

**BIULETYN POLSKIEGO TOWARZYSTWA
JĘZYKOZNAWCZEGO**

**BULLETIN DE LA SOCIÉTÉ POLONAISE
DE LINGUISTIQUE**

ZESZYT LXXIV – FASCICULE LXXIV

LEXIS

BIULETYN POLSKIEGO TOWARZYSTWA
JĘZYKOZNAWCZEGO

BULLETIN DE LA SOCIÉTÉ POLONAISE
DE LINGUISTIQUE

ZESZYT LXXIV – FASCICULE LXXIV

KOMITET REDAKCYJNY

Redaktor Naczelny: Renata Przybylska

Rada naukowa: Ireneusz Bobrowski (Uniwersytet Jagielloński),
Andrzej Bogusławski (Uniwersytet Warszawski), Barbara Dancygier (University
of British Columbia), Magdalena Danielewicz (Uniwersytet Warszawski), Katarzyna
Dziwirek (University of Washington), Laura Janda (Norges arktiske universitet,
Norwegia), Krystyna Kleszczowa (Uniwersytet Śląski), Igor Mel'čuk (Université de
Montréal, Kanada), Predrag Piper (Универзитет у Београду, Serbia),
Piotr Stalmaszczyk (Uniwersytet Łódzki), Dorota Szumska (Uniwersytet Jagielloński),
Branko Tošović (Universität Graz, Austria)

Sekretarz Redakcji: Patrycja Pałka

Redaktorzy językowi:
Krzysztof Porosło (język polski)
Alicja Witalisz (język angielski)

ADRES REDAKCJI

al. Mickiewicza 31
31-120 Kraków
e-mail: p.palka@uj.edu.pl

Nakład: 100

ISSN 0032-3802

© Copyright by Polskie Towarzystwo Językoznawcze

Kraków 2018

Wersją pierwotną czasopisma jest wersja elektroniczna

SPIS RZECZY – TABLE DES MATIÈRES

Wspomnienia

Zenon Lica, <i>Profesor Edward Breza (1932–2017)</i>	5
Barbara Kudra, <i>Profesor Władysław Cyran (1923–2016)</i>	7
Izabela Winiarska, <i>Dr Krystyna Długosz-Kurczabowa (1940–2016)</i>	11
Piotr Stalmaszczyk, Danuta Stanulewicz, <i>Profesor Roman Kalisz (1948–2017)</i>	15
Ewa Wolańska, <i>Docent doktor Maria Przybysz-Piwko (1950–2016)</i>	19

Artykuły

Źródła a językoznawstwo komputerowe

Mirosław Bańko, <i>Korzyści i problemy wynikające z „nadmiaru” danych językowych i z wielości metod ich analizy</i>	23
Mirosława Podhajcka, <i>Google Books jako korpus językowy</i>	31
Iwona Kosek, Sebastian Przybyszewski, Monika Czerepowicka, <i>Źródła w tworzeniu elektronicznego słownika paradygmatów frazeologizmów werbalnych</i>	47
Michail L. Kotin, <i>Möglichkeiten und Grenzen der korpusbasierten Grammatikforschung</i>	61

Źródła w badaniach historycznojęzykowych

Aleksandra Janowska, <i>Źródła a typy badań historycznojęzykowych</i>	71
Dorota Rojszczak-Robińska, <i>Łacińskie źródła jako kontekst badań historycznojęzykowych</i>	81
Joanna Kulwicka-Kamińska, <i>Piśmiennictwo religijne Tatarów Wielkiego Księstwa Litewskiego jako źródło badań historycznojęzykowych</i>	93

Źródła w badaniach dialektologicznych

Angelika Pawlaczyk, <i>Znaczenie metody biograficznej w badaniach gwary staroobrzędowców mieszkających w Polsce</i>	107
Maciej Rak, <i>Przydatność testamentów chłopskich w badaniach językoznawczych</i>	117

Źródła a leksykografia

Katarzyna Wojan, <i>Bibliografie słowników przekładowych jako narzędzie rekonstrukcji dziejów oraz oceny stanu i dorobku słownictwa narodowego</i>	129
--	-----

Monika Biesaga, <i>Poprawność polityczno-kulturowa w słowniku. Pola tematyczne a typy mikrostrukturalnych informacji leksykograficznych</i>	145
---	-----

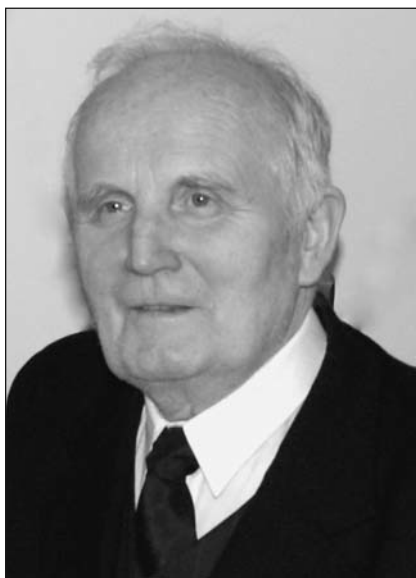
Zagadnienia szczegółowe

Barbara Lewandowska-Tomaszczyk, Paul A. Wilson, <i>Sources of data and methodological foundations of a contrastive linguistic analysis of emotion concepts</i>	157
Anna Kozłowska, <i>Źródła w badaniach idiolektów</i>	191
Ewa Ciszewska-Jankowska, <i>Le futur gnomique</i>	205

SPRAWY ADMINISTRACYJNE

LXXV Zjazd Polskiego Towarzystwa Językoznawczego w roku 2017	217
Autorzy	223
Lista recenzentów	229
Zalecenia edytorskie	231

Profesor Edward Breza
(24 IX 1932 – 12 X 2017)



Edward Breza, językoznawca, onomasta, dialektolog i historyk języka polskiego, urodził się 24 września 1932 r. w Kaliszu, w powiecie kościerskim. Po ukończeniu studiów polonistycznych w 1964 r. w Wyższej Szkole Pedagogicznej w Gdańsku rozpoczął tam pracę jako starszy asystent. Doktoryzował się w 1972 r. na Uniwersytecie Gdańskim, na którym również uzyskał habilitację w r. 1978. Rok później pracował na stanowisku docenta, a od 1988 r. na stanowisku profesora nadzwyczajnego, w 1993 r. został profesorem zwyczajnym. Profesor związany był również z Wyższą Szkołą Pedagogiczną w Bydgoszczy, w której był zatrudniony od 1 października 1978 do 30 czerwca 1994 r. Po przejściu na emeryturę nadal pracował w Bałtyckiej Wyższej Szkole Humanistycznej, Gdańskiej Wyższej Szkole Humanistycznej, w Instytucie Filologiczno-Historycznym w Mławie oraz w Zespole Kolegiów Nauczycielskich w Koszalinie.

Profesor Edward Breza wykazywał aktywność organizacyjną. Pełnił funkcję prodykanta do spraw dydaktycznych na Wydziale Humanistycznym Uniwersytetu Gdańskiego (1978–1981), dyrektora Instytutu Filologii Polskiej (1981–1991) oraz kierownika Zakładu Historii Języka Polskiego, Dialektologii i Onomastyki (1988–2002). Ponadto był przewodniczącym I Wydziału Gdańskiego Towarzystwa Naukowego (1980–

–1984), wieloletnim redaktorem „Zeszytów Naukowych Wydziału Humanistycznego Uniwersytetu Gdańskiego. Prac Językoznawczych”, „Rocznika Gdańskiego” i „Pomorskich Monografii Toponomastycznych”, członkiem kolegium redakcyjnego „Rocznika Gdańskiego” (1978–1992), a także Komitetu Redakcyjnego „Onomastyców” w Krakowie (od 1986), Rady Naukowej Muzeum Piśmiennictwa i Muzyki Kaszubsko-Pomorskiej (1989–1998), Międzynarodowego Komitetu Onomastycznego w Lowanium (Belgia), członkiem Komisji Onomastycznej Komitetu Językoznawstwa Polskiej Akademii Nauk oraz wieloletnim przewodniczącym Rady Języka Kaszubskiego.

Był również znakomitym dydaktykiem. Prowadził seminaria doktoranckie, magisterskie i licencjackie, wykłady i ćwiczenia z gramatyki opisowej i historycznej oraz kultury języka polskiego i dialektologii. Wypromował 22 doktorów, przeszło 600 magistrów i 200 licencjatów. Współpracował z ośrodkami zagranicznymi, m.in. uniwersytetem w Lipsku, Uniwersytetem Humboldta w Berlinie, Berlińską Akademią Nauk, Wyższą Szkołą Pedagogiczną w Magdeburgu, jak również Uniwersytetem Karola w Pradze, Komisjami Onomastycznymi Czeskiej i Słowackiej Akademii Nauk. Od 1993 r. wielokrotnie prowadził wykłady z zakresu historii języka polskiego oraz onomastyki na Uniwersytecie w Grodnie.

Jego zainteresowania badawcze obejmowały szerokie obszary wiedzy językoznawczej: onomastykę, kulturę języka, dialektologię, historię języka polskiego i kaszubskiego oraz dydaktykę języka polskiego. Wśród nich ważne miejsce zajmuje onomastyka. Jej poświęcił kilka książek. Są to m.in. *Toponimia powiatu kościerskiego* (1974), *Pochodzenie przydomków szlachty polskiej* (I wyd. 1978, 1986) oraz cztery tomy *Nazwisk Pomorzan. Pochodzenie i zmiany* (2000, 2002, 2004, 20018). W wydawnictwie Uniwersytetu Gdańskiego opracowywana jest obecnie druga część czwartego tomu.

Profesor Edward Breza ma także spore zasługi w badaniu kaszubszczyzny. Jest współautorem *Gramatyki kaszubskiej* (1981) oraz *Zasad pisowni kaszubskiej* (1975, II wyd. 1984). Pod Jego redakcją w 2001 r. ukazał się w serii *Najnowsze Dzieje Języków Słowiańskich* tom *Kaszubszczyzna*. Kaszubszczyźnie poświęcone są również rozliczne artykuły Profesora, wśród nich badające jej status oraz język pisarzy kaszubskich.

Nie sposób w tym miejscu omówić bogatego, liczącego ponad 2500 publikacji, dorobku naukowego Profesora. Do tego dochodzą liczne prace popularnonaukowe, które ukazywały się w regionalnych wydawnictwach i czasopismach. Świadczy to o Jego niezwyklej aktywności, za którą otrzymał liczne odznaczenia państwowe, resortowe i regionalne, m.in.: Złoty Krzyż Zasługi, Krzyż Kawalerski Orderu Odrodzenia Polski, Medal Komisji Edukacji Narodowej, odznakę „Zasłużony Ziemi Gdańskiej”, Medal Księcia Mściwoja II, Medal Uniwersytetu Gdańskiego i Medal Państwowej Wyższej Szkoły Zawodowej w Ciechanowie.

Był szanowanym autorytetem w świecie nauki polskiej i światowej, zawsze życzliwym Nauczycielem, Doradcą i Kolegą. Takim wspaniałym Człowiekiem na zawsze zostanie w naszej pamięci.

Profesor Władysław Cyran (16.03.1923 – 23.12.2016)



23 grudnia 2016 r. odszedł Profesor Władysław Cyran, emerytowany profesor zwyczajny Instytutu Filologii Polskiej UŁ, jeden z ostatnich przedstawicieli najstarszego i najznamienitszego pokolenia polskiego i łódzkiego językoznawstwa. Został pochowany w jednym grobie z żoną na łódzkim cmentarzu przy ul. Szczecińskiej.

Profesor Władysław Cyran był wybitnym językoznawcą, autorem licznych prac z historii języka polskiego, dialektologii, kultury języka, a także zagadnień z zakresu współczesnej polszczyzny. Wiedzą o tym dobrze dzisiejsi badacze języka, z których wielu powołuje się na prace Profesora.

Urodził się 16 marca 1923 roku w osadzie Liw (powiat węgrowski, woj. lubelskie) w rodzinie inteligenckiej; ojciec Józef Cyran był kierownikiem szkoły powszechnej w tej miejscowości. Profesor miał sześcioro rodzeństwa (trzech braci i trzy siostry). Sam nie miał dzieci, jego żona, Irena, zmarła w 1999 r.

Szkółę powszechną ukończył w rodzinnej miejscowości w 1935 r., a Gimnazjum Ogólnokształcące im. S. Żółkiewskiego już w Siedlcach w roku wybuchu II wojny światowej, czyli w 1939. W czasie okupacji pracował na kolei i jednocześnie uczęszczał do Szkoły Handlowej II stopnia. W roku szkolnym 1941/42 otrzymał świadec-

two jej ukończenia. Po wojnie na podstawie tego świadectwa został przyjęty do II klasy Liceum Handlowego w Siedlcach i zdał egzamin maturalny. W 1945 r. przybył do Łodzi i został przyjęty na Wydział Humanistyczny Uniwersytetu Łódzkiego. Poświęcił się studiom nad językiem ojczystym, które ukończył 16.12.1950 r., uzyskując stopień magistra filozofii w zakresie filologii polskiej na podstawie pracy magisterskiej *Ślady iloczasu w głównych zabytkach języka polskiego XIV i XV wieku*, napisanej pod kierunkiem profesora Zdzisława Stiebera. Wyniki tej pracy zostały opublikowane (bardzo pozytywnie artykuł ocenił prof. Kazimierz Nitsch), a wybitny polski językoznawca profesor Zdzisław Stieber w opinii z 1951 r. na potrzeby zatrudnienia w UŁ ówczesnego magistra Władysława Cyrana pisał tak: „Mgr Władysław Cyran jest wybitnie zdolnym pracownikiem naukowym. Jego praca magisterska o polskim iloczase w XV wieku ma niewątpliwą wartość naukową i jako taka została przyjęta w obszernym streszczeniu (arkusz druku) do głównego polskiego pisma językoznawczego, jakim jest «Biuletyn Polskiego Towarzystwa Językoznawczego». Powierzenie mu funkcji asystenta przy Zakładzie Języka Polskiego UŁ przysporzy Zakładowi bardzo dobrego współpracownika, a równocześnie pozwoli mgr. Cyranowi zajmować się nadal pracą naukową, na co w pełni zasługuje” [z teczki personalnej Profesora Władysława Cyrana].

Profesor Władysław Cyran pracował już w czasie studiów jako tzw. asystent wolontariusz (1.02.1949 – 31.03.1951), później jako kontraktowy zastępca asystenta (do 31.08.1951), a następnie od września 1951 r. jako asystent w Zakładzie Języka Polskiego UŁ. Pracę doktorską pt. *Gwary polskie w okolicach Siedlec* ukończył w 1957 r. W 1958 otrzymał stanowisko adiunkta. Praca doktorska została opublikowana w r. 1960, a w styczniu 1960 r. Profesor otrzymał stopień doktora nauk humanistycznych. Sześć lat później, w 1966 r., odbyło się kolokwium habilitacyjne Profesora na podstawie rozprawy habilitacyjnej, opublikowanej w 1967 r. pt. *Przysłówki polskie. Budowa słowotwórcza*. Praca ta została wyróżniona ministerialną nagrodą II stopnia. W grudniu 1966 r. Profesor uzyskał stopień naukowy docenta, a w 1967 otrzymał etat docenta na UŁ.

W październiku 1977 r. Rada Państwa nadała Władysławowi Cyranowi tytuł naukowy profesora nadzwyczajnego nauk humanistycznych. Jako profesor nadzwyczajny pracował do listopada 1990 r. W grudniu 1990 r. otrzymał od ministra edukacji narodowej nominację na stanowisko profesora zwyczajnego.

W latach 1972–1981 Profesor sprawował funkcję dziekana Wydziału Filologicznego, zaś w latach 1984–1987 był prorektorem UŁ do spraw nauki. Łącznie Profesor Władysław Cyran przez 15 lat sprawował funkcje dziekana, prodziekana (1969–1972) i prorektora.

W latach 1963–1969 Profesor był sekretarzem Wydziału I Łódzkiego Towarzystwa Naukowego, członkiem Redakcji Rozpraw Komisji Językowej ŁTN, a w latach 1978–1979 redaktorem Wydawnictw Filologicznych UŁ; przez dwie kadencje był członkiem Komitetu Językoznawstwa PAN.

Otrzymał wiele nagród naukowych i państwowych, nagród za działalność dydaktyczną, organizacyjną i wychowawczą – m.in. w 1967 r. (nagroda naukowa II stop-

nia), 1970 (Złota Odznaka ZNP), 1971 (Złoty Krzyż Zasługi, Złota Odznaka UŁ), 1973 (Honorowa Odznaka Województwa Łódzkiego), 1974 (Honorowa Odznaka Miasta Łodzi), 1975 (Medal w służbie społeczeństwu i nauce, Medal XXX-lecia UŁ), 1977 (Krzyż Kawalerski Orderu Odrodzenia Polski), 1978 (Medal Komisji Edukacji Narodowej), 1981 (Medal XXXV-lecia UŁ), 1982 (Zasłużony Nauczyciel PRL), 1985 (Krzyż Oficerski Orderu Odrodzenia Polski). W r. 1988 otrzymał nagrodę naukową I stopnia od Ministra Edukacji Narodowej.

Nigdy nie należał do żadnej partii politycznej.

Wypromował 130 magistrów, 5 doktorów (dwóch zostało profesorami), recenzował wiele prac doktorskich i habilitacyjnych – w UŁ i w innych ośrodkach uniwersyteckich w Polsce.

Jako dydaktyk Profesor Władysław Cyran prowadził zajęcia i wykłady z gramatyki języka staro-cerkiewno-słowiańskiego, ze wstępu do nauki o języku, językoznawstwa ogólnego, zagadnień kultury języka i poprawności językowej, gramatyki opisowej i historycznej języka polskiego.

W latach 1987–1993 Profesor sprawował funkcję kierownika Katedry Współczesnego Języka Polskiego. W 1993 przeszedł na emeryturę.

Zainteresowania naukowe Profesora Władysława Cyrana dotyczą przede wszystkim dialektologii oraz współczesnego języka polskiego z nachyleniem ku słowotwórstwu i poprawności językowej. Dialektologii dotyczyły materiały zbierane w latach 1953–1954 do *Małego atlasu gwar polskich*, a także praca doktorska (*Gwary polskie w okolicach Siedlec*) czy m.in. artykuły: *Procesy językowe w przejmowaniu wyrazów obcych przez gwary* (1975) – badania Profesora wykazały, że przy zapożyczaniu wyrazów obcych przez gwary zachodzą regularne zjawiska fonetyczne i morfologiczne, *Integracja gwar z językiem literackim* (1982) – o stosunku gwar do języka literackiego. Słowotwórstwa gwarowego dotyczyła książka pt. *Tendencje słowotwórcze w gwarach polskich* (1977) – zestawienie typowych dla gwar formantów słowotwórczych ze stanem ogólnopolskim i określenie na tej podstawie tendencji słowotwórczych w gwarach.

