

Jan Wawrzyńczyk

Od słownika ortograficznego
do „rejestru wyrazów
występujących w języku polskim”

Zamożniejsi użytkownicy języka polskiego interesujący się jego pisownią mogą dzisiaj kupić sobie w księgarniach aż cztery „wielkie” słowniki ortograficzne, z których każdy liczy ponad 100 000 haseł¹. Wydaje się, że ta niewyobrażalna kilkanaście lat temu obfitość (świadcząca o potędze wolnego rynku, który nie może pozostawić w spokoju także nauki), stanowi dla przeciętnego Polaka niezajmującego się zawodowo swoim językiem jako narzędziem komunikacji społecznej raczej *embarras de richesse* i jest przyczyną jego frustracji, kiedy odkryje rozmaite niezgodności między zaleceniami zawartymi w owych obszernych dziełach². Dla lingwisty natomiast ukazanie się w krótkich odstępach czasu wymienionych słowników ortograficznych, największych w dziejach naszej pisowni, to zapewne bardzo ciekawe, zasługujące na uwagę wydarzenie. Szczególnie opracowania te winny poruszyć historyka leksyki nowopolskiej, właśnie ze względu na ich objętość; niewątpliwie uderzający dlań będzie także podtytuł jednego z nich.

W pracach badawczych nad rozwojem polskiego słownictwa po II wojnie światowej prowadzonych w Polskiej Akademii Nauk³ nie wyzyskiwano słowników ortograficznych jako źródeł dokumentacji użycia, istnienia poszczególnych wyrazów w kolejnych okresach rozwoju polszczyzny. Trudno powiedzieć, czy była to w pełni świadoma decyzja badaczy, ponieważ we wzmiankowanych monografiach nie ma żadnej deklaracji na ten temat. Brak jej jest raczej świadectwem pewnej przypadkowości w doborze źródeł dokumenta-

¹ Zob.: Polański 1999; Markowski 2000; Janik-Płocińska, Sas, Turczyn 2001; Karpowicz 2001 (to ostatnie opracowanie oznaczam dalej skrótem SORW).

² Analizą owych rozbieżności nie będę się jednak zajmował; na pewno ktoś ją kiedyś podejmie.

³ Zob. powstałe publikacje: Smółkowa 1976; Smółkowa, Tekiel 1977; Tekiel 1988-1989; Smółkowa 1998-1999.

cyjnych. Na to, że tak było, że w selekcji nie kierowano się wyłącznie apriorycznymi kryteriami, sformułowanymi eksplicytnie i precyzyjnie, zdaje się wskazywać następujący fakt: w zestawie źródłowym znalazły się m.in. słowniki dwujęzyczne, angielsko-polskie // polsko-angielskie, francusko-polskie // polsko-francuskie i niemiecko-polskie // polsko-niemieckie, ale już nie rosyjsko-polskie // polsko-rosyjskie (por. Smółkowa 1976: 10, 109-112). Uznanie, że te ostatnie są tutaj merytorycznie zupełnie nieprzydatne, byłoby bezpodstawne; ich materiał leksykalny ma taką samą wartość dokumentacyjną, jak i słowniki obejmujące języki zachodnie⁴. Odrębnie należy się zastanowić, czy właśnie słowniki dwujęzyczne bezwzględnie zasługują na miejsce w zestawie źródeł dokumentacyjnych, a nie na przykład (i między innymi) słowniki ortograficzne języka polskiego. Oba rodzaje źródeł dostarczają informacji o tym, że poszczególne wyrazy były znane autorom danych słowników, przynajmniej biernie, że zostały przez nich (co najmniej przez nich) użyte w danym okresie. Nawet jeśli owo użycie polegało tylko na przepisaniu, przeniesieniu z innych słowników, którymi się posłużył dany autor, do słownika nowo tworzonego, to było to przecież użycie, czyli dany wyraz był w użyciu. Notabene, pojawienie się jakiegoś wyrazu w słowniku ortograficznym można uznać za zjawisko dokumentacyjnie donioślejsze, ponieważ zwykłe użytkowników słowników ortograficznych jest na pewno więcej, niż użytkowników słowników dwujęzycznych.

