

Łukasz Karpiński

Uniwersytet Warszawski

OD SŁOWNIKA DO SYSTEMU TERMINOLOGICZNEGO

„Koniec z wersją drukowaną encyklopedii Britannica:
online pokonał druk”

CHIP nr 3/2012

SUMMARY

From dictionary to terminological system

The article presents the range of lexicographic work in a historical context until the breakthrough moment in lexicography i.e. the decision of the publishers of the Encyclopedia Britannica to offer their product only in digital version. This fact closes certain period in the history of lexicography and it is a good starting point to present an original concept of «terminological system» as a universal database which may be the basis of a variety of dictionaries on different subjects. The article refers to earlier publication, it combines earlier presented concepts in order to offer a consistent and relatively comprehensive approach to the issues of contemporary lexicography.

Key words: lexicography, dictionary, terminological system, database, frequency list.

РЕЗЮМЕ

От словаря к терминологической системе

В первой части статьи указан очерк лексикографических работ в историческом контексте, законченный переломным моментом в лексикографии т. е. объявлением решения владельцами энциклопедии Britannica о переходе всех ее ресурсов исключительно в сеть Интернет. С настоящего момента энциклопедия будет доступна только в электронной версии. Этот факт символически подытоживает прежний период в истории лексикографии и представляет собой введение к авторской концепции

„терминологической системы”, т.е. универсальной базы данных, составляющей основу для разнообразных по тематике и конструкции словарей.

Статья ссылается непосредственно на предыдущие публикации, соединяя указанные тезисы в сплоченный обзор вопросов современной лексикографии.

Zadania leksykografii koncentrują się na działaniach związanych z metodyką gromadzenia i analizy materiału językowego, następnie jego sprawnego przetwarzania i systematyzowania w celu dalszego wykorzystania w różnego typu opracowaniach. Zazwyczaj są to różnorodne postacie słowników (językowych, encyklopedycznych, dydaktycznych czy terminologicznych)¹, jedno- lub wielojęzycznych, których różnorodność uwarunkowana jest najsilniej potrzebami rynkowymi oraz prowadzonymi badaniami lingwistycznymi. Całokształt tych działań wraz z podbudową teoretyczną oraz oceną wpływu słowników na kulturę i społeczeństwo stanowi zasadniczy trzon tej dziedziny, korzystającej przy tym z osiągnięć dziedzin pokrewnych, m.in. logiki, pedagogiki, teorii komunikacji czy informatyki.

Ponieważ każdy leksykon ma z definicji w jak najwierniejszy sposób odzwierciedlać zmiany zachodzące w języku (i jego odmianach specjalistycznych), a dynamika tych procesów nieustannie rośnie, również procesy leksykograficzne zaczęły korzystać z najnowszych technologii. Rozwój elektronicznych maszyn liczących, które z czasem przekształciły się w komputery sensu stricte, bardzo szybko został przeniesiony na grunt słownictwa. W połowie XX wieku pada znamienne pytanie A. Turinga *Can machine think?*², a niedługo później „prowadzone są szerokie prace mające na celu zbudowanie urządzeń czytających, które pozwoliłyby wprowadzać do maszyn teksty w języku naturalnym, oraz urządzeń zdolnych wprowadzać do maszyn mowę ludzką; (...) studiuje się zagadnienie maszynowego przekładu z jednego języka naturalnego na inny...”³. W tym samym okresie listy frekwencyjne, uzyskiwane na podstawie analizy określonych tekstów lub pierwszych, ograniczonych sprzętowo korpusów tekstowych zaczynają odgrywać coraz większą rolę⁴. Współcześnie analizy statystyczne tekstu stanowią nie tylko podstawę ekscerpowania terminologii, ale służą też do określania indywidu-

¹ Por. Ł. Karpiński, *Prezentacja terminologii w słowniku w ujęciu typologicznym*, „Linguodidactica” 2010, t. XIV, s. 87-100.

² A. M. Turing, *Computing machinery and intelligence*, „Mind”, October 1950, s. 433-460.

³ J. Lewin, *Znak, język, matematyka*, Warszawa, 1967, s. 13.

⁴ Zob. np. A. Lewicki, J. Sambor, *Projekt słownika frekwencyjnego współczesnego języka polskiego, sprawozdanie I* Wyd. PAN, z. 4, 1969, s. 117-129; J. Sambor, *Badania statystyczne nad słownictwem (na materiale Pana Tadeusza)*, Wrocław 1969, s. 163.

alnych cech danego stylu tekstu czy weryfikowania oryginalności autorstwa poszczególnych dzieł. W przypadku języków specjalistycznych analiza frekwencyjna może stanowić element pracy terminologicznej, charakteryzującej językowo daną dziedzinę wiedzy, szczególnie w zakresie takich rodzajów pracy terminologicznej jak: regulowanie (ustalenie udokumentowanej potrzebą zawodową normy zapożyczeń), standaryzacja (kojarzenie zbioru danych językowych z prototypem terminologicznym), harmonizacja (konfrontacja leksykonu z terminologią odpowiednich języków światowych) czy unifikacja (podporządkowanie zbioru słownictwa normom danego języka narodowego)⁵. Lista rangowa, rozumiana jako lista słów posortowanych malejąco względem liczby wystąpień w danym tekście (korpusie), z powodzeniem jest wykorzystywana także do wyodrębniania różnych typów jednostek słownictwa specjalistycznego⁶.