Słowotwórstwa historycznego, współczesnego i gwarowego dotyczył m.in. artykuł *Słowotwórstwo historyczne a słowotwórstwo współczesne i słowotwórstwo gwarowe* (1981), zaś słowotwórstwa historycznego książka z 1967 r. zatytułowana *Przysłówki polskie. Budowa słowotwórcza* – oparta na materiałach historycznych, poczynając od najstarszych zabytków językowych do współczesnych.

Kolejną dziedziną zainteresowań naukowych Profesora był współczesny język polski. Można tu wymienić studia wyrazowe prezentowane w najważniejszych czasopismach językoznawczych – m.in. *Dżesz czy dżasz?* (1959), *O tzw. wypowiedzeniach współrzędnych wynikowych we współczesnej polszczyźnie* (1970), *Pochodzenie samogłoski a w sufiksie nazwy Warszawa* (1973), *Ancymonek* (1974), *Krakersy* (1975), *O omyłkach graficznych* (1976), *Przekłamać. Ze zjawisk współczesnego języka* (1980) oraz studia składniowe: *Dwa drobiazgi ze składni polskiej...* (1961) – dotyczące połączeń typu *można zrobić, pięć jabłek, kilka koszy*; *Dlaczego giną w języku polskim formy czasu przyszłego złożone z bezokolicznikiem* (1961); *O częściach mowy w języku*

polskim (1970); *Przyczyny szerzenia się w l. poj. deklinacji męskiej form dopełniacza w funkcji biernika* (1984).

Profesor zajmował się także kwestiami kultury i poprawności językowej – na przykład artykuły: *Kilka uwag o poprawnej wymowie polskiej* (1972), *Błędy językowe a tendencje rozwojowe języka* (1982), *O błędach językowych* (1982), *Wydawnictwa poprawnościowe wobec zmian w języku polskim* (1988). Bardzo ważne i wciąż aktualne są spostrzeżenia Profesora zawarte w artykule *Mechanizm zapożyczania wyrazów w języku polskim* (1974).

Wszystkie prace Profesora są ważnymi, liczącymi się w dorobku polskiego językoznawstwa opracowaniami i do dziś aktualnymi.

Pan Profesor przy wszystkich swoich wielkich dokonaniach naukowych był człowiekiem skromnym, serdecznym, prawym, wrażliwym na piękno i o dużym poczuciu humoru. Był – o czym wie niewielu – melomanem, lubił słuchać muzyki poważnej, oraz zapalonym brydżystą. Takim pozostanie w pamięci tych, którzy się z Nim zetknęli – uczniów, studentów, kolegów i przyjaciół.

Barbara Kudra

Dr Krystyna Długosz-Kurczabowa



Krystyna Długosz-Kurczabowa urodziła się 13 kwietnia 1940 r. w Domaszowicach na Kielecczyźnie, zmarła 14 września 2016 r. w Warszawie. Choć mija już pierwsza rocznica jej śmierci, wciąż trudno pogodzić się z jej odejściem i pisać o niej w czasie przeszłym. Przez całe swoje zawodowe życie, tj. od 1962 r., kiedy ukończyła studia polonistyczne, była związana z Wydziałem Polonistyki Uniwersytetu Warszawskiego. W macierzystej jednostce w 1972 r. obroniła rozprawę doktorską poświęconą funkcjom formantów rzeczownikowych w *Leksykonie...* Jana Mączyńskiego, którą napisała pod kierunkiem prof. Przemysława Zwolińskiego. W 1976 r. – gdy na Wydziale Polonistyki Uniwersytetu Warszawskiego powołano do życia Instytut Języka Polskiego jako samodzielną jednostkę naukową – współtworzyła Zakład Historii Języka Polskiego i Dialektologii, była cenionym pracownikiem dydaktycznym, wymagającym, ale też lubianym i szanowanym nauczycielem akademickim. Dość powiedzieć, że obecna kadra średniego pokolenia warszawskiej polonistyki to w większości jej wychowankowie, studenci i magistranci, którym zaszczepiła swoje naukowe pasje.

Zakres pasji Krystyny Długosz-Kurczabowej był imponujący, jej naukowe zainteresowania dotyczyły słowotwórstwa historycznego, onomastyki, leksykologii i etymologii, języka religijnego, przekładów Biblii, w tym także onomastyki biblijnej.

Krystyna Długosz-Kurczabowa pracę naukową rozpoczęła, jak wielu ówczesnych adeptów polonistyki, od dialektologii, włączając się w prowadzone w latach powojennych przez ośrodek warszawski badania gwaroznawcze. Tematem jej pracy magisterskiej było *Porównanie dwóch wsi małopolskich w zakresie morfologii i słownictwa*, a tytułowe porównanie dotyczyło gwary jej rodzinnej miejscowości, Zagnańska, oraz odległej o ok. 100 km innej małopolskiej wsi – Książnice Wielkie. Do swojej małej ojczyzny w Górach Świętokrzyskich wracała wielokrotnie, były to zarówno powroty rodzinne i sentymentalne, jak i naukowe. Lubiła wspominać, że to nieopodal jej rodzinnych stron rośnie dąb Bartek – najstarsze drzewo w Polsce. O *Gwarze Zagnańska pół wieku później* pisała też w 2006 r. w artykule do książki ofiarowanej swemu bratu, Kazimierzowi Długoszowi, profesorowi Instytutu Filologii Polskiej Uniwersytetu Szczecińskiego, tak jak ona – poloniście, językoznawcy. Do problematyki gwaroznawczej powróciła również pod koniec życia. Od 2013 r. jako przewodnicząca warszawskiego oddziału Towarzystwa Kultury Języka wspierała program „Dialog pokoleń”, który Towarzystwo Kultury Języka prowadzi we współpracy z Narodowym Centrum Kultury. Jej osobistym wkładem w dialog międzypokoleniowy jest *Słownik gwary Zagnańska*. Słownika nie zdążyła opublikować, odeszła podczas ostatniego etapu pracy nad nim. W dniu śmierci do Towarzystwa Kultury Języka dotarł jego kosztorys z Wydawnictw Uniwersytetu Warszawskiego. Nie zdążyła się z nim zapoznać i nanieść ostatnich poprawek. Książka ukaże się staraniem rodziny.

Naukową pasją dr Krystyny Długosz-Kurczabowej było słowotwórstwo historyczne. Za kanoniczne należy uznać jej studium formacji rzeczownikowych ze słownika Jana Mączyńskiego opublikowane w „Pracach Filologicznych” (XXVII, 1977, s. 47–215). Jej podejście było nowoczesne, a ustalenia, zwłaszcza dotyczące funkcji specyfikującej w takich formacjach jak *glinka* (← *glina*), *mączka* (← *mąka*), *wódka* (← *woda*) itp., którą jako pierwsza opisała, zostały powszechnie przyjęte w słowotwórstwie historycznym. Równie znane i cenione są prace Krystyny Długosz-Kurczabowej poświęcone słowotwórstwu zapożyczeń. W studiach szczegółowych stawiała m.in. pytania, *Czy istnieje w polszczyźnie formant „-mistrz”* (1989). W cyklu artykułów omówiła m.in. formanty *-ada*, *-ajło*, *-alia*, *-ant*, *er*, *-eria*, *-inier/-ynier*, *-acja*, *-izm*, *-tor*, *-ulec*, *-us*. Problematyce słowotwórczej poświęcone są także studia *Sposoby przejmowania obcych formantów rzeczownikowych do polskiego systemu słowotwórczego* (1986), *Symbioza słowotwórcza* (1988) oraz napisana wspólnie ze Stanisławem Dubiszem *Gramatyka historyczna języka polskiego. Słowotwórstwo* (Warszawa 1999).

W dorobku Krystyny Długosz-Kurczabowej ważne miejsce zajmuje leksykologia i etymologia. Jest autorką cenionych słowników etymologicznych. W jednej z rozmów przyznała się, że do jej ulubionych książek należał *Słownik języka polskiego* Samuela Bogumiła Lindego. Zapewne dlatego w słownikach etymologicznych swego autorstwa również zastosowała układ gniazdowy, który jej zdaniem najlepiej ukazuje pokrewieństwo między wyrazami na pozór tak odległymi znaczeniowo jak np. *słodki* i *słony*. Uwagę zwraca także nieschematyczny sposób objaśnień, erudycyjne ciekawostki, nawiązania kulturowe i – podobnie jak u Lindego – bogata egzemplifikacja. Krystynie Długosz-Kurczabowej także w słownikach udało się zachować właściwy

całemu jej piśmownictwu naukowemu eseistyczny, erudycyjny, a jednocześnie przystępny i nade wszystko elegancki styl. Pracy nad popularyzowaniem wiedzy etymologicznej poświęciła wiele lat życia, rozpoczynając od popularnonaukowego i dydaktycznego *Słownika szkolnego. Etymologia* (WSiP 1998), który następnie został rozbudowany i ukazał się pt. *Nowy słownik etymologiczny języka polskiego* (PWN 2001), a w wersji zmienionej i poszerzonej jako *Słownik etymologiczny języka polskiego* (PWN 2003). Ze względu na duże zainteresowanie opracowaniami etymologicznymi Krystyny Długosz-Kurczabowej Wydawnictwo Naukowe PWN zamówiło u niej *Wielki słownik etymologiczno-historyczny języka polskiego*. Od pierwszego wydania w 2008 r. doczekał się licznych wznowień, ostatnie pochodzi z 2016 r.

W nurcie prac etymologicznych i popularnonaukowych należy także umieścić nietypowy słownik, który dowodzi, jak nietuzinkową osobowością była Krystyna Długosz-Kurczabowa. Otóż na próżno szukać w księgarniach liczącego 326 s. słownika pt. *Wyrazy, które intrygowały i które intrygują*, wydanego w 2015 r. przez „Wydawnictwo Kurczabowie”; to nietypowe dzieło ukazało się zaledwie w 75 egzemplarzach na 75. urodziny jego Autorki. Dedykowane jest wnukom – Krzysiovi i Kacperkowi, ale każdy z 75 egzemplarzy osobiście przez Autorkę był darowany konkretnym osobom na pamiątkę wspólnych chwil. Podstawą tego słownika były odpowiedzi, jakich udzielała przez lata w Poradni PWN. Praca ta stanowi swego rodzaju diachroniczne „tutti-frutti”, jest utworem napisanym z humorem i lekkością.

Bardzo ważną dziedzinę jej naukowej działalności (istotną także dla wspomnianej Krystyny Długosz-Kurczabowej) stanowią prace poświęcone językowi religijnemu oraz przekładowi Biblii. Wśród nich są studia z zakresu onomastyki biblijnej, jak *Biblijne imiona eponimiczne* (1988), *Onomastyka Nowego Testamentu* (1991), *Nowotestamentowe antroponimy w języku polskim* (1990), a przede wszystkim *Apelatywizacja biblijnych nazw własnych w języku polskim* (1990). Pani Krystyna należała do osób głęboko wierzących, o tym, jak ważną postacią była w macierzystej rzymskokatolickiej parafii przy ul. Gwiazdźistej na warszawskim Żoliborzu, mogliśmy przekonać się podczas jej pożegnania. Śmiało można powiedzieć, że współtworzyła atmosferę intelektualną tego miejsca. Choć na co dzień była osobą skromną, miała jednak świadomość swej wiedzy, była intelektualistką w najlepszym tego słowa znaczeniu – samodzielną, odważną, otwartą na inny punkt widzenia i gotową do dyskusji. Prowadziła otwarte wykłady, słynne były jej analizy filologiczne modlitw – *Ojciec nasz* (1997), *Zdrowaś Maryjo* (2003), część z nich szczęśliwie została opublikowana w formie artykułów. Jej niezwykła umysłowość, postawa intelektualna naznaczona wewnętrzną potrzebą prawdy, rozumnego i uczciwego używania słowa wyraźnie odbijała się na atmosferze umysłowej nie tylko jej macierzystej parafii, ale wszystkich zespołów, które tworzyła lub współtworzyła.

Krystyna Długosz-Kurczabowa była również niezwykle cenionym i trudnym do zastąpienia konsultantem językowym ekumenicznego przekładu Biblii. Jej współpraca z Towarzystwem Biblijnym zaczęła się od opublikowanej w „Przeglądzie Humanistycznym” krytycznej *Charakterystyki językowej ekumenicznego przekładu „Ewangelii św. Mateusza”* (1998). Jak podkreślają bibliści, początkowo była niechętna odchodzeniu od tradycyjnego stylu biblijnego w ekumenicznym przekładzie Pisma

Świętego, stopniowo jednak przekonywała się do idei translacji dynamicznej, języka współczesnego, ale – co podkreślała – dobrze przemyślanego i poprawnego.

Krystyna Długosz-Kurczabowa współpracowała przy wydaniu innych książek wyznaniowych Kościołów chrześcijańskich, *Konfesji sandomierskiej* (1995) oraz fragmentu *Postylli* M. Reja *Rozprawa o wieczerzy Pańskiej* (2005).

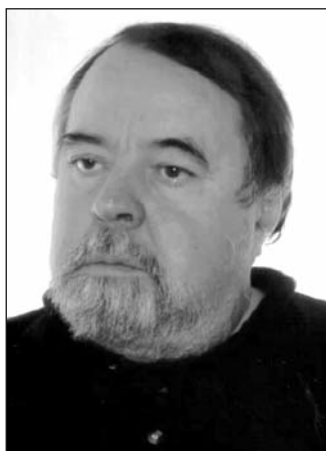
Na osobne omówienie zasługują podręczniki akademickie, których Krystyna Długosz-Kurczabowa jest współautorką ze Stanisławem Dubiszem. Najważniejszym dokonaniem tego „duetu diachronicznego” jest obszerny podręcznik *Gramatyka historyczna języka polskiego*, który znalazł trwałe miejsce w uniwersyteckiej dydaktyce polonistycznej. Pierwszy tom, obejmujący fonetykę, ukazał się w 1993 r. Od tego czasu podręcznik jest systematycznie wznawiany i – co ważniejsze – poszerzany systematycznie o opracowania kolejnych działów języka – fleksji, słowotwórstwa i składni (2006). Dzieło jest spójne koncepcyjnie i metodologicznie. Autorzy zawarli w nim własną wizję ewolucji polszczyzny, punktem wyjścia czyniąc pojęcie zmiany językowej. Opisy diachroniczne skonfrontowali z zestawieniami synchronicznymi. Uzupełnieniem podręcznika są ćwiczenia, czyli *Gramatyka historyczna języka polskiego w testach, ćwiczeniach i tematach egzaminacyjnych* (1999 i n.). Podręczniki akademickie wyrastają z doświadczeń dydaktycznych, stanowią ich ukoronowanie. Krystyna Długosz-Kurczabowa była typem nauczyciela – mistrza. Studenci cenili zajęcia z nią, podziwiali za nietuzinkową osobowość, błyskotliwy umysł, poczucie humoru i lakoniczne, acz trafne i niekiedy ironiczne puenty. Pomimo wielu obowiązków miała czas dla swoich magistrantów, była zawsze życzliwa i chętna do pomocy. Dr Krystyna Długosz-Kurczabowa zawodowe życie związała z Uniwersytetem Warszawskim, współpracowała także z kilkoma uczelniami, Wyższą Szkołą Humanistyczną w Pułtusku, Podyplomowym Studium Logopedycznym, Studium Wymowy Akademii Teatralnej w Warszawie.

Liczba publikacji naukowych dr Krystyny Długosz-Kurczabowej obejmuje ponad 100 pozycji, liczba artykułów popularyzujących wiedzę o dawnej polszczyźnie przekracza 300 pozycji. Przez lata recenzowała i opiniowała do druku podręczniki szkolne jako recenzentka Ministerstwa Edukacji Narodowej, była konsultantem językowym w Wydawnictwach Naukowo-Technicznych i Wydawnictwie Naukowym PWN.

Trudno nie wspomnieć o działalności społecznej dr Krystyny Długosz-Kurczabowej. Należała do Towarzystwa Miłośników Języka, była członkiem Polskiego Towarzystwa Językoznawczego, przez lata kierowała Oddziałem Warszawskim Towarzystwa Kultury Języka.

Dr Krystyna Długosz-Kurczabowa przez lata ciężko chorowała na serce, przyzwyczaiła nas do swojej choroby, bo tej nigdy nie traktowała jako przeszkody w pracy. Oszczędnie dawkowała informacje na ten temat. Dopiero gdy odeszła, uświadomiliśmy sobie, że naprawdę chorowała. Odeszła w pełni swego potencjału, pełna planów naukowych. Pamiętamy ją jako osobę szlachetną, niezwykle skromną, która często mawiała, że najlepiej czuje się w kameralnym gronie, a jednocześnie niezwykle błyskotliwą. Miała umysł otwarty, młodzieńczy do końca.

**Profesor Roman Kalisz
(1948–2017)**



14 stycznia 2017 r. zmarł w Gdyni prof. dr hab. Roman Kalisz, wybitny językoznawca, znakomity nauczyciel akademicki, współtwórca gdańskiej anglistyki.

Roman Kalisz urodził się 14 czerwca 1948 r. w Suwałkach, w rodzinie nauczycielki języka angielskiego i prawnika, dawnego więźnia obozów sowieckich. W r. 1965 ukończył w mieście rodzinnym Liceum im. Marii Konopnickiej. Po maturze zdecydował się kontynuować tradycje rodzinne, wybierając filologię angielską na Uniwersytecie im. Adama Mickiewicza w Poznaniu. Należał do pierwszego rocznika absolwentów poznańskiej anglistyki (w 1971 r.). Pracę magisterską pt. *The Grammar of the Verb in John Steinbeck's Cannery Row*, dotyczącą wybranych czasowników w prozie amerykańskiego pisarza, przygotował pod kierunkiem profesora Jacka Fisiaka. W latach 1971–1973 był zatrudniony w Pracowni Językoznawstwa Kontrastywnego w Instytucie Badań Literackich Polskiej Akademii Nauk w Poznaniu.

W 1973 r. rozpoczął pracę na Uniwersytecie Gdańskim, w nowo utworzonym wówczas Zakładzie Anglistyki (który później przekształcił się w Instytut). W roku następnym uzyskał stopień doktora nauk humanistycznych. Rozprawę doktorską, zatytułowaną *Syntax and Semantics of the Clauses of Result in English and Polish*, będącą studium kontrastywnym angielskich i polskich zdań skutkowych, napisał pod kierunkiem profesora Kazimierza Polańskiego. Rozprawa ta została wydana drukiem przez Wydawnictwo Uniwersytetu Gdańskiego w r. 1978.

