

PIOTR ŻMIGRODZKI, MIROSŁAW BAŃKO,
BOGUSŁAW DUNAJ, RENATA PRZYBYLSKA

Koncepcja *Wielkiego słownika języka polskiego* — przybliżenie drugie

Niniejszy tekst jest drukowaną wersją prezentacji, jaka odbyła się 4 grudnia 2006 roku na posiedzeniu Komitetu Językoznawstwa PAN w Warszawie. Stanowi on podsumowanie pewnego okresu pracy nad koncepcją *Wielkiego słownika języka polskiego* (dalej: WSJP), której wersja wstępna została wcześniej przedstawiona na posiedzeniu KJ PAN 3 października 2005 roku i następnie opublikowana (por. Dunaj, Przybylska, Żmigrodzki 2006). O kwestiach związanych z potrzebą podjęcia prac nad słownikiem, który objąłby opisem polszczyznę drugiej połowy XX i początku XXI wieku, traktował również artykuł: Żmigrodzki, Przybylska, Dunaj 2006.

1. Forma słownika

Słownik, który planujemy opracować, będzie słownikiem elektronicznym. Stoimy na stanowisku, że obecnie taka właśnie forma jego istnienia jest najwygodniejsza i najlepsza, zarówno z punktu widzenia autorów, jak i ze względu na potrzeby odbiorcy. Dla autorów istotne są przede wszystkim następujące zalety takiego słownika:

- ustrukturyzowanie zapisu danych leksykograficznych, umożliwiające ich przeszukiwanie i wydobywanie w razie potrzeby różnie scharakteryzowanych podzbiorów zarejestrowanego słownictwa, np. tylko potoczne, tylko przymiotniki itp.;
- możliwość modyfikacji hasła (zbioru haseł) w dowolnym momencie, co ułatwia wprowadzanie zmian do słownika i jego aktualizację na bieżąco;
- łatwa edycja — na podstawie jednego zbioru danych — różnych postaci słownika i słowników pochodnych.

Dla potencjalnego odbiorcy natomiast jako korzyści możemy wskazać m.in.:

- łatwość wyszukiwania informacji (nawet bez pełnej znajomości konwencji opisu leksykograficznego);
- dostęp za pośrednictwem Internetu do dowolnej liczby haseł bez konieczności kupowania całego słownika;
- możliwość uzyskiwania aktualizacji przez Internet.

Już w trakcie prac będzie możliwe udostępnianie części haseł w sieci; po ukończeniu zaś słownika można będzie go opublikować w postaci CD, DVD lub innego nośnika danych. Jeżeli będzie potrzeba, edycja książkowa również jest, rzecz jasna, możliwa.

2. Zasięg czasowy słownika

Opis objemnie jednostki leksykalne występujące w polszczyźnie (poświadczone w źródłach) po 1945 roku. Decyzję tę można uzasadnić na dwa sposoby. Po pierwsze, obraz polszczyzny pierwszych lat po drugiej wojnie światowej, jaki mamy udokumentowany w SJPDor¹ i słownikach późniejszych, jest niepełny ze względu na wybiórczą ekstrakcję danych, a także ograniczenia wynikające z istoty koncepcji leksykograficznej Doroszewskiego (zob. m.in. Piotrowski 1994: 55–64, Żmigrodzki 2005: 34–36), jak również z przyczyn pozamerytorycznych, związanych z okresem, w jakim słownik powstawał. Po drugie, uważamy, że polszczyzna okresu międzywojennego widziana dziś być może już jako polszczyzna historyczna, a jej opracowaniu leksykograficznemu należałoby poświęcić odrębne dzieło. Już zresztą okres po 1945 roku wydaje się w dziejach języka polskiego dość zróżnicowany i z pewnością w planowanym słowniku znajdą się jednostki leksykalne całkowicie współczesnemu (i przyszłemu) czytelnikowi nieznanne.

3. Charakter słownika

Planowany wielki słownik będzie miał charakter:

- dokumentacyjny — tzn. dążyć będziemy do rzetelnej ilustracji materiałowej haseł i ich dokumentacji cytatami pochodzącymi z autentycznych tekstów;
- deskryptywny, a nie preskryptywny — tzn. będziemy odnotowywać rzeczywisty stan języka, nie zaś stan pożądany.

Konsekwencją przyjęcia takiej opcji jest możliwość znalezienia się w opisie faktów, które w świetle obowiązującej normy skodyfikowanej uważane są za niepoprawne. Założenie to może być uznane za kontrowersyjne, jednak we współczesnej metaleksykografii jest przyjmowane raczej powszechnie. Ściślej mówiąc, będziemy praktykować tzw. normatywizm opisowy (zob. Bańko 2001: 39–49), tzn. odnotowywać formy uznane za niepoprawne², wyraźnie ich niepoprawność sygnalizując. Za podstawę werdyktów będziemy uznawać postać normy skodyfikowanej w momencie opracowywania danego hasła; za źródło tej normy uznajemy aktualnie *Wielki słownik poprawnej polszczyzny PWN* pod red. A. Markowskiego i *Wielki słownik ortograficzny PWN* pod red. E. Polańskiego.

