

WYKORZYSTANIE TERMINOLOGII W SYSTEMIE INFORMACYJNO-WYSZUKIWAWCZYM JĘZYKOZNAWSTWA SLAWISTYCZNEGO ISYBISŁAW

Paweł Kowalski*
Zofia Rudnik-Karwatowa**

W artykule omawiane jest wykorzystanie terminologii językoznawczej w bazie iSybisław, która prezentuje nowoczesny system informacyjno-wyszukiwawczy językoznawstwa slawistycznego. W języku tego systemu terminy językoznawcze pełnią specjalną funkcję: funkcję słów kluczowych. Szczególną uwagę autorzy zwracają na rozwiązanie problemów związanych z wielojęzycznością zbioru informacyjnego, synonimią i wieloznacznością terminów. Wskazują również na korzyści dla użytkowników płynące z wykorzystania terminologii językoznawczej w systemie iSybisław.

Słowa kluczowe: bibliograficzna baza danych językoznawstwa slawistycznego, iSybisław, język słów kluczowych, językoznawstwo slawistyczne, system informacyjno-wyszukiwawczy, terminologia, terminologia językoznawcza

1. Wstęp

Coraz powszechniejsze stosowanie w nauce zdobyczy technologicznych, zwłaszcza komputeryzacji, oraz upowszechnianie się internetu jako przestrzeni komunikacyjnej w środowisku naukowym wpływa na model funkcjonowania bibliografii. Przenoszenie dotychczasowych osiągnięć bibliograficznych do przestrzeni cyfrowej daje możliwość przekształcenia tradycyjnych bibliografii w bardziej lub mniej nowoczesne systemy informacyjno-wyszukiwawcze, które mają za zadanie jeszcze skuteczniej służyć użytkownikom. Praca nad każdym takim systemem wymaga od jego twórców rozstrzygnięcia wielu problemów metodologicznych. Najważniejszym zadaniem jest zaprojektowanie efektywnego systemu z jego narzędziem informacyjnym – językiem informacyjno-wyszukiwawczym. W systemach dziedzinowych do budowy systemu leksykalnego tego języka stosuje się terminy języka naturalnego (terminologię konkretnej dziedziny wiedzy).

W artykule ukazane zostanie wykorzystanie terminologii językoznawczej w systemie informacyjno-wyszukiwawczym iSybisław, który jest swoistą remediacją informacji slawistycznej, gromadzonej dotychczas w formie bibliografii

* Instytut Slawistyki PAN, Warszawa; kowalion@wp.pl

** Instytut Slawistyki PAN, Warszawa; z.karwatowa@gmail.com

papierowej, do przestrzeni cyfrowej¹. Terminologia językoznawcza wykorzystywana jest przede wszystkim w podstawowym narzędziu systemu – języku słów kluczowych, natomiast w tradycyjnej bibliografii oczywiście terminologii w żaden sposób nie wykorzystywano, ponieważ podstawowym narzędziem organizacji i prezentacji danych o dokumentach była klasyfikacja.

Język słów kluczowych to język informacyjno-wyszukiwawczy, czyli język paranaturalny, który odwzorowuje informacje zawarte w dokumentach, wykorzystywany jest do wyszukiwania informacji w zbiorach informacyjnych systemów informacyjno-wyszukiwawczych (Bojar 2005: 248). Niekiedy w literaturze przedmiotu jednostki języka słów kluczowych traktowane są jako podzbiór słownictwa języka naturalnego. Mają one spełniać określone, specjalne funkcje. W przypadku nauki podzbiór ten w znacznym stopniu pokrywa się z terminologią (zob. Babik 2010: 58). Pod pojęciem terminologii rozumiemy tutaj pewien podzbiór leksyki danego języka ogólnonarodowego, charakteryzujący się stosunkowo wąskim zakresem zastosowania i rozpowszechnienia oraz ograniczonym stopniem rozumienia (ograniczonym do grona specjalistów), innymi słowy, terminologia to pewien podzbiór leksyki danego języka ogólnonarodowego (zbiór terminów) o ściśle ustalonym znaczeniu w danej dziedzinie nauki lub techniki (zob. Bojar 2002: 277; zob. też Jadacka 2008: 141).

Odmienny jest status jednostek terminologicznych i jednostek języka słów kluczowych. Terminy denotują i konotują elementy rzeczywistości językowej (pozatekstowej), natomiast słowa kluczowe denotują bezpośrednio zbiory dokumentów zawartych w systemie i konotują bezpośrednio pewne ich podzbiory, zbiory cech dokumentów. Użytkownik może za pośrednictwem informacji uzyskanych z dokumentów poznać interesujący go obiekt czy fragment rzeczywistości pozadokumentacyjnej, wówczas dotyczy to denotacji pośredniej i konotacji pośredniej. Słowa kluczowe w takim systemie informacyjno-wyszukiwawczym pełnią dwie funkcje: metainformacyjną i wyszukiwawczą (por. Bojar: 2002, 2005).

