



Dorota Adamiec 
Instytut Języka Polskiego Polskiej Akademii Nauk, Kraków
dorota.adamiec@ijp.pan.pl

Magdalena Majdak 
Instytut Języka Polskiego Polskiej Akademii Nauk, Kraków
magdalena.majdak@ijp.pan.pl

Włodzimierz Gruszczyński 
Instytut Języka Polskiego Polskiej Akademii Nauk, Kraków
wlodzimierz.gruszczyński@ijp.pan.pl

Mateusz Żółtak 
Austrian Centre for Digital Humanities and Cultural
Heritage, Wien
mateusz.zoltak@oeaw.ac.at

BAROKOWA POLSZCZYŻNA W INTERNECIE, CZYLI ELEKTRONICZNY SŁOWNIK JĘZYKA POLSKIEGO XVII I XVIII WIEKU¹

Słowa kluczowe: leksykografia, historia leksykografii, słownik elektroniczny, korpus tekstów, polszczyżna barokowa

Keywords: lexicography, history of lexicography, electronic dictionary, text corpora, baroque Polish

W listopadzie 2021 r. *Elektroniczny słownik języka polskiego XVII i XVIII wieku* (e-SXVII)² został laureatem 4 edycji Nagrody Komitetu Językoznawstwa PAN za wybitne osiągnięcie naukowe w zakresie językoznawstwa. Nagroda została przyznana w kategorii zespołowej prof. dr. hab. Włodzimierzowi Gruszczyńskiemu wraz z kierowaną przez niego Pracownią Historii Języka Polskiego XVII i XVIII Wieku IJP PAN, w której powstaje e-SXVII. Słownik doceniono jako „osiągnięcie w zakresie wytworzenia zasobów lub narzędzi językowych”³. Jak się wydaje, takie wyróżnienie to dobry powód, aby poinformować środowisko językoznawcze o postępach

1 Wkład współautorów w powstanie artykułu wynosi: W. Gruszczyński – 35%, D. Adamiec – 35 %, M. Majdak – 20%, M. Żółtak – 10%.

2 Zob. <https://sxvii.pl/>.

3 Zob. https://komjezyk.pan.pl/index.php?option=com_content&view=article&id=132&Itemid=192.

prac nad słownikiem, o zastosowaniu nowych metod w jego opracowywaniu oraz przypomnieć zasady jego redagowania.

Pracownia Historii Języka Polskiego XVII i XVIII Wieku IJP PAN została utworzona w grudniu 1954 r. jako placówka Zakładu Językoznawstwa PAN w Warszawie, kierowanego przez prof. Witolda Doroszewskiego. Od początku celem działania pracowni były badania polszczyzny barokowej, skoncentrowane przede wszystkim na słownictwie epoki, które miało zostać opisane w wielotomowym *Słowniku języka polskiego XVII i 1. połowy XVIII wieku* (SXVII). Pracownią kierowali kolejno: prof. Halina Koneczna, prof. Stanisław Skorupka, dr Kazimierz Żelazko, dr Krystyna Siekierska, a od roku 2004 prof. Włodzimierz Gruszczyński. Zawsze szczupły zespół tworzy obecnie 9 osób: Dorota Adamiec, Renata Bronikowska, Włodzimierz Gruszczyński, Katarzyna Kryńska, Paweł Kupiszewski, Magdalena Majdak, Wiesław Morawski, Ewa Rodek, Aleksandra Wieczorek.

W początkowym okresie działania pracowni koncentrowały się na przygotowaniu koncepcji planowanego słownika i ustaleniu kanonu źródeł. Równolegle rozpoczęto ekscerpcję materiałową. Efektem tych prac jest kartoteka obejmująca około 2,8 mln kart z cytatami. W latach 1965–1973 został wydany dwutomowy SJPas, opracowany przez zespół pracowni. Słownik języka autora reprezentatywnego dla baroku stanowił przygotowanie do całościowego opracowania polszczyzny tej epoki. W roku 1996 ukazał się zeszyt próbny SXVII. Stał się on przedmiotem dyskusji na plenarnym posiedzeniu Komitetu Językoznawstwa PAN. W okresie 1999–2004 wydano pierwszy tom SXVII, na który złożyło się pięć zeszytów, obejmujących część wstępną słownika oraz około 1800 artykułów hasłowych na literę A.

W roku 2004 rozpoczęto prace nad e-SXVII i została podjęta decyzja o zawieszeniu dalszej publikacji SXVII w formie papierowej. Wprowadzenie tak rewolucyjnych (zwłaszcza na tamte czasy) zmian podyktowane było dążeniem do usprawnienia i przyspieszenia leksykograficznego opracowania słownictwa średniopolskiego oraz zachodzącymi ówczesnie zmianami technologicznymi związanymi z szybko postępującą cyfryzacją w świecie nauki. Oczywiście w pracach nad e-SXVII zamierzano wykorzystać doświadczenia i wcześniejszy dorobek pracowni, w szczególności kartotekę materiałową SXVII⁴ oraz SJPas. Koniecznym wstępem do zmian stało się przygotowanie przez kierownika pracowni, W. Gruszczyńskiego, założeń słownika elektronicznego. Jednocześnie pracownia nawiązała współpracę z informatykiem Mateuszem Żółtakiem⁵, który na podstawie tych założeń zaprojektował i stworzył cały system informatyczny e-SXVII, czyli bazę danych, formularz dla autorów haseł oraz program przekształcający dane z bazy na tekst artykułów hasłowych.