W 1981 r. opublikował rozprawę pt. *The Pragmatics, Semantics and Syntax of the English Sentences with Indicative That Complements and the Polish Sentences with Że Complements: A Contrastive Study* (Wydawnictwo UG), na podstawie której uzyskał stopień doktora habilitowanego, natomiast po wydaniu książki *Pragmatyka językowa* (Wydawnictwo UG, 1993) uzyskał w 1995 r. tytuł profesora nauk humanistycznych. We wspomnianej rozprawie Roman Kalisz zajął się angielskimi i polskimi zdaniami względnymi zawierającymi spójniki *that* i *że*. W tytule rozprawy Autor wysunął aspekty pragmatyczne i semantyczne tych zdań na plan pierwszy, przed aspektem składniowym. Ten porządek zapowiadał, na czym Roman Kalisz skoncentruje swoją uwagę w latach przyszłych – na pragmatyce i semantyce.

W monografii poświęconej pragmatyce językowej (była to pierwsza w języku polskim książka poświęcona temu zagadnieniu) Roman Kalisz zaproponował prakseologiczną definicję pragmatyki, rozumianej jako dziedzina językoznawstwa „zajmująca się skutecznym działaniem człowieka poprzez użycie środków językowych (jak również milczenia posiadającego wartość komunikatywną)” (s. 9). Warto podkreślić uwzględnienie w tej definicji pragmatycznego wymiaru milczenia mającego wartość komunikatywną. W tej monografii Roman Kalisz omawia wszystkie ważniejsze aspekty pragmatyki, takie jak akty mowy, reguły konwersacji, implikatura, referencjalność i presupozycja. Autor proponuje własne rozwiązania – jedną z takich propozycji jest dodanie do sześciu wyróżnionych przez Johna Searle’a aktów mowy (są to: asercje, dyrektywy, akty wyrażające podejmowanie odpowiedzialności przez mówiącego, akty illokucyjne ekspresywne, deklaracje i deklaracje asertywne) siódmego typu aktów mowy, a mianowicie ewaluatywnych aktów mowy (s. 48–49).

Profesor Roman Kalisz był też jednym z pierwszych badaczy w Polsce, którzy dostrzegli ogromny potencjał językoznawstwa kognitywnego. Poświęcił temu zagadnieniu liczne artykuły oraz wydaną w 2001 r. monografię *Językoznawstwo kognitywne w świetle językoznawstwa funkcjonalnego* (Gdańsk 2001). W zakończeniu tej książki napisał o językoznawstwie kognitywnym, że ma ono „już niebagatelne osiągnięcia i jest jedną z najburzliwiej rozwijających się dziedzin językoznawczych, mającą ogromne możliwości. Inne teorie jakby nieco wypaliły się intelektualnie i dlatego przewiduję w najbliższych latach dalszy rozkwit tej dziedziny” (s. 172) oraz „Językoznawstwo kognitywne jest częścią językoznawstwa funkcjonalnego, gdyż podziela większość postulatów funkcjonalistów. Jednak językoznawstwo kognitywne idzie jeszcze dalej w powiązaniu języka z cielesnymi i umysłowymi właściwościami człowieka. Można stwierdzić, że jest radykalną formą językoznawstwa funkcjonalnego” (s. 170).

W swoich pracach kognitywistycznych Profesor Kalisz zajmował się wieloma zjawiskami językowymi, takimi jak schematy wyobrazeniowe, metafory, metonimie, amalgamaty pojęciowe czy akty mowy. Nigdy nie traktował językoznawstwa i uprawianych teorii w sposób dogmatyczny, zawsze charakteryzowało go otwarte podejście do nowych zjawisk i propozycji teoretycznych. Sam nierzadko porównywał teorie i paradygmaty, m.in. różne teorie dotyczące aktów mowy (w książce z 1993 r.), kognitywizm i funkcjonalizm (w książce z 2001 r. oraz w broszurze opublikowanej

razem z Wojciechem Kubińskim w Amsterdamie, w serii „Working Papers in Functional Grammar”, w 1997 r.), kognitywizm i generatywizm (w artykule „Extended Standard Theory and Cognitive Linguistics: A Search for Common Ground”, w *Biuletynie Polskiego Towarzystwa Językoznawczego* 46, 1991), czy kognitywizm i relatywizm (w artykule „Językoznawstwo kognitywne a relatywizm”, w *Biuletynie Polskiego Towarzystwa Językoznawczego* 52, 1997).

Roman Kalisz starał się łączyć w swoich dociekaniach językoznawczych także pragmatykę z kognitywizmem, choć miał świadomość odrębności metodologicznej tych dziedzin. Jak zauważył w pracy pt. „Pojęcia pragmatyki językowej w świetle językoznawstwa kognitywnego” (opublikowanej w tomie *Językoznawstwo kognitywne II: Zjawiska pragmatyczne*, red. W. Kubiński, D. Stanulewicz, Gdańsk 2001): „Pragmatyczne podejście do języka i podejście kognitywne nie zawsze istnieją w najlepszej harmonii, gdyż dotychczasowe teorie pragmatyki językowej są teoriami socjologicznymi w swojej warstwie podstawowej, a teoria językoznawstwa kognitywnego jest teorią psychologiczną. Stąd mamy do czynienia z różnymi rozkładami akcentów w pracy badawczej”, ponadto „pragmatyka i językoznawstwo kognitywne nie są równoważne, jeśli chodzi o poziom badawczy. Pragmatyka językowa jako taka stanowi komponent języka tak jak semantyka czy składnia (według pewnych ujęć pragmatyka jest zjawiskiem pozajęzykowym, a nawet pozajęzykoznawczym). Językoznawstwo kognitywne jest teorią, pewnym sposobem podejścia do języka” (s. 13).

Roman Kalisz, we współpracy z Wojciechem Kubińskim i Andrzejem Bullerem, poszukiwał także powiązań gramatyki kognitywnej ze sztuczną inteligencją. Rezultatem tych badań była wspólna książka pt. *In Search of a Frame of Mind: An Introduction to Cognitive Linguistics and Artificial Intelligence* (Gdańsk 1996).

Profesor Kalisz był jednym z inicjatorów serii „Językoznawstwo kognitywne”, publikowanej przez Wydawnictwo Uniwersytetu Gdańskiego. W tomach tej serii publikowane są prace nie tylko polskich kognitywistów, lecz również tłumaczenia artykułów zagranicznych badaczy. Profesor Kalisz miał duże zasługi w pozyskaniu znakomitych autorów, m.in. Ronalda W. Langackera, Eugene’a Casada, Zoltána Kövecsesa, Tomasza P. Krzeszowskiego oraz Barbary Lewandowskiej-Tomaszczyk.

Roman Kalisz nierzadko podejmował polemikę z innymi językoznawcami, m.in. z Anną Wierzbicką na temat „alfabetu ludzkich myśli”, z Tomaszem P. Krzeszowskim na temat „podstawowej macierzy aksjologicznej”, a z Jadwigą Puzyrą na temat założeń semantyki aksjologicznej.

W ostatnich latach życia Profesor Kalisz pracował nad własną koncepcją znaczenia ogólnego. Główne założenia tej koncepcji przedstawił m.in. podczas konferencji „Relatywizm w języku i kulturze” (Sandomierz, 2008) oraz konferencji „International Conference on Philosophy of Language and Linguistics. PhiLang2009” (Łódź, 2009). Założenia tej koncepcji trafnie podsumowuje tytuł artykułu „Gramatyka prawdy, dobra i piękna” (opublikowany w tomie pod redakcją A. Kwiatkowskiej, *Przestrzenie kognitywnych poszukiwań*, Łódź 2012). W tej nowatorskiej koncepcji znaczenie wypowiedzenia plasuje się na trzech skalach: kontinuum PRAWDA – FAŁSZ, kontinuum DOBRO – ZŁO oraz kontinuum PIĘKNO – BRZYDOTA. Do

dwóch znanych i dobrze już opisanych skal mierzących prawdziwość wypowiedzeń oraz ich wartość aksjologiczną, Roman Kalisz dodaje trzecią, mierzącą właściwości estetyczne. W zakończeniu przywołanego artykułu autor stwierdził, że jego koncepcja „jest swoistą odwrotnością teorii rozwijanej przez Annę Wierzbicką, która szuka wyrażen podstawowych rozbijając znaczenia jednostek leksykalnych. My szukamy uogólnień, postulując skale podstawowe poznawczo w naszej kulturze” (s. 159). Zagadnieniu znaczenia ogólnego poświęcony był też jeden z ostatnich opublikowanych tekstów Romana Kalisza, „A Concept of General Meaning: Selected Theories in Comparison to Selected Semantic and Pragmatic Theories” (*Research in Language* 11: 3, 2013), w którym zarysował swoją koncepcję na tle szerokiego spektrum teorii semantycznych, pragmatycznych i aksjologicznych¹.

Roman Kalisz był związany z Uniwersytetem Gdańskim od 1973 r. Przez kilka kadencji sprawował funkcję dyrektora Instytutu Anglistyki. Angażował się również w organizację konferencji, zapraszał na wykłady licznych wybitnych naukowców. Wśród nich znalazł się współtwórca gramatyki kognitywnej, Ronald W. Langacker, który gościł na Uniwersytecie Gdańskim dwukrotnie.

Profesor Kalisz nie zapomniał o swoich rodzinnych Suwałkach, gdzie przez wiele lat sprawował opiekę naukowo-dydaktyczną i prowadził seminaria licencjackie dla studentów filologii angielskiej w Zespole Kolegiów Nauczycielskich. Przez kilka lat pracował także we Wszechnicy Mazurskiej w Olecku, Uczelni Techniczno-Lingwistycznej (wcześniej: Wyższej Szkole Języków Obcych) w Świeciu oraz w Ateneum – Szkole Wyższej w Gdańsku.

Profesor Kalisz wypromował przez lata 23 doktorów. Wśród nich są dziś doktorzy habilitowani oraz profesorowie uczelni polskich i zagranicznych. Był też recenzentem licznych rozpraw doktorskich i habilitacyjnych oraz w postępowaniach profesorskich. Zawsze chętnie podejmował się wszelkich recenzji, nie skąpił uwag, rad i wskazówek.

Profesor Roman Kalisz zostanie w naszej pamięci jako człowiek cieszący się życiem, rodziną, towarzystwem ludzi, językoznawstwem, dobrą kuchnią, muzyką i podróżami. Człowiek dobry, życzliwy i pogodny. Przyjaciół i Mistrz.

Piotr Stałmaszczyk
Danuta Stanulewicz

¹ Bibliografie prac Romana Kalisza, opracowane przez Danutę Stanulewicz, znajdują się w poświęconej mu księdze pamiątkowej, *Lingua Terra Cognita*, red. D. Stanulewicz, Gdańsk 2010: Wydawnictwo Uniwersytetu Gdańskiego, 37–44 (do 2009 r.), oraz w czasopiśmie *Beyond Philology* 14: 1, 2017, 15–25 (do 2017 r.).

**Docent doktor Maria Przybysz-Piwko
(1950–2016)**



Docent doktor Maria Przybysz-Piwko przez ponad 40 lat swojego zawodowego życia – jako badacz i dydaktyk – była związana z warszawską polonistyką i logopedią. Była cenionym nauczycielem akademickim, wychowawczynią kilku pokoleń logopedów, osobą powszechnie lubianą i szanowaną, oddaną ludziom i sprawie rozwoju logopedii.

W 1973 r. ukończyła filologię polską na Uniwersytecie Warszawskim. Jej praca magisterska, którą przygotowała pod kierunkiem prof. dra hab. Stanisława Skorupki, nosiła tytuł *Neologizmy słotwórcze i semantyczne w poezji Juliana Przybosa*. Jeszcze w tym samym roku podjęła pracę na stanowisku naukowo-technicznym w Instytucie Języka Polskiego PAN w Pracowni Badania Mechanizmów Mowy kierowanej przez prof. dr hab. Halinę Mierzejewską. W 1974 r. rozpoczęła studia doktoranckie na Wydziale Polonistyki UW. W 1976 r. została zatrudniona w Instytucie Języka Polskiego na stanowisku starszego asystenta. W 1986 r. obroniła rozprawę doktorską, napisaną pod kierunkiem prof. dr hab. H. Mierzejewskiej, zatytułowaną *Afatyczne*

zakłócenia leksemów polskich. Od 1986 r. pracowała na stanowisku adiunkta, a od 2007 r. – docenta. W 2011 r. odeszła na emeryturę, lecz nadal współpracowała z macierzystym wydziałem oraz Pomagisterskim Studium Logopedycznym UW.

Maria Przybysz-Piwko stale podnosiła kwalifikacje, biorąc udział w studiach podyplomowych, szkoleniach, seminariach i stażach. W latach 1973–1975 uczestniczyła w pracach interdyscyplinarnego zespołu ds. diagnozy i terapii osób z afazją, prowadzonych pod kierunkiem prof. H. Mierzejewskiej. W latach 1979–1981 odbyła studia podyplomowe w Pomagisterskim Studium Logopedycznym UW, gdzie uzyskała tytuł zawodowy logopedy. Uczestniczyła w seminarium neuropsychologicznym prowadzonym w Centralnym Laboratorium Badania Mózgu Bułgarskiej Akademii Nauk w Sofii (1977) oraz w stażu w Katedrze Lingwistyki i Fonetyki Uniwersytetu Karola w Pradze (1980). W latach 90. XX w. brała udział w licznych szkoleniach specjalistycznych.

Zainteresowania badawcze doc. Marii Przybysz-Piwko zawsze obejmowały zagadnienia z pogranicza lingwistyki i logopedii. Początkowo zajmowała się przede wszystkim problematyką afazjologiczną. W swoich dociekaniach nad stanem systemu językowego u osób z afazją reprezentowała podejście interdyscyplinarne. Zdawała sobie bowiem doskonale sprawę z tego, że badania afazji mają z natury rzeczy charakter kompleksowy. Wymagają porozumienia i współpracy kilku dyscyplin. A to z kolei oznacza konieczność stałej wymiany doświadczeń i przemyśleń, niezbędnych dla wspólnego poszukiwania najbardziej celowych i skutecznych rozwiązań metodologicznych. W swych publikacjach odwoływała się m.in. do ustaleń poczynionych na gruncie cybernetyki, neurochirurgii, neurofizjologii oraz neuropsychologii. Szczególnie bliska była jej neuropsychologiczna koncepcja afazji autorstwa rosyjskiego psychologa Aleksandra R. Łurii.

Drugi ważny obszar zainteresowań naukowo-badawczych doc. Marii Przybysz-Piwko to trudności w mowie, czytaniu i pisaniu dzieci dyslektycznych oraz dzieci z niedokształceniem mowy pochodzenia korowego. Opisując trudności w opanowywaniu mowy, czytania i pisania przez dzieci, skupiała się przede wszystkim na analizie osłabionych sprawności językowych, w tym przede wszystkim fonologicznych, które traktowała jako podstawę organizowania się systemu językowego u dziecka.

Szczególne miejsce w jej dociekaniach zajmowało zjawisko świadomości fonologicznej i fonemowej. Tę pierwszą definiowała jako umiejętność segmentacji strumienia mowy – ciągu fonicznego – na jednostki mniejsze niż wyraz, czyli sylaby i fonemy. Drugą z kolei rozumiała jako rozpoznawanie i odróżnianie elementarnych jednostek systemu fonologicznego. Na bazie tych rozróżnień teoretycznych stworzyła narzędzie logopedyczne służące jako pomoc w rozwoju umiejętności fonologicznych dziecka z trudnościami w czytaniu i pisaniu. Nosi ono nazwę *Fonologicznego alfabetu obrazkowego* (FAO) i rzeczywiście stanowi swoisty alfabet obrazkowy dla systemu fonologicznego, który pozwala „czytać” i „pisać” bez użycia liter.

Na dorobek publikacyjny doc. Marii Przybysz-Piwko składa się ponad 70 autorских i współautorskich rozpraw i artykułów naukowych, publikacji pod redakcją, recenzji, sprawozdań oraz publikacji popularyzatorskich i edukacyjnych z pogranicza

lingwistyki oraz logopedii i emisji głosu. Działalność naukowa obejmuje także wystąpienia na licznych krajowych i międzynarodowych konferencjach naukowych, także za granicą. W latach 1975–2011 doc. Maria Przybysz-Piwko wystąpiła z około 40 referatami, warsztatami i plakatami.

Docent Maria Przybysz-Piwko ma również bogaty dorobek dydaktyczny. Prowadziła różnorakie zajęcia dydaktyczne – ćwiczenia, warsztaty, konwersatoria, wykłady kursowe i monograficzne, proseminaria i seminaria (dyplomowe, licencjackie, magisterskie) dla słuchaczy Pomagisterskiego Studium Logopedycznego oraz studentów Wydziału Polonistyki. W bogatej ofercie dydaktycznej doc. Marii Przybysz-Piwko znajdowały się takie przedmioty, jak: lingwistyczne podstawy logopedii, wstęp do logopedii, fonetyka artykulacyjna i patoartykulacyjna, ortofonia i grafetyka (pismo jako wytwór czynności grafomotorycznych), autokorekcja, zaburzenia mowy, czytania i pisanie, afazja. W latach 1993–2013 wypromowała ponad 300 absolwentów Pomagisterskiego Studium Logopedycznego – logopedów dyplomowanych. Była też promotorem ok. 100 prac magisterskich na Wydziale Polonistyki UW. Tematyka tych prac była bardzo różnorodna, obejmowała niemal wszystkie zagadnienia wchodzące w zakres logopedii oraz pogranicza logopedyczno-lingwistycznego.

Działalność dydaktyczna doc. Marii Przybysz-Piwko wykraczała także poza granice Polski. Wielokrotnie prowadziła zajęcia fonetyczno-korekcyjne dla studentów polonistyki Uniwersytetu Amsterdamskiego (Faculteit der Geesteswetenschappen, Slavische talen en Culturen – sectie Pools). W 2004 r. wzięła udział w realizowanym w tamtejszej uczelni w ramach Slavisch Seminarium projekcie pod nazwą *Nowe materiały dydaktyczne do nauki wymowy i intonacji języka polskiego*. Efektem prac była płyta CD z materiałami językowymi do ćwiczenia artykulacji i intonacji polskiej.

Na uwagę zasługuje również bogata działalność organizacyjna doc. Marii Przybysz-Piwko. W trakcie pracy na UW sprawowała wielokrotnie funkcje kierownicze. W latach 1990–2003 oraz 2008–2010 pełniła funkcję kierownika Pomagisterskiego Studium Logopedycznego UW, a w latach 2002–2004 – kierownika Podyplomowego Studium Emisji Głosu UW. Dwukrotnie była kierownikiem Zakładu Fonetyki i Logopedii, działającego w ramach Instytutu Języka Polskiego, a następnie Instytutu Polonistyki Stosowanej Wydziału Polonistyki UW (2001–2003 i 2005–2006). Również dwukrotnie kierowała specjalnością „logopedia” na kierunku filologia polska.