2. Charakterystyka systemu iSybislaw

System iSybislaw powstał z inicjatywy polskich slawistów, a jego koncepcja została przedstawiona na XII Międzynarodowym Kongresie Sławistów w Krakowie w 1998 r. (Bojar –Rudnik-Karwatowa 1998: 41-47). Od początku projekt instytucjonalnie związany jest z Instytutem Sławistyki PAN i ma charakter międzynarodowy. Obecnie zespół realizujący projekt składa się z językoznawców slawistów z Białorusi, Chorwacji, Czech, Macedonii, Niemiec, Polski, Rosji, Sło-

¹ Remediację rozumiemy tutaj jako jedną z podstawowych cech mediów cyfrowych polegającą na reprezentowaniu jednego medium w drugim, przy czym w medium drugim niektóre cechy medium pierwszego zostają zachowane, inne ulegają zmianie (zob. Bolter-Grusin: 1999).

wacji i Ukrainy. W bazie iSybisław pod adresem www.isybislaw.ispan.waw.pl udostępnione są obecnie w wolnym dostępie i bezpłatnie informacje o ponad 22 tysiącach dokumentów slawistycznych wydanych od roku 1992 do dziś. Zbiór informacyjny cały czas jest powiększany, jednocześnie rozbudowywana jest też funkcjonalność systemu.

Tematyka gromadzonych dokumentów wyznaczana jest przez zakres wiedzy językoznawstwa slawistycznego, a więc zbiór informacyjny systemu obejmuje przede wszystkim dokumenty z zakresu językoznawstwa slawistycznego diachronicznego i synchronicznego, w tym językoznawcze prace konfrontatywne i kontrastywne słowiańsko-słowiańskie i słowiańsko-niesłowiańskie. Uwzględniane są, choć w mniejszym stopniu, prace z zakresu językoznawstwa teoretycznego i językoznawstwa ogólnego, istotne dla slawistyki językoznawczej. W systemie znajdują się informacje relewantne zarówno dla wszystkich tradycyjnych dyscyplin (poddyscyplin) językoznawstwa, jak: fonetyka, fonologia, morfologia, morfonologia, słowotwórstwo, składnia i inne, a także nowsze, w których strukturach nazewniczych ujawnia się interdyscyplinarny charakter: socjolingwistyka, psycholingwistyka, etnolingwistyka, lingwistyka tekstu (tekstologia lingwistyczna). Terminologia należąca do poszczególnych poddyscyplin językoznawstwa nie jest rozbudowywana równomiernie. Zasób terminów zależy w dużej mierze od rozwoju poszczególnych poddyscyplin i popularności podejmowanych w nich tematów. Największa liczba terminów w systemie iSybisław dotyczy słowotwórstwa, morfologii, składni, leksykologii, mniejsza – m.in. socjolingwistyki, etnolingwistyki, tekstologii, najmniejsza zaś np. morfonologii i morfotaktyki.

Zasięg językowy w systemie jest nieograniczony, co oznacza, że w bazie iSybisław znajdują się prace we wszystkich językach naturalnych: głównie w językach słowiańskich, co jest zdeterminowane zakresem tematycznym systemu i rozwojem językoznawstwa slawistycznego w krajach słowiańskich, a także w językach międzynarodowych i innych, takich jak np. języki węgierski, japoński, estoński. Aby usprawnić komunikację między użytkownikami, którzy są różnojęzyczni, a systemem, opracowano interfejs w trzech językach: polskim, rosyjskim i angielskim. Pozwala to każdemu użytkownikowi wybrać odpowiedni dla niego język przy korzystaniu z systemu.

3. Terminologia w systemie iSybisław

Terminologia językoznawcza w systemie iSybisław wykorzystywana jest przede wszystkim w podstawowym języku informacyjno-wyszukiwawczym tego systemu – języku słów kluczowych. Oprócz języka słów kluczowych stosuje się także język klasyfikacji i język tagów, które są traktowane tylko pomocniczo. Jednostki leksykalne języka słów kluczowych (funkcję słów kluczowych pełnią terminy językoznawcze) odzwierciedlają treści gromadzonych i opisywanych w

bazie dokumentów (monografii, prac zbiorowych, czasopism, artykułów, omówień itp.) w różnych językach. Obok językoznawczej terminologii sławistycznej (terminów szczegółowych) jako jednostki słów kluczowych systemu iSybislaw występują także terminy ogólnonaukowe, ogólnojęzykoznawcze. W zakres tej grupy wchodzi jednostki dotyczące teorii języka i językoznawstwa ogólnego, a także językoznawstwa kontrastywnego słowiańsko-niesłowiańskiego. Terminy stanowią zbiór otwarty, co oznacza możliwość jego rozbudowy dla potrzeb języka systemu zgodnie z odpowiednim stopniem szczegółowości (głębokości) i szerokości relewantnej informacji. Szczegółowość i kompletność odwzorowania treści dokumentów w charakterystykach wyszukiwawczych lub instrukcjach wyszukiwawczych podyktowane są przede wszystkim dążeniem do realizacji istotnych potrzeb informacyjnych użytkowników. Procedury z tym związane mają zapobiec stratom informacji i zapewnić uniknięcie szumu informacyjnego.