4 W ramach projektu Repozytorium Cyfrowego Instytutów Naukowych kartoteka została zdigitalizowana i jest dostępna pod adresem: <https://rcin.org.pl/ijp/dlibra/publication/20029>.

5 Współpraca ta (z małymi przerwami) trwa do dziś. M. Żółtak wprowadza zmiany służące ulepszeniu systemu i czuwa nad sprawnością jego działania.

Koncepcja pierwszego wówczas polskiego słownika powstającego wyjściowo w postaci elektronicznej była przez członków zespołu wielokrotnie prezentowana i poddawana pod dyskusję na konferencjach, seminariach, wykładach⁶. Konsultacje prowadzono zarówno ze środowiskiem językoznawczym, jak i informatycznym.

Początki e-SXVII wiązały się z różnymi trudnościami. Członkowie zespołu musieli, przełamując swoje wieloletnie przyzwyczajenie do tradycyjnych metod, nauczyć się nowego modelu pracy. Zadanie to było tym trudniejsze, że pracownia borykała się wtedy z podstawowymi brakami sprzętowymi: w 2004 r. dysponowano tylko jednym komputerem, brakowało szybkiego połączenia z internetem i serwera z prawdziwego zdarzenia. Koncepcja zmian spotkała się z różnymi reakcjami polemicznymi. Niektórzy specjaliści krytykowali zastosowane rozwiązanie informatyczne (bazę danych), sugerując przechowywanie danych w postaci dokumentu TEI-XML⁷. Zdystansowanie części środowiska językoznawczego wynikało z wykluczenia cyfrowego, które 15 lat temu zdarzało się jeszcze nierzadko wśród badaczy humanistów. Nieuchronnym skutkiem trudnych początków musiało być okresowe zahamowanie tempa prac. Statystyki z 2006 r. pokazują, że w bazie e-SXVII znajdowało się 1200 artykułów hasłowych. Na początku 2022 r. to już ponad 46 000 artykułów hasłowych (o różnym stopniu opracowania).

Elektroniczna forma słownika otworzyła nowe możliwości zarówno w zakresie pracy redaktora nad artykułem hasłowym, jak i w szerszej perspektywie ciągłego całościowego kształtowania e-SXVII. Fundamentalną zmianą jest uwolnienie od konieczności opracowywania artykułów hasłowych w porządku alfabetycznym, który narzucała edycja papierowa. Otwiera to perspektywę innego grupowania opracowywanych leksemów. Można seryjnie redagować hasła zbliżone pod względem na przykład semantycznym, funkcjonalnym bądź gramatycznym. Zasadniczą nowością w pracy redakcyjnej stała się możliwość wprowadzania zmian do już opracowanych haseł. Taka „otwartość” artykułu hasłowego sprawia, że w wypadku znalezienia nieznanego wcześniej kontekstów użycia, które wskazują na nowe znaczenia leksemu czy też jego nienotowane wystąpienia w połączeniach wyrazowych, autor może uzupełnić, rozwinąć opis leksykograficzny. Trzeba też przyznać, że „otwartość” hasła bywa swego rodzaju pułapką: trudno zakończyć pracę i uznać hasło za opracowane. Elektroniczna forma słownika sprzyja również zmianom na ogólniejszym poziomie, w zakresie założeń metodologicznych opisu słownikowego. Założenia te mogą być modyfikowane, precyzowane, gdy pojawia się taka potrzeba wynikająca z analizy

6 Warto wspomnieć m.in. artykuły naukowe W. Gruszczyńskiego (2005, 2010), wystąpienia na konferencjach, na przykład „Przyszłość słowników w Polsce” w Kamieniu Śląskim (2004), „Z zagadnień leksykologii i leksykografii języków słowiańskich” w Toruniu (2007), „II Konferencji Głosa do leksykografii polskiej” w Warszawie (2009), wykłady popularnonaukowe na Festiwalu Nauki w 2005 i 2010 r.

7 <https://tei-c.org/>.

materiału językowego⁸. Jedyne ograniczenie stanowi warunek zachowania spójności na poziomie makrostruktury słownika. Funkcjonalności techniczne elektronicznego słownika mogą być ciągle doskonałe. W ten sposób wprowadzono do e-SXVII mechanizm automatycznego umieszczania raz opracowanego frazeologizmu w artykułach hasłowych odnoszących się do wszystkich wyrazów wchodzących w skład tego związku. Podobnie z czasem pojawiła się możliwość dodawania ilustracji, szczególnie cenna w artykułach hasłowych dotyczących roślin lub archaizmów rzeczowych. Wydaje się, że zwłaszcza w świecie, w którym zmiana okazuje się jedyną stałą rzeczą, możliwość wprowadzania modyfikacji w słowniku sprawia, że staje się on słownikiem żywym, gotowym do reakcji na nowe, dziś jeszcze nawet nieznanne potrzeby.