¹ Rozwiązanie częściej używanych skrótów nazw słowników znajduje się na końcu tomu.

² Rzecz jasna, idzie tylko o formy wystarczająco rozpowszechnione w uzusie, czego znakiem będzie frekwencja korpusowa.

4. Źródła słownika

Podstawowym źródłem słownika będą komputerowe korpusy tekstów języka polskiego. Obecnie istnieje kilka takich zbiorów (por. Piotrowski 2003, Przepiórkowski 2004, Lewandowska-Tomaszczyk red. 2005: 17–24); największe nadzieje wiążemy wszak ze startującym aktualnie projektem Narodowego Korpusu Języka Polskiego, który miałby scalić i wzbogacić korpusy istniejące. Oczywiście, nie jest to podstawa wystarczająca. Ze względu na specyfikę i odrębność obu projektów można sobie wyobrazić sytuację, że istotne dla słownika materiały w scalonym korpusie się nie znajdują. W związku z tym planujemy ekscerpcję uzupełniającą. W istocie nie będzie to klasyczna ekscerpcja, ale digitalizacja określonych tekstów i przystosowanie ich do komputerowego przeszukiwania.

Objęłaby ona przede wszystkim następujące materiały:

- publikacje z lat 1945–1956;
- współczesne czasopisma hobbistyczne i inne o wąskim zasięgu;
- podręczniki szkolne³;
- witryny internetowe (oficjalne i prywatne).

Te właśnie zasoby w istniejących korpusach są najsłabiej reprezentowane, można natomiast podejrzewać, że są one źródłem wielu jednostek leksykalnych, wartych opisanie. Pożądane byłoby też uzupełnienie źródeł słownika o teksty mówione, zwłaszcza reprezentujące wypowiedzi spontaniczne w codziennych, przede wszystkim nieoficjalnych sytuacjach komunikacyjnych.

5. Jakie fakty językowe zostaną ujęte w słowniku?

Planujemy objęcie opisem możliwie szerokiego spektrum haseł, reprezentujących różne typy faktów językowych, tzn.:

- wyrazy „pospolite”;
- wybrane nazwy własne, tzn.:
 - ważniejsze nazwy geograficzne (wybór według frekwencji w korpusach),
 - nazwy własne używane w odniesieniach apelatywnych, np. *Belweder*, *Tworki*, *Janosik*, *Rapallo*;
- nieciągle jednostki leksykalne (stałe związki frazeologiczne);
- frazemy (w wyborze);
- „skrzydlate słowa”;
- przysłowia;
- morfemy:
 - prefiksy, prefiksoidy i pierwsze człony złożeń,
 - sufiksy, sufiksoidy i drugie człony złożeń — te, które tworzą długie serie i odznaczają się znacznym potencjałem słowotwórczym.

³ Ekscerpcja dotyczyć będzie podręczników kształcenia ogólnego.

6. Ogólne zasady hasłowania

Przy porządkowaniu materiału stosować się będzie następujące zasady ogólne:

- każdy obiekt leksykograficzny będzie stanowił osobny artykuł hasłowy, np. w odrębnych artykułach hasłowych będą opisywane jednostki leksykalne: *głowa*, *głowa nie od parady*, *mądrej głowie dość dwie słowie*, *ktos nie ma oleju w głowie*;
- artykuły hasłowe opisujące jednostki nieciągłe będą „dowiązane” do artykułów opisujących jednostki ciągłe za pomocą linków hipertekstowych, co ułatwi czytelnikowi orientację i nawigację w słowniku;
- tylko ciągi reprezentujące odrębne części mowy opisywane będą w kilku artykułach hasłowych, np.: **piec I** *rzecz.*; **piec II** *czas*. Inne przypadki homonimii będą opisywane w ramach jednego artykułu hasłowego⁴;
- pary aspektowe, zarówno prefiksalne, jak i sufiksalne, będą opisywane w osobnych artykułach hasłowych, powiązanych linkami. Pozwoli to uniknąć stosowania skomplikowanych nieraz rozwiązań technicznych, za pomocą których oddaje się w tradycyjnych słownikach podobieństwa i różnice struktury semantycznej czasowników różniących się aspektem.

7. Wydzielanie i hierarchizowanie znaczeń

W hasłach opisujących jednostki wieloznaczne jako zasadę przyjmujemy odnotowywanie najpierw znaczeń obecnych w języku ogólnym, w kolejności frekwencji korpusowej, a następnie znaczeń nacechowanych (stylistycznie, chronologicznie, regionalnie) w kolejności alfabetycznej kwalifikatorów. Od zasady tej dopuszczamy, oczywiście, wyjątki, uzasadnione np. chęcią lepszego ukazania struktury polisemii danej jednostki, struktury kognitywnej pojęcia itp. Numeracja podhasła będzie dwupoziomowa: kolejne znaczenia numerować się będzie liczbami arabskimi, literami zaś — kolejne znaczenia (odcienie znaczeniowe?) jednostki polisemicznej regularnie. Oto przykład:

SZKOŁA [...]