Użycie w systemie informacyjno-wyszukiwawczym iSybislaw terminów i innych wyrażen języków naturalnych (wszystkich indeksowanych języków słowiańskich i języka angielskiego) jako równokształtnych słów kluczowych (struktura formalna słów kluczowych i terminów w bazie jest tożsama) jest dogodnym rozwiązaniem dla użytkowników ze względu na ich kompetencję językową. W języku słów kluczowych wykorzystywane są jedynie relewantne i reprezentatywne dla językoznawstwa terminy, nie jest bowiem możliwe uwzględnienie terminologii językoznawczej pochodzącej z całego światowego piśmiennictwa sławistycznego. Zbiór słów kluczowych uzupełniany i weryfikowany jest na podstawie encyklopedii językoznawczych, gramatyk, leksykonów i słowników w poszczególnych językach słowiańskich.

Obok terminologii językoznawczej w języku słów kluczowych wykorzystywane są, w mniejszym zakresie, także inne wyrażenia języków naturalnych. Mają one za zadanie organizowanie informacji i wiedzy w systemie. Są to przede wszystkim nazwy własne, w tym m.in. nazwiska badaczy językoznawców, których prace są przedmiotem badań opisywanych w dokumentach, nazwiska autorów dzieł literackich, które stanowią przedmiot analizy językoznawczej, nazwy geograficzne, takie jak nazwy miejscowości i regionu, gdzie przeprowadzono badania zjawisk językowych opisywanych w dokumentach, miejsce i rok konferencji, której temat jest referowany i nazwy korporatywne, np. nazwy instytucji i towarzystw naukowych. Jako słowa kluczowe wykorzystywane są również nazwy pospolite, takie jak wyrażenia języka ogólnego stanowiące przedmiot badań, w tym leksemy, frazeologizmy, morfemy itp., wyrażenia mające dużą frekwencję w zbiorze dokumentów, np. nazwy barw w polszczyźnie, nazwy części ciała i inne.

Użycie notacji naturalnej w języku informacyjno-wyszukiwawczym jest dogodnym rozwiązaniem dla użytkowników. Pozwala im bowiem samodzielnie tworzyć strategie wyszukiwawcze, rozwiązywać problemy wyszukiwawcze oraz samodzielnie weryfikować pod kątem relewancji otrzymaną informację. Swobodne

i skuteczne przeszukiwanie bogatego zasobu terminologicznego i wyszukiwanie potrzebnych informacji wymusza jednak na użytkownikach posiadanie kompetencji językowej i językoznawczej. Upowszechnianie wiedzy terminologicznej w systemie powinno być dla użytkowników zrozumiałe i proste, dlatego też jednostki leksykalne języka słów kluczowych, tj. słowa kluczowe, muszą być przejrzyste. Zastosowanie terminologii danej dyscypliny, w tym wypadku językoznawstwa, pozwala utrzymać taką przejrzystość. Specjalistyczna terminologia danej dziedziny jest bowiem rozumiana i stosowana w komunikacji naukowej całego środowiska. Zastosowanie w systemie iSybislaw języka o notacji naturalnej jest korzystne także dla indeksatorów, ponieważ ułatwia im przygotowanie zarówno opisów formalnych, jak i opisów rzeczowych dokumentów.

4. Struktura formalna terminów

Jednostki terminologiczne wykorzystywane jako słowa kluczowe to struktury proste (jednoelementowe, jednowyrazowe) i struktury złożone (dwu- lub więcej elementowe, wielowyrazowe). Wśród struktur prostych można wyróżnić terminy proste (niepochodne), derywaty, złożenia (composita), złożeniowce, natomiast do struktur złożonych zalicza się zestawienia (nazwy zestawieniowe, skupienia). Największą frekwencją w zgromadzonym zbiorze odznaczają się rzeczowniki. Przykłady:

język polski: *czasownik, celownik, językoznawstwo, leksykologia, półustaw, paleografia, rzeczownik, zdanie, derywacja*;

język serbski: *глагол, дајтив, језикословље, лингвистика, лексикологија, њолу-устијав, њалеографија*;

język rosyjski: *аббревиатура, адъектив, апеллятивность, вербализация, вокализм, глагол, двуязычие, деривация, диатакис*.

Terminy złożone (dwu- i więcej elementowe) mają postać przede wszystkim syntagm rzeczownikowo-przymiotnikowych, rzadziej rzeczownikowo-rzeczownikowych:

język polski: *język polski, historia języka, sufiks zerowy, grupa leksykalno-semantyczna, słowotwórstwo diachroniczne, funkcjonalna perspektywa zdania*;

język serbski: *српски језик, историја језика, нулти суфикс, лексичкосемантичка група, слаби полугласник*;

język rosyjski: *абсолютный синоним, абстрактное существительное, аналитическая конструкция, близкородственный корень, диалектологическая карта, коммуникативная перспектива высказывания, лексико-семантическая группа*.

