

## II. RECENZJE I OMÓWIENIA

### CZY ISTNIEJE JĘZYK SŁÓW KLUCZOWYCH?

WIESŁAW BABIK: SŁOWA KLUCZOWE. KRAKÓW:  
WYDAW. UNIWERSYTETU JAGIELLOŃSKIEGO. 2010,  
242 S.

O słowach kluczowych mówi się od ponad pół wieku, zwłaszcza w kontekście wyszukiwania informacji w dokumentacyjnych systemach komputerowych, a od co najmniej kilkunastu lat także w związku z przeszukiwaniem zasobów sieciowych. Nie jest to termin jednoznaczny, co wynika stąd, że funkcjonuje nie tylko w dziedzinie informacji, ale także w informatyce, językoznawstwie, literaturoznawstwie, neurobiologii. Można wręcz powiedzieć, że współcześnie poszerza swój zakres i zasięg, zwłaszcza w związku z zastosowaniami sieciowymi.

Każdy komputerowy system informacyjno-wyszukiwawczy daje możliwość wyszukiwania przez słowa kluczowe, a dokładniej indeks słów kluczowych (*key-words*), ale nawet pobieżna obserwacja tego sposobu wyszukiwania pozwala stwierdzić, że czasem chodzi o pojedyncze słowa języka naturalnego (przejęte z opisów bibliograficznych i innych tekstów) z zachowaniem ich form gramatycznych, czasem natomiast są to wyrażenia wielowyrazowe (frazy), a czasem mamy do czynienia z wykorzystaniem słowników terminologicznych. W pierwszych dwóch przypadkach mówi się o języku swobodnych słów kluczowych, w drugim – o kontrolowanym (znormalizowanym) języku słów kluczowych. To rozróżnienie i niejednoznaczność terminu *słowo kluczowe* prowadzą do pytania o to, czy możemy mówić o jednym języku i jego dwóch odmianach, a nawet dalej, czy możemy mówić o języku sztucznym, czy też tylko o słowach języka naturalnego, które nazwaliśmy słowami kluczowymi, ponieważ w pewnych warunkach korzystamy z nich jako „kluczy”, które pozwalają dotrzeć do jakiegoś tekstu/informacji. Takie wątpliwości pojawiają się, zwłaszcza gdy korzystamy z zupełnie dowolnych słów w przeglądarkach sieciowych.

Podstawowy więc problem to ustalenie zakresu terminu *słowo kluczowe*, z jednej strony w kontekście języka naturalnego, z drugiej zaś jako jednostki leksykalnej sztucznego języka paranaturalnego, wykorzystywanego do indeksowania i wyszukiwania w systemach informacyjno-wyszukiwawczych i zasobach sieciowych. Jest to punkt wyjścia do dalszych rozważań:

- jaki jest związek słów kluczowych z językiem naturalnym, jaki z językami sztucznymi paranaturalnymi,
- jakie są cechy charakterystyczne języka słów kluczowych,
- jakie ma wersje,
- jak funkcjonuje,
- jak się mają słowa kluczowe do słownika mentalnego człowieka.

Na te pytania próbuje odpowiedzieć Wiesław Babik w swojej książce, która jest pierwszą monografią tego zagadnienia na gruncie polskim. Celem głównym, który sobie postawił autor, była próba określenia tożsamości języka słów kluczowych, jego charakterystyki (natury) poprzez określenie granic pomiędzy nim a językiem naturalnym oraz innymi paranaturalnymi językami informacyjno-wyszukiwawczymi. Inaczej mówiąc, zbadanie, czy język słów kluczowych jest językiem sztucznym, czy naturalnym.

Książka składa się z trzech części, z których dwie (druga i trzecia) poświęcone są językowi słów kluczowych jako językowi informacyjno-wyszukiwawczemu, a dokładniej słowom kluczowym jako jednostkom leksykalnym tego języka, natomiast pierwsza stanowi wprowadzenie do problematyki szeroko rozumianych słów kluczowych.

W części pierwszej (*Słowa kluczowe w różnych kontekstach*, s. 23-78) autor przedstawia definicje słów kluczowych spotykane w piśmiennictwie różnych dziedzin, skupiając się bardziej na rozumieniu słów kluczowych w wyszukiwarkach internetowych (przeglądarkach sieciowych) i systemach hipertekstowych, rozpatruje słowa kluczowe w kontekście języka naturalnego – m.in. związek słów kluczowych ze słownikami mentalnymi, kojarzenie terminów i ustalanie ich powiązań semantycznych (gniazda semantyczne). Są to pytania o to, jak myślimy, jakie znaczenie ma język naturalny w wyszukiwaniu informacji, jaka jest językowa reprezentacja świata w umyśle, wreszcie jak poszukujemy informacji? Ta różnorodność aspektów badawczych słów kluczowych sprawia, że stają się one przedmiotem zainteresowania interdyscyplinarnego (językoznawców, psychologów, pracowników informacji, informatyków).