System informatyczny ma typową architekturę sieciowych systemów zarządzania treścią. Składa się z relacyjnej bazy danych, formularza do wprowadzania danych dla redaktorów haseł oraz komponentu wyszukiwania i wizualizacji haseł na potrzeby użytkownika. Dwa ostatnie elementy wykonane są w postaci stron internetowych. Z uwagi na to, że w momencie rozpoczynania prac nad systemem e-SXVII gotowe systemy zarządzania treścią dopiero raczkowały i nie dostarczały wystarczających funkcjonalności⁹, komponenty e-SXVII zostały zaprogramowane od podstaw. O wyborze języka PHP zdecydowały indywidualne preferencje programisty. Jako system bazy danych wybrano PostgreSQL. Był to jedyny dostępny w 2004 r. system baz danych dostarczający wymaganych funkcjonalności, a jednocześnie darmowy. Zdecydowano się na bardzo szczegółowy i restrykcyjny model danych, z odrębnymi tablicami i polami dla każdego elementu opisu hasła.

Zawartość kanonu źródeł e-SXVII odzwierciedla historię zmian, które zaszły na przestrzeni 70 lat prac nad leksyką barokową prowadzonych przez IJP PAN. Początkowy trzon bazy źródeł stanowiły teksty ekscerpowane na potrzeby wersji drukowanej SXVII. Materiały te (zaczepnięte pierwotnie z 276 pozycji o różnej objętości – por. SXVII: XX) zgromadzono w papierowej kartotece. Stosowano różny zakres ekscerpji: wobec części tekstów posłużono się ekscerpją pełną, która objęła wszystkie wyrazy z całego zabytku, z innych źródeł wybrano obszerny fragment i poddano ekscerpji pełnej. Najczęstszą metodą była ekscerpja wybiórcza (rozumowana), obejmująca wybrane wyrazy. Już w trakcie wydawania na papierze kolejnych zeszytów SXVII prowadzono uzupełniającą ekscerpję dodatkową – zwięźliła koncepcja słownika z zawsze otwartym kanonem źródeł, który warto stale rozszerzać o dopiero odnalezione czy udostępnione teksty¹⁰.

8 Przykładem takiej zmiany jest opis klasy liczebników oraz imiesłówów.

9 Na przykład pierwsze wydanie systemu zarządzania treścią WordPress datowane jest na maj 2003 r.

10 Z tego względu zapadła decyzja o rezygnacji z podawania informacji statystycznych, które były obowiązkową częścią artykułu hasłowego w wersji papierowej słownika.

Rozwój bibliotek cyfrowych otworzył nowe możliwości materiałowe dla e-SXVII. Cyfryzacja oznaczała nieporównywalnie większą dostępność tekstów z epoki – najpierw głównie starodruków, obecnie także zdigitalizowanych rękopisów w bibliotekach cyfrowych. Rozwój technologiczny umożliwił również prace nad Elektronicznym korpusem tekstów polskich z XVII i XVIII w. (do 1772 r.) (KorBa), które zostały podjęte przez Pracownię Historii Języka Polskiego XVII i XVIII Wieku IJP PAN w ramach dwóch projektów Narodowego Programu Rozwoju Humanistyki¹¹. Efektem tych przemian i przedsięwzięć jest spektakularny wzrost liczby źródeł e-SXVII – aktualnie jest to prawie 2500 pozycji bardzo zróżnicowanych pod względem objętości. W bazie danych obejmującej autorów, tłumaczy, drukarzy i wydawców z epoki znajduje się ponad 1100 rekordów. Zmianom ilościowym zasobu źródeł e-SXVII towarzyszą zmiany jakościowe – coraz łatwiejszy w realizacji staje się obowiązujący od początku postulat dążenia do korzystania z tekstów oryginalnych, a nie z wydań XIX-, XX- czy XXI-wiecznych. Podstawowy typ źródeł stanowią starodruki, w większości będące pierwodrukami. Wzrasta zakres włączenia do materiałów rękopisów, dawniej znacznie ograniczony ze względu na konieczność pozyskania fotokopii. Możliwość wykorzystania rękopisów rośnie również dzięki rozwijającym się technologiom automatycznego odczytywania i przetwarzania pisma ręcznego (Handwritten Text Recognition – HTR). W wypadku tekstów niedostępnych w oryginale wciąż podstawą cytatów materiałowych są krytyczne wydania współczesne (rzadko z XIX w. ze względu na nadmierne ingerencje wydawców). Jednak często zostają one z czasem zastępowane przez edycje z epoki, pojawiające się sukcesywnie w bibliotekach cyfrowych.