Struktura formalna poszczególnych terminów w różnych językach zależy od specyficznych cech systemowych danego języka. Najczęstsze w bazie iSybislaw terminy złożone to struktury dwuelementowe składające się z jednego elementu rzeczownikowego i jednego przymiotnikowego (Adj + N lub N + Adj), rzadziej reprezentowany jest model rzeczownik + rzeczownik (N + N). Na tle języków słowiańskich polszczyznę charakteryzuje odmienny od pozostałych języków słowiańskich nienacechowany szyk takich jednostek. Rzeczownik jako człon określany umieszczany jest w prepozycji, przymiotnik (lub rzeczownik) jako człon określający znajduje się w postpozycji (np. *język polski*, *lingwistyka korpusowa*). Polskie terminy dwuczłonowe reprezentują więc model N + Adj. W innych językach słowiańskich element określający człon rzeczownikowy znajduje się najczęściej na miejscu pierwszym (np. serb. *српски језик*, *нулџи суфикс*, ros. *русский язык*, *корпусная лингвистика*). Terminy te reprezentują więc model Adj + N.

Warto odnotować, że wśród wielowrazowych jednostek terminologicznych bardzo liczną grupę tworzą terminy stanowiące wyrażenie węższe niż odpowiednie wyrażenie (podstawa motywacyjna) o strukturze prostej, pozostające względem niego w paradygmatycznej relacji podrzędności zakresowej (hiponimii), np. *genologia lingwistyczna* – *genologia*, *sufiks zerowy* – *sufiks*, *spółgłoska palatalna* – *spółgłoska*. Organizacja paradygmatyczna tych struktur jest tu przejawem dążenia do systemowości jednostek terminologicznych. Obserwowana we współczesnej terminologii językoznawczej tendencja do wzrostu terminów motywowanych, w danym przypadku derywowanych składniowo, paradoksalnie prowadzi do wydłużenia terminów wbrew zasadzie operatywności, czyli ekonomiczności jednostek terminologicznych.

5. Wykorzystanie terminologii w systemie iSybislaw.

Procedury dostosowujące

Wykorzystanie terminów języków naturalnych (wszystkich języków słowiańskich i angielskiego) w funkcji słów kluczowych wiąże się z wieloma problemami natury metodologicznej. Chodzi przede wszystkim o rozwiązanie kwestii synonimii (i bliskoznaczności) oraz wieloznaczności, zjawisk niepożądanych w systemie informacyjno-wyszukiwawczym, a „przeniesionych” z języków naturalnych. Dodatkowo wielojęzyczność zbioru dokumentów i różnojęzyczność terminów w funkcji słów kluczowych zmusza do poszukiwania skutecznego dostępu do zbioru informacji i skutecznego wyszukiwania potrzebnych informacji. W tym celu zastosowano rozwiązanie polegające na ustanowieniu klas ekwiwalentów, grupujących głównie terminy równoznaczne z jednego języka naturalnego (ekwiwalenty wewnątrzjęzykowe), jak i z wielu języków naturalnych i w wielu alfabetach (ekwiwalenty międzyjęzykowe, zewnętrzjęzykowe).

Budowanie klas ekwiwalentów wiąże się z transformacjami informacji wyjściowych z informacji źródłowych (zob. Rudnik-Karwatowa, Banasiak, Mikos 2013: 632). Tworzenie takich klas stanowi wyzwanie dla twórców systemu iSybislaw oraz całego międzynarodowego zespołu partnerów i współpracowników. Na etapie przygotowania terminów do łączenia w klasy ekwiwalentów należy także korygować brak precyzji, niepoprawność i błędy merytoryczne w terminologii stosowanej w pracach włączanych do systemu. Struktura klasy ekwiwalentów składa się z różnojęzycznych jednostek – terminów ekwiwalentnych z różnych języków słowiańskich i języka angielskiego oraz różnojęzycznych alfabetów. Prace nad konstruowaniem takich klas mają charakter wieloetapowy. Na początku terminy odwzorowujące treść dokumentów tworzą zbiór jednojęzyczny (łączone są pary synonimiczne), na dalszym etapie prac do klas ekwiwalentów wprowadzane są terminy równoznaczne z różnych języków.

Najczęściej źródłem synonimii i bliskoznaczności terminów rejestrowanych w dokumentach jest zjawisko internacjonalizacji, por. ros. *двухязычие – билингвизм, языкознание – лингвистика, производное слово – дериват*, słoweń. *na-rečje – dialekt, glasoslovje – fonetika, besedje – leksika*, pol. *językoznawstwo – lingwistyka, zmiękczenie – palatalizacja, słownictwo – leksyka* i wiele innych. Innym źródłem synonimii w terminologii językoznawczej jest wariantywność użycia jednostki syntetycznej i jednostki analitycznej, por. pol. *alfabet cyryliccki – cyrylica, ekologia języka – ekolingwistyka, akt illokucyjny – illokucja, język polski – polszczyzna*, słoweń. *latinska abeceda – latinica, slovanska abeceda – slovenica, slovenski jezik – slovenščina, tekstna lingvistika – besediloslovje*. W przypadku takich par równoznacznych terminów procedura tworzenia klas ekwiwalencji w systemie iSybislaw polega na łączeniu tych jednostek terminologicznych z jednego języka, a następnie dołączaniu do nich odpowiedników innojęzycznych.