Działalność organizacyjna doc. Marii Przybysz-Piwko obejmowała też aktywność w licznych towarzystwach naukowych i zawodowych oraz innych gremiach. Najważniejsze z nich to Polskie Towarzystwo Logopedyczne oraz Polskie Towarzystwo Językoznawcze. Była również członkiem Komisji Rozwoju i Zaburzeń Mowy Komitetu Językoznawstwa Polskiej Akademii Nauk, Komisji Certyfikacyjnej Polskiego Związku Logopedów, Zespołu Konsultacyjnego ds. Logopedii przy Ministerstwie Edukacji Narodowej i Sportu, Zespołu Kwalifikacyjnego Głównej Komisji ds. Stopni Specjalizacji Zawodowej Nauczycieli (w zakresie logopedii), Komisji Egzaminacyjnej do Przeprowadzania Egzaminu Specjalizacyjnego w Dziedzinie „Neurologopedia” Ministerstwa Zdrowia. Częstokroć pełniła we władzach organizacji odpowiedzialne funkcje, m.in. przewodniczącej Polskiego Kolegium Logopedycznego

oraz wiceprezesa Towarzystwa Kultury Języka, wiceprzewodniczącej, a następnie przewodniczącej Sekcji Logopedycznej Towarzystwa.

Doc. Maria Przybysz-Piwko zaangażowana była nie tylko w pracę naukową, dydaktyczną i organizacyjną na rzecz logopedii, lecz także w działalność samorządu lokalnego oraz służbę społeczną. W latach 1998–2016 była radną w rodzinnych Markach, mieście położonym w aglomeracji warszawskiej w powiecie wołomińskim. Przez trzy kolejne kadencje (1998–2010) przewodniczyła Radzie Miasta. W uznaniu zasług dla rodzinnego miasta i jego mieszkańców otrzymała pośmiertnie honorowy tytuł „Zasłużony dla Miasta Marki”.

Doc. Maria Przybysz-Piwko zmarła 16 października 2016 r. w Markach. Odeszła przedwcześnie i – mimo ciężkiej choroby, z którą w ostatnich latach heroicznie walczyła – niespodziewanie. Jej uczniowie i przyjaciele zachowają Ją we wdzięcznej pamięci.

Ewa Wolańska

doi: <http://dx.doi.org/10.17651/bptj.74.1>

MIROŚLAW BAŃKO

Uniwersytet Warszawski, Warszawa

Korzyści i problemy wynikające z „nadmiaru” danych językowych i z wielości metod ich analizy

Streszczenie

Ilość i różnorodność danych, z jakich może korzystać językoznawca dziś, jest nieporównanie większa niż w przeszłości. Nowe dane wpisują się w schemat znany od dawna: obejmuje on dane systemowe, meta-językowe, elicytacyjne, introspekcyjne, tekstowe, a także inne przekazy kulturowe istotne dla badań nad językiem, np. dane ikonograficzne. Nowe źródła danych i nowe narzędzia do ich analizy znacząco jednak wzbogacają warsztat badawczy językoznawcy.

Pociąga to za sobą korzyści, ale też rodzi problemy, m.in. problem „nadmiaru” danych, których nie sposób zinterpretować tradycyjnymi metodami, oraz problem interpretacji wyników, gdy dane czerpane z różnych źródeł prowadzą do różnych wniosków. Obfitość danych, zwłaszcza cyfrowych, jest też wyzwaniem dla dydaktyki akademickiej. Niektóre umiejętności wymagane dziś do badań nad językiem daleko odbiegają od tradycyjnego modelu kształcenia filologów.

Słowa kluczowe: dane językowe, korpusy językowe, humanistyka cyfrowa, edukacja językoznawcza

The “overload” of linguistic data and the abundance of data analysis methods – benefits and problems

Abstract

The amount and variety of data that linguists have at their disposal are incomparably vaster now than in the past. New data fit into the schema known for a long time: this encompasses systemic data, metalinguistic data, elicitation data, introspective data, textual data, and other cultural messages relevant to the study of language, e.g. iconographic data. However, new data sources and new analytic tools significantly enrich linguistic research methodology.

This brings benefits but also causes problems: the problem of data “overload” where traditional analytic techniques prove inadequate, and the problem of data interpretation in cases where data drawn from different sources yield different results. The abundance of data is also a challenge to academic training: some research skills now necessary in linguistic studies are far from the traditional model of philological training.

Keywords: linguistic data, language corpora, digital humanities, linguistic training

Słowo *dane* – kalka łacińskiego *data* – jest mylące. Sugeruje, że dane są zawsze *da ne*, gotowe do użycia, podczas gdy w rzeczywistości trzeba je najpierw *pozyskać*, *wyszukać*, *zgrupować*, potem zaś *opracować*, *przetworzyć*, *sprawdzić*, *zinterpretować*. Zakres czynności, którym podlegają dane, można ustalić, analizując kolokacje tego słowa w korpusie językowym, przy czym analiza taka jest sama w sobie przykładem pozyskiwania i opracowywania danych. Wśród kolokacji słowa *dane* zwracają uwagę połączenia świadczące o tym, że dane często podlegają obróbce komputerowej, por. *bank danych*, *baza danych*, *przetwarzanie danych*, *transmisja danych*.

Czasem dane istotnie są dane, jak w szkolnych zadaniach z matematyki i fizyki, które kazano nam rozwiązywać według pewnego schematu: z lewej strony pod nagłówkiem *Dane* należało zapisać posiadane informacje, np. odległość i prędkość jazdy, z prawej zaś pod nagłówkiem *Szukane* – informację poszukiwaną, np. czas przejazdu. Znacznie częściej jednak dane są ukryte, czekają dopiero na ujawnienie. Przykładem ujawniania danych są techniki analityczne stosowane do dużych baz danych w ramach tzw. *data mining* – procedur określanych po polsku m.in. jako *eksploracja danych* lub *wydobywanie danych* (np. Larose 2006). Ostatnia z tych nazw, podobnie jak oryginalna, angielska, ma motywację metaforyczną, związaną z górnictwem, czyli przemysłem wydobywczym, tak bowiem jak górnik wydobywa spod ziemi coś, co nie jest widoczne, tak specjaliści w zakresie *data mining* starają się ujawnić fakty, które w bazach danych są obecne *implicite*, choć pozostają na razie niewidoczne.

W artykule tym będą mnie interesować tylko dane istotne dla badań nad językiem, przede wszystkim – dla badań dotyczących słownictwa. Nie muszą one wcale pochodzić ze źródeł *stricte* językowych, zestawianych przez językoznawców do celów badawczych. Na przykład biblioteki cyfrowe, archiwa prasowe, a nawet katalogi biblioteczne mogą być źródłem informacji świadczących o używaniu języka w określony sposób.

Znakiem czasu jest obecnie obfitość danych: nigdy dotąd nie dysponowaliśmy tak wieloma źródłami danych istotnych językowo oraz narzędziami do ich wyszukiwania i przetwarzania (dotyczy to oczywiście tylko języków żywych, zwłaszcza mających wielu użytkowników; liczba zabytków języków martwych może być bardzo skromna). Znakiem czasu jest także tempo zmian: danych szybko przybywa i pojawiają się nowe narzędzia do ich analizy. Na przykład odkąd zaczęły powstawać komputerowe korpusy językowe do celów badawczych, ich wielkość rosła o rząd wielkości niemal z dekady na dekadę: w latach 60. kresem możliwości był korpus wielkości jednego miliona słów (np. Brown Corpus), w latach 80. wydawnictwo Collins zbudowało na potrzeby leksykograficzne korpus COBUILD wielkości 20 milionów słów, w latach 90. siłami kilku instytucji powstał British National Corpus, obejmujący 100 milionów słów, a dziś np. wspomniany COBUILD chwali się objętością przeszło 4 miliardów słów.

W Polsce wydarzenia przebiegały podobnie, choć z pewnym opóźnieniem. Ich początek wyznaczają listy frekwencyjne *Słownictwo współczesnego języka polskiego* (Kurcz i in. 1974–1977), obejmujące pięć stylów funkcjonalnych, po 100 tysięcy słów z każdego. Następnym etapem, już w latach 90., był nie ukończony korpus polszczyzny przygotowywany w Instytucie Języka Polskiego PAN oraz duży na tamte lata i dobrze zrównoważony Korpus Języka Polskiego PWN, w całości (ok. 100 mi-

lionów słów) dostępny tylko w wydawnictwie do prac słownikowych, w części (ok. 40 milionów słów) udostępniony publicznie w Internecie (od niedawna jednak w taki sposób, który czyni go mało użytecznym do badań). W latach 2007–2010, na podstawie wcześniej powstałych zasobów i nowych, pozyskanych w tym czasie, powstał Narodowy Korpus Języka Polskiego (NKJP), obejmujący ponad miliard słów, z czego ok. 250 milionów w części zrównoważonej (zob. Przepiórkowski i in. 2012, gdzie omówiono też krótko wcześniejsze projekty korpusowe w Polsce i na świecie). Trwa dyskusja nad rozbudową NKJP, przy czym rozwiązania brane pod uwagę to nie tylko powiększenie istniejącego korpusu bez zmiany jego architektury, lecz także udostępnienie za pomocą tych samych wyszukiwarek szeregu korpusów specjalistycznych, co pozwoliłoby użytkownikom zestawiać własne zbiory danych, odpowiednio do celów badawczych (Ogrodniczuk i in. 2017).

Obfitość łatwo dostępnych danych wpłynęła na zmianę zapatrywań wielu lingwistów na to, czym właściwie są dane w opisie języka i na jakich danych opis ten powinien bazować. Zmiana dotyczy głównie opisu języka współczesnego badaczom, o ile bowiem minione stany języka opisywano na podstawie dawnych tekstów (jeśli zostały zachowane), o tyle w badaniach nad współczesnym językiem – zwłaszcza mieszczących się w nurcie generatywnym lub przezeń inspirowanych – przeceniano rolę intuicji i kompetencji językowej badacza (Stubbs 1996: 22–50). Zdarzało się, że zdania preparowane uważano za lepsze do celów badawczych niż przykłady o genezie naturalnej, a na zarzut, że dane językowe powinny być autentyczne i udokumentowane, można było usłyszeć ironiczną odpowiedź w stylu: „Proszę mój przykład włączyć do korpusu, od razu będzie udokumentowany”. Ten przerost instrospekcjonizmu, charakterystyczny dla lingwistyki lat 60., 70. i 80. XX wieku, mógł mieć swoje źródło w psychologicznym mechanizmie racjonalizacji: językoznawcy tamtego czasu, nie mając łatwego dostępu do wystarczająco dużych zasobów danych, zaczęli deprecjonować to, co było trudno dostępne, a wychwalać to, co było w ich zasięgu. Czynili tak, aby nadać sens swoim działaniom (Bańko 2001: 26–29, Bańko 2012: 286).

Rosnąca ilość łatwo dostępnych danych i narzędzi do ich analizy sprawiła, że wielu dawnych stronników lingwistyki instrospekcyjnej z czasem zaczęło doceniać dane empiryczne, a nawet głosić poglądy zaprzeczające wcześniej przez nich głoszonym. Obfitość danych – pozwalająca mówić wręcz o „nadmiarze” danych, z którymi trudno sobie poradzić – rodzi jednak problemy, do których chciałbym się tutaj odnieść: po pierwsze, dane różnego rodzaju mogą być ze sobą niezgodne; po drugie, obfitość danych, zwłaszcza cyfrowych, wymaga od badaczy nowych kwalifikacji i jest wyzwaniem dla dydaktyki akademickiej¹.

¹ Nadmiar, o którym piszę, nie jest pozorny. Bierze się z dysproporcji między ilością danych obecnie dostępnych a możliwościami ich analizy, zwłaszcza gdy badacze próbują stosować tradycyjny warsztat filologiczny w nowej sytuacji. Pisząc „nadmiar”, nie używam więc cudzysłowu na znak dystansu ani ironii, nie jest to też cudzysłów cytacyjny, gdyż nie nawiązuję nim do żadnej wcześniejszej publikacji. Warto jednak dodać, że o nadmiarze danych i szerzej – o nadmiarze informacji wiele pisano, a samo pojęcie (ang. *information overload*) stało się popularne dzięki głośnej książce Alvina i Heidi Tofflerów *Szok przyszłości* (ang. *Future shock*, 1970).

Zajmijmy się pierwszą kwestią. Problem niezgodności danych można rozwiązać najprościej poprzez decyzję, aby badania opierać wyłącznie na jednym źródle. Postawa taka – rodzaj metodologicznego puryzmu – wydaje się ryzykowna, ale chyba nie jest pozbawiona podstaw, skoro w programie LXXV Zjazdu PTJ znalazł się referat, który jej bronił (dr Iwony Dembowskiej-Wosik). Poprzestając na jednym źródle danych, unikamy trudnych pytań, które pojawiają się, gdy dopuścimy wykorzystanie wielu źródeł, ale ponosimy ryzyko niedostrzeżenia czegoś, co istotnie zmieniłoby obraz rzeczy ukształtowany na takiej podstawie.

W niedawno ukończonym projekcie badawczym APPROVAL (<http://www.approval.uw.edu.pl>, zob. też Bańko i in. 2016) porównywaliśmy wyrazy zapożyczone i ich rodzime synonimy przy użyciu trzech metod: po pierwsze, za pomocą szczegółowych analiz opartych na danych słownikowych i tekstowych różnego rodzaju (w pewnym stopniu też na danych ikonograficznych i dokumentach życia społecznego); po drugie, za pomocą ankiet; po trzecie, za pomocą kwerend w przestrzeni semantycznej zbudowanej na Narodowym Korpusie Języka Polskiego, implementującej pewien model znaczenia leksykalnego należący, ogólnie biorąc, do semantyki dystrybucyjnej (szczegóły znajdują się w witrynie internetowej projektu i w publikacjach tam wymienionych). Nie zauważyliśmy, aby wyniki osiągnięte różnymi metodami były sprzeczne, ale zdarzało się, że tendencja ujawniona za pomocą jednej metody nie znajdowała potwierdzenia ani zaprzeczenia w wynikach osiągniętych inną metodą. Można widzieć w tym argument za tym stosowaniem różnych metod i różnych danych: gdy jedno z nich są mało diagnostyczne, inne okazują się bardziej przydatne.

Ten sam wniosek nasuwa analiza ok. 30 prac licencjackich i magisterskich przygotowanych ostatnio pod moim kierunkiem, opartych na metodologii wypracowanej w projekcie APPROVAL: w analizie porównawczej synonimów dane pochodzące z różnych źródeł nie są, a przynajmniej nie były dotychczas sprzeczne, mają jednak nierówną moc diagnostyczną, gdyż różnice między badanymi wyrazami raz lepiej widać w danych jednego rodzaju, innym razem lepiej w danych innego rodzaju. Na przykład jeden rzut oka na obrazy zindeksowane przez wyszukiwarkę Google z wyrazami *spektakl* i *przedstawienie* ujawnia, że spektakle są zwykle profesjonalne, a przedstawienia – amatorskie i szkolne. Analiza kolokacji, z różnych powodów bardziej wiarygodna jako podstawa porównywania synonimów, potwierdza ten wniosek, ale wymaga więcej pracy (stosunkowo najłatwiej osiągnąć cel za pomocą bazy danych HASK, obejmującej kolokacje z NKJP i wyposażonej w narzędzia ułatwiające porównywanie kolokacji różnych wyrazów, zarówno w formie tabelarycznej, jak i graficznej, zob. Pęzik 2013).

Zróżnicowanie danych wykorzystywanych w projekcie APPROVAL i w innych badaniach nad słownictwem zachęca do próby ich uporządkowania. Typologia, którą przedstawię, dotyczy wyłącznie danych użytecznych w badaniach leksykalnych. Oparta będzie na dwóch kryteriach, które można wyrazić za pomocą pytań: po pierwsze, czego dane dotyczą? po drugie, skąd pochodzą?

Ze względu na pierwsze kryterium należy odróżnić dane stricte językowe, dotyczące struktury i cech funkcjonalnych wyrazu (nie wyłączając jego frekwencji, ety-

mologii ani historii) od danych zasadniczo pozajęzykowych, np. informacji odnoszących się do tła historycznego, społecznego i kulturowego. Piszemy „zasadniczo”, gdyż granica między danymi jednego i drugiego typu nie jest wyraźna i nie tylko znaczenie desygnacyjne, ale zwłaszcza znaczenie konotacyjne wyrazu w wielkim stopniu polega na danych, które pochodzą spoza języka. Zauważmy, że dane pozajęzykowe mogą mieć charakter werbalny lub niewerbalny, do tych ostatnich należy m.in. ikonografia.

Ze względu na pochodzenie należy wyróżnić przede wszystkim dane systemowe, czyli należące do systemu języka: są to m.in. znaczenia przenośne wyrazów, utarte związki frazeologiczne, wyrazy pochodne, w sumie – dane, na których opiera się badanie konotacji semantycznych i językowych obrazów świata (np. Tokarski 1988, Wierzbicka 1993). Do danych systemowych można zaliczyć też wyrazy swoiste, takie jak *rzeń*, *galop*, *stajnia* lub *podkowa*, gdyż samo ich istnienie wskazuje na ważne dla mówiących cechy desygnatu, w tym wypadku konia. Do systemowych należą ponadto przynajmniej niektóre dane z zakresu prozodii i tzw. mowy ciała, towarzyszące treściom werbalnym.

Dane niesystemowe można podzielić na szereg mniejszych grup. Pierwszą stanowią dane metajęzykowe, czerpane z takich źródeł jak słowniki, poradniki językowe i inne publikacje językoznawcze. Tu należą też językowe bazy danych, np. SłowoSieć, czyli polska wersja WordNetu, Narodowy Fotokorpus Języka Polskiego lub przestrzeń semantyczna zbudowana w projekcie APPROVAL.

Druga grupa danych pochodzących spoza samego języka to dane elicytacyjne, pozyskiwane dzięki ankietom i badaniom eksperymentalnym. Częstsze są w warsztacie psychologów i socjologów niż humanistów, ale właśnie dlatego w humanistyce stanowią pewien problem (o czym dalej).

Trzecia grupa to dane instrospekcyjne, które – ze względu na psychologiczny aspekt języka jako obiektu badań – są wciąż wykorzystywane w pewnym zakresie, np. do falsyfikowania hipotez semantycznych poprzez sprowadzanie ich do sprzeczności.

Czwarta grupa to dane tekstowe, dawniej pochodzące z kartotek słownikowych, dziś częściej z korpusów i innych zasobów Internetu, np. ze stron WWW, bibliotek cyfrowych, archiwów prasowych online i internetowych katalogów bibliotecznych (katalogi dają wgląd w to, jak słowa funkcjonują w tytułach książek). Godne uwagi, że dane tekstowe coraz częściej wykorzystywane są zamiast metajęzykowych, przy czym czynią to zarówno językoznawcy, jak i laicy. Tradycyjnie słowniki były źródłem danych – by tak rzec – pierwszego kontaktu, obecnie zaś, wahając się między *z* a *rz* w zapisie jakiegoś słowa, niejedna osoba wykona dwie kwerendy na stronach WWW, porówna wyniki i wyciągnie wnioski, że poprawna jest forma częstsza. Ktoś bardziej świadomy zajrzy w tej sytuacji raczej do słownika ortograficznego, ale gdy zadanie polega na ustaleniu np. typowej łączliwości jakiegoś słowa, wybór korpusu językowego jako pierwszego źródła danych jest racjonalny, gdyż w nim prędzej znajdzie się odpowiedź niż w słownikach.