Wraz z doskonaleniem pamięci i nośników danych otwierają się nowe możliwości w zakresie gromadzenia danych słownikowych w formie elektronicznej. Regały pełne kart z artykułami hasłowymi zostają sukcesywnie wyparte przez zapisane cyfrowo dane. K. Leski zwracał uwagę, iż jest niewiele miejsca na zapisy na takiej karcie, „w razie potrzeby można wykorzystać dwie lub więcej kart, nie jest to jednak wygodne. (...) W przypadku tezaurusów technicznych i innych, w których występuje np. wiele szczegółowych typów urządzeń o przeznaczeniu identycznym lub bardzo sobie bliskie, a także w których bezpośrednie powiązania hierarchiczne i kojarzeniowe są liczne – powierzchnie poszczególnych rubryk tej karty będą zbyt małe. Dla zbiorów terminów dla takich tezaurusów stosować trzeba karty większe. Mogą to być więc karty np. formatu A5 porubrykowane w sposób identyczny. W niektórych przypadkach trzeba będzie stosować nawet karty w formacie A4⁷. Obecnie tradycyjnie pojmowana karta terminologiczna nie stanowi wyjściowego punktu gromadzenia materiału leksykograficznego, a może być jedynie pochodną komputerowej bazy danych – wydrukiem wykonanym w zadanym formacie, o ile jest współcześnie jeszcze taka potrzeba. Zamiast rubrykować kartki A6 proces gromadzenia danych do słownika rozpoczyna się od zaprojektowania bazy danych i określenia ilości rekordów (komórek składających się na kompletny opis danego hasła), które będą potrzebne.

5 Przykład wnioskowania na podstawie listy frekwencyjnej zob. Ł. Karpiński, P. Michałowski, *Wybrane metody analizy terminologii specjalistycznej (na przykładzie technoloektu geografii)*, „Edukacja dla Przyszłości” 2012, t. IX, Białystok, s. 23-26; podane pojęcia stosowane w rozumieniu prezentowanym w pracy: J. Lukszyn, W. Zmarzer, *Teoretyczne podstawy terminologii*, Warszawa 2001, s. 86-89.

6 Przykład wyodrębniania hipoterminów oraz kryteria tego doboru zob. Ł. Karpiński, *W kwestii statusu terminologicznego hipoterminu w językach specjalistycznych*, „Linguodidactica” 2011, t. XV, s. 89-100.

7 K. Leski, *Zasady budowy tezauryusa*, Warszawa 1978, s. 52.

Kolejny krok w kierunku cyfryzacji procesu leksykograficznego to tworzenie korpusów tekstowych – zbiorów tekstów pokazujących poszukiwane wyrazy w rzeczywistym, używanym przez ludzi kontekście. Korpusy językowe czerpią z artykułów prasowych i popularnonaukowych, prac naukowych, opracowań, dokumentów oficjalnych, fragmentów współczesnych dzieł literackich, podręczników, mogą również zawierać transkrypcje zwykłych rozmów aby rejestrować również język potoczny. Korpus tekstów musi być odpowiednio zrównoważony gatunkowo, chronologicznie, stylowo, terytorialnie i demograficznie, aby jak najlepiej odzwierciedlać realny stan rozwoju danego języka. To właśnie założona uprzednio struktura różni korpusy choćby od ogólnych zasobów sieci. Obecnie każde znane na rynku wydawnictwo informuje, iż aktualizacja leksyki została dokonana na podstawie aktualnych korpusów tekstowych.

Rozwój technik komputerowych stał się także w pewnym momencie szansą dla drukowanych prac leksykograficznych. Skład komputerowy przestał być czasochłonnym procesem, przez co wydawnictwa redukowały koszty materiałowe, kładąc większy nacisk na szybkość aktualizacji materiału językowego. Na marginesie można zauważyć, że gdy redukcja kosztów dotykała skład osobowy, przygotowujący dany leksykon, natychmiast odbijało się to na jego jakości, a bezkrytyczne i tanie kopiowanie wydawnictw konkurencyjnych w prosty sposób prowadziło do pozwów, rozpraw i wyroków sądowych. Wydaje się jednak, że apogeum tego typu działalności miało miejsce w latach minionych.

Ostatecznym ciosem zadany tradycyjnym technikom leksykograficznym był rozwój komunikacji sieciowej. Różnego rodzaju leksykony dostępne bezpłatnie w sieci Internet przewartościowały rynek słownikarski. Dla renomowanych wydawnictw stało się jasne, iż nawet kompozycja „tom książkowy + nośnik danych (CD, DVD, pen-drive)” nie wystarczy, a przyszłość leksykografii leży w sieci.