Istnieje obszerna literatura dotycząca internacjonalizacji i internacjonalizmów (jednostek internacjonalnych, jednostek międzynarodowych), w mniejszym stopniu interesowano się dotychczas internacjonalizmami terminologicznymi (zob. m.in. Gajda 1990: 98–106; Lukszyn, Zmarzer 2006; Гринев-Гриневич 2008; Mazurkiewicz-Sułkowska 2014, 2015: 166–180).

W słowiańskiej terminologii językoznawczej bogatą reprezentację mają terminy przejmowane, po odpowiedniej adaptacji (czasem też w postaci zapożyczeń cytatów), z języków klasycznych: greki i łaciny, a ostatnio też z języka angielskiego, oraz terminy pochodzące z elementów (morfemów) grecko-łacińskich, których forma i znaczenie są identyczne lub bliskie w wielu językach. O uznaniu za internacjonalizmy decyduje obecność ich przynajmniej w trzech językach odmiennych typologicznie. Dużą część starej terminologii językoznawstwa sławistycznego stanowią internacjonalizmy leksykalne i morfemowe (zwłaszcza terminy gramatyczne) zaadaptowane w języku biorcy lub przejęte jako terminy „cytaty”, obok których w późniejszym okresie pojawiły się ekwiwalentne terminy rodzime: pol.

dualis – liczba podwójna, hiatus – rozziw, nominativus – mianownik, genetivus – dopełniacz, czes. derivace – odvozování, enklitikon – příklonka, konsonant – souhláska, numerale – číslovka, słoweń. morfologija – oblikoslovje, semantika – pomenoslovje, sintaksa – skladnja, tekstna lingvistika – besediloslovje i wiele innych.

We współczesnej terminologii językoznawczej obserwuje się odwrotną tendencję: tworzenie internacjonalnych terminów alternatywnych do będących w użyciu terminów rodzimych. Odnotowujemy to na przykład w polskiej literaturze onomastycznej, gdzie do niedawna tradycyjnie stosowano przeważnie terminy rodzime, jak *nazwa własna (nazwy własne)* lub wymiennie termin rodzimy albo międzynarodowy, nie preferując tego drugiego, np. *nazwy własne* i *onimia*, *nazwy osobowe* i *antroponimia*, *nazwy wodne* i *hydronimia*. W najnowszych pracach onomastycznych uwidacznia się tendencja do zastępowania terminów rodzimych internacjonalizmami, por. *nazwa własna* → *onim*, *nazwy własne* → *onimia*, *nazwy terenowe* → *mikrotoponimy*, *mikrotoponimia*. Obecnie w pracach onomastycznych używa się terminu *chrematonim*, a także *chrematonimy*, *chrematonimia* 1 ‘nazwy obiektów i instytucji związanych z nowoczesną cywilizacją’, *chrematonimia* 2 ‘subdyscyplina onomastyczna zajmująca się chrematonimami’ (Breza 1998: 343). Nieaktualny jest postulat rodzimości terminologii, nawet w ograniczonych warunkach, zgłoszony kilkanaście lat temu przez polską onomastkę: „Jeśli natomiast brak wyrazów pokrewnych, lepiej nazwy nie wprowadzać lub wprowadzić drugoplanowo, jako odpowiednik nazwy rodzimej, np. *nazwa wyrobu przemysłowego (chrematonim)*” (Kaleta 1998: 78).

W ciągu ostatnich ponad trzydziestu lat proces internacjonalizacji terminologii językoznawczej uintensywnił się, co uwidacznia się we wzroście aktywności szeregu morfemów afiksalnych i podstaw słotwórczych mających charakter internacjonalizmów. Przejawem procesu internacjonalizacji jest także przejmowanie modeli tworzenia struktur z członami obcymi w niespokrewnionych ze sobą językach, co pokazują przykłady bardzo produktywnych w ostatnich dziesięcioleciach, a zwłaszcza od przełomu XX/XXI wieku, nazw dyscyplin, poddyscyplin i kierunków naukowych, por. klasy ekwiwalentów międzyjęzykowych: ang. *text linguistics* z niem. *Textlinguistik*, pol. *lingwistyka tekstu / tekstologia lingwistyczna*, czes. *lingvistika textu / textová lingvistika*; ang. *pragmatics / pragmalinguistics*, pol. *pragmatyka językowa / pragmalingwistyka*, czes. *pragmatická lingvistika / pragmalingvistika*; ang. *conversation analysis*, pol. *analiza konwersacyjna*, czes. *konveršační analýza*; ang. *linguistic picture of the world*, pol. *językowy obraz świata*, czes. *jazykový obraz světa* i in. W niektórych klasach ekwiwalencji występują pary lub nawet trójki terminów równoznacznych z jednego języka (z każdego języka), np. pol. *lingwistyka kognitywna / językoznawstwo kognitywne*, *lingwistyka tekstu / tekstologia lingwistyczna*, czes. *lingvistika textu / textová lingvistika*, ang. *pragmatics / pragmalinguistics*, pol. *pragmatyka językowa / pragmatyka lingwistyczna / pragmalingwistyka*, czes. *pragmatická lingvistika / pragmalingvistika*.