Odmienność kryteriów doboru haseł w słownikach dwujęzycznych (koncentracja na jednostkach sprawiających trudności translatoryczne) i w słownikach ortograficznych (koncentracja na jednostkach, których pisownia może nastroczać trudności) nie zmusza autora danego słownika do preferowania w jego siatce haseł wyrazów mało używanych czy nieużywanych, zapomnianych itp. Wypadki występowania w słownikach dwujęzycznych leksemów nieobecnych w żadnym słowniku definicyjnym, czyli semantycznym, języka polskiego, jak też fakty wcześniejszej rejestracji leksemów w słownikach przekładowych niż jednojęzycznych, nie są w polskiej leksykografii czymś zupełnie wyjątkowym⁵. Z tego względu słuszne jest traktowanie słowników przekładowych jako użytecznych źródeł dokumentacyjnych, potrzebnych przy ustalaniu chronologii słownictwa polskiego.

⁴ Znajomość dorobku polskiej leksykografii przekładowej pozwala mi twierdzić, że poziom jakościowy słowników rosyjsko-polskich/polsko-rosyjskich nie odbiegał nigdy, na całej przestrzeni jej dziejów, od poziomu naszych słowników angielsko-polskich, niemiecko-polskich itd. Wszystkie one miały i mają podobne zalety, jak też obciążone są identycznymi typologicznie wadami, skazami tradycji słownikarskiej, ukształtowanej zasadniczo w XIX w.

⁵ Zob. m.in. Mędelska, Marszałek 2000: 39-49 czy wcześniejszą informację w moim artykule (Wawrzyńczyk 1990: 58-60).

Analogicznie słowniki ortograficzne: także z nich historyk polskiego słownictwa może wyekscerpować niejedyn wyraz, który został w ogóle pominięty przez słowniki semantyczne, albo pojawił się w nich później, po zaistnieniu w słowniku ortograficznym.

Tak na przykład dwa dawniejsze popularne w swoim czasie słowniki ortograficzne (Arct 1936; Jodłowski, Taszycki 1968) rejestrują wiele wyrazów, które w literaturze leksykologicznej zostały uznane za neologizmy powojenne, z dokumentacją wyłącznie prasową pochodzącą z okresu 1972-1992. Należą tu m.in. twory (rodzime derywaty i zapożyczenia): *antyhigieniczny*, *antykatolik*, *antyreligijny*, *autohipnoza*, *challenge*, *chłoporobotnik* itd.⁶

Słowniki ortograficzne języka polskiego niewątpliwie rejestrują jego wyrazy, niekiedy, jak wskazałem, wyprzedzając w czasie słowniki semantyczne, ale nie znaczy to, że są właśnie rejestrami leksyki. Jest to ich wtórna funkcja; im mniejszy jest słownik ortograficzny, tym mniejsze ma on znaczenie jako rejestr wyrazów i odwrotnie: im jest większy, tym ważniejszy, przydatniejszy jest także jako „rejestr wyrazów występujących w języku polskim”. Przytoczone w cudzysłowie znamienne określenie znalazło się tylko w jednym spośród dzieł ortograficznych wymienionych w przyp. 1 jako podtytuł opracowania, które nie jest najobszerniejsze, gdyż liczy ok. 160 000 haseł; jest to słownik T. Karpowicza. Wyprzedza go wyraźnie słownik ortograficzno-fleksyjny B. Janik-Płocińskiej, M. Sas i R. Turczyna z 182 154 hasłami, w związku z czym można by rzec, że zasługuje on na wyróżniające miano rejestru wyrazów języka polskiego bardziej niż SORW. Podtytuł *Rejestr wyrazów występujących w języku polskim* jest zwodniczy. Jego autor wprowadzie tłumaczy, na skrzydełku obwoluty, że nie zawarł w SORW wszystkich polskich wyrazów i podaje (niektóre) przyczyny niekompletności swojego rejestru, ale nie każdy czytelnik zwróci uwagę na tekst tego wyjaśnienia *de facto* sprostowania na nietrwałej obwolucie; obwoluty zwykle szybko się niszczy czy giną, a podtytuł na karcie tytułowej pozostaje i wprowadza w błąd. Jeśli jest jedynie chwytem reklamowym, to raczej niefortunnym.

