ukr)	⊙	[Wyświetl]
словотворення	ukraiński (ukr)	⊙	[Wyświetl]
[Wyświetl dokumenty]			

Рис. 3. Фрагмент мови ключових слів з класом еквівалентів до польського терміна *slowotwórstwo 1*

а згодом виникла *zoonomastyka* [8: 330] і *chrematonomastyka*⁹. У структурі створюваної інформаційно-пошукової мови ці неологізми поєднано з надалі уживаними термінами у відповідні класи еквівалентів. Можна прогнозувати, що розвиток ономастики і потреби найменування сприятимуть виникненню нових піддисциплін, що вивчають астроніми, етніоніми, фітоніми, космоніми і це знайде своє відображення в термінології.

Розв'язання проблеми багатозначності термінів, принаймні, т.зв. класичної багатозначності, яка полягає у повторенні того самого типу відношень між двома значеннями багатозначних термінів [9: 71], збагатить семантичну структуру створюваної інформаційно-пошукової мови, усуне інформаційний «шум» і забезпечить повноту пошуку інформації в системі iSybislaw¹⁰.

Актуалізація термінології. Важливий чинник впливу на ефективність системи становить якість інформаційно-пошукової мови, що виявляється в актуальності термінів, які виконують функцію ключових слів. У зв'язку з цим термінологія, включена до структури створюваної мови сис-

⁹ Вихідним зразком була, ймовірно, чеська модель, що має інтернаціональний характер, пор.: *antroponomastika*, *toponomastika*, *chrematonomastika*, а також німецька: *Anthroponomastik*, *Toponomastik*.

¹⁰ Докладніше про розв'язання проблеми синонімії та багатозначності термінів у зв'язку зі створюваною інформаційно-пошуковою мовою див.: [14: 631–646].

тематично актуалізується. Праця в цьому напрямку охоплює як заповнення термінологічних прогалів, так і виявлення альтернативних інновацій до вже наявних термінів. Наприклад, до розбудовуваної лексичної системи мови ключових слів протягом останніх кількох років введено такі польські терміни, об'єднані з іншослов'янськими відповідниками в класи еквівалентів: пол. *chrematonim*, чес. *chrématonýmim*, рос. *хрематоним*; пол. *chrematonimia 1* 'сукупність, система хрематонімів', чес. *chrématonymie*, рос. *хрематонимия*; пол. *chrematonimia 2* 'дисципліна' і синонімічний термін пол. *chrematonomastyka* 'те саме', чес. *chrématonomastika*; пол. *gramatyka komunikacyjna*, чес. *komunikační gramatika* і *komunikativní gramatika*, рос. *коммуникативная грамматика*; пол. *lingwistyka genderowa* і *lingwistyka płci*, чес. *genderová lingvistika*, рос. *гендерная лингвистика*; пол. *lingwistyka korpusowa*, чес. *korpusová lingvistika*, рос. *корпусная лингвистика*; пол. *lingwistyka kulturowa*, чес. *kulturní lingvistika*, рос. *лингвокультурология* та багато інших термінів з не згаданих тут слов'янських мов. Частина з них це нові чи порівняно нові терміни, які виникли разом з новими явищами та проблемами або внаслідок нового, іншого погляду на явища. Найчисленніше серед різномовних термінів представлені польські терміни, що є передусім результатом праці над польськомовною частиною створюваної інформаційної мови. Значну групу серед неологізмів становлять терміни, що називають піддисципліни та наукові теорії. Це структури, характерні для нашого часу. Слов'янська мовознавча термінологія реагує на зміни шкіл, напрямів і теорій, на розширення меж мовознавства.

Упорядкування слов'янської мовознавчої термінології для потреб інформаційно-пошукової системи iSybislaw має багатовимірний, багатоаспектний і постійний характер. Метою такої праці є оптимізація системи та її найважливішого складника — інформаційно-пошукової мови, а також ефективна стратегія пошуку для задоволення інформаційних потреб користувачів та поглиблення знань у галузі слов'янського мовознавства. З огляду на обмежений обсяг статті тут представлено лише найважливіші здобутки колективу розробників iSybislaw у цій царині, в основі яких лежить переконання, що користувач інформації і його потреби мають бути в центрі інформаційних процесів.