Zakres chronologiczny e-SXVII obejmuje lata 1601–1800. Rok 1600 to najpóźniejsza data wydania dla tekstów źródłowych wykorzystywanych w SPXVI. SXVII obejmował zasadniczo tylko 1. połowę XVIII stulecia (choć od początku korzystał z materiałów ekscerpowanych z 3 tomu słownika Michała Abrahama Trotza, wydanego w 1764 r.). Jak wiadomo, wszelkie cezury czasowe funkcjonują w historii języka jedynie jako umowne granice. Rozszerzenie chronologii e-SXVII o 2. połowę XVIII w. wynikało z faktu, że źródła z tego półwiecza uwzględniono w SJPD z założenia w bardzo ograniczonym zakresie¹², więc nadal celowe wydaje się szczegółowe opracowanie leksykograficzne ówczesnego materiału językowego. Podstawę

11 Projekt „Elektroniczny korpus tekstów polskich z XVII i XVIII wieku (do 1772 r.)” realizowany był w latach 2013–2018 (https://korba.edu.pl/query_corpus/), a realizacja projektu „Rozbudowa Elektronicznego korpusu tekstów polskich XVII i XVIII wieku i jego integracja z Elektronicznym słownikiem języka polskiego XVII i XVIII wieku” rozpoczęła się w roku 2019 i planowana jest do końca 2023 r.

12 „[...] po dłuższych dyskusjach zgodzono się, że w Słowniku uwzględniona zostanie pewna ściśle określona liczba źródeł z drugiej połowy XVIII wieku, które ekscerpowane będą nie dla dania wyczerpującej charakterystyki językowej tego okresu, ale przede wszystkim jako materiał oświetlający dalszy rozwój do dziś żywych wyrazów polskich” (SJPD: XIV).

włączenia do kanonu źródeł e-SXVII stanowi datacja konkretnego tekstu. W ten sposób trafił do kanonu na przykład późny utwór Piotra Skargi *Mesjasz nowych arianów* (Kraków 1612) oraz wczesny tekst Jana Śniadeckiego *Prospekt dzieła pod tytułem Rachunku algebraicznego teoria* (Kraków 1781). Słowniki historycznojęzykowe tworzą tu kontinuum: starsze teksty P. Skargi są uwzględnione w SPXVI, zaś późniejsze prace J. Śniadeckiego znalazły się wśród źródeł SJPD.

Zakres geograficzny e-SXVII uwzględnia teksty polskojęzyczne w pierwszej kolejności z wszystkich regionów I Rzeczypospolitej (Małopolska, Mazowsze, Wielkopolska, Podlasie, Pomorze i Prusy, Wielkie Księstwo Litewskie, Ziemie Ruskie, Inflanty), a także teksty wydane po polsku poza tym terenem, przede wszystkim na Śląsku (który mimo że w ciągu tych dwóch stuleci znajdował się poza granicami państwa, to zachowywał wspólnotę językową), ale również w dalszej zagranicy (Lipsk, Amsterdam, Rzym, Nancy, Antwerpia).

Podstawowym kryterium przy rozstrzygnięciach odnoszących się do makrostruktury e-SXVII było przyjęcie takich zasad doboru haseł, które pozwoliłyby na opis możliwie pełnego zasobu leksykalnego polszczyzny w badanym okresie historycznym. W e-SXVII oprócz artykułów hasłowych poświęconych klasom jednostek leksykalnych, standardowo opisywanych w słownikach ogólnych, są także artykuły hasłowe poświęcone mniej typowym grupom słownictwa, przede wszystkim przymiotnikom odnazwowym oraz nazwom mieszkańców. W słownikach języka ogólnego zazwyczaj nie uwzględnia się nazw własnych, natomiast w e-SXVII ze względu na liczne, charakterystyczne dla epoki wystąpienia nazw własnych w użyciach przenośnych podjęta została decyzja o włączaniu tej grupy nazw własnych do zasobu haseł słownika. Por. hasło:

ACHILLES [...]

1. mit. nazwa własna »bohater grecki słynny z męstwa i urody«: Onego Hectora Konmi Koło Troi Włoczeł Achilles mezny na hanbe Oycowską. *OpalK-Sat 81*. [...];
2. przen. »mężny rycerz«: Te starożytne limfy w sercach udzielnych książąt [...] osobliwe dotąd rościły swych godności delektacje [...] ile kropel, tyle mężnych Herkulesów i odważnych wydając Achillesów [mowa o rodzie Wielopolskich]. *HugLacPrag 20*.

Ze względu na bardzo liczną reprezentację materiałową oraz fakt, że część skrótów ma nieoczywiste rozwinięcia i użytkownicy słownika mogą ich poszukiwać, do e-SXVII włączono hasła opisujące skróty (np. O. 'ojciec', ICHMĆ 'ichmość'). W zasobie haseł e-SXVII obecne są również zasymilowane wyrazy obce, szczególnie licznie zapożyczane ówczesnie z łaciny (np. ASTUS). Kryterium asymilacji stanowią cechy świadczące o włączeniu wyrazu w polski system gramatyczny (polskie końcówki fleksyjne, zaimki potwierdzające adaptację rodzajową).