W czasopiśmie komputerowym „CHIP” nr 3/2012 ukazał się artykuł pod znamienym tytułem: „Koniec z wersją drukowaną encyklopedii Britannica: online pokonał druk”. Podane w artykule uzasadnienie decyzji brzmi: „Szef Encyklopedii Britannica, Jorge Cauz, zdaje sobie sprawę, jak twierdzi, że nie wszystkim ta decyzja się spodoba. Jego zdaniem pozwoli to jednak jego firmie zainwestować więcej czasu i nakładów w edycję online. Encyklopedia ma być jeszcze częściej aktualizowana, ma być w niej jeszcze więcej artykułów, a edycja cyfrowa daje, jego zdaniem, dużo szersze możliwości z uwagi na brak limitu objętościowego”. W stosunku do pierwszego wydania tej encyklopedii z lat 1768–1771 składającej się z 3 tomów jest to więcej niż rewolucja. Transformacja encyklopedii, która do tej pory sprzedała ponad 7 mln drukowanych zestawów, ma jeszcze jeden wymiar. Do tej pory media chętnie porównywały ją z Wikipedią. Obok tej nowoczesnej internetowej bazy wiedzy, która zawsze jest w zasięgu ręki, Britannica wyglądała na przestarzałą i niewygodną w użyciu. Jednak teraz obie encyklopedie będą dostęp-

ne przez Internet, a konkurencja między nimi sprowadzi się do wyboru pomiędzy płatnym i darmowym, profesjonalnym i amatorskim serwisem⁸. Na elektronicznym rynku dojdzie więc do bezpośredniego starcia dwóch filozofii leksykograficznych. Warto w tym momencie wspomnieć, iż sama Wikipedia powstała jako uzupełnienie profesjonalnego projektu „Nupedia”⁹, a więc już raz w swojej historii wyparła bardziej renomowany i redagowany przez ekspertów leksykon.

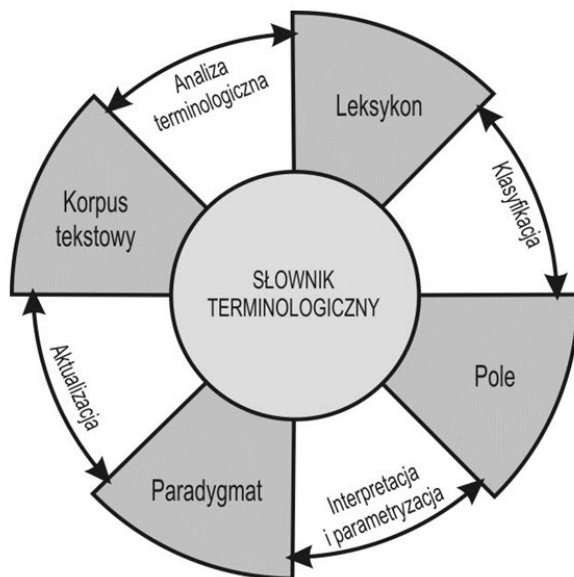
Słowniki w tradycyjnym ujęciu, tj. zbiory wyrazów ułożonych zazwyczaj alfabetycznie i opracowanych w formie artykułów hasłowych według pewnych założonych zasad i skierowane do określonych użytkowników przestają decydować obecnie o sytuacji na rynku leksykograficznym. Jak wspomniano wcześniej, za podstawę prac przyjmuje się obecnie stworzenie odpowiedniej bazy danych, które mogą być udostępniane poprzez sieć czy fizyczne nośniki danych na różne sposoby¹⁰.

W obecnej sytuacji za pełnoprawnego następcę tradycyjnego słownika terminologicznego można uważać „system terminologiczny”, rozumiany jako uniwersalna baza danych terminologicznych, składająca się z uszeregowanych tematycznie parametrów w celu uzyskania optymalnej możliwości selekcji informacji w zależności od potrzeb użytkowych, jak najdokładniej odzwierciedlających strukturę leksykalną, gramatyczną i systemową danego języka (specjalistycznego). Podstawowe założenia koncepcyjne tego systemu można przedstawić w następujący sposób:

⁸ Źródła wypowiedzi: <http://www.chip.pl/news>; <http://wyborcza.pl>; <http://corporate.britannica.com/about/today/>, dostęp: lipiec-wrzesień 2012.

⁹ Internetowa encyklopedia, istniejąca od 9 marca 2000 do 26 września 2003. Cechą charakterystyczną Nupedii był wielostopniowy proces redakcyjny, którego celem było stworzenie haseł o jakości porównywalnej z komercyjnymi encyklopediami. Przez cały okres swojego istnienia Nupedia stworzyła 24 artykuły, które przeszły przez proces redakcyjny aż do wersji ostatecznej, a także 3 artykuły w wersjach zbliżonych do finalnych. 74 artykuły Nupedii były nieukończone i znajdowały się w wersjach roboczych o różnym stopniu zaawansowania [źródło: pl.wikipedia.org, dostęp: 20.10.2012].