Niekompletność SORW widać nie tylko na tle wydanego mniej więcej w tym samym czasie, obszerniejszego słownika ortograficzno-fleksyjnego B. Janik-Płocińskiej, M. Sas i R. Turczyna. Nie jest na przykład jasne, dlaczego w siatce haseł SORW brak wielu pozycji figurujących chociażby w małym

⁶ Szczegółów bibliograficznych nt. ich dokumentacji nie podaję tu z oszczędności miejsca. Tymczasem gromadzę je w formie elektronicznej, planując w przyszłości 10-tomowe wydanie książkowe pt. *Słownik bibliograficzny języka polskiego*. Ukazał się w druku dopiero t. 1 (Wawrzyńczyk 2000).

słowniczku S. Jodłowskiego i W. Taszyckiego (takich, jak: *abachit*, *achrodekstryna*, *adiuretyna*, *agrobiotyk*, *alantoina*; to tylko wypadki wybrane na chybił trafił z obrębu litery A, dokładne porównanie całego materiału ujawniłoby ich prawdopodobnie kilkaset). Nie wiadomo też, co było powodem niez uwzględnienia znacznej liczby haseł zarejestrowanych w fundamentalnym słowniku pod red. W. Doroszewskiego⁷; brak w SORW m.in. wyrazów *babczyn*, *babczyny*, *babokracja*, *babosz*, *babuńcia*, *bachancki*, *bachandria*, *bagienność*, *bajbak* (znowu wrywkowe przykłady z jednego tylko przedziału literowego), opisanych w słowniku pod red. W. Doroszewskiego⁸. Oczywiście tylko część tych przykładów jest godna uwagi z ortograficznego punktu widzenia, ale przecież autor SORW oświadczył, że w swej pracy odnotowuje także „takie wyrazy, których postać graficzna nie powinna budzić niczyich wątpliwości” (SORW, s. X). Notabene, wcześniej zaznaczył, że jego rejestr został opracowany na podstawie dotychczas wydanych n a j w i ę k s z y c h⁹ słowników jednojęzycznych (por. op. cit., s. III), ale zaraz zastrzegł, że najbardziej dlań pomocny okazał się *Nowy słownik ortograficzny PWN z zasadami pisowni i interpunkcji* pod red. E. Polańskiego, a spośród słowników semantycznych słownik języka polskiego pod red. M. Szymczaka oraz słownik pod red. B. Dunaja (por. tamże)¹⁰.

SORW zatem to bardzo wybiórczy rejestr wyrazów „występujących w języku polskim”. Tym razem, przesunawszy do przodu lewą część cudzysłowu w kontrowersyjnym określeniu, chciałbym zwrócić uwagę na to, że informując, „jak wiele różnorodnych wyrazów składa się na całokształt polszczyzny”¹¹, T. Karpowicz najprawdopodobniej ma jednak na myśli nie cały język polski, lecz niemal wyłącznie język polski opisany w słownikach. Wyrazy zaś występujące w słownikach a wyrazy występujące poza nimi, używane, brzmiące, utrwalane we wszelkich tekstach (nie tylko w słownikach, leksykonach, encyklopediach) pisanych ręcznie, drukowanych, rejestrowanych elektronicznie, to ciągle jeszcze dwa różne, bardzo nierówne pod względem

⁷ Zob. Doroszewski 1958-1969. Jest to główne osiągnięcie polskiej leksykografii 2. połowy XX w., do dziś słusznie zachowujące wielką rangę naukową i praktyczną.

⁸ Słownika tego nie ma na liście źródeł bibliograficznych zamykającej tekst SORW. Można mówić jedynie o pośrednim udziale materiału hasłowego dzieła W. Doroszewskiego w SORW; jednoczesna nieobecność wyrazów *babczyn*, *babczyny*, *babokracja*, *babosz*, *babuńcia*, *bachancki* itp. bardzo licznych grup leksyki w słowniku pod red. E. Polańskiego zdaje się wskazywać, że pośrednikiem był właśnie słownik ortograficzny PWN.

⁹ Podkreślenie moje – J.W.

¹⁰ Słownik ortograficzny pod red. E. Polańskiego ma ok. 125 000 haseł, słownik pod red. M. Szymczaka tylko ok. 80 000 haseł, a słownik ostatni jeszcze mniej.

¹¹ Ten cytat pochodzi z lewego skrzydełka obwoluty SORW.

rozmiarów świata; badaczom narzuca się pierwszy, leksykograficzny świat polszczyzny, wciąż na poznanie czeka świat drugi. Ilu różnych wyrazów polskich mogli użyć Polacy tylko na przestrzeni minionego właśnie XX w.? Użyć jeden raz, dwa razy, trzy razy, trzysta razy, trzy tysiące razy itd.? Taka statystyka przekracza ludzkie możliwości. Można jedynie snuć mniej lub bardziej uzasadnione przypuszczenia, poprzedzone refleksją teoretyczną, nie tylko zresztą nad kluczowym tutaj pojęciem wyrazu, co do wielkości materiału niezarejestrowanego w słownikach języka polskiego, ale na rejestrację zasługującego, wciąż czekającego.