Na tym lista się nie kończy, gdyż oprócz danych tekstowych – w szczególności werbalnych przekazów kulturowych, utrwalonych w pamięci społecznej, takich jak

wiersze, piosenki, dowcipy – źródłem danych istotnych dla opisu znaczeń mogą być nie mniej ważne i porównywalne z nimi przekazy niewerbalne, np. rysunki, obrazy, filmy, a także społeczne rytuały towarzyszące słowom, np. zachowania etykietałne, obrzędy świąteczne, przedstawienia karnawałowe, zabawy dziecięce. W zasadzie każdy rodzaj ludzkiej działalności i każdy jej efekt, jeśli tylko ma charakter znakowy, jest potencjalnie istotny dla opisu znaczenia słów.

Klasyfikować dane można na więcej sposobów, w nauce znany jest np. podział na dane ilościowe i jakościowe. Powyższy przegląd nie wyczerpuje więc możliwości, ale uświadamia, że samo porównanie wyników badań opartych na danych różnego rodzaju może być trudne, nie mówiąc już o ocenie stopnia ich zgodności. Jest jednak też drugi problem, wspomniany wyżej: obfitość danych, zwłaszcza danych cyfrowych, wymaga od badaczy nowych kwalifikacji i stanowi wyzwanie dla dydaktyki akademickiej. Nieobce nam wrażenie „nadmiaru” danych, czasem poczucie bezsilności wobec ich masy, bierze się stąd, że stare przyzwyczajenia przenosimy na nowy grunt i nowe dane próbujemy badać starymi metodami. Przyjrzymy się temu problemowi nieco dokładniej.

Do niedawna warsztat badawczy filologa obejmował wykonywanie kwerend bibliotecznych, sporządzanie bibliografii, znajomość dostępnych słowników i innych publikacji referencyjnych oraz umiejętność posługiwania się nimi. Do tego, w zależności od dziedziny, a więc nie zawsze, dochodziły inne umiejętności, np. w zakresie analizy mowy lub prowadzenia badań terenowych metodą kwestionariuszową. Dziś natomiast powinniśmy oczekiwać, aby uczestnicy seminariów językoznawczych potrafili posługiwać się nowymi źródłami danych i nowymi narzędziami służącymi do prowadzenia badań, takimi jak wyszukiwarki danych, narzędzia wspomagające gromadzenie danych oraz pakiety statystyczne. W programach kształcenia językoznawczego powinny znaleźć się więc zajęcia poświęcone metodologii badań ankietowych, wyszukiwaniu danych w zasobach elektronicznych, a nawet tworzeniu takich zasobów na własne potrzeby, np. zbudowaniu własnego korpusu. Niektóre z tych kwalifikacji daleko odbiegają od tradycyjnego modelu kształcenia filologów.

Zwróć tu uwagę na dwa problemy szczegółowe, które nie wyczerpują zagadnienia. Pierwszy dotyczy konstruowania ankiet oraz sposobu prezentacji i oceny ich wyników. Mało kto uczy humanistów, jak to robić, w związku z czym nie tylko w pracach licencjackich i magisterskich, ale nawet w doktorskich spotyka się dane czerpane z ankiet źle skonstruowanych lub źle przeprowadzonych, opisane nie dość dokładnie i zbyt łatwo uogólnione, najwidoczniej bez świadomości, że wyniki ankiety wymagają testów statystycznych w celu sprawdzenia, które z zaobserwowanych różnic są istotne, a które mogą być przypadkowe. Jedyna zmiana ostatnio polega na tym, że studenci chętnie przeprowadzają ankiety internetowe za pomocą narzędzi Google, nie wpływa to jednak na jakość ankiet, a jedynie ułatwia pozyskanie dużej liczby wyników (przy bardzo ograniczonej kontroli, co jest odrębnym problemem)².

² Przyrost danych lingwistycznych zawdzięczamy przede wszystkim korpusom, które są znakiem czasu. Ankiety nie stanowią *signum temporis*, ale dzięki nowym narzędziom do ich przeprowadzania oraz pakietom statystycznym, które pomagają w opracowaniu i interpretacji wyników, stały się źródłem danych dziś łatwiej dostępnych niż w przeszłości i dlatego mają swój udział w tworzeniu „nadmiaru” danych, por. przypis 1.

Drugi przykład na to, że za przyrostem danych nie nadążają kompetencje ich użytkowników, dotyczy wyszukiwania w korpusach internetowych. Obsługi wyszukiwarki PELCRA w NKJP (Pęzik 2012), mającej interfejs okienkowy, łatwo się nauczyć, ale zakres filtrowania wyników jest w niej ograniczony do tego, co przewidziano na etapie jej budowy. Więcej możliwości oferuje wyszukiwarka IPI PAN, udostępniająca kwerendy oparte na formalizmie tzw. wyrażeń regularnych, jest jednak trudniejsza w użyciu i u osób nienawykłych do matematycznych formalizmów oraz do pracy z wierszem poleceń może wywoływać odruch niechęci (w rezultacie zdarza się, że osoby takie ręcznie selekcionują i porządkują setki, a nawet tysiące wyników kwerendy, choć zadanie takie dałoby się przynajmniej częściowo zautomatyzować). Rozwiązaniem problemu mogłby być interfejs okienkowy do wyszukiwarki IPI PAN: kilka lat temu na zlecenie prof. Marka Łazińskiego interfejs taki opracował Jarosław Strojek, wdrożono go m.in. w Korpusie Języka Młodzieży XXI wieku, ale ponieważ nie został udostępniony na stronie NKJP, szerszego rozpowszechnienia nie uzyskał.

Ogólnie biorąc, aby wygodnie korzystać z dostępnych dziś danych językowych, potrzebne jest pewne obycie informatyczne i znajomość użytecznych narzędzi, innych niż programy biurowe typu Microsoft Word lub Excel. Czasami przydaje się też znajomość podstaw programowania.

Na pytanie, czy w takim razie należy uczyć humanistów statystyki, wyrażeń regularnych, a może nawet programowania (np. w języku R), nie ma łatwiej odpowiedzi, skłaniałbym się jednak ku temu, że owszem, należy takie umiejętności przekazywać stopniowo, na początku na kursach obieralnych, później chyba też na obligatoryjnych, przynajmniej w ramach niektórych specjalności lub specjalizacji akademickich. Osobna kwestia to kto miałby to robić: zewnętrzny specjalista świadczący zajęcia usługowe na takiej zasadzie jak lektor języka obcego, czy własny pracownik, np. polonista, który zdobył potrzebne kwalifikacje na kursach, warsztatach, regularnych studiach lub w inny sposób? Zauważmy, że skoro np. prace dyplomowe polonistów nie powstają pod kierunkiem lektorów, to nie będą też powstawać pod kierunkiem specjalistów z innej dziedziny, zapraszanych na zajęcia z innej jednostki uczelnianej. Wcześniej czy później kadra dydaktyczna kierunków humanistycznych musi sama posiłkować kwalifikacje niezbędne do korzystania z nowych danych i nowych narzędzi.

Na razie można postępować metodą małych kroków: wprowadzać do programów uniwersyteckich zajęcia dotyczące nowych danych językowych i sposobów korzystania z nich, przygotowywać przewodniki i inne publikacje dydaktyczne, pomocne w nabywaniu nowych umiejętności (np. Kwaśnicka-Janowicz, Pałka 2017), a także oferować szkolenia osobom zainteresowanym. Działania takie już trwają i na szczęście istnieje zrozumienie dla potrzeby wsparcia informatycznego w dziedzinie badań humanistycznych, czego wyrazem jest np. powstanie i działalność Laboratorium Cyfrowego Humanistyki UW.

Bibliografia

- APPROVAL: Adaptacja, percepcja i recepcja wyrazów obcych w języku polskim i czeskim. Uwarunkowania językowe, psychologiczne i historyczno-kulturowe, <http://www.approval.uw.edu.pl>.
- BAŃKO Mirosław (2001): *Z pogranicza leksykografii i językoznawstwa*. – Warszawa: Uniwersytet Warszawski, Wydział Polonistyki.
- BAŃKO Mirosław (2012): NKJP w oczach leksykografa. – [w:] Adam PRZEPIÓRKOWSKI i in. (red.): *Narodowy Korpus Języka Polskiego*. – Warszawa: PWN, 283–290.
- BAŃKO Mirosław, SVOBODOVÁ Diana, RĄCZASZEK-LEONARDI Joanna, TATJEWSKI Marcin (2016): *Nie całym obce. Zapożyczenia wyrazowe w języku polskim i czeskim*. – Warszawa: PWN.
- Korpus Języka Młodzieży XXI wieku, <http://kjm.polon.uw.edu.pl/>.
- KURCZ Ida, LEWICKI Andrzej, SAMBOR Jadwiga, WORONCZAK Jerzy (1974–1977): *Słownictwo współczesnego języka polskiego. Listy frekwencyjne*. – Warszawa: PAN.
- KWAŚNICKA-JANOWICZ Agata, PAŁKA Patrycja (2017): *Przewodnik po elektronicznych zasobach językowych dla polonistów (słowniki, kartoteki, korpusy, kompendia)*. – Kraków: Wydawnictwo JAK, Towarzystwo Miłośników Języka Polskiego, http://tmjp.pl/images/pdf/przewodnik_po_elektronicznych_zasobach.pdf.
- Laboratorium Cyfrowe Humanistyki UW, <http://lach.edu.pl/>.
- LAROSE Daniel T. (2006): *Odkrywanie wiedzy z danych. Wprowadzenie do eksploracji danych*. – Warszawa: PWN.
- Narodowy Fotokorpus Języka Polskiego, <http://nfjp.pl/>.
- Narodowy Korpus Języka Polskiego, <http://nkjp.pl>.
- NKJP, zob. Narodowy Korpus Języka Polskiego.
- OGRODNICZUK Maciej, DERWOJEDOWA Magdalena, ŁAZIŃSKI Marek, PĘZIK Piotr (2017): Narodowy Korpus Języka Polskiego – co dalej? – *Prace Filologiczne* LXXI, 237–246.
- PĘZIK Piotr (2012): Wyszukiwarka PELCRA dla danych NKJP. – [w:] Adam PRZEPIÓRKOWSKI i in. (red.): *Narodowy Korpus Języka Polskiego*. – Warszawa: PWN, 253–273.
- PĘZIK Piotr (2013): Paradygmat dystrybucyjny w badaniach frazeologicznych. Powtarzalność, reprodukcja i idiomatyzacja. – [w:] Piotr STAŁMASZCZYK (red.): *Metodologie językoznawstwa. Ewolucja języka, ewolucja teorii językoznawczych*. – Łódź: Wydawnictwo UŁ, 143–160.
- PRZEPIÓRKOWSKI Adam, BAŃKO Mirosław, GÓRSKI Rafał, LEWANDOWSKA-TOMASZCZYK Barbara, red. (2012): *Narodowy Korpus Języka Polskiego*. – Warszawa: PWN.
- SłowoSieć, <http://plwordnet.pwr.wroc.pl/wordnet/>.
- STUBBS Michael (1996): *Text and Corpus Analysis*. – Oxford: Blackwell Publishers.
- TOKARSKI Ryszard (1988): Konotacja jako składnik treści słowa. – [w:] Jerzy BARTMIŃSKI (red.): *Konotacja*. – Lublin: UMCS, 35–54.
- WIERZBICKA Anna (1993): Nazwy zwierząt. – [w:] Jerzy BARTMIŃSKI, Ryszard TOKARSKI (red.): *O definicjach i definiowaniu*. – Lublin: UMCS, 251–267.

doi: <http://dx.doi.org/10.17651/bptj.74.2>

MIROŚŁAWA PODHAJECKA

Uniwersytet Opolski, Opole

Google Books jako korpus językowy

Streszczenie

Artykuł poświęcony jest omówieniu Google Books, dostępnej przez Internet biblioteki wirtualnej, obejmującej skany 30 milionów książek. Jest to aktualnie najbogatsze na świecie źródło danych tekstowych w postaci cyfrowej. Zbiory Google Books można nazwać korpusem, ale zasadniczo różnią się one od tradycyjnych korpusów językowych. Kłopoty klasyfikacyjne wynikają z konkretnych ograniczeń, z jakimi trzeba się zmierzyć w trakcie badań. Między innymi część źródeł to wersje pełnotekstowe, a część – wersje z ograniczonym podglądem, dane bibliograficzne są nierzadko błędne, a jakość optycznego rozpoznawania tekstu, zwłaszcza w przypadku starszych tekstów, jest daleka od doskonałości. Referat omawia krótko problemy badawcze dotyczące Google Books.

Słowa kluczowe: Google Books, korpus, analiza, problemy badawcze

Google Books as a language corpus

Abstract

This article concerns Google Books, a digital library available on the Internet, which contains scans of 30 million books. At present, it is the largest source of textual data in digital format worldwide. Google Books may be called a corpus, but it is markedly different from traditional language corpora. Classification difficulties arise from specific limitations encountered during research. Among other things, some sources are available as full texts, while others offer limited preview; bibliographic metadata are often wrong; and the quality of optical character recognition is far from perfect, especially when applied to older texts. The article briefly discusses research problems involved in using Google Books.

Keywords: Google Books, corpus, analysis, research problems

Wprowadzenie

Niniejszy artykuł poświęcony jest charakterystyce Google Books, internetowej biblioteki wirtualnej, udostępniającej użytkownikom skany książek w postaci plików

PDF¹. Jest to dziś najbogatsze na świecie źródło tekstów w postaci cyfrowej, które określa się czasem mianem współczesnej „Biblioteki Aleksandryjskiej” [M.P.] (Roberts 2017). Zasoby Google Books są przechowywane na serwerach firmy Google, a głównym celem projektu, przynajmniej według oficjalnych przekazów (Brin 2009), jest zachowanie dla potomności dziedzictwa kulturowego ludzkości.

Repozytorium Google Books różni się znacząco od tradycyjnych korpusów. Większość korpusów stanowią bazy zamknięte, których zawartość się nie zmienia, natomiast pozostałe są bazami otwartymi, obejmującymi tzw. korpusy monitorujące (zob. Waliński 2005: 29–33), do których dane zbiera się w sposób ciągły². Status Google Books jest w tym względzie niejasny – nie wiadomo, czy mamy do czynienia z bazą zamkniętą czy otwartą – ale zasoby te stanowią szczególną bazę danych ze względu na modyfikowaną politykę dostępności³. Być może z tej przyczyny jeszcze do niedawna nie uwzględniano ich w opracowaniach z zakresu językoznawstwa korpusowego⁴.

Cel tego artykułu jest dwojaki. Po pierwsze, chcę odpowiedzieć na pytanie, czy Google Books można nazwać korpusem, a jeśli tak – dlaczego. Po drugie, zamierzam krótko omówić zarówno możliwości badawcze, jak i wady tych zasobów. Szczególną uwagę poświęcę tym drugim, ponieważ praca z materiałami Google Books niesie ze sobą trudności, których świadomi powinni być badacze⁵.

Korpus w badaniach językoznawczych

Choć językoznawstwo korpusowe rozwija się już od prawie stu lat, a pierwszy korpus elektroniczny – Brown Corpus – powstał ponad pięćdziesiąt lat temu (Landau 2001: 273, 278), w ramach analizy Google Books warto ponownie przyjrzeć się definicji terminu *korpus*. Według Piotrowskiego (2003: 146), jest to

zbiór tekstów zapisanych w postaci cyfrowej i stanowiących reprezentatywną próbkę danego języka. Korpus różni się od dowolnego zbioru tekstów tym, że stanowi pewną przemyślaną całość. Cechą wyróżniającą korpusów komputerowych jest też ich rozmiar.

Reprezentatywność oznacza, że zebrane teksty stanowią próbkę charakterystyczną dla języka w danym okresie. Łatwo ją określić dla korpusów języka współczesnego, takich jak British National Corpus⁶, natomiast w mniejszym stopniu dla korpusów

¹ Projekt został zainicjowany w 2004 roku jako Google Print, ale już rok później zaczął funkcjonować pod nazwą Google Books.

² Korpus Monco (<http://monco.frazeo.pl/>), na przykład, powiększa się codziennie o 8 milionów wyrazów.

³ Zob. podrzdział 1. *Zmienne dane ilościowe*.

⁴ Przykładowo, tom *Corpus linguistics* (2009) pod redakcją Anke Lüdelinga i Merji Kytö ani razu nie wspomina zasobów Google Books.

⁵ Wszystkie cytaty z anglojęzycznych źródeł przetłumaczone zostały przeze mnie i opatrzone są moimi inicjałami w kwadratowych nawiasach.

⁶ Nie zawsze określenie zawartości korpusu współczesnego jest jednak bezproblemowe (zob. Górski 2003).

historycznych. Równie ważnym parametrem jest równowaga pomiędzy różnymi gatunkami tekstów, dzięki której korpus tworzy „przemyślaną całość”. Mówiąc inaczej, im więcej gatunków będzie reprezentowanych w korpusie, tym mniejsze prawdopodobieństwo wystąpienia idiosynkrazji badawczych⁷. Znaczenie ma też rozmiar, ponieważ tylko odpowiednio duże korpusy wykazują dystrybucję różnych cech językowych (Biber 1993: 243).

Trzeba przyznać, że Google Books trudno dopasować do powyższej definicji. Zbiory odpowiadają co prawda kryterium rozmiaru i – częściowo – formatu⁸, ale teksty zbierano bez żadnej dbałości o reprezentatywność i równowagę między próbkami. Jak ujęli to wprost Davis i Chapman (2016: 132), „pracownicy Google Books zeskanowali wszystko, co znaleźli w kilkunastu dużych bibliotekach uniwersyteckich” [M.P.]. Z punktu widzenia metodologii językoznawstwa korpusowego jest to czysta herezja. Okazuje się jednak, że dziś nie jest to wcale tak oczywiste, jak wcześniej wydawało się językoznawcom.

Rozmiar jest bowiem kryterium najważniejszym. Mayer-Schönberger i Cukier (2013: 19) sugerują, że dotychczasowe metody badań w świecie „małych danych” nie mają zastosowania do świata „wielkich danych”. Dodają także:

wielkie dane wpływają na to, jak rozumiemy i badamy świat. W świecie małych danych często stawialiśmy hipotezy dotyczące funkcjonowania świata, które następnie próbowaliśmy dowodzić poprzez zbieranie i analizowanie danych. W przyszłości nasze rozumienie problemów będzie związane w większym stopniu z obfitością danych niż hipotez [M.P.] (Mayer-Schönberger, Cukier 2013: 70).