W okresie opracowywania pierwszego tomu SXVII wydanego drukiem zapadła decyzja o wyłączeniu z opracowania w porządku alfabetycznym przyimków, partykuł, spójników i zaimków. Te kategorie haseł miały zostać opracowane oddzielnie i opublikowane w specjalnych tomach. Słownik elektroniczny odchodzi od tych ograniczeń.

Odrębnymi artykułami hasłowymi są hasła homonimiczne, np. ALBA I »długa biała szata liturgiczna« i ALBA II »miara objętości płynów i ciał sypkich, pół kwarty; naczynie o takiej objętości«. Dla użytkownika słownika te dwa artykuły wyświetlają się razem, jeden pod drugim. Dzięki temu ułatwiony jest jednoczesny ogląd obydwu homonimów.

Makrostruktura e-SXVII została uwolniona (jak już wyżej wspomniano) od obligatoryjnego dla formy papierowej układu haseł. Prace leksykograficzne mogą przebiegać w porządku alfabetycznym, gniazdowym bądź pojęciowym.

Na poziomie makrostruktury e-SXVII możliwe są również zmiany w zakresie wyodrębniania haseł i podhaseł. Imiesłowcy przymiotnikowe w SXVII były opisywane jedynie jako formy gramatyczne czasowników, w żaden sposób niewyróżnione. Żywe procesy adiektywizacji zachodzące ówczesnie wśród imiesłowców, powodujące znaczący brak stabilizacji tych jednostek, skłoniły jednak twórców e-SXVII do zmian sposobu ich opisu leksykograficznego. Teraz imiesłowcy przymiotnikowe są podhasłami albo samodzielnymi hasłami (jeśli nie ma w materiale form finitywnych lub bezokolicznika). Dzięki temu opis imiesłowców i czasowników zyskał na spójności i jednorodności.

Słownik tworzony elektronicznie pozwala na wprowadzenie rozbudowanej i spójnej mediostruktury¹³. Wewnętrzna sieć połączeń pomiędzy poszczególnymi hasłami powstaje równocześnie z kolejnymi artykułami hasłowymi, nie ma potrzeby zestawiania jak dawniej papierowych kartotek odsyłaczy, które nigdy nie były kompletne i nie oddawały całej sieci powiązań między artykułami hasłowymi.

Mediostruktura e-SXVII jest wielowymiarowa. Przy każdym artykule hasłowym istnieje możliwość dodawania odsyłaczy o różnym charakterze: odsyłacze łączące wyraz hasłowy z jego zaprzeczeniem (AFEKT I p. NIEAFEKT I; BEZPIECZNY p. NIEBEZPIECZNY), różnorodne odsyłacze semantyczne i słowotwórcze (BIBLIOTEKARZ p. BIBLIOPOLA; BAŻANTNIK p. BAZANTARNIK). Odsyłacze takie pozwalają na bezpośrednie przejścia między dwoma artykułami hasłowymi. W e-SXVII występują również hasła odsyłaczowe. Mogą one być dwojakiego rodzaju. Kierują użytkownika do właściwego artykułu hasłowego w sytuacji nieoczywistej

13 Pojęcie *mediostruktura* występuje w tekstach dotyczących słowników terminologicznych (np. Łukasik 2014; Szemińska 2014) w odniesieniu do sieci powiązań międzyhasłowych, z zaznaczeniem różnego rozumienia terminu. W słowniku terminologicznym *Leksykografia. Słownik specjalistyczny* (Bielińska 2020: 181) *mediostruktura* została zdefiniowana jako 'system odsyłaczy w publikacji leksykograficznej'.

formy hasłowej (LZA p. NIE LZA) albo wskazują drogę od formy współczesnej hasła (która jest zasygnalizowana nawiasem kwadratowym) do średniopolskiej postaci ([KORAN] p. ALKORAN). Wewnętrzna sieć powiązań między artykułami hasłowymi w słowniku elektronicznym obejmuje też definicje semantyczne. Ten mechanizm jest wykorzystywany w definicjach wyrazów pochodzących od archaizmów rzeczowych, np. w hasle ARKABUŻNIK zapis definicji: »żołnierz ciężkiej jazdy uzbrojony w →arkabuz« umożliwia bezpośrednie przejście do hasła ARKABUZ. Jak już wspomniano wyżej, mediostruktura e-SXVII umożliwia włączenie wielowrazowych struktur (połączeń wyrazowych, związków frazeologicznych, przysłów) do artykułów hasłowych odnoszących się do wszystkich wyrazów wchodzących w ich skład. Taka jednostka jest opracowana w bazie danych słownika tylko raz, więc ewentualne modyfikacje w jej charakterystyce są jednolite we wszystkich powiązanych artykułach, np. identyczny opis frazeologizmu *wdziać cudzy bot na swoje kopyto* będzie się znajdował w artykułach hasłowych: WDZIAĆ, CUDZY, BOT, KOPYTO. Różnokierunkowe powiązania dotyczą również ilustracji materiałowej zamieszczanej przy każdym znaczeniu – lokalizacja cytatu pełni jednocześnie funkcję odnośnika prowadzącego do wykazu wszystkich cytatów z danego źródła obecnych w słowniku oraz do skanu tekstu źródłowego, jeśli jest on dostępny w bibliotece cyfrowej. W odwrotnym kierunku każdy cytat na liście cytatów jest opatrzony odesłaniem do odpowiedniego hasła. Niewidoczne bezpośrednio w artykule hasłowym, ale również obecne w strukturze e-SXVII, są powiązania obejmujące dane tekstów źródłowych: opis bibliograficzny zawiera odesłanie do notki biograficznej autora.

