

pracy – Danuta Wańka – we wstępie do książki poinformowała, że referaty zostały uszeregowane według kolejności wystąpień, zabieg ten spowodował, że opracowania o zbliżonej tematyce zostały umieszczone w różnych częściach publikacji, co nie obniża jednak wartości całego wydawnictwa.

Reasumując, należy stwierdzić, że *Biblioteki pedagogiczne wobec wyzwania współczesnej oświaty* to interesujący i wielowątkowy zbiór artykułów, napisanych jasnym i zrozumiałym ję-

zykiem oraz wyposażonych w bogatą bazę źródłową. Publikacja wydaje się o tyle potrzebna i cenna, że niejednokrotnie wskazuje na pewne praktyczne rozwiązania przyjęte w działalności wielu bibliotek pedagogicznych, które mogą stanowić źródło inspiracji dla pracowników innych tego typu placówek, stale poszukujących nowych kierunków rozwoju.

Marzena Jarocka

Biblioteka Pedagogiczna w Toruniu

W roku 2008 do rąk czytelników trafiła pozycja ze wszech miar niezwykła, która potwierdza znaną tezę, że rozwój w nauce to (dramatyczne) obalenie istniejących teorii. W tym stwierdzeniu mieści się pierwsze dokonanie jej autora. Drugim – z pewnością nie mniej ważnym – jest udowodnienie, na czym polega interdyscyplinarność w nauce. Spróbujemy wskazać na oba walory recenzowanego dzieła. Uczynimy to jednak w odwrotnej kolejności, ponieważ z punktu widzenia środowiska, do którego są adresowane „Toruńskie Studia Bibliologiczne”, wykorzystanie osiągnięć współczesnej nauki o informacji w innych dyscyplinach wydaje się ważniejsze niż wartość – nie ulega wątpliwości, że dla filologa olbrzymia – samych ustaleń *stricte* językoznawczych. (*Notabene* autor dedykuje swe dzieło głównie bibliotekarzom: „Książka ta jest skromnym hołdem filologa dla tych, którzy *skanują*, a w szerszej perspektywie – dla tych wszystkich, któ-

rym nie jest obojętny los *starych gazetek*” [z okładki]).

Często powtarzamy, że w procesie cyfryzacji, a szerzej – w narzędziach całej kultury cyfrowej, drzemią wręcz nieograniczone możliwości wykorzystania naukowego w różnych dziedzinach, oczywiście także w humanistyce. Rzadziej jednak dajemy dowody potwierdzające tę tezę. Nie dopatrzyłem się, aby autor recenzowanej pozycji szczególnie często wspominał o „nieograniczonych możliwościach badawczych, jakie daje cyfryzacja”, „nowej jakości warsztatu badawczego,

Piotr Wierzchoń: Fotodokumentacja. Chronologizacja. Emendacja. Teoria i praktyka weryfikacji materiału leksykalnego w badaniach lingwistycznych, Poznań: Instytut Językoznawstwa UAM 2008, 494 s., ISBN 978-83-926446-9-9

która wiąże się z cyfryzacją” etc. On po prostu wykorzystał osiągnięcia cyfryzacji w badaniach językoznawczych, uzyskując w efekcie coś, co Karl Popper nazywał „odkryciem naukowym” (rzecz w humanistyce, nie tylko zresztą polskiej, jak się okazuje, rzadka). Piotr Wierzchoń nie formułuje po raz kolejny postulatów, że „warto podjąć badania w zakresie...”, „należy poddać analizie kompleks faktów...”, „trzeba jak najszybciej wyjaśnić zjawisko...”, okraszając je przy okazji własnymi poglądami na pewne kwestie, które nijak mają się do rzeczywistości, a nieraz charakteryzują się wręcz kontryfaktyczną naturą. P. Wierzchoń najpierw opracowuje metodę, a następnie podejmuje badania, poddaje materiał analizie i dokonuje jego eksplikacji – (aż) tyle.