1. a) ‘instytucja oświatowa....’
b) ‘budynek....’
c) ‘ogół uczniów....’
d) ‘czas spędzony na lekcjach w szkole....’
2. ‘kierunek w nauce, sztuce lub filozofii...’ [...]

Jak z powyższego przykładu wynika, nie akceptujemy postulatów, wysuwanych przez niektórych językoznawców teoretyków (np. Grochowski 1982: 81–87, Grzegorzczkowska 1990: 42–47), aby przypadki polisemii regularnej z hasła słownikowych usuwać i informować o nich w jakiejś „części ogólnej, opisowej” (postulatów zresztą

⁴ W dyskusji na posiedzeniu KJ PAN postulat ten wywołał szczególnie dużo uwag krytycznych, w związku z czym najprawdopodobniej zasady hasłowania homonimów zostaną zmienione; szczegółowe rozwiązania techniczne będą przedmiotem dalszej dyskusji.

niespełnionych w żadnym ogólnym słowniku języka polskiego). Szczegółowe rozważenie tej sprawy zostawiamy na inną okazję⁵.

8. Mikrostruktura słownika

Przewidujemy następujące składniki artykułu hasłowego (ich kolejność w zasadzie jest ustalona, niektóre jednak występować będą w zależności od specyfiki hasła):

- | | |
|---------------------------------|--|
| 1. wyrażenie hasłowe; | 9. informacja syntaktyczna; |
| 2. wariant wyrażenia hasłowego; | 10. kolokacje; |
| 3. informacja o wymowie; | 11. przykłady pełnozdaniove; |
| 4. etymologia; | 12. jednostki nieciągłe zawierające wyrażenie; |
| 5. informacja fleksyjna; | 13. frazemy; |
| 6. kwalifikatory; | 14. „skrzydlate słowa”; |
| 7. definicja; | 15. przysłowia. |
| 8. relacje semantyczne; | |

Ich krótkie omówienie przedstawiamy poniżej.

8.1. Wyrażenie hasłowe. Stosowane tu będą tradycyjne konwencje, obowiązujące w słownikach polskich przynajmniej od początku XX wieku.

8.2. Wariant wyrażenia hasłowego. Będzie to informacja o istnieniu obocznej fonetycznie lub ortograficznie postaci wyrażenia hasłowego, a jednocześnie odsyłacz do hasła opisującego tę postać oboczną. Przykłady:

ŚPIESZYĆ, też: **spieszyc** [...]
SPIESZYĆ, też: **śpieszyc** [...]
SZUKAĆ WZROKIEM, też: **szukać oczami** [...]
SZUKAĆ OZAMI, też: **szukać wzrokiem**
HIPIKA, też: **hippika** [...]
HIPPIKA, też: **hipika** [...]
WIDEO, też: **video** [...]
VIDEO, też: **wideo** [...]

Każdy z tych wariantów będzie opisany w osobnym haśle o pełnej strukturze, z materiałem przykładowym ilustrującym użycie wariantu.

W jednym wspólnym haśle/podhaśle będą natomiast opisywane tylko warianty ortograficzne różniące się pisownią wielką lub małą literą, jak np.:

STÓŁ PAŃSKI [...] ‘ołtarz w kościele’[...] ORT.: *stół pański* a. *stół Pański*, a. *Stół Pański* [...]
OKRĄGLY STÓŁ [...] 1. ‘rozmowy prowadzone w poszukiwaniu kompromisu między skonfliktowanymi stronami’ [...] 2. ‘rozmowy między opozycją a rządem w Polsce w 1989 roku’ [...] ORT.: *Okrągły Stół* [...]

⁵ Zob. też artykuł A. Fedorowicz w niniejszym tomie.

8.3. Informacja o wymowie. Będzie podawana tylko wówczas, gdy wymowa wyrazu/wyrażenia odbiega od ogólnych zasad wymowy w języku polskim, a więc przede wszystkim w odniesieniu do jednostek obcego pochodzenia. Stosowane będą dwie konwencje: ilustracja wymowy alfabetem łacińskim oraz transkrypcja międzynarodowa (według standardu IPA⁶). W wersji elektronicznej towarzyszyć temu będą pliki dźwiękowe z wzorcowymi realizacjami. W przyszłości można będzie uzupełnić o taką informację wszystkie hasła, być może także dotyczącą słowoform nastrożających problemów normatywnych, np. *wziąłem, muszę* itp.

8.4. Etymologia będzie podawana przy wyrazach obcego pochodzenia. Zasadą ogólną jest podawanie tylko języka, z którego zapożyczenie przeniknęło bezpośrednio, np.:

KETCHUP, keczup, wym. [keczup] (ang. ketchup) [...]