Ciekawe badania na materiale Google Books przeprowadzili Davies i Chapman (2016: 131–152). Ich zdaniem ten korpus o niejednolitej strukturze, skompilowany w sposób całkowicie przypadkowy, przedstawia w zasadzie te same wyniki co korpus w pełni zrównoważony i reprezentatywny, w tym przypadku Corpus of Historical American English (COHA). Ustalili przy tym nie tylko korelację pomiędzy wystąpieniami wybranych przypadkowo wyrazów (np. *teenager*) i konstrukcji gramatycznych (np. *you needn't worry*) w Google Books i COHA, ale dowiedli również, że korelacja pomiędzy nimi jest systemowa⁹. Warto zacytować zwłaszcza konkluzję, jakie wynikają z przeprowadzonej analizy. Otóż Davies i Chapman (2016: 147) doszli do następującego wniosku:

⁷ Istnieje wiele dowodów wskazujących na fakt, że frekwencja wyrazu jest ściśle powiązana z gatunkiem tekstu, w którym on występuje (Diller 2013: 63).

⁸ Przytoczona definicja korpusu zakłada, że teksty dostępne są w formacie cyfrowym. W przypadku korpusów tradycyjnych oznacza to, że teksty z różnych formatów wyjściowych zostają przekonwertowane do wspólnego formatu (np. XML), a następnie anotowane morfoskładniowo. Tymczasem zasoby Google Books dostępne są wyłącznie w formacie PDF, który nie pozwala na skuteczne wykorzystanie narzędzi korpusowych.

⁹ Autorzy wybrali wyrazy, które pojawiły się w COHA przynajmniej 15 razy. Było ich 116 630. Następnie 10% potraktowali jako próbkę badawczą (= 11 663 wyrazy), sprawdzając frekwencję każdej jednostki w poszczególnych dekadach pomiędzy rokiem 1870 i 1950. Korelacja pomiędzy COHA i Google Books wyniosła 0.51.

kryterium reprezentatywności zakłada, że powinniśmy odpowiednio „wymodelować” całą docelową populację tekstów w „świecie rzeczywistym”. Ale w przypadku, gdy mamy do własnej dyspozycji całą populację, a przynajmniej jej odpowiednio dużą część, modelowanie nie jest tak istotne [M.P.].

Jeśli przyjmiemy, że zasoby Google Books obejmują odpowiednio dużą część populacji tekstów, na co wskazują podane poniżej dane liczbowe, będzie można z powodzeniem traktować je jako korpus językowy. Jest on przy tym na tyle duży, aby zapewnić – mimo swoich niedoskonałości – różnorodność gatunkową tekstów¹⁰.

Google Books

Szacuje się, że od czasu wynalezienia druku na świecie ukazało się około 130 milionów publikacji zwartych (Skipworth 2010). W 2007 roku Marissa Mayer, ówczesna wiceprezes Google, zapowiedziała, że zostaną one zeskanowane w ciągu 10 lat (Wu 2015). W październiku 2010 roku dyrektor techniczny firmy podał informację, że Google Books obejmuje już 15 milionów książek, pochodzących ze 100 krajów i opublikowanych w ponad 400 językach (Crawford 2010), a w 2015 roku liczbę skanów oszacowano na ponad 30 milionów (Lewis 2015). Dziś znalezienie informacji na temat aktualnego stanu projektu jest trudne (zob. Rosenberg 2017), stąd należy założyć, że zbiory obejmują wciąż tę samą liczbę zeskanowanych druków.

Książki Google Books, które można przeszukiwać w 20 wersjach językowych¹¹, pochodzą z dwóch źródeł: projektu bibliotecznego, za pośrednictwem którego firma Google nawiązała współpracę z 40 bibliotekami z całego świata i konsorcjum 12 amerykańskich uniwersytetów badawczych (Fischer 2007), oraz programu partnerskiego, w którym uczestniczy 35 tysięcy wydawców z całego świata, przekazujących swoje publikacje do umieszczenia ich w formie podglądu¹².

Wyszukiwarka Google Books pozwala na wyszukiwanie proste i zaawansowane. W drugim przypadku, po zastosowaniu opcji Narzędzia, wyszukiwanie można prowadzić we wszystkich zasobach tekstowych, w książkach z ograniczonym podglądem lub wyłącznie w wersjach pełnotekstowych. Wyniki wyszukiwania umożliwiają dostęp do materiałów w pełnym widoku, w ograniczonym podglądzie, w niewielkim fragmencie tekstu lub w podglądzie niedostępnym¹³.

Interfejs użytkownika pozwala na zastosowanie trzech filtrów. Są to: typ dokumentu (np. książka lub czasopismo, ale już nie gatunki tekstowe); daty publikacji (wiek XXI, XX i XIX, a także wybrane zakresy dat), a także kolejność sortowania wyników (według trafności lub chronologii). Konsekwencją zastosowania którego-

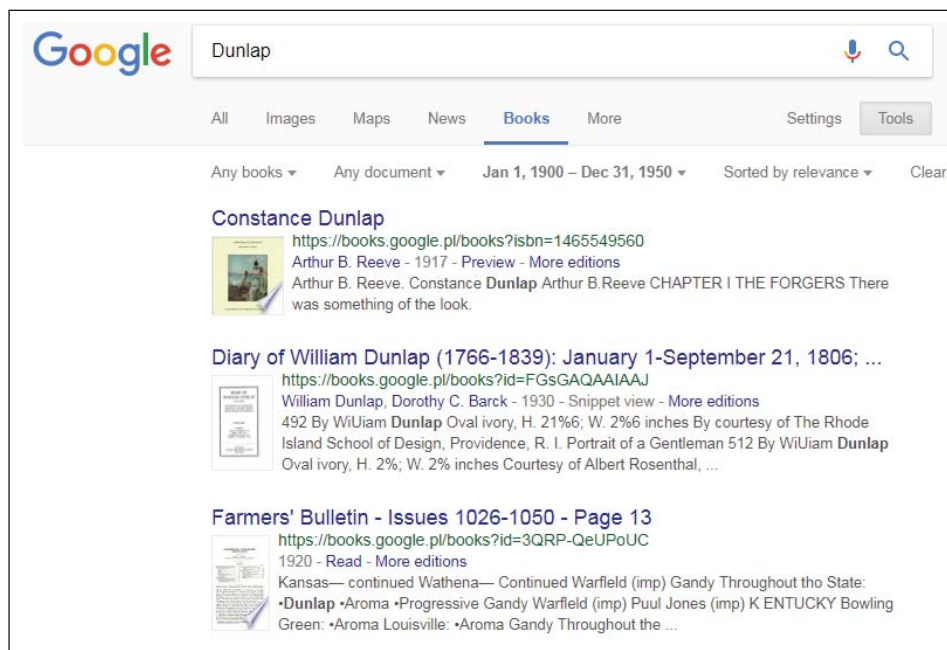
¹⁰ Należy zauważyć, że definicja terminu *korpus* ewoluuje: korpusem coraz częściej określa się ogromne zbiory tekstów, m.in. zasoby internetowe (np. Kilgariff 2003), zasoby bibliotek cyfrowych (np. Graliński 2013), a także elektroniczne zasoby słownikowe (np. Żmigrodzki 2005).

¹¹ W poprzedniej wersji możliwe było wyszukiwanie w tekstach dla 46 języków.

¹² Publikacje takie często nie mają numeracji stron, co utrudnia m.in. ich cytowanie (zob. Parks 2014).

¹³ Podgląd niedostępny dotyczy zwykle książek, które nie zostały jeszcze zeskanowane i dla których podaje się wyłącznie dane bibliograficzne.

kolwiek filtru jest jednak brak informacji o częstości wystąpień wyszukiwanej jednostki (zob. Rys. 1).



Rys. 1. Wyniki wyszukiwania zaawansowanego z zastosowaniem filtru (okres: 1900–1950).

Korpus Google Books można wykorzystać na wiele sposobów. Umożliwia on jakościowe analizy językowe, takie jak porównanie wystąpień synonimów (np. Diller 2013), analizy ilościowe, które mogą mieć zastosowanie w praktyce pedagogicznej czy translatorskiej (np. Green 2014)¹⁴, a także m.in. poszukiwania zmian leksykalnych, autorstwa cytatów czy poświadczeń tekstowych (np. Podhajecka 2015)¹⁵.

Zdaniem Mays (2015), nieocenioną zaletą pracy z zasobami Google Books jest wygoda. Omawiając wyniki swoich poszukiwań, stwierdza ona:

[...] przed erą Google Books takie badania wymagałyby wyprawy do innych miast i spędzenia kilku tygodni na przeglądaniu zasobów archiwalnych i sądowych. Moje badania na materiale Google Books [...] były szybsze i pełniejsze, niż gdybym wyjechała na kwerendę [M.P.].

¹⁴ Przykładowo, chcąc ustalić najwłaściwszy ekwiwalent anglojęzyczny dla frazy *nieoczekiwana zmiana pogody*, możemy posłużyć się frekwencją wariantów ("unexpected change in the weather" F 966 / "unexpected weather change" F 113 / "unexpected change of the weather" F 8).

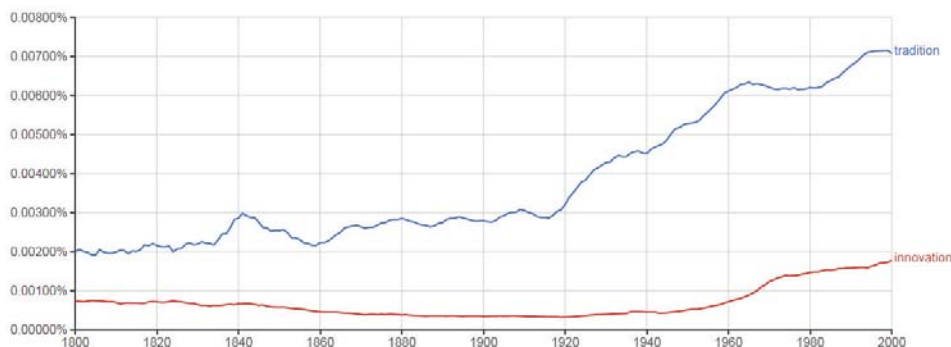
¹⁵ Kilka lat temu szukałam poświadczeń tekstowych, antydatujących wyrazy hasłowe w angielskim słowniku historycznym *Oxford English Dictionary* (OED3). Mimo że redaktorzy słownika dysponują wieloma bazami danych, udało się mi się znaleźć wcześniejsze przykłady użycia dla 50% próbki, co świadczy o ogromnym potencjale badawczym Google Books (zob. Podhajecka 2011).

Korpus Google Books jest dostępny przez Internet, co oznacza, że można z niego korzystać niezależnie od tego, gdzie się znajdujemy. Amerykański historyk Anthony Grafton (2012: 234) dowodzi, na przykład, że „Google Books każdego dnia udowadnia nam coraz bardziej, że z równym powodzeniem można zajmować się różnymi aspektami francuskiej myśli i literatury, mieszkając w Nowym Jorku lub Paryżu, i to o wiele bardziej efektywnie”.

Istnieje wiele repozytoriów internetowych, publikujących treść całych książek i czasopism, np. Project Gutenberg, HathiTrust, Internet Archive¹⁶ czy Early English Books Online (EEBO)¹⁷, które świetnie wpisują się w szeroko rozumianą definicję korpusu. Należy jednak pamiętać, że nie każda biblioteka wirtualna (np. Library of Congress American Memory¹⁸) posiada funkcję wyszukiwania w pełnej bazie tekstów.

Google Books Ngram Viewer

Jak twierdzi Levy (2011: 32, 67), celem projektu Google Books było nie tyle udostępnienie użytkownikom milionów książek, co wyekstrahowanie zgromadzonej w nich wiedzy, stanowiącej dorobek ludzkości. Jednym z flagowych narzędzi służących temu celowi w zakresie badań językowych (raczej niż językoznawczych) jest serwis Ngram Viewer, umożliwiający tworzenie wykresów graficznych na podstawie danych frekwencyjnych jednostek leksykalnych na przestrzeni 200 lat (1800–2000) (zob. Rys. 2).



Rys. 2. Porównanie wystąpień wyrazów *tradition* i *innovation* (Google Books Ngram Viewer).

Ngram Viewer bazuje na blokach tekstu: jeden wyraz (np. *economy*) to w terminologii wyszukiwarki 1-gram, sekwencja dwóch wyrazów (np. *innovative economy*)

¹⁶ Porównanie zasobów Google Books, HathiTrust i Archive.org przedstawiono na stronie <https://blog.gdeltproject.org/the-world-as-seen-through-books-comparing-the-internet-archive-hathitrust-and-google-books-ngrams-collections/>.

¹⁷ W odróżnieniu od wielu innych repozytoriów dostęp do EEBO wymaga opłacenia subskrypcji.

¹⁸ Jest to repozytorium cyfrowe Biblioteki Kongresu w Waszyngtonie, obejmujące szereg podzbiorów, które należy przeszukiwać indywidualnie.

¹⁹ Pięć milionów książek stanowi około 16% zasobów Google Books (Michel et al. 2011: 176).

to 2-gram itd. Google Ngram wyszukuje sekwencje do pięciu wyrazów (5-gram) bez żadnego kontekstu. Materiał tekstowy wersji Ngram Viewer 1.0 (2009) obejmuje pięć milionów książek¹⁹, a wersji Ngram Viewer 2.0 (2012) – osiem milionów książek, charakteryzujących się wysoką jakością metadanych.

Narzędzie to stało się podstawą tzw. kulturomiki (ang. *culturomics*) – ilościowej analizy terminów kulturowych w perspektywie diachronicznej (Michel et al. 2011: 176). Nie wszyscy jednak są zgodni co do jakości otrzymanych wyników. Zimmer (2012) podkreśla, że Ngram Viewer przetwarza gigantyczne ilości chaotycznego tekstu w uporządkowane zbiory danych językowych, a Skiena i Ward (2013: 68) uznają serwis za „fantastyczny teleskop do patrzenia w przeszłość” [M.P.]. Do obozu krytyków należą z kolei Pechenick, Danforth i Sheridan Dodds (2015), twierdzący, że „znaczna część korpusu stanowią teksty naukowe. Efektem tego jest wzrost ilości fraz typowych dla języka akademickiego, nie języka ogólnego” [M.P.].

Niezależnie od opinii krytycznych korpus, na którym bazuje Ngram Viewer 2.0 (2012), poddano anotacji morfoskładniowej dla ośmiu języków, co – zdaniem Zimmera (2012) – stanowi ważne osiągnięcie w dziedzinie przetwarzania języka naturalnego (zob. Rys. 3).



Rys. 3. Porównanie frekwencji rzeczownika *attempt* i czasownika *attempt* (Google Books Ngram Viewer).

Amerykański językoznawca Mark Davies stworzył interfejs użytkownika do korpusu Google Books, podzielonego na dwa podkorpusy: Google Books American Corpus (155 miliardów wyrazów) oraz Google Books British Corpus (34 miliardy wyrazów). Umożliwia on wyszukiwanie w tych samych zasobach tekstowych, na których bazuje Ngram Viewer. Davies (2014: 402) twierdzi, że jego interfejs ma szersze zastosowanie, ponieważ pozwala m.in. na wyszukiwanie kolokatów, ale uwagi te mają odniesienie tylko do wersji Ngram Viewer 1.0.

Problemy w korzystaniu z Google Books

Poniżej omówię krótko problemy, z jakimi można spotkać się w trakcie analizy Google Books. Litera F stanowi skrót of terminu *frekwencja*.

1. Zmienne dane ilościowe

Charakterystyczne dla Google Books jest nie tylko to, że przybywa tekstów, ale także to, że co jakiś czas zmienia się ich status w zakresie dostępności. Mówiąc dokładniej, dostęp bywa „aktywowany”, gdy książka uprzednio niedostępna staje się dostępna, lub „dezaktywowany”, gdy książka dostępna – w formacie pełnotekstowym lub z ograniczonym podglądem – staje się niedostępna. Jest to często przyczyną frustracji badaczy. Dodatkowo Google Books stosuje nierówną politykę udostępniania materiałów w trybie pełnotekstowym²⁰.

Nawet jeśli Google Books jest korpusem zamkniętym, z uwagi na zmiany w dostępności tekstów dane frekwencyjne wyszukiwanych jednostek są różne. Co ciekawe, oprócz fluktuacji danych (np. *caddis fly*) można znaleźć także przypadki stałego obniżania się częstości wystąpień (np. *Clinton*).

"caddis fly"

F 31 700 (1.02.2017)

F 31 800 (12.09.2017)

F 31 300 (16.11.2017)

"Clinton"

F 1 990 000 (3.02.2017)

F 1 820 000 (11.09.2017)

F 1 570 000 (16.11.2017)

Wyszukiwarka Google Books podaje wyniki w ułamku sekundy, ale dane liczbowe stanowią wartości przybliżone²¹. Google stosuje w tym celu złożone i często zmieniające się algorytmy wyszukiwania, które stanowią starannie chronione *know-how* firmy. Tak więc w sytuacji, gdy otrzymujemy rozbieżne wyniki wyszukiwania, należy podawać datę dostępu.

2. Problemy z obliczaniem frekwencji

Ważną cechą korpusów tradycyjnych jest lematyzacja, która polega na „redukcji różnych form fleksyjnych tego samego wyrazu do jego formy podstawowej” (Waliński 2005: 35), czyli tzw. lemmy. Ponieważ Google Books nie pozwala na wyszukanie lemm, czasem konieczne staje się „ręczne” zliczenie frekwencji poszczególnych słowoform. W zależności od potrzeb badawczych może to dotyczyć form wariantywnych (np. *cashier's check / cashier's cheque*), morfologicznych (np. *doubting Thomas / doubting Thomases*) lub historycznych (np. *tsar / czar / czar*)²². Podstawowym problemem podejścia ilościowego jest to, że wyszukiwarka umożliwia weryfikację jedynie części wyników. Przykładowo, dla frazy *doubting Thomas* Google Books podaje 72 600 wystąpień, ale wgląd uzyskujemy jedynie do 442. Niestety, nie dysponujemy żadnymi narzędziami do weryfikacji podanej frekwencji.

²⁰ Wiele książek wydanych w Stanach Zjednoczonych dostępnych jest w trybie pełnotekstowym, a w innych krajach – w trybie ograniczonym.

²¹ Dane frekwencyjne pojawiają się wraz z przysłówkiem *about* (około).

²² Określenie frekwencji wszystkich form graficznych może być przydatne z diachronicznego punktu widzenia.

"cashier's check" (F 173 000) + "cashier's cheque" (F 1 780)
 "doubting Thomas" (F 72 600) + "doubting Thomases" (F 33 200)
 "tsar" (F 491 000) + "tzar" (F 291 000) + "czar" (F 501 000)

Wiadomo, że badanie korpusowe wychodzi od formy, nie od znaczenia. Podstawową metodą w wyszukiwaniu wyrazów użytych w konkretnym znaczeniu są wyrażenia boolowskie (ciągi wyrazów z operatorem np. AND lub OR)²³.

Clinton 'Geol. A formation of the Silurian system in the central and eastern U.S; also, the corresponding epoch' [formacja sylurska w środkowych i wschodnich Stanach Zjednoczonych, a także właściwa epoka geologiczna – M.P.]