¹⁰ To właśnie posiadanie udokumentowanej bazy danych, a nie struktura haseł słownikowych były podstawą do wydania wyroku w sprawie o naruszenie praw autorskich twórców *Słownik języka polskiego*. W wyroku Sądu Najwyższego w tej sprawie pada znamienne zdanie: „Nie ulega wątpliwości, że (...) stworzenie słownika języka polskiego – oprócz wiedzy z zakresu gramatyki i leksykografii – wymaga zebrania tzw. korpusu, czyli grupy tekstów źródłowych, w których wystąpiło dane hasło. Jak stwierdził biegły, bez stworzenia tzw. korpusu, który w przypadku słownika języka polskiego określono na około 30 milionów jednostek tekstowych, językoznawca nie jest w stanie opracować słownika, któremu nie można zasadnie postawić zarzutu naśladownictwa” [Wyrok z dnia 15 listopada 2002 r., II CKN 1289/00, OSNC 2004/3/44, Biul. SN 2003/6/10].



Rys. 1. Układ koncepcyjny systemu terminologicznego (schemat autora)

Z powyższego schematu wynika, iż zasadnicze prace leksykograficzne można podzielić na 8 części, z których połowa odnosi się do zbieranych i wytwarzanych w bazie terminologicznej danych, a druga część oznacza procesy weryfikujące i porządkujące. Całokształt prac stanowi spójny układ działań, które można stosować w zależności od potrzeb.

1) W schemacie przyjęto, iż początek procesu leksykograficznego znajduje się na styku pojęć „Aktualizacja” i „Korpus tekstowy”. W przypadku przejścia od bardziej tradycyjnych metod tworzenia słowników, wprowadzenia lub przetworzenia posiadanych drukowanych danych leksykograficznych weryfikacja ogólnej koncepcji słownika (makrostruktury) i postawienie konkretnych celów przed kompilowanym korpusem tekstowym będzie podstawowym procesem leksykograficznym¹¹. Do szeroko pojętych zadań zaliczanych do makrostrukturalnego projektowania można zaliczyć:

- określenie kompetencji autora (zespołu autorskiego) słownika;
- określenie typu odbiorcy i optymalnej objętości słownika;
- określenie zakresu tematycznego, a w związku z tym ukierunkowania bazy konceptualnej;
- określenie źródeł pozyskiwania terminologii dla potrzeb słownika;

¹¹ Por. Ł. Karpiński, *Zarys leksykografii terminologicznej*, Warszawa 2008, s. 72-73.

- zaplanowanie zamieszczenia w bazie danych odpowiednich ekwiwalentów w językach obcych;
- opracowanie elementów uzupełniających wartość merytoryczną słownika;
- stworzenie bazy danych dla potrzeb gromadzenia i analizy materiału leksykalnego.

Znaczenie korpusu tekstowego zostało dość szczegółowo opisane w pierwszej części artykułu i w związku z tym, na zakończenie można dodać, iż efektem końcowym tej części pracy będzie zebrany odpowiedni korpus tekstowy.

2) Analiza terminologiczna – wielość typów konkretnej pracy terminologicznej w literaturze jest określana następującymi parametrami: porządkowanie, harmonizacja, regulowanie, normalizacja, standaryzacja, unifikacja, kodyfikacja i reglamentacja¹²:

- najbliższa poprzedniemu etapowi w procesie leksykograficznym jest kodyfikacja, gdyż określa obecność danych jednostek terminologicznych w autorytatywnym korpusie tekstów; w jej wyniku weryfikowany jest zebrany materiał językowy, rejestrowane są zmiany, nowe wyrażenia, zmiana zakresu znaczeniowego itp.;
- porządkowanie jest procesem określania miejsca danego leksykonu terminologicznego w układzie dostępnych języków specjalistycznych, odzwierciedlających wszystkie dostępne rodzaje wiedzy fachowej; ponieważ żadna dziedzina wiedzy nie funkcjonuje w oderwaniu od innych – pokrewnych, należy skorelować wielkość zebranego korpusu tekstowego z założoną ilością rekordów (w tradycyjnej leksykografii rozumianych jako hasła słownikowe);
- harmonizacja terminologii rozumiana jest jako porównanie danych terminologicznych zawartych w słowniku z tzw. prymarnym systemem terminologicznym odpowiedniego języka światowego, w rezultacie czego następuje weryfikacja słownictwa, oceniana jest aktualność korpusu i zgodność znaczeniowa, leksykalna i gramatyczna z nadrzędnymi normami branżowymi i terminologicznymi; ważnym elementem jest także określenie stopnia terminologizacji danego technolektu i rozróżnienie jednostek języka na terminy, quasi-terminy, hipoterminy i in.¹³
- regulowanie zakłada stosowanie rekomendacji terminologicznych dotyczących stosowania zapożyczeń, neosemantyzmów i neologizmów terminologicznych, a także jednostek z pogranicza języków specjalistycznych

¹² Por. J. Lukszyn, W. Zmarzer, *Teoretyczne podstawy terminologii*, Warszawa 2001, s. 86-89.

¹³ Por. Ł. Karpiński, *Jednostki języka specjalistycznego w aspekcie statusu terminologicznego*, „Linguodidactica” 2007, t. XI, s. 63-74 oraz Ł. Karpiński, *W kwestii statusu terminologicznego hipotermynu w językach specjalistycznych*, „Linguodidactica”, XV, 2011, s. 89-100.