Niekiedy podawane będzie również źródło wcześniejsze, zwłaszcza jeśli źródło bezpośrednie jest trudne do określenia.

8.5. Informacja fleksyjna obejmować będzie:

- określenie części mowy;
- rodzaj gramatyczny rzeczownika;
- aspekt czasownika;
- formę stopnia wyższego przymiotników i przysłówków stopniowalnych;
- paradygmat odmiany (jednostki odmienne).

W słowniku elektronicznym pierwsze cztery z wymienionych elementów będą umieszczone przy wyrażeniu hasłowym. Paradygmat odmiany będzie się natomiast pojawiał w osobnym okienku, otwieranym na życzenie użytkownika. Nie przewidujemy wypracowywania na potrzeby tego słownika oryginalnego modelu opisu fleksyjnego; zamierzamy skorzystać z niemałych przecież osiągnięć polskich badaczy fleksji, dostosowując je do charakteru dzieła. Niewątpliwie powinien to być opis nowoczesny metodologicznie, w miarę możliwości pełny i łatwy do digitalizacji. Szczegóły zostaną ustalone w późniejszym terminie.

8.6. Kwalifikatory. Zestaw kwalifikatorów w słowniku powinien być zhierarchizowany. Opracować trzeba możliwie jasne kryteria ich stosowania, tak, by były one łatwe w użyciu. Przykładowo, zastosowanie kwalifikatora *daw.* byłoby uzasadnione w stosunku do jednostki, która posiada synonim współczesny, jak w wypadku leksemów:

arendarz — współczesny odpowiednik: *dzierżawca*;

austeria — współczesny odpowiednik: *zajazd*.

⁶ Alfabet ten jest powszechnie stosowany w świecie, jego użycie zatem uczyniłoby opis przystępnym również dla cudzoziemców. Strona internetowa IPA: <http://www2.arts.gla.ac.uk/IPA/ipa.html>.

Próbie wypracowania takiej siatki kwalifikatorów i kryteriów ich stosowania przynosi artykuł J. Kurkiewicza, publikowany w tym tomie.

8.7. Definicja. Przyjmujemy następujące założenia ogólne:

- definicje powinny mieć charakter językowy, a nie encyklopedyczny. Jest to jeden z najstarszych postulatów językoznawców w stosunku do definicji leksykograficznych (por. m.in. Grochowski 1982: 122–124, Bańko 2001: 111–122, Żmigrodzki 2005: 85–88), dlatego czujemy się zwolnieni z obowiązku jego objaśniania w tym miejscu. Komentarza wymaga natomiast problem różnic między językowym a encyklopedycznym (specjalistycznym) rozumieniem niektórych leksemów. Część z tych różnic jest znana niemal każdemu (np. *delfin* — ryba czy ssak?), inne mniej się rzucają w oczy (np. dla botanika kwiat gerbery to kwiatostan, kaktus zaś posiada ciernie, nie kolce), z niektórych w ogóle mało kto sobie zdaje sprawę (np. pojęcie *stacja* w języku kolejarzy lub takie pojęcia prawne⁷, jak *bójka* czy *zgwalcenie*). Słownik języka polskiego nie może zdawać sprawy ze wszystkich tych różnic, jednak najważniejsze będziemy w definicjach sygnalizować. Rozważamy też zamieszczenie — tam, gdzie będzie to dla nas uchwytne — swoistych „ostrzeżeń”, że dany leksem w języku naukowym, czy np. prawnym ma inne znaczenie. Byłoby to jednocześnie odesłanie czytelnika do innego typu źródeł wiedzy;
- w miarę możliwości należy unikać błędów w definiowaniu, jednak w praktyce wyeliminowanie tych błędów okazuje się trudne, jeśli nie niemożliwe (por. Bańko 2006);
- dążyć będziemy do większego niż w dotychczasowych słownikach wykorzystania dorobku semantyki językoznawczej w definiowaniu jednostek języka.

Wypracowanie metod i zasad definiowania poszczególnych typów jednostek leksykalnych w słowniku jest odrębnym zadaniem. W pracy Dunaj, Przybylska, Żmigrodzki 2006 zostały sformułowane pewne wskazówki w tej dziedzinie i, jak sądzimy, mogą one posłużyć za podstawę dalszej dyskusji nad zagadnieniami opisu semantycznego w słowniku.

8.8. Relacje semantyczne. Planuje się podawanie w haśle opisującym każdą jednostkę leksykalną (rozumianą tu jako uporządkowana para: forma — znaczenie) jednostek pozostających z nią w relacji:

- synonimii;
- antonimii;
- sprzeczności;
- hiperonimii.

⁷ W wypadku pojęć prawnych (a w pewnej mierze też naukowych) dochodzi do tego zmienność ich znaczenia, związana ze zmianami ustaw czy zmianami stanu wiedzy lub koncepcji teoretycznych.