Warsztatem badawczym dla autora są dokumenty ujęte w bibliotekach akademickich i regionalnych, przechowywane w systemie dLibra, zazwyczaj w formacie djvu. Z punktu widzenia prowadzonych przez Piotra Wierzchoń badań istotna jest warstwa tekstowa takich plików oraz – co także nie jest bez znaczenia – ich wysoka „czytelność” maszynowa. Jest to najważniejszy warunek, jaki musi spełniać zeskanowany zasób, w przeciwnym wypadku jego wykorzystanie w fotodokumentacji językoznawczej będzie prawie niemożliwe, a w każdym razie wyjątkowo utrudnione. Dlatego też w procesie skanowania należy pamiętać, aby:

- nie zniekształcać obrazu tekstu, dbając o jego jak najbardziej płaskie odwzorowanie,

- unikać przebiccia rewersów skanowanych stron,
- zapewniać optymalną rozdzielczość skanowania.

Teksty niespełniające tych wymagań będą mniej czytelne, przez co mniej przydatne w procesie ekscerpcji.

Z punktu widzenia badań językoznawczych za najbardziej wartościowe zbiory cyfrowe autor recenzowanej publikacji uznał zasoby następujących bibliotek cyfrowych: Dolnośląskiej Biblioteki Cyfrowej, Jeleniogórskiej Biblioteki Cyfrowej, Kujawsko-Pomorskiej Biblioteki Cyfrowej, Małopolskiej Biblioteki Cyfrowej, Podlaskiej Biblioteki Cyfrowej, Śląskiej Biblioteki Cyfrowej, Wielkopolskiej Biblioteki Cyfrowej, Biblioteki Cyfrowej Uniwersytetu Wrocławskiego, Zielonogórskiej Biblioteki Cyfrowej, Biblioteki Cyfrowej Politechniki Łódzkiej, Biblioteki Cyfrowej Politechniki Warszawskiej, Cyfrowej Biblioteki Narodowej Polona. Oczywiście, autor zaznacza, że w przypadku nad wyraz dynamicznego rozwoju bibliotek cyfrowych dane faktygraficzne tego typu podlegają natychmiastowej dezaktualizacji, ponieważ – zresztą na szczęście – praktycznie każdego miesiąca w Polsce powstają nowe biblioteki cyfrowe.

Zdigitalizowane zasoby prasy krajowej i emigracyjnej (głównie polskiej) oraz w mniejszym stopniu innych typów wydawnictw zostały przez Piotra Wierzchoń wykorzystane przede wszystkim do datacji neologizmów w języku polskim, a mówiąc bardziej precyzyjnie: do rewizji istniejących w tym zakresie usta-

leń. Z punktu widzenia stopnia złożoności zastosowanej metody zrobił to w najprostszy z możliwych sposobów, wykorzystując technikę optycznego rozpoznawania pisma (OCR). Można nawet pokusić się o stwierdzenie, że posunięcie badawcze autora jest czynnością wręcz trywialną. Problem w tym, że przed Piotrem Wierzcho-niem nikt tej „oczywistości” nie zdołał zauważyć i wyzyskać naukowo. Autor podważył dotychczasowe datacje neologizmów oraz neosemantyzmów dla języka polskiego (głównie neologizmów rzekomo powojennych), obalając lwia część dotychczasowych data-wań momentu pojawienia się danego wyrazu w języku. Przesunął datę ich pierwszego wystąpienia (mówiąc ściślej: pierwszej odnalezionej, poświad-czonej filologicznie rejestracji) w nie-

których wypadkach o dziesiątki lat wstecz.

Rdzeniem recenzowanej książ-ki jest fotodokumentacja, dlatego re-cenzentowi nie wolno unikać poda-nia przykładów redatacji dokonanych na podstawie ekscerpcji z tekstów cy-frowych. Przyjrzyjmy się zatem kilku efektem tej metody.