Część druga dotyczy słów kluczowych jako jednostek leksykalnych języka informacyjno-wyszukiwawczego (*Słowo kluczowe jako jednostka leksykalna języka słów kluczowych*, s. 79-120). Autor daje przegląd historii i genezy języka słów kluczowych, sięgając do czasów bardzo odległych (słowa z tytułu), a we współczesnym znaczeniu odwołuje się do systemu unitermów Mortimera Taubego (połowa XX w.), podaje definicje, wymienia cechy istotne, wskazuje typy języków słów kluczowych. Mamy tu charakterystykę funkcjonalną i strukturalną języka słów kluczowych rozpatrywanego jako pewnego rodzaju model, omówienie systemu leksykalnego i gramatycznego, porównanie z językiem naturalnym i językami sztucznymi paranaturalnymi – językiem deskryptorowym i językiem haseł przedmiotowych. Autor przyjmuje definicję zawartą w *Słowniku encyklopedycznym informacji, systemów i języków informacyjno-wyszukiwawczych* (B. Bojar, 2002), mianowicie: „język słów kluczowych – język informacyjno-wyszukiwawczy o notacji paranaturalnej, bez określonej explicite w słowniku paradygmatyki, w którym jednostki słownika nazywane są słowami kluczowymi, a na gramatykę składają się reguły indeksowania współrzędnego”. Definicję tę nieco poszerza, nie uwzględniając cechy paradygmatyki (s. 85).

Część trzecia poświęcona jest wykorzystaniu słów kluczowych (nie języka słów kluczowych) w wyszukiwaniu informacji (*Słowa kluczowe w systemach wyszukiwania informacji*, s. 121-216). Są tu uwzględnione słowniki słów kluczowych, polskie i obce, z różnych dziedzin, których wykaz (37) autor zamieścił i poddał analizie w swojej publikacji (s. 235-237). Są to wykazy słów kluczowych m.in. z zakresu medycyny, farmakologii, zdrowia publicznego, teologii,

religioznawstwa, pedagogiki, psychologii, logopedii, językoznawstwa, literaturoznawstwa, marksizmu-leninizmu, nauk politycznych, socjologii, ekonomii politycznej, filozofii, etnografii, botaniki. W przypadku słowników mamy do czynienia ze słowami kluczowymi odpowiednio dobranymi i uporządkowanymi, a więc terminologią danej dziedziny. Ten aspekt jest czasem wyraźnie podkreślany w podtytułach słowników (wykaz podstawowych terminów z...). Jak pisze autor, celem tego rozdziału była weryfikacja tezy o końcu języka słów kluczowych.

Zacznijmy od tytułu i części wstępnej. Autor nadał swojej książce tytuł *Słowa kluczowe*, nie: *Język słów kluczowych*. Jest to pewnego rodzaju dysonans, ponieważ już w pierwszym akapicie wstępu metodologicznego pisze: „Przedmiotem rozprawy są słowa kluczowe używane w indeksowaniu i wyszukiwaniu informacji. Jak wszystkie języki informacyjno-wyszukiwawcze stanowią one lingwistyczny instrument służący do organizacji wiedzy o świecie, utrwalonej w postaci dokumentów i udostępnianej za pomocą dokumentacyjnych systemów informacyjno-wyszukiwawczych, zarówno w tradycyjnych dokumentacyjnych systemach wyszukiwania informacji online, jak i zasobach Internetu. Zostały przedstawione na szerokim tle ich wykorzystania w różnych dziedzinach, nazywanych tu kontekstami. Rozważania dotyczą przede wszystkim ich statusu jako jednostek leksykalnych języka informacyjno-wyszukiwawczego” (s. 9, podkr. JS).

Z tekstu tego możemy wnioskować, że:

a) autor stawia znak równości między terminami *słowa kluczowe* i *język słów kluczowych* w odniesieniu do zastosowania w systemach informacyjno-wyszukiwawczych; pytanie, czy jest to tylko zmiana terminologiczna, czy także zakresowa?

b) język słów kluczowych traktuje jako język informacyjno-wyszukiwawczy, w którym słowa kluczowe są jednostkami leksykalnymi; pytanie, czy we wszystkich przypadkach i zastosowaniach?