Idzie tu o relacje interpretowane w sposób restryktywny, ścisły, zgodnie z zasadami wyłożonymi np. w pracach: Grochowski 1982: 71–80; Bednarek, Grochowski 1993 (a więc np. synonimia rozumiana będzie jako równoznaczność, nie bliskoznaczność). Decydujemy się na tak wąskie rozumienie relacji ze względu na intencjonalnie naukowy charakter słownika. Wydaje się bowiem, że liczba dostępnych w Polsce popularnych słowników wyrazów bliskoznacznych jest wystarczająca.

8.9. Informacja syntaktyczna. Będzie podawana odrębnie dla każdego znaczenia jednostki leksykalnej. W hasłach opisujących jednostki czasownikowe będzie podawany pełny schemat składniowy zdania minimalnego (w 3 os. lp), np.:

SZUKAĆ [...] 1. starać się znaleźć [czynić coś, aby dowiedzieć się, gdzie jest] przedmiot, którego w danej chwili potrzebujemy, albo osobę, z którą chcemy mieć kontakt [...]
KTOŚ szuka **KOGOŚ/CZEGOŚ GDZIEŚ** [...]

W hasłach opisujących inne jednostki, które posiadają wymagania syntaktyczne, informacja będzie zawierać formę przyłączanego podrzędnika, np.:

SKŁONNY [...]
 skłonny **DO CZEGOŚ/DO TEGO, ŻEBY** + zdanie

8.10. Kolokacje. Pod pojęciem kolokacji rozumiemy tutaj (za Cruse'em, 1986) każde statystycznie częste połączenie jednostki hasłowej z innymi jednostkami. Właśnie na kolokacjach planujemy oprzeć ilustrację użycia opisywanej jednostki w tekstach. Kolokacje będą uszeregowane według schematów składniowych. Zagadnieniom związanym z kolokacjami w słowniku są poświęcone w niniejszym tomie dwa artykuły: R. Przybylskiej i T. Rokickiej, nie będziemy więc w tym miejscu powtarzać zawartych tam twierdzeń. Przykładowo, propozycja układu kolokacji w hasle rzeczownikowym jest następująca:

1. grupy nominalne z rzeczownikiem hasłowym w nadrzędniku:
 - a) połączenia z przymiotnikami,
 - b) połączenia z podrzędnym rzeczownikiem,
 - c) połączenia z frazą przyimkową;
2. grupy nominalne z rzeczownikiem hasłowym w podrzędniku;
3. kolokacje z czasownikami:
 - a) rzeczownik jako podmiot,
 - b) rzeczownik hasłowy akomodowany przez czasownik — w kolejności przypadków;
4. frazy okolicznikowe;
5. wyrażenia porównawcze.

Odpowiada to następującemu, przykładowemu zestawowi kolokacji hasła **stół** w znaczeniu 'mebel'⁸:

⁸ Nie jest to, oczywiście, definicja, jaką chcielibyśmy widzieć w naszym słowniku, ale tylko określenie identyfikujące znaczenie na potrzeby tej egzemplifikacji.

STÓL [...]

- 1a. kwadratowy, prostokątny, owalny, okrągły, podłużny, długi, szeroki, wąski ~; ~ drewniany, kamienny, marmurowy a. z drewna, z kamienia, z marmuru; ~ dębowy; ~ na jednej nodze, na trzech nogach; ~ rozkładany, rozsuwany; ~ kuchenny, jadalny, ogrodowy; ~ konferencyjny, prezydialny, negocjacyjny; ~ do pracy; ~ na (*ileś*) osób; główny ~; 1b. ~ obrad, rokowań
2. noga od stołu; krawędź, blat, róg, szczyt, koniec stołu; dekoracja stołu, akcesoria do stołu, obrus, serweta na~; miejsce za stołem/przy stole; zachowanie przy stole
- 3a. ~ stoi (*gdzieś*), chwieje się; gnie się, ugina się od czegoś/pod ciężarem czegoś; 3b. siadać, u/siąść, zasiąść/zasiadać, za/prosić/zapraszać do stołu; pchać się do stołu, podać/podawać do stołu; wstać/wstawać od stołu, odejść od stołu; sprzątnąć/sprzątać ze stołu, zebrać/zbierać ze stołu; rozłożyć, złożyć ~, nakryć ~, rozstawić, ustawić, zsunąć, rozsunąć stoły, zastawić ~ czymś; położyć/kłaść, postawić, stawiać coś na stole a. na ~; siedzieć za stołem; siadać, zasiąść/zasiadać do stołu/za stołem; siedzieć przy stole; posadzić/sadzać przy stole/za stołem; z/gromadzić się, zebrać/zbierać się przy stole; położyć/kłaść, postawić/stawiać (coś) na stole a. na ~; ucztować przy stole
4. coś stoi, leży na stole
5. gładki, równy jak ~; równo jak po stole

8.11. Przykłady pełnozdaniove w słowniku będą miały przede wszystkim funkcję dokumentacyjną i chronologiczną, a ponadto będą ilustrować te aspekty użycia jednostki leksykalnej, których nie można zobrazować kolokacjami. Umieszczane będą w kolejności chronologicznej, z lokalizacjami tak dokładnymi, jak to tylko możliwe. Ideałem byłoby umieszczanie danych o autorze tekstu, tytule, dacie wydania i stronie, z której cytat zaczerpnięto. Najprawdopodobniej jednak podanie tej ostatniej informacji w odniesieniu do tekstów z korpusów komputerowych będzie niemożliwe, gdyż pliki w korpusach z reguły są przechowywane bez informacji o podziale na strony wydań drukowanych.