- wyrażen żargonowych, które z czasem mogą wyjść z użycia, ale niewykluczone, że mogą uwiarygodnić swój status terminologiczny w języku specjalistycznym wypierając dotychczasowy termin;
- normalizacja uwzględnia międzynarodowe standardy i wytyczne w zakresie terminologii ustalone przez międzynarodowe organizacje, a w kraju przez Polski Komitet Normalizacyjny. Działania w ramach tego procesu mają na celu aktualizację systemu terminologicznego danego technolektu, co bezpośrednio przekłada się na analizę internacjonalizmów w danej dziedzinie oraz sprawdzenie zgodności z na bieżąco wydawanymi normatywami ISO, IEN czy ITU;
- standaryzacja odnosi się do procedur porównania zbioru terminologii z właściwym prototypem terminologicznym, tj. modelem tworzenia i/lub modyfikowania nowych jednostek wyrazowych – listy frekwencyjne lub indeksy a`tergo mogą w tym wypadku wesprzeć proces wyszukiwania dotychczas stosowanych w danej terminologii prefiksów czy sufiksów oraz sposobów tworzenia neologizmów;
- unifikacja polega na podporządkowaniu zebranego materiału regułom i normom danego języka narodowego zarówno pod kątem formalnym, jak i znaczeniowym, a więc ustalenia wzoru odmiany, sposobu pisowni i in.;
- reglamentacja polega na wyłonieniu zbędnych synonimów i dubletów, niejednoznacznych i okazjonalnych jednostek terminologicznych, czasowo pojawiających się w językach specjalistycznych.

W wyniku analizy terminologicznej zebrany materiał językowy powinien zostać uporządkowany i uaktualniony, co pozwoli przejść do dalszego etapu kompilowania bazy danych i wywoływania zawartych w niej danych w różnorodnych leksykonach.

3) Pojęcie leksykonu w kontekście systemu terminologicznego jest ściśle uwarunkowane założeniami makrostruktury słownika (jego ogólnej koncepcji użytkowej) oraz przyjętej typologii słowników¹⁴. W przedstawionej wyżej koncepcji systemu terminologicznego słownik nie jest stricte podmiotem działań leksyko-graficznych, a jego typ, stopień dokładności¹⁵ czy planowane zastosowanie de-

¹⁴ Leksykony ze względu na rodzaj prezentacji wiedzy można podzielić na następujące kategorie: słowniki językowe (SJ), słowniki encyklopedyczne (SE), słowniki dydaktyczne (SD) i słowniki terminologiczne (ST), zob. Ł. Karpiński, *Prezentacja terminologii w słowniku w ujęciu typologicznym*, „Linguodidactica” 2010, t. XIV, s. 89-91.

¹⁵ Por. P. Michałowski, *W poszukiwaniu idealnego modelu słownika terminologii przedmiotowej*, „Języki specjalistyczne” 2004, t. 4, s. 196-197 – podział słowników pod względem stopnia dokładności.

terminują jedynie wywołanie poszczególnych elementów zebranych dla potrzeb danego rekordu (hasła słownikowego). W zależności od potrzeb edytorskich z bazy będą pobierane dane odpowiadające np. jednojęzycznemu słownikowi objaśniającemu, encyklopedii branżowej czy słownikowi przekładowemu według określonego klucza tak, aby mikrostruktura docelowego słownika (wyświetlanego w formie elektronicznej czy dostępnego w wersji drukowanej) odpowiadała przyjętym normom. Budowa każdego rekordu bazy, zawierającego praktycznie wszystkie funkcjonujące w literaturze parametry, sprowadzi słownik w formie drukowanej do dobranego tematycznie przez użytkownika wydruku z bazy, a słownik elektroniczny – do interfejsu odwołującego się do danych zawartych na nośniku lub witryny internetowej, umożliwiającej dostęp do wykupionych z góry opcji dostępu. Ta sama baza danych powinna być kompatybilna z algorytmami translatorów maszynowych, a więc jedna baza stworzona według wszelkich zasad leksykografii i inżynierii może znaleźć kilka równoległych zastosowań. W Tabeli 1 przedstawiono zestawienie parametrów mikrostruktury, stanowiących podstawę projektowania baz danych¹⁶:

Tabela 1

Lp.	Parametry mikrostruktury	Wybrane przykłady
1	Rejestracyjne	numer hasła, źródło, komentarz, data rejestracji terminu
2	Formalne	ortografia, transkrypcja, część mowy, warianty gramatyczne, skróty
3	Etymologiczne	język źródłowy, język pośrednik, formy paralelne
4	Leksykalne	ugrupowanie tematyczne, typ terminu, rejestr stylistyczny
5	Interpretacyjne	różne typy definicji
6	Asocjacyjne	synonimy, antonimy, homonimy termin nadrzędny, podrzędny, terminy skojarzeniowe
7	Pragmatyczne	neologizmy, błędne znaczenia, terminy autorskie, terminy standardowe
8	Ilustracyjne	diagramy, rysunki, wykresy
9	Graficzne	liternictwo, symbole, wyróżnienia tekstowe
10	Ekwiwalencyjne	ekwiwalenty w językach obcych
11	Użytkowe	kolokacje, teksty standardowe, korpusy
12	Klasyfikacyjne	położenie terminu w klasyfikacji

¹⁶ Por. С. В. Гринев, *Введение в терминологию*, Москва 1995, s. 48-51; J. Lukszyn, W. Zmarzer, *Teoretyczne podstawy terminologii*, Warszawa 2001, s. 136, Ł. Karpiński, *Zarys leksykografii terminologicznej*, Warszawa 2008, s. 111.