F 1 820 000 (11.09.2017)

↓

"Clinton" AND "geology" (F 155 000)
 "Clinton" "geology" OR "geological" (F 300 000)

Dunlap 'Hort. A well-known American variety of strawberry' [znana amerykańska odmiana truskawek – M.P.]

F 611 000 (11.09.2017)

↓

"Dunlap" "strawberry" (F 28 300)
 "Dunlap" "strawberries" (F 26 000)

Z praktyki wynika, że najlepiej sprawdzają się takie wyrażenia boolowskie, w których przynajmniej jeden z elementów określa pożądaną cechę semantyczną (np. "geology" w przypadku nazwy własnej *Clinton* albo "strawberry" – w przypadku *Dunlap*)²⁴.

W sytuacji, gdy nieznana pozostaje forma, pomocne będzie wyszukiwanie za pomocą wieloznacznika *, np. "to * a decision". W tym przypadku pomoże on znaleźć czasowniki, które tworzą kolokacje z podanym rzeczownikiem (np. *to make a decision, to take a decision, to draw a decision, to formulate a decision* itd.).

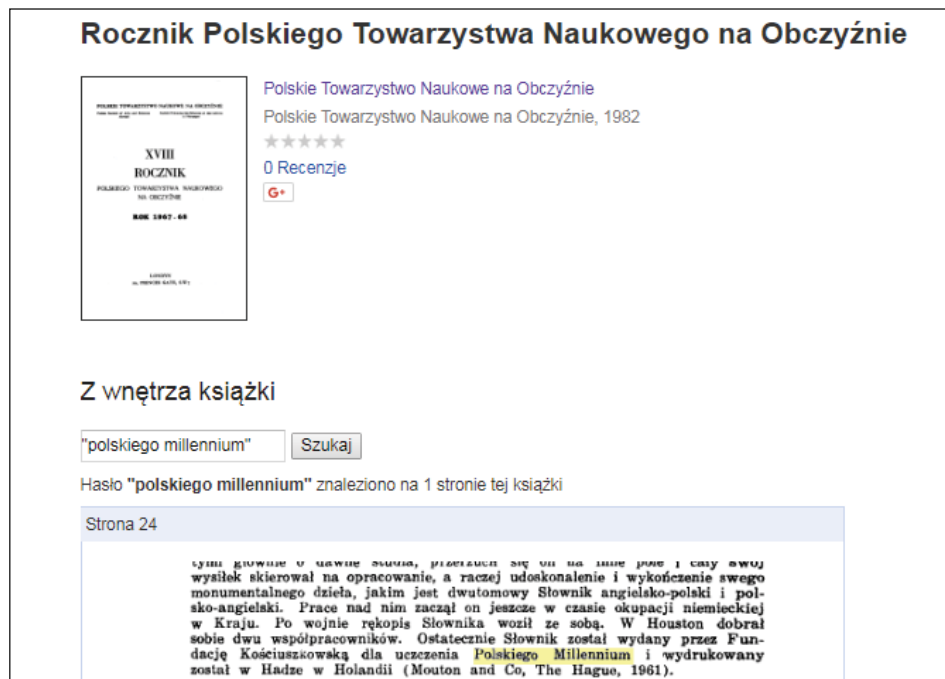
3. Niewiarygodne metadane

Książki Google Books opatrzone są informacją bibliograficzną (metadanymi). Zalicza się do nich imię i nazwisko autora, tytuł, datę wydania, nazwę wydawcy, miejsce publikacji i klasyfikację przedmiotową. Jak wynika z badań, wiele metadanych podanych w korpusie jest jednak nieprawidłowych. James i Weiss (2012) szacują skalę błędów na poziomie 36%.

²³ Operator AND jest opcjonalny, a więc indentyczne wyniki wyszukiwania zapewni wyrażenie boolowskie "Dunlap" AND "strawberry" oraz "Dunlap" "strawberry".

²⁴ Doskonale nadają się do tego celu wyrazy kluczowe z definicji słownikowych. Przedstawione w artykule anglojęzyczne definicje zaczerpnęłam ze słownika Webstera (WNID2, 1934).

Rzeczywiście, w trakcie analizy źródeł wychodzą na jaw częste błędy. Można je odkryć m.in. w datowaniu książek (np. książka *Ellis's Catalogue of Rare Books* jest datowana na rok 1728, ale jest to data założenia firmy Ellis), a także czasopism, którym notorycznie przypisuje się datę wydania oryginalnego (np. tom 155 czasopisma „The New Republic” datowano na 1946 rok, ale pojawiają się w nim odniesienia do lat 1947–1949).

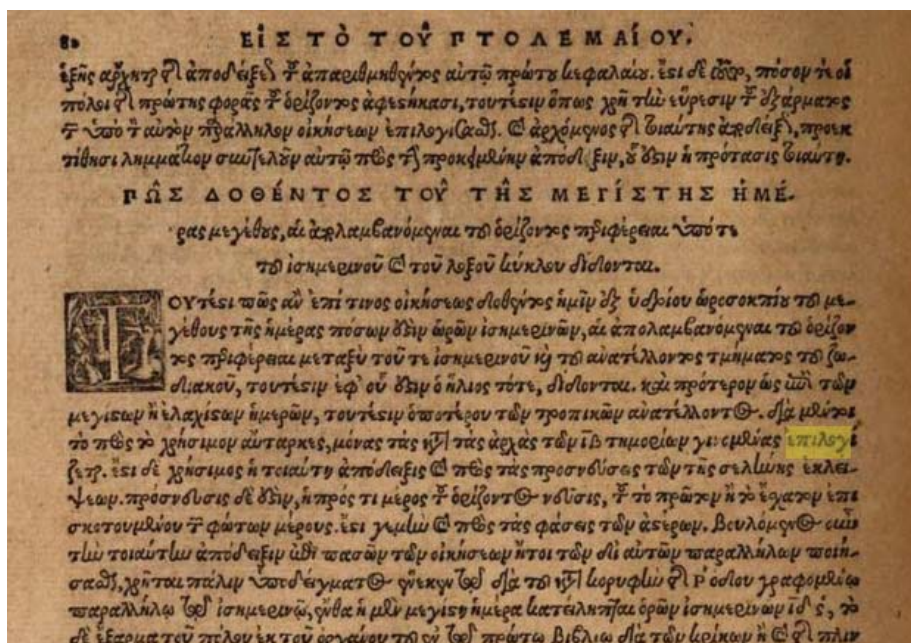


Rys. 4. Błędne metadane *Rocznika Polskiego Towarzystwa Naukowego na Obczyźnie*.

Jak pokazuje Rys. 4, według wyszukiwarki Google Books frazę "polskiego millennium" odnotowuje *Rocznik Polskiego Towarzystwa Naukowego na Obczyźnie* nr 18 z roku 1982. Rzecz w tym, że rocznik nr 18 wydano w 1967 roku, w 1982 roku ukazał się rocznik nr 24, a szukana fraza pochodzi z rocznika nr 21, wydanego na przełomie lat 1970/71.

4. Zła jakość optycznego rozpoznawania znaków (OCR)

Firma Google wprowadziła do obiegu ogromne ilości tekstu, poddawszy go technice optycznego rozpoznawania znaków (OCR), które jednak nie przeszły weryfikacji człowieka, mającej na celu usunięcie błędów. Pełnotekstowe wersje książek w Google Books można, z większym lub mniejszym powodzeniem, przeszukiwać online. Ściągnięcie skanów na własny komputer okazuje się jednak mało przydatne, ponieważ są to pliki niepoddane technice OCR, mające jednocześnie status plików zabezpieczonych.



Rys. 5. Przykład źle rozpoznanych znaków (*extasy*) w tekście greckim²⁵.

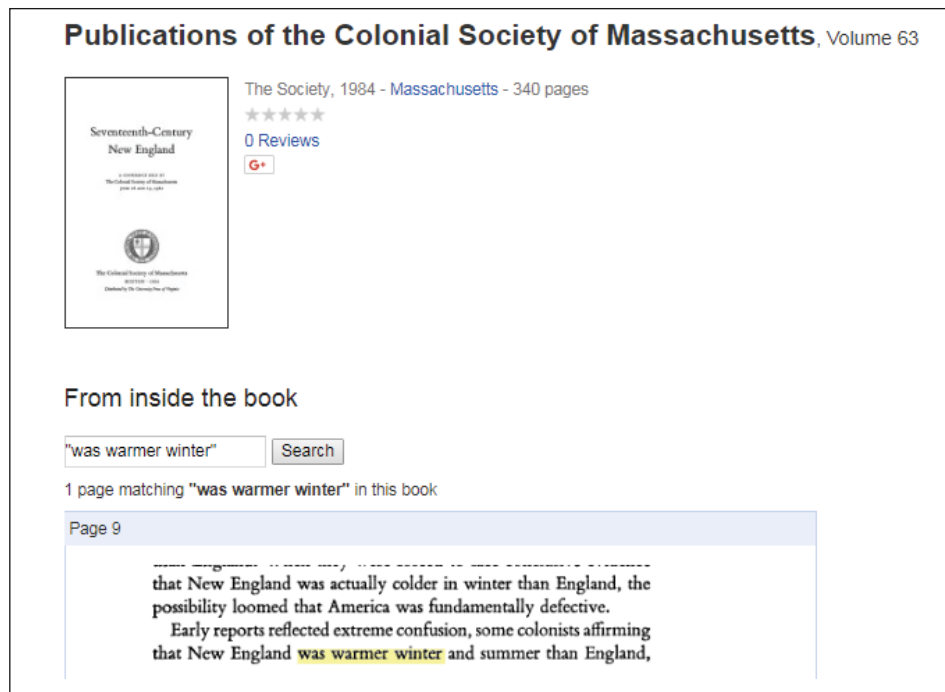
Problem optycznego rozpoznawania znaków dotyczy również tekstów z użyciem starodawnych kształtów czcionek i tekstów w innym alfabecie niż łaciński. Jak widać na Rys. 5, angielska forma historyczna *extasy* została znaleziona w greckim tekście dzieła Ptolemeusza, *Magnae Constructionis*..., wydany w Bazylei w 1538 roku.

5. Niska użyteczność fragmentów tekstów

Z uwagi na prawa autorskie Google Books często umożliwia jedynie dostęp do niewielkiego fragmentu książki (zob. Rys. 6). Przy założeniu, że wiele metadanych jest błędnych, taki format stanowi niewiarygodne źródło danych. Ograniczony kontekst może też być niewystarczający, aby określić, czy dana jednostka występuje we właściwym znaczeniu.

Jedyną szansą potwierdzenia danych jest weryfikacja w źródle pełnotekstowym i/lub w bibliotece tradycyjnej.

²⁵ Powyższy przykład pochodzi ze strony <https://books.google.pl/books?id=J9gBfUNAnV6C&pg=RA4-PA82&dq=%22extasy%22&hl=pl&sa=X&ved=0ahUKEwiRpoKxyo3bAhUCJFAKHf6VDc8Q6AEIJzAA#v=onepage&q=%22extasy%22&f=false>.



Rys. 6. Przykład fragmentu tekstu udostępnionego w Google Books.

6. Brak możliwości zapytań o części mowy

Najpoważniejszym problemem z punktu widzenia językoznawstwa korpusowego jest fakt, że korpus Google Books nie zawiera znaczników morfoskładniowych, co znacząco ogranicza możliwości analizy językoznawczej tekstów. Przykładowo, szukając słowoformy *crowns*, otrzymujemy zarówno wystąpienia rzeczownikowe, jak i czasownikowe. Próba uporządkowania wyników oznacza żmudną, a przy tym nie zawsze satysfakcjonującą, weryfikację danych.

Spójrzmy także na poniższy przykład:

dragon '... a fabulous animal, generally represented as a monstrous winged and scaly serpent ...' [bajkowy stwór, zwykle przedstawiany jako ogromny, skrzydlaty, pokryty łuskami wąż – M.P.] / **dragon** 'to guard zealeausly' [strzec gorliwie – M.P.]

"to dragon" (F 2 150 000) / "dragon" (F 1 030 000) / "dragons" (F 562 000) / "dragoned" (F 2 530) / "dragoning" (F 259)

F ?

Jak podaje WNID2, obok rzeczownika *dragon* w angielszczyźnie funkcjonowała także homonimiczna forma czasownikowa. Ustalenie frekwencji czasownika, gdy forma trzeciej osoby czasu teraźniejszego *dragons* może oznaczać rzeczownik w liczbie mnogiej, forma bezokolicznika *to dragon* – frazę przyimkową, a forma imiesłowiu

przymiotnikowego czynnego *dragoning* – rzeczownik odsłowny, jest praktycznie niemożliwe.

Podsumowanie

Jak starałam się pokazać, Google Books można nazywać korpusem, mimo że został on skompilowany wbrew wszelkim zasadom – a także celom – językoznawstwa korpusowego. Jego najważniejszą zaletą jest rozmiar, ponieważ 30 milionów książek to około 2 bilionów wyrazów²⁶, natomiast do wad zaliczyć należy błędne metadane, niezadawalające wyniki optycznego rozpoznawania znaków oraz problemy z weryfikacją frekwencji. O ile korpusy tworzone w oparciu o rygorystyczną procedurę kompilacji zapewniają wyniki dla różnych typów zapytań, wiele z nich nie jest technicznie możliwych w przypadku Google Books²⁷. W tej sytuacji musimy przyjąć za pewnik, jak sugeruje Kilgariff (2007: 151), że „jeśli mamy do dyspozycji olbrzymią ilość danych, uzyskujemy lepsze rezultaty” [M.P.].

Zdaniem Diemera (2012) do tej pory istnieje wciąż niewiele analiz bazujących na narzędziach i zasobach Google, ponieważ badaczy zniechęciły negatywne opinie takich autorytetów jak Geoffrey Nunberg (np. Nunberg 2009). Wierzy jednak, że mogą – a nawet powinny – mieć one zastosowanie, nawet mimo braku dokładnych danych frekwencyjnych. Jego zdaniem kończy się era zamkniętych korpusów, jeśli więc ograniczymy się tylko do nich, stracimy „szansę rozwoju” [M.P.]. Co więcej, pojawiają się nawet głosy, że korpus Google Books przyczynił się do całkowitej transformacji humanistyki cyfrowej (Howard 2017).

Warto dodać, że już od dekady trwa dyskusja naukowa dotycząca ogromnych, zmiennych i różnorodnych zbiorów danych (ang. *big data*)²⁸, których analiza jest trudna, ale wiąże się z zupełnie nowymi perspektywami (Aiden, Michel 2013). Jak twierdzą Mayer-Schönberger i Cukier (2013: 7), „prawdziwa rewolucja nie kryje się w maszynach obliczających, ale w danych i sposobach ich wykorzystania” [M.P.]. To

²⁶ Z badań statystycznych wynika, że średnia objętość książki wynosi około 64 000 wyrazów (Habash 2012), co po pomnożeniu przez 30 milionów daje 1 920 000 000 wyrazów. Przy obliczeniach nie należy się kierować danymi z korpusu Daviesa, ponieważ zawiera on zaledwie część zasobów Google Books od roku 1810. Przykładowo, Google Books American corpus, obejmujący 155 miliardów wyrazów, zaczerpnięto z niespełna półtora miliona książek (<https://eadh.org/news/2011/05/13/google-books-corpus>).

²⁷ Jak słusznie zauważył recenzent niniejszego artykułu, w przypadku tzw. korpusów „zanieczyszczonych” (ang. *dirty corpora*), do których należy z pewnością sieć stron WWW (zob. Kilgariff 2003, Kilgariff 2007), można stosować szereg procedur w celu oszacowania błędu statystycznego. Warto jednak zaznaczyć, że korpus Google Books uniemożliwia eksplorację danych przy pomocy standardowych metod korpusowych (zob. Oakes, Krakowian, Uzar 2005: 116–132), choćby dlatego, że teksty są dostępne w formacie PDF, a wyszukiwanie jest możliwe wyłącznie przy użyciu wyszukiwarki Google.

²⁸ Terminy kluczowe określające te zbiory to: *volume* (rozmiar), *velocity* (zmiennosc), *variety* (różnorodność), *veracity* (prawdziwość, wiarygodność) i *value* (wartość) [M.P.] (Dhaenens, Jourdan 2016: 9).

właśnie ogromne ilości danych, zdaniem Halevy’ego, Norviga i Pereiry (2009: 8), mają być w przyszłości źródłem wiedzy.

Zasoby cyfrowe

COHA = Corpus of Historical American English. <https://corpus.byu.edu/coha/>.
 EEBO = Early English Books Online. <https://eebo.chadwyck.com/home>.
 Google Books. <https://books.google.com/>.
 Google Books Corpus. <https://googlebooks.byu.edu/>.
 HathiTrust Digital Library. <https://www.hathitrust.org/>.
 Internet Archive. <https://archive.org/>.
 Project Gutenberg. <https://www.gutenberg.org/>.
 The Library of Congress American Memory. <https://memory.loc.gov/ammem/index.html>.

Bibliografia

- AIDEN Erez, MICHEL Jean-Baptiste (2013): *Uncharted: Big data as a lens on human culture*. – New York: Riverhead Books.
- BIBER Douglas (1993): Representativeness in corpus design. – *Literary and Linguistic Computing* 8 (4), 243–257.
- BRIN Sergey (2009): A library to last forever. – *The New York Times*, 8.10.2009, <http://www.nytimes.com/2009/10/09/opinion/09brin.html>.
- CRAWFORD James (2010): On the future of books. – 14.10.2010. <http://booksearch.blogspot.com/2010/10/on-future-of-books.html>.
- DAVIES Mark (2014): Making Google Books Ngrams useful for a wide range of research on language change. – *International Journal of Corpus Linguistics* 19 (3), 401–416.
- DAVIES Mark, CHAPMAN Don (2016): The effectiveness of representativeness and size in historical corpora: An empirical study of changes in lexical frequency. – [w:] Don CHAPMAN, Colette MOORE, Miranda WILCOX (red.): *Studies in the history of the English language* 7, 131–152.
- DHAENENS Clarisse, JOURDAN Laetitia (2016): *Metaheuristics for big data*. – London: John Wiley and Sons.
- DIEMER Stefan (2012): Corpus linguistics with Google? – <https://www.bu.edu/isle/files/2012/01/Stefan-Diemer-Corpus-Linguistics-with-Google.pdf>.
- DILLER Hans-Jürgen (2013): Culturomics and genre: *Wrath and anger* in the 17th century. – [w:] R.W. MCCONCHIE [et al.] (red.): *Selected proceedings of the 2012 Symposium on New Approaches in English Historical Lexis (HEL-LEX 3)*. – Somerville: Cascadilla Proceedings Project, 54–65.
- FISCHER Karen (2007): Committee on Institutional Cooperation (CIC) joins Google’s Library Project. – *The University of Iowa Libraries*. – September 6, 2007. <https://blog.lib.uiowa.edu/transitions/?p=68>.
- GÓRSKI Rafał L. (2003): Korpus współczesnego języka polskiego IJP PAN, tzw. korpus krakowski. – [w:] Stanisław GAJDA (red.): *Językoznawstwo w Polsce. Stan i perspektywy*. – Opole: Wydawnictwo Uniwersytetu Opolskiego, 158–161.
- GRAFTON Anthony (2012): Kodeks w kryzysie. Dematerializacja książki (tłum. Michał Choptiany). – *Wielogłos: Pismo Wydziału Polonistyki UJ* 3 (13), 229–264.
- GRALIŃSKI Filip (2013). Polish digital libraries as a text corpus. – [w:] Zygmunt VETULANI, Hans USZKOREIT (red.): *Proceedings of the 6th Language and Technology Conference*. – Poznań: Fundacja Uniwersytetu im. Adama Mickiewicza, 509–513.