8.12. Jednostki nieciągłe. W naszych rozważaniach nie posługujemy się terminem *związek frazeologiczny* ze względu na jego niejednoznaczność i mglistość (por. Grochowski 1982: 23–27, Miodunka 1989: 76–80, Bańko 2001: 149–169). Przedmiotem opisu w słowniku są jednostki leksykalne (por. Grochowski 1982: 28–34), które mogą być ciągłe (stanowiąc w zapisie nieprzerwaną sekwencję znaków — takie zwykle są wyrazami hasłowymi w polskich słownikach) lub nieciągłe. Termin *jednostka leksykalna* rozumiemy zgodnie z definicją Grochowskiego (1982: 28–29)⁹. Jak wspomniano wcześniej, w słowniku elektronicznym (nie przesadzamy w tym miejscu struktury ewentualnych słowników drukowanych, opracowywanych na podstawie naszego) każda taka jednostka będzie miała osobne hasło o pełnej strukturze, zawierającej wszystkie składniki, łącznie z charakterystyką w zakresie części mowy¹⁰.

⁹ Inne problemy związane z odróżnianiem jednostek leksykalnych od połączeń jednostek (kolokacji) podejmuje w tym tomie K. Węgrzynek.

¹⁰ Z przypisywaniem nieciągłym jednostkom leksykalnym odpowiednich wartości części mowy spotkaliśmy się już w ISJP.

W hasle natomiast opisującym poszczególne jej składniki jednostki te będą tylko wymienione, w postaci odnośników hipertekstowych, rozwijających się po kliknięciu do postaci pełnego hasła.

Przykładowo, jednostka *okruchy z pańskiego stołu* będzie opisana w hasle, do którego odsyłacze znajdziemy w hasłach: **stół**, **okruchy** i **pański**.

Samo hasło może mieć postać:

OKRUCHY Z PAŃSKIEGO STOLU rz, blp2 ‘rzeczy niepotrzebne osobom wyżej postawionym, chwilowo dostępne ludziom nie mającym na co dzień do nich dostępu’

KOLOK.: zbierać ~

PRZYKŁ.: *Jak się nachapiał wszyscy dookoła, jakieś okruchy z pańskiego stołu skapną i artyscie.*

8.13. Frazemy. Pod pojęciem frazemu (zob. np. Chlebda 1993: 328) rozumie się „względnie trwałą formę językową, która w danej sytuacji stała się — niezależnie od swych cech strukturalnych i semantycznych — przyjętym (nierzadko jedynym) sposobem wyrażania danego potencjału językowego”. Zasadniczą cechą frazemu jest jego „odtworzalność w danej sytuacji i dla nazwania danej wiązki sensów”. Tak rozumiane frazemy stanowią szeroki pas przejściowy między wielowyzrazowymi jednostkami leksykalnymi a doraźnie tworzonymi połączeniami jednostek języka. Zaliczają się tu m.in. takie obiekty, jak: *świeżo malowane* (wywieszka); *moje uszanowanie*; *a to ci dopiero; że się tak wyrażę; ... i odtąd żyli długo i szczęśliwie; byliśmy — jesteście — będziemy*. Zasoób tak pojmywanych frazemów jest bardzo liczny, dlatego w słowniku proponujemy odnotowywanie tylko tych najbardziej upowszechnionych (na podstawie materiałów korpusowych), wraz z definicjami pragmatycznymi¹¹ i ilustracją materiałową.

8.14. Skrzydlate słowa. Czym są skrzydlate słowa, określić nie jest łatwo. Najlepszy dowód, że nawet obszerna monografia W. Chlebdy (2005)¹² nie przynosi ścisłej definicji. Można by powiedzieć, że są to wypowiedzi (bądź fragmenty wypowiedzi) jakiejś konkretnej osoby, które utrwaliły się w językowym uzusie i są spontanicznie odtwarzane, czasem przyjmując inne znaczenie niż wypowiedź źródłowa. Tak rozumiane skrzydlate słowa mogą zostać uznane za typ frazemów, czy nawet — jak to zdaje się sugerować sam W. Chlebda — za podgrupę frazeologizmów, ich cechą wyróżniającą jest wszak „autorskość”, czyli możliwość wskazania autora, źródła, osoby, od której dana wypowiedź pochodzi (choć źródło to nie musi być znane użytkownikom). Chodziłoby zatem o takie wyrażenia i zwroty, jak np.: *skończyły się żarty, zaczynają się schody; gruba kreska; puścić kogoś w skarpetkach; grupa trzymająca władzę*¹³. We współczesnym języku polskim wydają się one stosunkowo rozpowszechnione nie

¹¹ Według Bednarka i Grochowskiego (1993) definicja pragmatyczna informuje o sytuacji i/lub intencji, w jakiej dana jednostka może być użyta.