Mikrostruktura e-SXVII obejmuje bloki informacji zwykle uwzględniane w słownikach języka ogólnego. Wyraz hasłowy zapisany jest zgodnie ze współczesną ortografią i opatrzony podstawową charakterystyką gramatyczną (czyli określeniem części mowy), czasem wskazane zostaje też pochodzenie słowotwórcze (zwłaszcza przy przymiotnikach odnazwowych i derywatach homonimicznych, np. AUGSBURSKI przym. od AUGSBURG [miasto w dzisiejszej Bawarii]; CELNY I przym. od CŁO i CELNY II przym. od CEL). W dalszym ciągu artykułu hasłowego wymienione są wszystkie poświadczone w materiale warianty fonetyczne hasła. Kolejne bloki informacji w artykule hasłowym obejmują: notowanie w słownikach (czyli podstawową informację o występowaniu wyrazu na przestrzeni dziejów polszczyzny, od SStp po SJPD), zestawienie występujących w materiale form gramatycznych, etymologię dla wyrazów obcego pochodzenia, definicje znaczeń opatrzonych (w miarę potrzeb) kwalifikatorami i ilustracją materiałową w postaci cytatów źródłowych ułożonych w kolejności chronologicznej (od najstarszych). W obrębie poszczególnych znaczeń opracowywane są połączenia wyrazowe, związki frazeologiczne i przysłowia; przytaczane są także cytaty z użyciami przenośnymi wyrazu hasłowego w danym znaczeniu. Po znaczeniach podajemy wystąpienia metafrazy wyrazu hasłowego, czyli przytoczenia formy wyrazu hasłowego w XVII-wiecznych gramatykach i innych tekstach odnoszących się do języka polskiego jako obiektu. W tego typu

cytatach nie ma żadnych wskazówek co do znaczenia, np. „Singulariter N. Tá Bábá G. tey Bábý [...] Pluraliter N. te Bábý G. tych Bab. *WojnaInst 34*”. Kolejnymi elementami artykułu hasłowego mogą być podhasła. Na końcu umieszczamy odsyłacze. Pod każdym artykułem znajdują się inicjały autora i data ostatniej modyfikacji. Określamy również stopień zaawansowania prac nad hasłem za pomocą umownych określeń: „załączek artykułu hasłowego” lub „[artykuł] w trakcie opracowania”; brak tych określeń oznacza, że artykuł został uznany za zatwierdzony, co jednak nie wyklucza możliwości zmian i uzupełnień, jeśli zajdzie taka potrzeba.

Rozwój humanistyki cyfrowej umożliwił powstanie sieci różnych zasobów elektronicznych polszczyzny XVII i XVIII w. Jednym z obecnie realizowanych kierunków działań pracowni są dążenia do rozwoju powiązań e-SXVII z tymi zasobami (Gruszczynski, Ogrodniczuk 2019). Z jednej strony ma to na celu pobieranie z nich nowych danych i usprawnienie dzięki temu prac nad e-SXVII. Z drugiej – potrzebne i celowe wydaje się wykorzystanie informacji zgromadzonych w e-SXVII do wzbogacania innych zbiorów. Do obecnie konsolidowanych zasobów należą: wspomniany już Elektroniczny Korpus Tekstów Polskich z XVII i XVIII w. (do 1772 r.) (KorBa), Cyfrowa Biblioteka Druków Ulotnych Polskich i Polski Dotyczących z XVI, XVII i XVIII Wieku (CBDU)¹⁴ oraz zdigitalizowana kartoteka SXVII. Prace obejmują również tworzenie relacji z różnymi zewnętrznymi bazami danych bio- i bibliograficznych, np. Wikidata, VIAF¹⁵. E-SXVII pozostaje wierny idei otwartości także w tym zakresie i zapewne wciąż będą powstawały nowe powiązania z kolejnymi zasobami.

Działające już połączenia pomiędzy e-SXVII i KorBa obejmują dwa mechanizmy. Po pierwsze, w formularzu redaktorskim e-SXVII możliwe jest automatyczne pobieranie form gramatycznych opracowywanego leksemu występujących w korpusie. Po drugie, działa również automatyczne wyszukiwanie wystąpień form wyrazu hasłowego w korpusie. Ten mechanizm bardzo usprawnia proces doboru cytatów źródłowych. Użytkownik słownika może również zobaczyć szerszą (niż przywołana w kolejnych znaczeniach) dokumentację materiałową dla wyrazu hasłowego, jeśli skorzysta z wyboru opcji: „Więcej cytatów w Korpusie Barokowym” (Wieczorek 2021).