Nawet cząstkowy rejestr wyrazów występujących w języku polskim XX w., utworzony w wyniku zsumowania haseł słowników istniejących (z której to sumy opracowując SORW skreślono, na podstawie niejasnych kryteriów, sporą część słownictwa), ale sporządzony w sposób systematyczny, tzn. uwolniony od niekonsekwencji charakterystycznych dla wszystkich naszych słowników języka polskiego, musi mieć więcej haseł niż SORW.

Najczęściej spotyka się w siatkach haseł luki w prezentacji grup wyrazów jednorodnych derywacyjnie. Oto w SORW znalazły się złożenia liczebniko-przymiotnikowe: *jednogodzinny*, *dwugodzinny*, *trzygodzinny* itd. do *trzynastogodzinny*, zabrakło tworów *czternastogodzinny* i *piętnastogodzinny*, dalej uwzględniony został *szesnastogodzinny*, pominięty *siedemnastogodzinny* i *osiemnastogodzinny*; z „wyższych” złożzeń odszukałem przymiotniki: *dziewiętnastogodzinny*, *dwudziestogodzinny*, *dwudziestoczworogodzinny*, *trzydziestogodzinny* i *czterdziestogodzinny*. Z podgrupy osobliwszej, o strukturze semantycznej ‘x-ipółgodzinny’, wprowadzone zostały tylko dwa derywaty: *ośmioipółgodzinny* oraz *dziewięcioipółgodzinny*. W ogóle można tu mówić o następującej negatywnej prawidłowości cechującej porównywane siatki haseł: jeśli w danym słowniku figuruje jakikolwiek derywat z dłuższej nieco¹² serii, to jest pewne, że zabraknie w jego siatce okazów równie użytecznych, co te, których nie pominięto. Tak na przykład w tymże SORW jest przymiotnik *czterocyfrowy*, ale brak formacji *pięciocyfrowy*, jest *sześcioczłonowy*, ale brak *pięcioczłonowy*, jest *siedmiodniowy*, brak *sześciodniowy*, jest *siedmiojęzyczny*, brak *sześćojęzyczny*; są przymiotniki *krakowski*, *warszawski*, ale nie ma *poza-krakowski*, lecz jedynie *pozawarszawski*; jest *poodcinać*, ale brak *poodrzynać*;

¹² Nie mówiąc już o seriach nieskończenie długich, których okazów nie da się wylczyć; mamy wtedy do czynienia w istocie rzeczy z formami fleksyjnymi, a nie słowotwórstwem (por. zwłaszcza odśłowniki na *-anie*, *-anie się*, *-enie*, *-enie się*, przysłówki na *-o*). Te trudne kwestie o nader istotnym znaczeniu dla sprawy rejestru wyrazów występujących w języku polskim naświetla m.in. A. Bogusławski (1992). Materiał hasłowy SORW nie został uporządkowany do końca także pod tym względem.

obok *deflacja* podano *deflacyjny*, ale przy *reflacja* zabrakło *reflacyjny*; do *San Marino* dorobiono *sanmaryński*, ale obok *Sankt Petersburg* zabrakło *sankt-petersburski*. Itd. itp.

Mniejsze lub (raczej) większe luki w hasłach istniejących słowników ortograficznych i // lub semantycznych, niemające nic wspólnego ze znanym problemem¹³ *cases vides*, należy tropić myśląc o opracowaniu słownika, którego główną (lub lepiej: jedyną) funkcją byłaby rejestracja polskich wyrazów nie tylko z punktu widzenia relacji derywacyjno-fleksyjnych. Nader ważna jest tutaj także perspektywa onomazjologiczna. Kontrola onomazjologiczna ujawnia w SORW m.in. brak informacji na temat wielu oznaczeń kobiet (por. „samotne” masculina *bibliofil*, *filister*, *gestapowiec*, *polonofil*, *przedrzeźniacz* itp.); brak wielu rzeczownikowych i przymiotnikowych wykładników treści antonimicznych, zwłaszcza wśród terminów (por. obecność m.in. wyrażeń *anielskość*, *rusofil*, *przednionapędowy*, *tylnojęzykowość*, ale brak odpowiednich nazw *diabelskość*, *rusofob*, *tylnonapędowy*, *przedniojęzykowość* itd.); odczuwa się też nieobecność różnych hiperonimów, kohiponimów czy konwersów.