4) Klasyfikacja terminologii stanowi bezpośrednie odwołanie do systemowości języków specjalistycznych wyrażanej poprzez zestaw relacji semantycznych w układzie nadrzędności, podrzędności i wielostronnego kojarzenia. W autorskiej koncepcji proponowana jest klasyfikacja modułowa¹⁷, wywodząca się z metodologicznych założeń porządkujących procesy decyzyjne oraz teorii (wielkich) systemów. Zastosowanie teorii systemów w pracach terminograficznych wydawało się oczywiste ze względu na:

- bezpośrednią zależność systemu terminologicznego od rozwoju danej dyscypliny;
- uwzględnianie w teorii systemów wielu czynników zewnętrznych, odzwierciedlających się w terminologii poprzez wpływ dziedzin pokrewnych na zakres tematyczny oraz potrzeb potencjalnych użytkowników słownika;
- każdy system musi uwzględniać nanoszenie poprawek i stałe aktualizowanie elementów;
- każdy system powinien być w spójny i logiczny sposób dzielony na mniejsze elementy, współdziałające ze sobą.

Można stwierdzić, iż każda logicznie spójna metoda podziału danego zbioru terminologii na mniejsze podzbiory powinna usystematyzować daną terminologię, lecz klasyfikacja modułowa proponuje w tym względzie 3-stopniową stałą, wyczerpującą i rozłączną strukturę. Każdy termin, który wejdzie w jej skład, znajdzie swoje jedyne i ściśle określone miejsce, a wewnętrzna jednorodność struktury umożliwi porównywanie danych z wielu klasyfikacji tematycznych, poszerzając możliwości w zakresie sortowania i obróbki materiału terminologicznego.

5) W wyniku klasyfikacji terminologii otrzymana struktura pozwoli w prosty sposób wyodrębnić poszczególne pola terminologiczne rozumiane jako podzbiory terminów powiązanych ze sobą tematycznie w ramach większego zbioru. „Przeciętnie objętość pola terminologicznego waha się w granicach od kilkudziesięciu do 150–200 jednostek terminologicznych. Wyraźne zwiększenie tej liczby prowadzi do przekształcenia się pola terminologicznego w odrębny leksykon”¹⁸. Wyodrębnienie pola terminologicznego stanowi bezpośrednie wprowadzenie do weryfikacji definicji wszystkich jego elementów.

6) Struktura znaczeniowa pola terminologicznego może być poddawana wnikliwej analizie w celu doprecyzowania znaczenia każdego elementu. Zmiany, wykryte podczas wykonywania prac terminologicznych, mają bezpośredni

¹⁷ Zasadnicze założenia klasyfikacji modułowej i sposoby jej kształtowania przedstawiono m.in. w: Ł. Karpiński, *Model kompilacji modułowego teaurusu terminologii specjalistycznej*, „Linguodidactica” 2006, t. X, s. 57-67.

¹⁸ J. Lukszyn, W. Zmarzer, *Teoretyczne podstawy terminologii*, Warszawa 2001, s. 56.

wpływ na aktualizację znaczenia danego wyrażenia. W przypadku uniwersalnego systemu terminologicznego ważne jest odpowiednie konstruowanie definicji każdego pojęcia. Aby otrzymać zestaw cech, charakteryzujących dane pojęcie, należy rozważyć wszelkie możliwe sposoby i okoliczności występowania danego zjawiska (obiektu lub abstraktu), w czym pomocne będzie „umieszczenie” go w kolejnych przestrzeniach pojęciowych¹⁹. Proponowane etapowe „umieszczanie” definiowanego pojęcia pozwoli określić elementy składowe końcowej definicji.