Można dalej zapytać, dlaczego autor unika w tytule terminu *język słów kluczowych*:

– Czy dlatego, że omawia słowa kluczowe, wykraczając poza zastosowania w systemach informacyjno-wyszukiwawczych?

– Czy dlatego, że ma wątpliwości, czy język słów kluczowych istnieje, choć przecież opisuje go, podając jego cechy konstytutywne (słownictwo, gramatykę), zajmuje się jego semantyką, typami, wykorzystaniem?

– Czy dlatego, że termin 'język słów kluczowych' jest niejednoznaczny (język swobodnych słów kluczowych lub kontrolowanych słów kluczowych), co jest wynikiem niejednolitej natury słów kluczowych?

Zatrzymajmy się przy języku swobodnych słów kluczowych, który jest prototypem właściwego (kontrolowanego) języka słów kluczowych. Język swobodnych słów kluczowych znany jest nam z wyszukiwania w Internecie. Są to słowa wybierane z tekstów języka naturalnego i, jak spostrzega słusznie autor, „Cały aparat indeksowania i wyszukiwania nie należy do języka, lecz do systemu i został tu rozbudowany o wiele dodatkowych opcji, zależnych od aktualnych możliwości technologii komputerowej, a nie od samego języka” (s. 157). Te obserwacje prowadzą do stwierdzenia: „Języki słów kluczowych w odmianie języków swobodnych słów kluczowych stanowią więc zredukowaną

do potrzeb optymalizacji procesów informacyjnych odmianę języka naturalnego” (s. 119). W zakończeniu książki autor zajmuje zdecydowane stanowisko wobec języka swobodnych słów kluczowych, stwierdzając, że „nigdy nie był i nie jest językiem sztucznym, podobnie jak swobodne słowa kluczowe w Internecie” (s. 204). Jeśli nie był i nie jest językiem sztucznym, to znaczy, że jest językiem naturalnym. I to jest bardzo ważna konkluzja opracowania Wiesława Babika, z którą należy się zgodzić.

Co zatem z językiem słów kluczowych w wersji kontrolowanej, czyli nieco znormalizowanej? Autor uznaje, że słownictwo tej wersji języka jest słownictwem naturalnym, tylko w niewielkim stopniu poddającym się normalizacji. Zwraca też uwagę na brak *explicite* relacji paradygmatycznych w słownikach języków słów kluczowych (inaczej niż w językach deskryptorowych i językach haseł przedmiotowych), co oznacza właśnie brak normalizacji. (Co do relacji paradygmatycznych, to może dopuszczalne jest uznanie, że istnieją one w sposób niejawni, potencjalni, podobnie jak słownik tego języka, a ujawniają się dzięki możliwościom kojarzeniowym umysłu). Wszystkie jednostki leksykalne są samodzielne, podobnie jak w językach deskryptorowych. Autor stwierdza, że pierwotną wersją tego języka był język swobodnych słów kluczowych ze słownictwem niekontrolowanym. „Później próbowano z tego zrobić język kontrolowanych (znormalizowanych) słów kluczowych” (s. 203). Czy zrobiono? Można wątpić, skoro autor pisze: „Nazwy 'język (swobodnych, kontrolowanych) słów kluczowych' utrzymują się przede wszystkim ze względów marketingowych. Po prostu przyzwyczajono się do stosowania tych nazw na określenie słów kluczowych używanych w wyszukiwaniu” (s. 204). To potwierdzałoby zmianę zakresową terminów 'język słów kluczowych' i 'słowa kluczowe' a jednocześnie uzasadniało tytuł monografii. Ostatecznie, po opisanu tego języka w części drugiej i trzeciej, autor dochodzi do wniosku, że język słów kluczowych nie istnieje jako sztuczny język informacyjno-wyszukiwawczy. Jest to tylko termin, usankcjonowany tradycją, który nie ma uzasadnienia w rzeczywistości, bowiem właśnie w rzeczywistości są tylko słowa języka naturalnego, które bywają używane w funkcji wyszukiwawczej, metainformacyjnej. „Obecnie słowa kluczowe są formą wykorzystywania języka naturalnego w procesach zautomatyzowanego wyszukiwania informacji” (s. 204). W zakończeniu autor pisze, że nie udało mu się udzielić pełnej odpowiedzi na postawione w rozprawie pytanie o tożsamość i granice tego języka, bo wszystko tu jest „rozmyte”, a więc jest to trochę „dziwny język”, stwierdza (s. 203). Stąd już tylko krok, aby uznać, że język taki nie istnieje obiektywnie, istnieją tylko słowa języka naturalnego, używane w funkcji wyszukiwawczej, nazywane słowami kluczowymi.