1. Bibliograficzna baza danych światowego językoznawstwa slawistycznego iSybislaw, b.d.; Режим доступу: <http://www.isybislaw.ispan.waw.pl>
2. *Vańko M.* (Red.), *Inny słownik języka polskiego*. Warszawa: Wydawnictwo Naukowe PWN.
3. *Bojar B.* (Red.) (2002), *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych*. Warszawa: Wydawnictwo SBP.
4. *Апресян Ю.Д.* (1974). *Лексическая семантика. Синонимические средства языка*. М.: Наука.
5. *Babik W.* (2010), *Słowa kluczowe*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
6. *Blicharski M.* (1999), *Słowotwórstwo nazw osób w rosyjskich żargonach przestępczych*. In: *Blicharski M., Fontański H.* (Red.), *Słowotwórstwo, semantyka i składnia języków słowiańskich*. T. 1. Katowice: Wydawnictwo Uniwersytetu Śląskiego, s. 38—52.
7. *Bojar B., Rudnik-Karwatowa Z.* (1998), *Koncepcja nowoczesnego systemu informacji slawistycznej*. In: *Dalewska-Greń H., Rusek J., Siatkowski J.* (Red.), *Językoznawstwo. Prace na*

- XII Międzynarodowy Kongres Słowistów w Krakowie 1998. Warszawa: Energeia (Z Polskich Studiów Słowistycznych; seria 9) s. 42–49.
8. *Breza E.* (1998), Nazwy obiektów i instytucji związanych z nowoczesną cywilizacją (chrematonymy). In: Rzetelska-Feleszko E. (Red.), *Polskie nazwy własne. Encyklopedia*. Warszawa; Kraków: Wydawnictwo Instytutu Języka Polskiego PAN, s. 343–361.
 9. *Gajda S.* (1990), *Wprowadzenie do teorii terminu*. Opole: WSP im. Powstańców Śląskich w Opolu. (Studia i Monografie; 162).
 10. *Jadacka H.* (2006), Skróty [словникова стаття (artykuł hasłowy)]. In: *Jadacka H., Markowski A., Zdunkiewicz-Jedynak D.* (Red.), *Poprawna polszczyzna. Hasła problemowe*. Warszawa: Wydawnictwo Naukowe PWN.
 11. *Kaproń-Charzyńska I.* (2005), Derywacja ujemna we współczesnym języku polskim. *Rzeczowniki i przymiotniki*. Toruń: Wydawnictwo TOP KURIER.
 12. *Nagórko A.* (2010), *Podręczna gramatyka języka polskiego*. Warszawa: Wydaw. Naukowe PWN.
 13. *Rangelova A., Tichá Z.* (2012), Dynamika abreviace a abreviačních útvarů (na českém jazykovém materiálu). In: *Rudnik-Karwatowa Z.* (Red.), *Specyfika leksyki i słowotwórstwa języków słowiańskich na przełomie XX i XXI wieku*. Warszawa: Słowistyczny Ośrodek Wydawniczy, s. 171–188.
 14. *Rudnik-Karwatowa Z., Banasiak J., Mikos Z.* (2013), Językowe problemy optymalizacji wyszukiwania informacji w systemie iSybislaw. „*Slavia Orientalis*”, R. 62, nr 4, s. 631–646.
 15. *Rudnik-Karwatowa Z., Karpińska H.* (1999), *Słownik słów kluczowych językoznawstwa słowistycznego*. Warszawa: Słowistyczny Ośrodek Wydawniczy.
 16. *Rudnik-Karwatowa Z., Karpińska H.* (2006), *Słownik słów kluczowych językoznawstwa słowistycznego*. [Вид. виправл. і доповн.]. Warszawa: Słowistyczny Ośrodek Wydawniczy [CD-ROM].
 17. *Topolińska Z.* (2009), Kilka słów o naszej siatce pojęciowej i terminologicznej. In: *Mindak-Zawadzka J., Dolinski I.M.* (Red.), *Językowy świat Słowian. Zjawiska, interpretacje, znaki zapytania. Księga jubileuszowa na 70-lecie Profesora Włodzimierza Pianki*. Warszawa: Instytut Słowistyki Zachodniej Uniwersytetu Warszawskiego.
 18. *Wróbel H.* (2001), *Gramatyka języka polskiego*. Kraków: „OD NOWA”.

Статтю отримано 19.04.2016

Zofia Ewa Rudnik-Karwatowa

Institute of Slavic Studies of the PAS, Warsaw

TERMINOLOGY ARRANGEMENT FOR INFORMATION RETRIEVAL NEEDS IN ISYBISLAW SYSTEM

The use of a keyword language in the information retrieval system (IRS) of Slavic linguistics iSybislaw (see Bibliographic Database of World Slavic Linguistics, www.isybislaw.ispan.waw.pl) equals assigning the function of keywords to linguistic terms that represent the content of the documents contained in the information set of the system. The adopted solution is advantageous for users, however, poses a number of problems in the IRS such as the phenomenon of synonymy and ambiguity of terms, typical for natural languages.

In the multilingual system used in iSybislaw an original methodological approach involving the creation of classes of equivalent terms used currently in the linguistic literature in a given language, as well as combining them in multilingual equivalent classes was implemented. This solution is intended to provide efficiency of the multilingual system in terms of meeting the information needs of cross-lingual users. The paper presents both theoretical issues and experiences from the work on the arrangement of terminology in the iSybislaw system.

Key words: *terminology, terminology arrangement, linguistic terminology, equivalence, information-retrieval system, iSybislaw, information retrieval system efficiency*