Drugą część procesu definiowania stanowi tzw. parametryzacja danego pojęcia, będąca de facto próbą stworzenia cyfrowej definicji, ukierunkowanej na zastosowanie jej w translatorach maszynowych, korzystających poprzez swoje algorytmy z zamieszczonych w bazie danych językowych. Parametryzacja pojęć jest autorskim opracowaniem powstałym na bazie arystotelesowskiej zasady percepcji świata na klasy (akcydensy), uwzględniającej późniejszą krytykę filozoficzną oraz współczesne możliwości w zakresie informatycznych technik obliczeniowych. W każdej klasie na obecnym etapie wyodrębniono trzy podklasy, a każda z nich może przyjmować różne wartości liczbowe, uzależnione od natężenia danego zjawiska. Wielopłaszczyznowy opis parametryczny w procesie translacji zbliża bazę danych do uproszczonego modelu myślowego: algorytm programu (translatora maszynowego) wyodrębniając dane słowo (wyrażenie), odnajduje jego opis w bazie danych, odczytuje składowe parametry kategoryzujące dane pojęcie (co w procesach myślowych człowieka można by określić jako określenie kontekstu), wyszukuje ekwiwalent w innym języku o możliwie najbliższej strukturze analogicznych parametrów. Alternatywną propozycją jest śledzenie zmian wartości parametrów w całym tekście, a następnie zastosowanie lub sugerowanie przez algorytm translatora maszynowego takiego doboru ekwiwalentów, aby odwzorowywały one w przybliżeniu zmiany parametrów w języku docelowym. W takim wypadku mocniejszy nacisk kładziony jest na tłumaczenie kontekstowe, które w podejściu parametrycznym nabiera jakościowo nowego znaczenia²⁰.

¹⁹ Metoda zaadaptowana z: A. Kaufmann, M. Fustier, A. Drevet, *Inwentyka. Metody poszukiwania twórczych rozwiązań*, Warszawa 1975, s. 93-95, rozwinięcie procesu definiowania przedstawiono w: Ł. Karpiński, *Od cechy dystynktywnej do definicji realnej*, „Linguodidactica” 2009, t. XIII, s. 123-136. Trzy przywołane przestrzenie to: *przestrzeń afektywna* – otoczenie społeczne, w którym rozpatrujemy dane pojęcie, *przestrzeń funkcjonalna*, która rozpatruje pojęcie w kontekście użycia, funkcji czy możliwości, jakie oferuje oraz *przestrzeń technologiczna*, która ujmuje zjawisko z pozycji określenia jego elementów składowych, sposobu jego wykonania czy realizacji określonych działań.

²⁰ Obecnie definicje cyfrowe obejmują 27 parametrów i opisują takie części mowy jak spójniki, wykrzykniki i modulanty, gdyż grupa tych wyrażeń zawiera dość mocno zróżnicowane znaczeniowo elementy, pozwalające uzyskać przy komputerowej analizie tekstów wyraźne dane. Badania

7) Paradygmat terminologiczny oznacza „podzbiór terminów, charakteryzujący się spójnością, która polega na zachodzeniu między tymi terminami stałych relacji semantycznych”²¹. W układzie koncepcyjnym systemu terminologicznego paradygmat, oprócz podanych cech, odnosi się także do realizowanych w tekstach stałych lub nowopowstałych kolokacji. Relacje semantyczne rejestrowane w bazie danych pod wpływem analiz zebranego i aktualizowanego korpusu tekstowego mogą ulegać pewnym modyfikacjom. Również w wyniku prowadzenia aktualizacyjnych prac terminologicznych mogą zaistnieć przesłanki do weryfikacji relacji semantycznych, co może pociągnąć za sobą zmiany w definiowaniu i parametryzowaniu oraz potrzebę korekty pola terminologicznego i ewentualnie pozostałych elementów systemu terminologicznego.

8) Aktualizacja stanowi uzupełnienie spójności systemu terminologicznego. Na podstawie zmian językowych rejestrowanych w korpusach tekstowych lub w tekstach normalizacyjnych można wnosząc na bieżąco poprawki do bazy danych, co ponownie podkreśla przewagę elektronicznych systemów terminologicznych nad słownikami drukowanymi. Obecność procesu aktualizacji w układzie prac leksykograficznych jest oczywista, stąd postulatem wydaje się jedynie doskonalenie oprogramowania pozwalającego w sprawny sposób wychwytywać wszelkie zmiany w strukturach znaczeniowych lub formalno-gramatycznych uzupełnianego korpusu tekstowego. Zazwyczaj zmiany te wynikają z indywidualnych doświadczeń (interpretacji) użytkowników danego języka, a specyfika bezpośrednich kontekstów czy szeroko pojętej konsytuacji komunikacyjnej stanowi dodatkowy element twórczo wpływający na rozwój lub zmiany w języku. Stąd też system leksykograficzny powinien uwzględniać również czynniki wpływające na jakość komunikacji międzyludzkiej²². W tym momencie obieg informacji w systemie terminologicznym, przedstawiony na rysunku 1, wykonał pełny cykl, co można wiązać w tradycyjnym ujęciu z procesem wydawniczym słownika, a współcześnie – z przygotowaniem aktualnej bazy danych, mogącej stanowić podstawę różnorodnych publikacji leksykograficznych i oprogramowania tłumaczeniowego.

skuteczności definicji cyfrowych obecnie trwają i będą przedmiotem osobnej publikacji. Więcej w: Ł. Karpiński, *Semantyka komunikatu a maszynowy zapis parametryzacji tekstu: Pantext 2.0*, „Komunikacja Specjalistyczna” 2012, t. 5, s. 21-32.

²¹ *Słownik terminologii przedmiotowej*, J. Lukszyn (red.), Warszawa 2005, s. 77.