Tworzenie klas ekwiwalentów terminologicznych złożonych z jednostek różnojęzycznych nie jest łatwym zadaniem, ponieważ wymaga od indeksatorów poza dobrą znajomością języków obcych także kompetencji językoznawczej, zwłaszcza systemów (podsystemów) terminologicznych w języku dokumentu. Terminy innojęzyczne (jednostki pojedyncze, dublety, trójki itd.) włączane są do pola klasy terminów z jednego języka, por. pol. *językoznawstwo, lingwistyka*, serb. *језикословље, лингвистика*, słoweń. *jezikoslovje, lingvistika*, głuź. *rečespyt, brus. мовознаўства, лінгвістыка*, ros. *лингвистика, языкознание*, ukr. *мовознавство, лінгвістика*, ang. *linguistics*. Innym przykładem takiej wielojęzycznej klasy w systemie iSybislaw jest klasa terminów składniowych: pol. *struktura tematyczno-rematyczna, aktualne rozczłonkowanie zdania, funkcjonalna perspektywa zdania*, ros. *актуальное членение, коммуникативная перспектива высказывания, функциональная перспектива высказывания*, czes. *aktuální větné členění, funkční větná perspektiva, větná perspektiva*, słoweń. *funkcionalna stavčna perspektiva, tematsko-rematska struktura*, ukr. *актуальне членування*. W klasie tej uwydatnia się wieloaspektowość obiektów badań, różne perspektywy metodologiczne i szkoły badawcze.

Przy gromadzeniu terminów do słowników słów kluczowych w każdym z języków używana jest metoda indukcyjno-dedukcyjna, co oznacza, że jednostki te są gromadzone z tekstów, a następnie indeksowane i sprawdzane oraz porównywane z istniejącymi reprezentatywnymi pracami, takimi jak: słowniki, tezaury, gramatyki, encyklopedie.

Prace przy tworzeniu klas ekwiwalentów przebiegają wielokierunkowo, ponieważ każdy język uwzględniany w systemie jest traktowany równorzędnie, co pozwala uniknąć błędów wynikających z asymetrii terminologicznej. Takie podejście ma zastosowanie praktyczne dla użytkowników systemu, którzy poszukują relewantnej informacji. Jednocześnie jest to element organizacji informacji mającej na celu optymalizację strategii wyszukiwawczych. Jak pokazuje praktyka pracy nad systemem, wykorzystanie terminologii w funkcji słów kluczowych w poszczególnych językach słowiańskich nie jest jednak równomierne. W wielojęzycznym zbiorze przy budowaniu klas ekwiwalentów pierwszeństwo zostało przyznane językowi polskiemu, ponieważ to w polskim środowisku językoznawczym zainicjowano prace nad systemem informacji sławistycznej i nad bazą dziedzinową. Dlatego też najliczniejszą reprezentację wśród różnojęzycznych terminów mają terminy polskie.

Prace nad polskojęzyczną częścią języka informacyjno-wyszukiwawczego zaowocowały wydaniem w 1999 roku pierwszego słownika słów kluczowych językoznawstwa sławistycznego autorstwa Zofii Rudnik-Karwatowej i Hanny Karpińskiej (Rudnik-Karwatowa-Karpińska 1999). Praca liczy około 2500 haseł (terminów językoznawczych, które są równokształtne z jednostkami języka słów kluczowych) reprezentatywnych dla polskiego językoznawstwa sławistycznego.

Słownik stanowi podstawowe narzędzie organizacji i prezentacji leksyki języka słów kluczowych równokształtnych z terminami językoznawczymi. W roku 2006 ukazała się wersja cyfrowa słownika, w której zbiór jednostek został zaktualizowany i uzupełniony o ok. 500 nowych jednostek hasłowych nieuwzględnionych w wydaniu pierwszym.