Niewątpliwie wnioski autora są wynikiem nie tylko analizy piśmiennictwa na ten temat, ale także wynikiem obserwacji indeksowania tekstów i wyszukiwania w zasobach sieciowych oraz dokumentacyjnych bazach danych, a także wpływu technologii na sposoby wyszukiwania. Obserwacje te rzeczywiście prowadzą do wniosku, że albo musimy uznać, że język słów kluczowych jest „dziwny” zarówno w odniesieniu do słownictwa, jak i gramatyki (zerowej), albo jest to język naturalny. Słownik języka słów kluczowych jest otwarty i tożsamy ze słownictwem języka naturalnego (w zasadzie nie podlega normalizacji, każde wyrażenie jest dozwolone z wyjątkiem tzw. stop-listy, czyli pewnej grupy

wyrażeń wykluczonych, nieznaczących, np. przyimki, rodzajniki), znaczenia słów kluczowych pokrywają się ze znaczeniami w języku naturalnym, a nawet gdyby było inaczej, to o tym nie wiemy, bo nie ma objaśnień explicite. Natomiast w przypadku zasobów internetowych i niektórych systemów dokumentacyjnych słowa kluczowe są wybierane z tekstów dokładnie w takiej formie, w jakiej zostały zapisane w tekście. Jeśli odwołamy się do kompetencji użytkownika w korzystaniu z tego języka, to musimy stwierdzić, że jest to wyłącznie kompetencja wynikająca z kompetencji języka naturalnego. Znaczeń słów kluczowych nie uczymy się, ponieważ znamy je skądinąd. A zatem:

- Czy nie popełniamy nadużycia, nazywając coś językiem sztucznym, co nim nie jest?
- Czy nie przypisaliśmy językowi naturalnemu funkcji języka sztucznego tylko dlatego, że odnosi się do rzeczywistości tekstowej?
- Czy mechaniczne wybieranie słów z tekstu może być uznane za proces indeksowania?

Są to pytania, które nasuwają się podczas lektury książki Babika.

W tym kontekście pojawia się pytanie o przyszłość języków informacyjno-wyszukiwawczych<sup>1</sup>, zwłaszcza wobec rozwoju technologicznego sprawiającego, że coraz rzadziej w wyszukiwaniu informacji pośredniczy jakiś „bezpośredni pomocnik”. Użytkownik jest raczej skazany na siebie, a to oznacza, że najchętniej będzie korzystał z języka naturalnego, który jest mu najbliższy mentalnie. Czy potrzebne są nam języki typu klasyfikacji, skoro, aby je zrozumieć, potrzebny jest język przekładowy paranaturalny (czy może naturalny?) w postaci indeksu przedmiotowego, prowadzącego do odpowiednich miejsc (klas) struktury klasyfikacyjnej, oznaczonych symbolicznie? Wreszcie, trzeba pamiętać o najnowszym wpływie technologii na systemy informacyjne, mianowicie tzw. tagowaniu, czyli nadawaniu przez czytelników/użytkowników całkowicie dowolnych określeń (nazywanych jednak słowami kluczowymi) tekstom dostępnym przez przeglądarki sieciowe. Wszystko to na pewno powinno skłaniać do spojrzenia na języki informacyjno-wyszukiwawcze z jednej strony w kontekście rozwoju technologicznego, z drugiej zaś z punktu widzenia roli języka naturalnego w wyszukiwaniu informacji. Wydaje się, że takiego właśnie spojrzenia na język słów kluczowych dokonał Wiesław Babik w swojej książce, dowodząc, że coś, co zwyczajowo określamy jako język słów kluczowych, jest co najwyżej zbiorem zwyczajowo nazywanych słów kluczowych, a te z kolei w rzeczywistości są tylko słowami języka naturalnego. Z punktu widzenia teorii języków informacyjno-wyszukiwawczych jest to spojrzenie rewolucyjne, ale może należy uznać je tylko za racjonalne.

*Jadwiga Sadowska*

---

<sup>1</sup> Zob. B. Bojar: *Języki informacyjno-wyszukiwawcze wczoraj, dziś... czy jutro?* „Zagadnienia Informatyki Naukowej” 2009, nr 1, s. 3-24.