Tylko niektóre uzupełnienia luk wskazanych w hasłowniku SORW można znaleźć w dziele B. Janik-Płocińskiej, M. Sas i R. Turczyńska, którzy jednak nie reklamowali się jako twórcy leksykograficznej nowości „rejestru wyrazów występujących w języku polskim”.

Źródło podstawowe, z którego należy czerpać materiał do budzącego jak najmniej zastrzeżeń, wciąż nieułożonego rejestru tego rodzaju, stanowią oczywiście teksty, a nie słowniki. Zasoby leksykalne utrwalone w tekstach, ale nie wyekscerpowane z nich i nie opisane w słownikach języka nowopolskiego, są nadal wielkie, wciąż jest ogromnie dużo do zrobienia na tym polu. Orientując się (być może nieco dokładniej niż inni badacze naszego słownictwa) zwłaszcza w problematyce leksyki używanej w okresie 1918-1939 i do dziś pozostającej poza okładkami słowników semantycznych czy ortograficznych języka polskiego, szacuję, że rejestr wyrazów polszczyzny XIX-XX w. może liczyć co najmniej 400 000 pozycji.

Bibliografia

- Arct M., 1936: *Słownik ortograficzny języka polskiego*, Warszawa.
Bogusławski A., 1992: *Two essays on inflection*, Warszawa.
Doroszewski W. (red.), 1958-1969: *Słownik języka polskiego*, t. 1-11, Warszawa.

¹³ Zdaniem W. Mańczaka: językoznawczym mitem (por. Mańczak 1996: 120).

- Janik-Płocińska B., Sas M., Turczyn R., 2001: *Wielki słownik ortograficzno-fleksyjny*, Warszawa.
- Jodłowski S., Taszycki W., 1968: *Słownik ortograficzny i prawidła pisowni polskiej*, Wrocław.
- Karpowicz T., 2001: *Słownik ortograficzny. Rejestr wyrazów występujących w języku polskim*, Warszawa.
- Mańczak W., 1996: *Problemy językoznawstwa ogólnego*, Wrocław.
- Markowski A. (red.), 2000: *Ortograficzny słownik języka polskiego*, Warszawa.
- Mędelska J., Marszałek M., 2000: *Twórzmy słownik „polskiego języka radzieckiego”*, *Poradnik Językowy*, z. 8, s. 39-49.
- Polański E. (red.), 1999: *Nowy słownik ortograficzny PWN z zasadami pisowni i interpunkcji*, Warszawa.
- Smólkowa T., 1976: *Nowe słownictwo polskie. Badania rzeczowników*, Wrocław.
- Smólkowa T., Tekiel D., 1977: *Nowe słownictwo polskie. Przymiotniki i przysłówki*, Wrocław.
- Smólkowa T. (red.), 1998-1999: *Nowe słownictwo polskie. Materiały z prasy lat 1985-1992*, cz. I-II, Kraków.
- Tekiel D. (red.), 1988-1989: *Nowe słownictwo polskie. Materiały z prasy lat 1972-1981*, cz. I-II, Wrocław.
- Wawrzyńczyk J., 1990: *O pewnym źródle dokumentacyjnym słownictwa polskiego*, *Poradnik Językowy*, z. 1, s. 58-60.
- Wawrzyńczyk J., 2000: *Słownik bibliograficzny języka polskiego*, t. 1 (A-Ć), Warszawa.

Рэзюме

В статье высказываются замечания о новых больших орфографических словарях современного польского языка. Составители некоторых из них стремятся зафиксировать значительное количество лексических единиц, до сих пор еще не отмеченных ни в одном словнике не только орфографических, но и толковых словарей. Более пристально рассматривается объемистый труд Т. Карповича (Т. Karpowicz, *Słownik ortograficzny. Rejestr wyrazów występujących w języku polskim*, Warszawa 2001), название которого выделяется уникальным подзаголовком, отражающим идею орфографического словаря как списка слов того или иного языка. Выдвигаются критические соображения по поводу реализации данной идеи и указываются методологические и практические недостатки работы Т. Карповича.