6. Wielojęzyczność systemu iSybisław

Ze względu na wielojęzyczny charakter systemu iSybisław istotne jest rozróżnienie wieloznaczności jednostek terminologicznych w obrębie jednego języka, a także wielu języków jednocześnie. Terminy wieloznaczne są prezentowane przez jednostki leksykalne, które są jednoznaczne, zaś jednostki języka słów kluczowych uzupełniane dodatkową informacją zamieszczaną w nawiasie i numeracją 1, 2, 3. W języku polskim termin wieloznaczny *język* reprezentują dwie klasy jednostek w różnych językach słowiańskich i polszczyźnie z dopowiedzeniem w nawiasie i numeracją: *język 1 (system znaków)* i *język 2 (narząd mowy)*, ros. *язык 1*, słoweń. *jezik 1*, chorw. *jezik 1*, ros. *язык 2*, słoweń. *jezik 2*, chorw. *jezik 2*. Wieloznaczny termin w języku polskim *leksykon* prezentują trzy klasy jednostek ekwiwalentnych w językach słowiańskich i polszczyźnie: *leksykon 1* ‘słownik encyklopedyczny’ *leksykon 2 (w gramatyce generatywnej)* i *leksykon 3* ‘słownictwo, leksyka’, który wchodzi w klasę ekwiwalentów z sześcioma synonimami w języku polskim: *leksykon*, *słownictwo*, *leksyka*, *słownik 2*, *zasób leksykalny*, *zasób słownikowy* i ich odpowiednikami w innych językach słowiańskich: ukr. *лексикон*, *словниковий склад*, *лексика*, słowac. *lexika*, *slovná zásoba*, czes. *lexikon 2*, *slovní zásoba*, ros. *лексика*, *лексикон*. Podobną strukturę prezentują nazwy poddyscyplin językoznawczych, por. *słowotwórstwo 1 (tworzenie wyrazów)* i *słowotwórstwo 2 (dyscyplina)*, *semantyka 1 (znaczenie znaków językowych)* i *semantyka 2 (dyscyplina)*, *fonetyka 1 (cechy fizyczne dźwięków mowy)* i *fonetyka 2 (dyscyplina)*. *słowotwórstwo 1 (tworzenie wyrazów)* i *słowotwórstwo 2 (dyscyplina)*, serb. *уворба речи 1* i *уворба речи 2*, ukr. *словотвір 1*, *словотворення* i *дериватологія*, *словотвір 2 (дисципліна)*, słoweń. *besedotvorje 1*, *besedotvorje 2*, pol. *semantyka 1 (znaczenie znaków językowych)* i *semantyka 2 (dyscyplina)*, chorw. *semantika 1 (značenje jezičnih znakova)* i *semantika 2 (disciplina)*, białorus. *семантыка 1 (значэнне моўных знакаў)* i *семантыка 2 (дысцыпліна)*, pol. *fonetyka 1 (cechy fizyczne dźwięków mowy)* i *fonetyka 2 (dyscyplina)*, słoweń. *glasoslovje 1*, *fonetika 1* i *fonetika 2*, czes. *fonetika 1*, *hláskoslovi 1*. Powyższe pary ujawniają charakterystyczny model semantyczny, w którego strukturze dochodzi do relacji ‘zjawisko’ – ‘dyscyplina (poddyscyplina) zajmująca się badaniem danego zjawiska’ (Апресян 1974: 204). Na dalszym etapie opracowywania wielojęzycznej klasy ekwiwalentów powinny być uzupełniane dopowiedzenia także w innych językach, ponieważ sama numeracja 1, 2, 3 zamieszczana przy słowach kluczowych w

innych niż polski językach słowiańskich może być dla użytkowników, którzy nie znają polszczyzny, niezrozumiała.

7. Podsumowanie

W systemie dziedzinowej informacji slawistycznej iSybislaw terminologia językoznawcza pełni funkcję słów kluczowych. Wykorzystywane w systemie jednostki terminologiczne to terminy relewantne dla współczesnego językoznawstwa slawistycznego, obecne w tekstach opisywanych w iSybislawie dokumentów. Przedstawiony w artykule model strukturalny analizy pozwala zauważyć, że w systemie w funkcji słów kluczowych przede wszystkim odnotowywane są jednostki jedno- i dwuelementowe, rzadziej występują jednostki wieloelementowe.

Wykorzystanie terminologii językoznawczej nie tylko ułatwia użytkownikom dostęp do informacji i wiedzy, ale zapewnia także kompletność i dokładność wyszukiwania relewantnych informacji zawartych w zbiorze informacyjnym systemu (przy spełnieniu przez system kryterium szczegółowości informacji). Uwzględnianie w systemie iSybislaw terminów językoznawczych wymaga rozwiązywania problemów dotyczących słownictwa języka naturalnego, jak synonimia i wieloznaczność, które są niepożądane jako jednostki leksykalne języka słów kluczowych.

Wartością systemu jest ustanowienie klas ekwiwalentnych terminów umożliwiających użytkownikom dotarcie do wszystkich pożądaných informacji, jednocześnie we wszystkich językach słowiańskich i języku angielskim, przy wyszukiwaniu za pomocą tylko jednego wybranego języka. Docelowo w systemie iSybislaw planowane jest stworzenie klas ekwiwalentów terminologicznych we wszystkich językach słowiańskich i języku angielskim.