¹² Por. omówienie tej książki w niniejszym tomie.

¹³ Jednostki te wiążą się, kolejno, z osobami: Bolesława Wieniawy-Długoszowskiego, Tadeusza Mazowieckiego, Lecha Wałęsy, Lwa Rywina.

tylko w odmianie oficjalnej, ale również w potocznej. Odnotowywanie tych manifestacji językowych w słowniku jest więc uzasadnione. Za kryterium doboru znów wypada uznać frekwencję tekstową, opis zaś takiego hasła powinien objąć:

- znaczenie utrwalone w uzusie językowym;
- ilustracje przykładową;
- dodatkowe informacje dotyczące warunków pierwszego użycia jednostki i jej autorstwa.

Przykładem może być poniższy opis wyrażenia *grupa trzymająca władzę*:

GRUPA TRZYMAJĄCA WŁADZĘ *pejor.* ‘grupa osób, kontrolująca, jawnie bądź w sposób ukryty, działanie państwa lub instytucji’

PRZYKŁ.: *Wspierana przez Moskwę ukraińska grupa trzymająca władzę jest gotowa na wiele, byle tylko tej władzy nie oddać.* Gaz.Wyb. 26.10.2004

ŹRÓDŁO: Użyte przez Lwa Rywina w rozmowie z Adamem Michnikiem w lipcu 2002 roku, której nagranie i publikacja dały początek tzw. „aferze Rywina” (w rzeczywistości: *Jest grupa, która zawołała mnie: [...] I ta grupa po prostu znaczy z jednej strony gwarantuje ci ustawę i gwarantuje ci akceptację zakupu [Polsatu]. [...] Ponieważ ta grupa trzyma w rękę władzę [...]*).

8.15. Przysłowia będą w słowniku notowane w wyborze. Chcielibyśmy odejść od tradycji nasycania hasel w polskich słownikach ogólnych przysłowiami, które pozostają poza czynnym i biernym zasobem przeciętnego Polaka, a skoncentrować się na opisie tych rzeczywiście używanych. W tym jednak wypadku odwołanie się do danych korpusowych może nie wystarczyć, gdyż przysłowia są w tekstach, zwłaszcza pisanych, reprezentowane dość ubogo. Dlatego przewiduje się wyłonienie podstawowego zasobu przysłów do opisu w słowniku przez odwołanie się do badań ankietowych. Opis przysłowia będzie obejmował, podobnie jak frazemu, przede wszystkim warunki, sytuację i cel jego użycia. Ambicją autorów jest wyposażenie również tych hasel w ilustrację materiałową, jednak zdajemy sobie sprawę, że może się to okazać technicznie trudne ze względu na słabą dostępność odpowiednich przykładów. Problemem może być również konieczność odwoływania się do szerokich kontekstów, jak np.:

UDERZ W STÓŁ, A NOŻYCE SIĘ ODEZWA ‘używane jako komentarz w sytuacji, gdy ktoś bierze do siebie to, o czym mowa’

PRZYKŁ.: *Watykański dokument jest krytyczny wobec niektórych nurtów współczesnej teologii religii. Choć autorzy deklaracji nie wymieniają z imienia i nazwiska krytykowanych myślicieli, to adresaci szybko się odnaleźli na kartach „Dominus Iesus”. Zgodnie z zasadą: uderz w stół, a nożyce się odezwą.* W drodze, 382

9. Wielkość słownika

Wielkość słownika jest trudna do oszacowania, szczególnie jeśli chcielibyśmy stosować tradycyjne miary, używane w leksykografii. Docelowym zadaniem jest objęcie opisem w zasadzie wszystkich jednostek języka, możliwych do odnalezienia w tekstach deklarowanego przedziału czasowego. Można więc stwierdzić, że praca nad słownikiem właściwie nie skończy się nigdy, ponieważ materiał będzie przyrastać,

a możliwość uzupełniania haseł jest technicznie łatwa. Paradoksalnie, w tym punkcie forma słownika książkowego okazałaby się korzystniejsza, gdyż wymusiłaby w którymś momencie zakończenie prac.