Dalsze plany obejmują uruchomienie podobnych powiązań ze zdigitalizowaną kartoteką SXVII. Automatyczne otwieranie zbioru elektronicznych fiszek odnoszących się do wybranego wyrazu bardzo ułatwiłoby proces przeszukiwania kartoteki (Bilińska-Brynk, Rodek 2020). Projektowane połączenie między e-SXVII i CBDU

¹⁴ <https://cbdu.ijp.pan.pl/>.

¹⁵ Wikidata (pol. Wikidane) to „projekt internetowy mający na celu stworzenie wolnej, otwartej, wielojęzycznej bazy różnorodnych danych” (zob. <https://pl.wikipedia.org/wiki/Wikidane>). Virtual International Authority File to międzynarodowy projekt bibliotek narodowych oraz wybranych konsorcjów bibliotek regionalnych i ponadnarodowych (zob. <http://viaf.org/>).

będzie z kolei służyło wykorzystaniu informacji ze słownika do rozszerzenia funkcji biblioteki o możliwość automatycznego sprawdzenia znaczeń wyrazów występujących w tekstach druków ulotnych (Ogrodniczuk, Gruszczyński 2019). Dalsze plany przewidują powiązanie e-SXVII ze zdigitalizowanymi słownikami z epoki (Grzegorza Knapiusza, Michała Abrahama Trotza), a w późniejszej perspektywie być może również z kolejnymi wielkimi opracowaniami leksykograficznymi polszczyzny (SPXVI, słownik Samuela Bogumiła Lindego, słownik wileński, słownik warszawski). Już obecnie formularz redakcyjny umożliwia automatyczne sprawdzenie, czy opracowywane hasło jest poświadczony w SStp i SJPD.

Opracowywanie słownika obejmującego swym zakresem dwa stulecia rozwoju języka stanowi poważne wyzwanie naukowe, metodologiczne i organizacyjne. Często wymaga zmian w zakresie sposobu opracowania tak, aby opis leksykograficzny bogatego i skomplikowanego materiału językowego był możliwie pełny i precyzyjny. Wprowadzane modyfikacje są na bieżąco prezentowane w literaturze naukowej (np. Adamiec, Bronikowska 2016; Majdak 2018; Bronikowska et al. 2020). Nowoczesna formuła słownika „narodzonego elektronicznie” (czyli tworzonego cyfrowo z zastosowaniem różnorodnych funkcji niemożliwych do wprowadzenia w słowniku papierowym) pozwala na znaczące usprawnienie prac. Największymi jednak zaletami słownika elektronicznego pozostają jego otwartość oraz stała dostępność dla użytkownika w trakcie procesu powstawania. Obiecującą perspektywę otwiera tworzona obecnie platforma integrująca zasoby i narzędzia ułatwiające badania nad polszczyzną XVII i XVIII w., w której e-SXVII będzie stanowił centralną pozycję.

Literatura

- ADAMIEC D., BRONIKOWSKA R., 2016, *Wybór formy hasłowej jako problem opisu w Elektronicznym słowniku języka polskiego XVII i XVIII wieku (na tle tradycji leksykograficznej)*, „Prace Filologiczne” LXXVIII, s. 13–26.
- BIELIŃSKA M. (red.), 2020, *Leksykografia. Słownik specjalistyczny*, Kraków.
- BILIŃSKA-BRYNK J., RODEK E., 2020, *Paper Quotation Slips to the Electronic Dictionary of the 17th- and 18th-Century Polish – Digital Index and its Integration with the Dictionary*, [w:] Z. Gavriilidou, M. Mitsiaki, A. Fliatouras (red.), *Proceedings of XIX EURALEX Congress: Lexicography for Inclusion*, t. 1, Alexandroupolis, s. 465–470.
- BRONIKOWSKA R., MAJDAK M., WIECZOREK A., ŻÓŁTAK M., 2020, *The Electronic Dictionary of the 17th- and 18th-Century Polish – Towards the Open Formula Asset of the Historical Vocabulary*, [w:] Z. Gavriilidou, M. Mitsiaki, A. Fliatouras (red.), *Proceedings of XIX EURALEX Congress: Lexicography for Inclusion*, t. 1, Alexandroupolis, s. 471–475.
- CBDU: Cyfrowa Biblioteka Druków Ulotnych Polskich i Polski Dotyczących z XVI, XVII i XVIII Wieku, [on-line:] <https://cbdu.ijp.pan.pl>.
- e-SXVII: W. Gruszczyński (red.), *Elektroniczny słownik języka polskiego XVII i XVIII wieku*, [on-line:] <https://sxvii.pl>.