W pracy zatytułowanej *System słowotwórczy polszczyzny (1945–2000)*⁴ Hanna Jadacka podaje dużą grupę neologizmów, jakie jej zdaniem poja-wiły się w polszczyźnie po 1945 r. We-dług autorki, są to m.in. takie wyrazy, jak: „autokontrola”⁵, „odmianoznaw-stwo”⁶, „ofensywność”⁷, „podatniczy”⁸, „podkategoria”⁹. Zobaczymy, jaki efekt przyniosło poszukiwanie tych jedno-stek metodą ekscerpcji fotodokumen-tacyjnej w zasobach dLibry:

równo w rytmice, jak obrazowaniu. Helsztyński okazuje się tu poetą o oryginalnem, wybitnie zarysowanem obliczu własnem. Patos, bujność metafor, spletających się z sobą, zahaczających o siebie nawzajem, zawiła składnia, częsta inwersja, śmiałe nowotwory językowe, bogactwo obrazów i myśli, wreszcie niekiedy pewien brak autokontroli pisarskiej – wszystko to ma coś z siedemnastowiecznego b a r o k u,

Skan 1. Kopia cyfrowa tekstu z roku 1933 z poświadczeniem rzeczownika „autokontrola”

Źródło: *Kronika Gostyńska* (1933.10.01, Ser. 5, nr 8) [on-line]. Wielkopolska Biblioteka Cyfrowa [dostęp 15 lipca 2009]. Do-stępny w World Wide Web: <http://www.wbc.poznan.pl/dlibra/doccontent?id=16077&dirids=1>.

⁴ H. Jadacka, *System słowotwórczy polszczyzny (1945–2000)*, Warszawa 2001.

⁵ Tamże, s. 97.

⁶ Tamże, s. 58.

⁷ Tamże, s. 187.

⁸ Tamże, s. 104.

⁹ Tamże, s. 90.

3 trym. Odmianoznawstwo, 2 g. t., tamże, obow. dla IV r. roln. gr. B.
1, 2, 3 trym. Hodowla roślin, 2 g. t., obow. dla IV r. roln. gr. B., tamże.

Skan 2. Kopia cyfrowa tekstu z roku 1932 z poświadczeniem rzeczownika „odmianoznawstwo”

Źródło: *Spis wykładów i skład Uniwersytetu na rok akademicki 1932/33* [on-line]. Wielkopolska Biblioteka Cyfrowa [dostęp 15 lipca 2009]. Dostępny w World Wide Web: <http://www.wbc.poznan.pl/dlibra/doccontent?id=53851&dirids=1>.

Rozgłos literacki zdobył Lam głównie dzięki ciętym dowcipom i publicystycznej ofensywności — co „podało się publice“, lecz nie przeszkodziło, aby go rychło zapomniano. Przyjmowały się dowcipy Lamy, lecz poszły w zapomnienie głębokie myśli i uwagi tego prawdziwego filozofa życiowego.

Skan 3. Kopia cyfrowa tekstu z roku 1938 z poświadczeniem rzeczownika „ofensywność”

Źródło: *Prosto z Mostu: tygodnik literacko-artystyczny* (R. 4 (1938), nr 30) [on-line]. e-Biblioteka Uniwersytetu Warszawskiego [dostęp 15 lipca 2009]. Dostępny w World Wide Web: <http://ebuw.uw.edu.pl/dlibra/doccontent?id=998&dirids=1>.

W tych dniach ukażą się przepisy wykonawcze do ustawy o podatku majątkowym
Warszawa, 15. 11. (PAT.) „Gazeta Warszawska“ donosi, że w najbliższych dniach ukażą się w „Dzienniku Ustaw” instrukcje wykonawcze do ustawy o podatku majątkowym. Instrukcje te będą zawierały przepisy o normach szacunkowych oraz o spisach płatników i zeznaniach podatkowych.