10. Przewidywany plan prac

Ze względów praktycznych, także związanych z zasadami finansowania projektów naukowych, zadanie to musi być wykonywane etapami. W pierwszej kolejności planuje się zatem wypracowanie ścisłych i szczegółowych zasad opisu, jego przygotowanie pod względem teoretycznym, technicznym¹⁴ i materiałowym wraz z opracowaniem podstawowego zasobu leksykalnego polszczyzny. Składać się nań będzie 15 tysięcy „gniazd leksykalnych”, pod którym to pojęciem rozumiemy leksem i wszystkie hasła, jakie w związku z tym leksemem powinno się opracować, biorąc pod uwagę wyżej przedstawione zasady. Wbrew pozorom nie wydaje się to mało. Ów podstawowy zasób leksykalny stanowią wyrazy najczęściej w polszczyźnie występujące (ich wyłonienia dokonamy na podstawie słowników frekwencyjnych i frekwencji w korpusach) i najbogatsze w znaczenia, reprezentowane także w wielu jednostkach nieciągłych, frazematycznych i przysłowiach. Przykładowo, gdyby pojedynczy artykuł hasłowy **głowa** z USJP tylko przekształcić zgodnie z zaprezentowanymi tu zasadami, powstałoby 186 odrębnych haseł (możliwe, że dałoby się znaleźć jeszcze więcej jednostek godnych odnotowania), a np. w wypadku hasła **stół** byłoby ich 27.

Na wykonanie tego zadania przewidujemy pięć lat. W latach następnych powstały zbiór haseł mógłby być sukcesywnie uzupełniany, w tempie zależnym od możliwości technicznych i finansowych zespołu wykonawców. Kwestią natomiast udostępniania przygotowywanej bazy danych, warunkami technicznymi i finansowymi tego udostępniania, także publikacji ewentualnych wydań książkowych lub tworzenia na podstawie bazy słowników pochodnych zajmą się najprawdopodobniej inne, powołane do tego organy.

11. Zakończenie

Prezentacja niniejsza przynosi uściślenie wielu ustaleń uczynionych we wcześniejszej wersji koncepcji. Rzecz jasna, nie jest ono jednak wystarczające. Bardzo wiele zagadnień zostało przedstawionych tylko szkicowo; właściwie prawie wszystkie wymagają szczegółowego namysłu. Prace teoretyczne będą więc kontynuowane tak długo, aż zaowocują rozwiniętą, pełną koncepcją, której manifestacją będzie instrukcja redakcyjna przyszłego słownika.

¹⁴ W niniejszej prezentacji świadomie pomijamy techniczno-informatyczne aspekty przedsięwzięcia, które swoją drogą stanowią nie lada wyzwanie.

Bibliografia

- Bańko M., 2001, Z pogranicza leksykografii i językoznawstwa. Studia o słowniku jednojęzycznym, Warszawa.
- 2006, *Circulus vitiosus — circulus odiosus*. Uwagi o cyrkularności w definicjach słownikowych, [w:] *Munuscula linguistica in honorem Alexandrae Cieślíkowa oblata*, red. K. Rymut i in., Kraków, s. 21–34.
- Bednarek A., Grochowski M., 1993, Zadania z semantyki językoznawczej, Toruń.
- Bogusławski A., 1988, Język w słowniku. Desiderata semantyczne do wielkiego słownika polszczyzny, Wrocław.
- Chlebda W., 1993, Frazematyka, [w:] *Encyklopedia kultury polskiej XX wieku. Współczesny język polski*, red. J. Bartmiński, Wrocław, s. 327–334.
- 2005, Szkice o skrzydlatych słowach. Interpretacje lingwistyczne, Opole.
- Cruse D.A., 1986, *Lexical Semantics*, Cambridge University Press, Cambridge.
- Dunaj B., Przybylska R., Żmigrodzki P., 2006, Zarys koncepcji wielkiego słownika języka polskiego, *Polonica XXVI–XXVII*, s. 5–16.
- Grochowski M., 1982, *Zarys leksykologii i leksykografii*, Toruń.
- Grzegorzczkowska R., 1990, *Wprowadzenie do semantyki językoznawczej*, Warszawa.
- Lewandowska-Tomaszczyk B., red., 2005, *Podstawy językoznawstwa korpusowego*, Łódź.
- Miodunka W., 1989, *Podstawy leksykologii i leksykografii*, Warszawa.
- Piotrowski T., 1994, *Z zagadnień leksykografii*, Warszawa.
- 2003, Językoznawstwo korpusowe. Wprowadzenie do problematyki, [w:] *Językoznawstwo w Polsce. Stan i perspektywy*, red. S. Gajda, Opole, s. 143–154.
- Przepiórkowski A., 2004, *Korpus IPI PAN. Wersja wstępna. The IPI PAN Corpus. Preliminary version*, Warszawa.
- Żmigrodzki P., 2005, *Wprowadzenie do leksykografii polskiej*, wyd. II, Katowice.
- 2006, Leksyka związana z koleją w *Uniwersalnym słowniku języka polskiego*. Wybrane problemy metaleksykograficzne, *Polonica XXVI–XXVII*, s. 17–36.
- Żmigrodzki P., Przybylska R., Dunaj B., 2006, O potrzebie wielkiego słownika języka polskiego, *LingVaria*, nr 1, s. 171–179.