- GRUSZCZYŃSKI W., 2005, *O przyszłości Słownika języka polskiego XVII i 1. połowy XVIII wieku*, „Poradnik Językowy” nr 7, s. 48–61.
- GRUSZCZYŃSKI W., 2010, *Jednostka opisu leksykograficznego w słowniku historycznym na przykładzie Słownika języka polskiego XVII i 1. połowy XVIII wieku*, „Poradnik Językowy” nr 4, s. 26–40.
- GRUSZCZYŃSKI W., OGRONICZUK M., 2019, *Projekt zintegrowanego systemu informatycznego do studiowania polszczyzny XVII i XVIII wieku*; prezentacja na konferencji „Tradycja i nowoczesność w badaniach języków słowiańskich” Toruń, 5–6 września 2019 r., [on-line:] https://www.researchgate.net/publication/341827719_Projekt_zintegrowanego_systemu_informatycznego_do_studiowania_polszczyzny_XVII_i_XVIII_wieku.
- KORBA: Elektroniczny Korpus Tekstów Polskich z XVII i XVIII w. (do 1772 r.), [on-line:] <https://korba.edu.pl>.
- ŁUKASIK M., 2014, *Spójność w słowniku terminologicznym*, [w:] M. Kornacka (red.), *Spójność tekstu specjalistycznego*, Warszawa, s. 69–79.
- MAJDAK M., 2018, *Elektroniczny słownik języka polskiego XVII i XVIII wieku IJP PAN*, [w:] M. Pastuch, M. Siuciak (red.), *Historia języka w XXI wieku. Stan i perspektywy*, „Prace Naukowe Uniwersytetu Śląskiego w Katowicach”, nr 3670, Katowice, s. 176–182.
- OGRONICZUK M., GRUSZCZYŃSKI W., 2019, *Connecting Data for Digital Libraries: The Library, the Dictionary and the Corpus*, [w:] A. Jatowt, A. Maeda, S.Y. Syn (red.), *Digital Libraries at the Crossroads of Digital Information for the Future. 21st International Conference on Asia-Pacific Digital Libraries, ICADL 2019, Kuala Lumpur, Malaysia, November 4–7, 2019, Proceedings*, „Lecture Notes in Computer Science”, t. 11853, Cham, s. 125–138, https://doi.org/10.1007/978-3-030-34058-2_13.
- SJPAs: H. Koneczna, W. Doroszewski (red.), *Słownik języka Jana Chryzostoma Paska*, t. 1–2, Wrocław – Warszawa – Kraków 1965–1973.
- SJPD: W. Doroszewski (red.), *Słownik języka polskiego*, t. I–XI, Warszawa 1958–1969.
- SPXVI: *Słownik polszczyzny XVI wieku*, t. I–IV, red. komitet redakcyjny, t. V–XVII, red. M.R. Mayenowa, t. XVIII–XXXIV, red. F. Peplowski, t. XXXV–XXXVI, red. K. Mrowcewicz, P. Potoniec, Wrocław – Warszawa – Kraków, 1956–2012, [on-line:] <http://spxvi.edu.pl/>.
- SSTP: S. Urbańczyk (red.), *Słownik staropolski*, t. I–XI, Wrocław – Warszawa – Kraków 1953–2002.
- SXVII: K. Siekierska (red.), *Słownik języka polskiego XVII i 1. połowy XVIII wieku*, t. 1, Kraków 1999–2004.
- SZEMIŃSKA W., 2014, *Struktura słownika jako narzędzia pracy lingwistycznej*, [w:] W. Zmazer (red.), *Narzędzia pracy lingwistycznej*, Warszawa, s. 240–256.
- WIECZOREK A., 2021, *Integracja Elektronicznego słownika języka polskiego XVII i XVIII wieku i Elektronicznego Korpusu Tekstów Polskich z XVII i XVIII Wieku okiem użytkownika i redaktora*, [w:] E. Horyń, E. Młynarczyk, P. Żmigrodzki (red.), *Język polski – między tradycją a współczesnością. Księga jubileuszowa z okazji stulecia Towarzystwa Miłośników Języka Polskiego*, Kraków, s. 547–560, <https://doi.org/10.24917/9788380846258.44>.

Baroque Polish in the Internet or the Electronic Dictionary of Polish in the 17th and 18th Centuries

Abstract

In 2021, the *Electronic Dictionary of Polish Language of the 17th and 18th Centuries* (e-SXVII) was given the 4th edition Award of the Linguistics Committee of the Polish Academy of Sciences for outstanding scientific achievements in the field of linguistics. The award was granted in the category of team work carried out by Prof. Włodzimierz Gruszczyński PhD together with the Studio of the History of Polish Language of the 17th and 18th Centuries IJP PAN under his management, where e-SXVII is being created. The dictionary has been recognized as “an achievement in producing linguistic resources or tools”. The article reports on the history of the studio and introduces the evolution of conceptual works, especially after the decision to suspend further publication of SXVII in paper form (2004) and transform it into the first Polish dictionary created initially in the electronic form. It also deals with detailed issues regarding the canon of sources, structure and content (shape) of individual elements of the entry and media structure, and relations with other digital collections (especially the *Electronic Corpus of Polish Texts from the 17th and 18th Centuries*). The authors of the article also describe an important IT tool created for the needs of this dictionary.