BIBLIOGRAFIA

- Babik 2010: W. Babik, *Słowa kluczowe*, Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- Bojar–Rudnik-Karwatowa 1998: B. Bojar, Z. Rudnik-Karwatowa, *Koncepcja nowoczesnego systemu informacji slawistycznej*, Językoznawstwo : prace na XII Międzynarodowy Kongres Sławistów w Krakowie 1998, Warszawa: Wydawnictwo Energeia.
- Bojar 2002: B. Bojar (red.), *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych*, Warszawa: Wydawnictwo SBP.
- Bojar 2005: B. Bojar, *Językoznawstwo dla studentów informacji naukowej*, Warszawa: Wydawnictwo SBP.
- Bolter-Grusin 1999: J.D. Bolter, R. Grusin, *Remediation. Understanding new media*, Cambridge: The MIT Press.
- Breza 1998: E. Breza, *Nazwy obiektów i instytucji związanych z nowoczesną cywilizacją (chrematonymy)*, Polskie nazwy własne. Encyklopedia, red. E. Rzetelska-Feleszko, Warszawa; Kraków: Wydawnictwo Instytutu Języka Polskiego PAN.

- Gajda 1990: S. Gajda, *Wprowadzenie do teorii terminu*, Opole: Wyższa Szkoła Pedagogiczna im. Powstańców Śląskich w Opolu.
- Jadacka 2008: H. Jadacka, *Terminologia* [Hasło], H. Jadacka, A. Markowski, D. Zdunkiewicz-Jedynak, *Poprawna polszczyzna : hasła problemowe*, Warszawa: Wydawnictwo Naukowe PWN.
- Kaleta 1998: Z. Kaleta, *Kierunki i metodologia badań. Terminologia*, Polskie nazwy własne. Encyklopedia, red. E. Rzetelska-Feleszko, Warszawa; Kraków: Wydawnictwo Instytutu Języka Polskiego PAN, 45–81.
- Lukszyn-Zmarzer 2006: J. Lukszyn, W. Zmarzer, *Teoretyczne podstawy terminologii*, Warszawa: Wydział Lingwistyki Stosowanej i Filologii Wschodniosłowiańskich.
- Mazurkiewicz-Sułkowska 2014: J. Mazurkiewicz-Sułkowska, *Słowiańska terminologia techniczna na materiale polskim, rosyjskim i bułgarskim*, Łódź: Wydawnictwo UŁ.
- Mazurkiewicz-Sułkowska 2015: J. Mazurkiewicz-Sułkowska, Internacjonalizmy jako źródło ekonomiczności w słowiańskiej terminologii technicznej (na materiale polskim, rosyjskim i bułgarskim), *Studia z Filologii Polskiej i Słowiańskiej* 50, 166–180.
- Rudnik-Karwatowa et al. 2013: Z. Rudnik-Karwatowa, J. Banasiak, Z. Mikos, Językowe problemy optymalizacji wyszukiwania informacji w systemie iSybislaw, *Slavia Orientalis* 62, 631–646.
- Rudnik-Karwatowa 2016: Z. Rudnik-Karwatowa, Упорядкування термінології для потреб інформаційно-пошукової системи iSybislaw, *Українська мова* 3, 105–117.
- Rudnik-Karwatowa-Karpińska 1999: Z. Rudnik-Karwatowa, H. Karpińska, *Słownik słów kluczowych językoznawstwa slawistycznego*, Warszawa: Slawistyczny Ośrodek Wydawniczy.
- Апресян 1974: Ю. Д. Апресян, *Лексическая семантика. Синонимические средства языка*, Москва: Наука.
- Гринев-Гриневиц 2008: С.В. Гринев-Гриневиц, *Терминоведение. Учебное пособие для студентов высших учебных заведений. Учебное издание*, Москва: Издательский центр „Академия“.

Paweł Kowalski
Zofia Rudnik-Karwatowa

THE USAGE OF TERMINOLOGY IN THE ISYBISLAW INFORMATION RETRIEVAL SYSTEM OF THE SLAVIC LINGUISTIC BIBLIOGRAPHY

Summary

The paper describes the usage of linguistic terminology in the online iSybislaw bibliographic database which represents the information retrieval system of the Slavic linguistics bibliography (www.isybislaw.ispan.waw.pl). The concept of the system was presented at the International Congress of Slavists in Krakow in 1998. iSybislaw consists of works in the field of Slavic linguistics, contrastive studies, Slavic – non-Slavic contrastive studies, and to a lesser degree, general and theoretical linguistics. The linguistic terminology reflected in the keyword language is the core of one of the information retrieval tools (IR tool) in the iSybislaw system: Its primal function is the retrieval information function. The iSybislaw system is multilingual, and it provides information in all Slavic languages as well as in English. The languages of the gathered documents represent all natural languages. Due to the multilinguality of the system, methodological problems such as synonymy

and polysemy need to be resolved. The constructors of the system use the original methodological approach to create classes of equivalent terms that are relevant for current Slavic linguistics. The classes of terms in one language are combined with those in another and are joined in a multilingual class of equivalence. The use of linguistic terminology in the function of keywords in the iSybislaw system provides efficiency of information retrieval for cross-lingual searches and multilingual users. The paper focuses on the theoretical and practical usage of linguistic terminology in the iSybislaw system.

Keywords: the bibliographical database of Slavic linguistics, iSybislaw, language of keywords, Slavic linguistics, information and retrieval system, terminology, linguistic terminology