Skan 4. Kopia cyfrowa tekstu z roku 1923 z poświadczeniem przymiotnika „podatniczy”

Źródło: *Gazeta Bydgoska* (1923.11.17, R. 2, nr 264) [on-line]. Kujawsko-Pomorska Biblioteka Cyfrowa [dostęp 15 lipca 2009]. Dostępny w World Wide Web: <http://kpbc.umk.pl/dlibra/doccontent?id=3294&dirids=1>.

wieki wieków; to oczywista konieczność! Ale czym jest przeszłość, tem jest co do joty terażniejszość i przyszłość, bo terażniejszość jest rodzącą się, a przyszłość mającą się urodzić przeszłością! A więc we wszystkich stajach czasu panuje konieczność, jeżeli Bóg Bogiem, jeżeli myśl jego prawdą, jeżeli słowo jego świętem prawem! Tegoż samego uczy nas loika w tak zwanych trzech modalności kategoriach, czyli właściwie w podkategorjach, t. j., w możliwości, rzeczywistości i konieczności. Co jest istotną

Skan 5. Kopia cyfrowa tekstu z roku 1842 z poświadczeniem rzeczownika „podkategoria”

Źródło: *Orędownik Naukowy* (1842.02.14, nr 7) [on-line]. Wielkopolska Biblioteka Cyfrowa [dostęp 15 lipca 2009]. Dostępny w World Wide Web: <http://www.wbc.poznan.pl/dlibra/doccontent?id=61730&dirids=1>.

Jakikolwiek komentarz w kontekście przytoczonego powyżej materiału jest z pewnością zbędny.

Omawiana książka skłania w końcu do ogólnej refleksji na temat dalszego wykorzystania zdigitalizowanych tekstów (których korpusy przecież nieustannie „pączkują”, przynosząc efekty w postaci kolejnych bibliotek cyfrowych). Pewnym problemem jest brak standaryzacji notacji lingwochronologicznej. Wydaje się, że autor słusznie postuluje, wspierając wcześniejsze apele, stworzenie bazy danych „dokumentacji wyrazowej dla języka polskiego” (dotychczas do miana takiej bazy mógłby pretendować słownik bibliograficzny autorstwa Jana Wawrzyńczyka⁴, jednak jest to rejestr opra-

cowany nie w postaci fotodokumentacyjnej, co – biorąc pod uwagę jego rozmiary – nie umniejsza rangi przedsięwzięcia). Jego zdaniem, wystarczy, że rekord takiej bazy będzie zawierać tylko kilka pól (wyraz, data pierwszego wystąpienia, miejsce oraz fotodokumentacja w postaci skanu fragmentu tekstu, w którym wystąpił), a stanie się niezastąpionym warsztatem dla badaczy językoznawców. Autor idzie dalej, twierdząc, że „Sumę przedstawionych rozwiązań teoretycznych (jak i ilustracji praktycznych) określić można mianem *ekscerptologii*. Ekscerptologia tak pojęta będzie nauką-teorią opisującą parametry czynności i wyniki ekscerpacji lingwistycznej [...]. Przy czym mówiąc o ekscerptologii, chcemy dookreślić: nie mamy na uwadze przypadkowości ekscerpacji (okazjonalności); ekscerpacja manualna (o ile nie jest tzw. ekscerpacją pełną) zawsze będzie czyn-

⁴ J. Wawrzyńczyk, *Słownik bibliograficzny języka polskiego* [on-line]. Warszawa 2004–2009 [dostęp 15 lipca 2009]. Dostępny w World Wide Web: <http://www.leksykapolska.pl/sbjp.html>.

nością przypadkową, lecz opis jednostek wyekscerpowanych z danych tekstów winien mieć charakter sparymetryzowany” (s. 460). Cóż, pozostaje życzyć językoznawcom spełnienia tych zamierzeń, autorowi zaś kolejnych równie wartościowych, nietuzinko-

wych (i chyba nawet bez przesady: rewolucyjnych) dzieł.

Piotr Nowak

*Instytut Językoznawstwa Uniwersytetu
im. Adama Mickiewicza w Poznaniu*