²² Założenia przebiegu wydarzenia komunikacji językowej (WKJ) zostały przedstawione szerzej w artykułach: Ł. Karpiński, *Aby zrozumieć... – mechanizmy dyskursu specjalistycznego*, „Linguodidactica” 2006, t. X, s. 69-82 oraz Ł. Karpiński, *Rozumienie a zakłócenia w kanale informacyjnym*, „Linguodidactica” 2007, t. XI, 2007, s. 75-82.

Zamierzeniem autora było przedstawienie współczesnych prac leksykograficznych w postaci systemu wzajemnie przenikających się elementów. Te części składowe w większości były już stosowane w praktyce słownikarskiej, jednak ich umiejętnie skorelowanie wraz z unifikacją tworzenia baz danych terminologicznych i rozwojem parametryzacji cyfrowej powinny doprowadzić w przyszłości do stworzenia uniwersalnej bazy danych, z której zasobów korzystaliby zarówno ludzie, jak i translatory maszynowe. A może będą one również zapleczem dla wymarzonych przez ludzi „sztucznych inteligencji”?

Bibliografia

- Hetmański M., *Umysł a maszyny. Krytyka obliczeniowej teorii umysłu*, Lublin 2000.
- Kardela H., Schemat i prototyp w morfologii kognitywnej, „Kognitywistyka. Problemy i perspektywy”, Kardela H., Muszyński Z., Rajewski M. (red.), Lublin 2005.
- Karpiński Ł., *Model kompilacji modułowego teaurusu terminologii specjalistycznej*, „Linguodidactica” 2006, t. X, s. 57-67.
- Karpiński Ł., *Aby zrozumieć... – mechanizmy dyskursu specjalistycznego*, „Linguodidactica” 2006, t. X, s. 69-82.
- Karpiński Ł., *Rozumienie a zakłócenia w kanale informacyjnym*, „Linguodidactica” 2007, t. XI, s. 75-82.
- Karpiński Ł., *Jednostki języka specjalistycznego w aspekcie statusu terminologicznego*, „Linguodidactica” 2007, t. XI, s. 63-74.
- Karpiński Ł., Ruciński M., *Niektóre problemy tłumaczeniowe w pracach legislacyjnych Unii Europejskiej na przykładzie terminologii z zakresu rybactwa i polityki morskiej*, „Linguodidactica” 2008, t. XII, s. 69-74.
- Karpiński Ł., *Od cechy dystynktywnej do definicji realnej*, „Linguodidactica” 2009, t. XIII, s. 123-136.
- Karpiński Ł., *Prezentacja terminologii w słowniku w ujęciu typologicznym*, „Linguodidactica” 2010, t. XIV, s. 87-100.
- Karpiński Ł., *W kwestii statusu terminologicznego hipotermu w językach specjalistycznych*, „Linguodidactica” 2011, t. XV, s. 89-100.
- Karpiński Ł., *Semantyka komunikatu a maszynowy zapis parametryzacji tekstu: Pantext 2.0*, „Komunikacja Specjalistyczna” 2012, t. 5, s. 21-32.
- Karpiński Ł., Michałowski P., *Wybrane metody analizy terminologii specjalistycznej (na przykładzie technoloektu geografii)*, „Edukacja dla Przyszłości” 2012, t. IX, Białystok, s. 23-26
- Karpiński Ł., *Zarys leksykografii terminologicznej*, KJS UW, Warszawa 2008.
- Kaufmann A., Fustier M., Drevet A., *Inwentyka. Metody poszukiwania twórczych rozwiązań*, Warszawa 1975.
- Leski K., *Zasady budowy teaurusu*, Warszawa 1978.
- Lewicki A., Sambor J., *Projekt słownika frekwencyjnego współczesnego języka polskiego, sprawozdanie I Wydz. PAN*, z. 4, 1969, s. 117-129.

Lewin J., *Znaki, język, matematyka*, [w:] *Język, matematyka, cybernetyka*, J. Lewin, J. Gastiew, J. Rozanow, Warszawa 1967, s. 7-57.

Lukszyn J., Zmarzer W., *Teoretyczne podstawy terminologii*, Warszawa 2001.

Marciszewski W., *Sztuczna Inteligencja*, Znak, Kraków 1998.

Mecner P., *Elementy gramatyki umysłu*, Kraków 2007.

Michałowski P., *W poszukiwaniu idealnego modelu słownika terminologii przedmiotowej*, „Języki specjalistyczne” 2004, t. 4, J. Lewandowski (red.), Warszawa, s. 193-201.

Sambor J., *Badania statystyczne nad słownictwem (na materiale „Pana Tadeusza”)*, Wrocław 1969, s. 163.

Słownik terminologii przedmiotowej, J. Lukszyn (red.), Warszawa 2005.

Turing A. M., *Computing machinery and intelligence*, „Mind”, October 1950, s. 433-460.

Zmysłowski A. J. i in., *Ontologiczne uwarunkowania nauki*, Zabrze 2002.

Гринев С. В., *Введение в терминологию*, Москва 1995.

<http://www.chip.pl/news>

<http://wyborcza.pl>

<http://corporate.britannica.com/about/today>