- GREEN Eugene (2014): The diffusion of *I need you to* + infinitive in world Englishes. – [w:] Eugene GREENE, Charles F. MAYER (red.): *The variability of current world Englishes*. – Berlin: de Gruyter, 257–284.
- HABASH Gabe (2012): Average book length: Guess how many words there are in a novel. PWxyz (blog). Przedruk: *Huffington Post*, 3.09.2012. https://www.huffingtonpost.com/2012/03/09/book-length_n_1334636.html.
- HALEVY Alon, NORVIG Peter, PEREIRA Fernando (2009): The unreasonable effectiveness of data. – *IEEE Intelligent Systems* 24 (2), 8–12.
- HOWARD Jennifer (2017): What happened to Google's effort to scan millions of university library books? – *EdSurge*, 10.08.2017, <https://www.edsurge.com/news/2017-08-10-what-happened-to-google-s-effort-to-scan-millions-of-university-library-books>.
- JAMES Ryan, WEISS Andrew (2012): An assessment of Google Books' metadata. – *Journal of Library Metadata* 12 (1), 15–22.
- KILGARRIFF Adam (2003): The Web as corpus. – [w:] Paul RAYSON [et al.] (red.): *Proceedings of the Corpus Linguistics 2001 Conference. Lancaster University (UK) 29 March – 2 April 2001*, 342–344, <http://ucrel.lancs.ac.uk/publications/CL2003/CL2001%20conference/papers/kilgarri.pdf>.
- KILGARRIFF Adam (2007): Googleology is bad science. – *Computational Linguistics* 3 (1), 147–151.
- LANDAU Sidney I. (2001): *Dictionaries: The art and craft of lexicography*. – Cambridge: Cambridge University Press.
- LEVY Steven (2011): *In the plex: How Google thinks, works, and shapes our lives*. – New York: Simon & Schuster.
- LEWIS Danny (2015): Google Books isn't copyright infringement. – *Simthsonian.com*, 20.10.2015, <https://www.smithsonianmag.com/smart-news/court-ruling-legalizes-google-books-180956997/>.
- LÜDELING Anke, KYTÖ Merja (red.) (2009): *Corpus linguistics*, t. 2. – Berlin: De Gruyter Mouton.
- MAYER-SCHÖNBERGER Viktor, CUKIER Kenneth (2013): *A revolution that will transform how we live, work, and think*. – Boston: Houghton Mifflin Harcourt.
- MAYS Dorothy A. (2015): Google Books: Far more than just books. – 20.10.2015, <http://publiclibraries-online.org/2015/10/far-more-than-just-books/>.
- MICHEL Jean-Baptiste et al. (2011): Quantitative analysis of culture using millions of digitized books. – *Science* 331 (176), 176–182.
- NUNBERG Geoffrey (2009): Google's book search: A disaster for scholars. – *Chronicle of Higher Education, the Chronicle Review*, 31.8.2009, <http://chronicle.com/article/Googles-Book-Search-A/48245/>.
- OAKES Michael, KRAKOWIAN Przemysław, UZAR Rafał (2005): Testy statystyczne w językoznawstwie korpusowym. – [w:] Barbara LEWANDOWSKA-TOMASZCZYK (red.): *Podstawy językoznawstwa korpusowego*. – Łódź: Wydawnictwo Uniwersytetu Łódzkiego, 116–132.
- OED3 = SIMPSON John, PROFFITT Michael, red. (2000–): *Oxford English dictionary*, wyd. 3, <http://www.oed.com/>.
- PARKS Tim (2014): References, please! – *The New York Review of Books*, 13.9.2014, <http://www.nybooks.com/daily/2014/09/13/references-please/>.
- PECHENICK Eitan A., DANFORTH Christopher, SHERIDAN DODDS Peter (2015): Characterizing the Google Books corpus: Strong limits to inferences of socio-cultural and linguistic evolution. – *Plos One*, 7.10.2015, <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0137041>.
- PIOTROWSKI Tadeusz (2003): Językoznawstwo korpusowe – wprowadzenie do problematyki. – [w:] Stanisław GAJDA (red.): *Językoznawstwo w Polsce. Stan i perspektywy*. – Opole: Wydawnictwo Uniwersytetu Opolskiego, 143–154.
- PODHAJECKA Mirosława (2011): Can Google Books collection complement traditional corpora? – [w:] Stanisław GÓZDŹ-ROSKOWSKI (red.): *PALC 2009: Explorations across languages and corpora*. – Frankfurt am Main: Peter Lang, 529–546.

- PODHAJECKA Mirosława (2015): Google Books as a source of historical data: The entry for *macaroni* in OED3. – [w:] Elżbieta MAŃCZAK-WOHLFELD, Barbara PODOLAK (red.): *Words and dictionaries. A Festschrift for Professor Stanisław Stachowski on the occasion of His 85th birthday*. – Kraków: Jagiellonian University Press, 247–263.
- [PTOLEMY] (1538): *Clavdii Ptolemaei Magnae Constructionis, Id est Perfectae Coelestium motuum pertractationis*, Lib. XIII. – Basilaee: Apud Ioannem Vvaldervm.
- ROBERTS Jeff J. (2017): Why Google Books deserves better than these obituaries. – *Fortune*, 24.4.2017, <http://fortune.com/go/tech/google-books-future/>.
- ROSENBERG Scott (2017): How Google Books got lost. – *Backchannel*, 11.4.2017, <https://www.wired.com/2017/04/how-google-book-search-got-lost/>.
- SKIENA Steven, WARD Charles B. (2013): *Who's bigger? Where historical figures really rank*. – Cambridge: Cambridge University Press.
- SKIPWORTH Hunter (2010): Google counts total number of books in the world. – *The Telegraph*, 6.8.2010, <http://www.telegraph.co.uk/technology/google/7930273/Google-counts-total-number-of-books-in-the-world.html>.
- WALIŃSKI Jacek (2005): Typologia korpusów oraz warsztat informatyczny lingwistyki korpusowej. – [w:] Barbara LEWANDOWSKA-TOMASZCZYK (red.): *Podstawy językoznawstwa korpusowego*. – Łódź: Wydawnictwo Uniwersytetu Łódzkiego, 27–33.
- WNID2 – William A. NEILSON, Thomas A. KNOTT, Paul W. CARHART (red.) (1934): *Webster's new international dictionary of the English language*, wyd. 2. – Springfield, MA: Merriam-Webster.
- WU Tim (2015): What ever happened to Google Books? – *The New Yorker*, 11.9.2015, <http://www.newyorker.com/business/currency/what-ever-happened-to-google-books>.
- ZIMMER Ben (2012): Bigger, better Google Ngrams: Brace yourself for the power of grammar. – *The Atlantic*, 18.10.2012, <https://www.theatlantic.com/technology/archive/2012/10/bigger-better-googe-ngrams-brace-yourself-for-the-power-of-grammar/263487/>.
- ŻMIGRODZKI Piotr (2005): Słownik jako korpus tekstów – korpus tekstów jako słownik. Perspektywy polskiej leksykografii naukowej. – *Poradnik Językowy* 6, 3–14.

doi: <http://dx.doi.org/10.17651/bptj.74.3>

IWONA KOSEK

UWM w Olsztynie

SEBASTIAN PRZYBYSZEWSKI

UWM w Olsztynie

MONIKA CZEREPOWICKA

UWM w Olsztynie

Źródła w tworzeniu elektronicznego słownika paradygmatów frazeologizmów werbalnych¹

Streszczenie

Celem artykułu jest wskazanie typów źródeł i sposobów ich wykorzystania przy opracowywaniu słownika paradygmatów frazeologizmów czasownikowych. Autorzy charakteryzują wykorzystywane źródła, omawiają użyteczność źródeł słownikowych i korpusowych (w zależności od etapu prac nad słownikiem) oraz ich ograniczenia. Szczególną uwagę poświęcają roli introspekcji w tworzeniu słownika gramatycznego.

Wyrazy kluczowe: leksykografia, frazeologizmy werbalne, korpus, introspekcja

Language data sources in an electronic dictionary of paradigms of verbal phraseological units

Abstract

The article discusses types of language data sources and the ways they may be used in the compilation of a dictionary of paradigms of verbal phraseological units. The authors describe the sources they used and discuss the usefulness of lexicographic and corpus sources and their limitations. Special attention is paid to the role of introspection in the process of making a grammatical dictionary.

Keywords: lexicography, verbal phraseological units, corpus, introspection

¹ Praca naukowa finansowana ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2013/09/B/HS2/01222.

1. Uwagi wprowadzające²

Tworząc słownik, korzysta się z trzech typów informacji: własnej kompetencji językowej, wcześniejszych słowników danego typu (jeśli takie istnieją) oraz tekstów dokumentujących zastosowanie (użycie) jednostek (zob. Bańko 2001: 13; Kilgarriff 2013: 77–78)³. Wszystkie wymienione elementy można rozumieć jako źródła w leksykografii. W tradycji określenie *źródła* interpretowano zwykle wężiej, odnosząc je do oryginalnych tekstów, na podstawie których opracowywano listę jednostek i ilustracje ich użycia. Oś rozważań leksykografów stanowił przy takim rozumieniu źródeł dobór tekstów i jego konsekwencje dla ustalenia listy opisywanych jednostek (zob. np. Lubaś 1988). Dyskusję nad znaczeniem tekstów źródłowych dla leksykografii i gramatyki, nad stopniem ich wykorzystania w polskich słownikach, a także relacje (różnice) między korpusem a kartoteką opisuje Mirosław Bańko (2001: 25–49, zob. też Bańko 2012), nie będziemy jej zatem w tym miejscu przywoływać. Niewątpliwie o przydatności tekstów źródłowych nie trzeba nikogo przekonywać, zwłaszcza w dobie rozwijającego się językoznawstwa korpusowego, a istniejące zbiory tekstów (przede wszystkim NKJP) stanowią współcześnie istotną, jeśli nie podstawową leksykograficzną bazę materiałową (zob. np. Żmigrodzki 2015).

W prezentowanym artykule jako źródła rozumiemy zarówno zbiory tekstów (oryginalne przykłady użycia), jak i dostępne słowniki. Te drugie można traktować jako metaźródła, kryją się za nimi bowiem konkretne autorskie interpretacje i decyzje. Mówimy zatem w istocie o informacjach z dwu poziomów – „czystego” i „przefiltrowanego”. Chcemy wskazać, jak te informacje są i mogą być przydatne w tworzonym słowniku fleksyjnym frazeologizmów werbalnych (szerzej o słowniku piszemy w p. 2.), a także zwrócić uwagę na pewne ogólniejsze problemy związane z wykorzystaniem różnych typów źródeł.

Stopień przydatności źródeł i ich waga w sposób oczywisty zależą od typu słownika⁴, przewidywanego odbiorcy (zob. Sinclair 1995, Čermák 2003), ale też od

² Wkład każdego z autorów w powstanie niniejszego artykułu jest równy (33,33%). Wspólnie opracowano koncepcję i założenia tekstu, zebrano materiał przykładowy, przeprowadzono analizę i nadano pracy ostateczny kształt.

³ Kilgarriff (2013: 77–78) mówi o trzech sposobach tworzenia słownika, są to: kopiowanie, introspekcja i ogląd danych. Pierwszy, rozumiany jako kwerenda po innych, istniejących już słownikach, nie może się z oczywistych względów ograniczać do samego przepisania rozwiązań, o ile dany słownik ma wniesić jakąś nową jakość do leksykografii. Centralna jest, zdaniem Kilgarriffa, introspekcja, choć i ona ma swoje ograniczenia: badacz, opierając się na swojej kompetencji językowej, może pewne rzeczy sobie uświadamiać, lecz inne mogą ująć jego uwadze. W tym względzie pożyteczny jest ogląd danych (korpus).

⁴ Twórca słownika gramatycznego, podobnie jak każdy leksykograf, musi wybrać między postawą preskryptywną i deskryptywną. Współczesna leksykografia odeszła od preskryptywności właściwej XVII- i XVIII-wiecznym słownikom, jak choćby *Dictionnaire de l'Académie française* stanowiącemu wzór dla wielu późniejszych opracowań. Jednakże postawa deskryptywna, dowartościowująca rzeczywiste użycie języka, nie zwalnia całkowicie leksykografa od zadania wskazywania odbiorcy form preferowanych – my również, choć zajmujemy postawę deskryptywną, dokonujemy niekiedy arbitralnych, opartych na naszej kompetencji, decyzji, odrzucając pewne formy, mimo że są reprezentowane w tekstach.

zakładanej roli korpusów. Jak wiadomo, wyróżnia się 3 podejścia w badaniach z ich wykorzystaniem: badania ilustrowane przez korpus (*corpus-illustrated / informed approach*), oparte na korpusie (*corpus-based approach*) i sterowane korpusem (*corpus-driven approach*), zob. np. Tummers, Heylen, Geeraerts 2005, Górski 2012, Grabowski 2015. W tworzonym słowniku prezentujemy podejście drugie, wykorzystujemy zatem korpus nie tylko jako źródło oryginalnych przykładów, ale przede wszystkim do weryfikowania / falsyfikowania hipotez i sprawdzania ich pełności. Trzecie, najbardziej radykalne, podejście w tworzonym słowniku jest badawczo mniej przydatne. Po pierwsze z tego względu, że widzimy frazeologię jako element systemu (pewną grupę jednostek leksykalnych), nie jako zjawisko tekstowe; po drugie – bardzo liczne związki frazeologiczne zostały już wydobyte z tekstów, zarejestrowane, nie opisano natomiast ich paradygmatów i na tym zadaniu się skupiamy; po trzecie wreszcie, jak pokazują dotychczasowe badania (zob. np. Moon 1998; Kilgarriff 2013: 80), najlepiej frekwencyjnie wyodrębniają się z tekstów regularne semantycznie kolokacje, a ich fleksja jest wysoce przewidywalna, inaczej niż jednostek o zidiomatyzowanym znaczeniu. Przedstawione względy decydują o tym, że rolę korpusów w tworzonym słowniku można postrzegać jako drugoplanową: listę jednostek ustalamy głównie na podstawie istniejących słowników i własnej kartoteki, a podstawę opisu paradygmatów stanowi własna kompetencja językowa. W takich prostych stwierdzeniach roli korpusów tekstowych w słowniku gramatycznym zamknąć się jednak nie da, przynajmniej z dwu powodów. Przede wszystkim specyfika jednostek frazeologicznych decyduje o tym, że nie ma całkowitej odpowiedniości między fleksją (paradygmatem) związku a paradygmatem homonimicznego z głównym komponentem leksemu. W przypadku frazeologii zależność gramatyki od semantyki jest bardzo widoczna, wręcz wyostrowana w zestawieniu z leksemami. Nigdy nie można założyć, że związek będzie miał taki sam paradygmat jak jego główny komponent – homonimia o niczym nie przesądza. Decydujące są własności znaczeniowo-funkcjonalne frazeologizmu i wynikająca z nich jego tekstowa łączliwość. Z tego względu słownik gramatyczny leksemów (w przypadku tworzonego słownika to SGJP oraz opracowanie *Czasownik polski. Odmiana, słownik* Saloniego, 2001) jest w pracy nad paradygmatyką frazeologii niewystarczający. Ponadto, choć oryginalne teksty dla omawianego typu słownika nie są tak istotne jak dla leksykografii ogólnej, niewątpliwie mogą pomóc w rozwiązaniu kilku typów kwestii szczegółowych⁵.

2. Typy źródeł i ich wykorzystanie w pracy nad słownikiem fleksyjnym frazeologizmów

Niniejsze refleksje związane są z pracą nad elektronicznym *Słownikiem paradygmatów polskich frazeologizmów czasownikowych* VERBEL (słownik w postaci pliku jest dostępny na stronie www.uwm.edu.pl/verbel/; zob. też Czerepowicka,

⁵ Na temat możliwości wykorzystania korpusów w badaniu gramatyki por. m.in.: Conrad 2010; Górski 2012.

Kosek, Przybyszewski 2014). Frazeologizmy traktujemy jako jednostki języka (zgodnie z koncepcją A. Bogusławskiego (1976)), spośród nich w słowniku ujmujemy tylko te, które w funkcjonalnej klasyfikacji związków frazeologicznych Andrzeja Marii Lewickiego (1986) określane są jako zwroty (zob. też Lewicki 1976; Lewicki, Pajdzińska 2001). Słownik, w postaci specjalnie do tego celu opracowanej aplikacji Werbosław⁶, podaje wprost paradygmaty około 5000 zwrotów. Punkt wyjścia badań stanowi założenie, że fleksja (zasób form) jednostek frazeologicznych jest oparta na fleksji ich komponentów i grupy syntaktycznej, której schemat realizują (Lewicki, Pajdzińska, Rejakowa 1987: 20). Zwroty odmieniają się więc tak, jak ich główny czasownikowy komponent (zestaw kategorii fleksyjnych czasownika przyjmujemy za Zygmuntem Salonim, zob. Saloni 2001; SGJP). Ich paradygmat nie jest jednak prostym odbiciem paradygmatu podstawowego składnika. Nieregularne znaczeniowo frazeologizmy werbalne, jak można wnioskować z dotychczasowych badań, mają więcej ograniczeń aniżeli czasowniki, na których są oparte (zob. Bąba 1998; Kosek 2013; Czerepowicka, Kosek, Przybyszewski 2014; Przybyszewski 2015). Podstawowe zadanie przy opisie paradygmatu zwrotu polega zatem na określeniu, czy podlega on ograniczeniom fleksyjnym, a jeśli tak, to w jakim zakresie.

Prace nad omawianym słownikiem można podzielić na trzy etapy: 1) zbieranie materiału i ustalanie postaci jednostek; 2) określanie paradygmatów zgromadzonych zwrotów; 3) wprowadzanie paradygmatów do aplikacji Werbosław. Źródła były przez nas wykorzystywane na dwóch pierwszych etapach.

2.1. Gromadzenie materiału i ustalanie postaci jednostek

Źródłami pomocnymi podczas pierwszego etapu prac były przede wszystkim słowniki ogólne języka polskiego (ISJP, SJPD, SJPSz, SWJP, USJP, WSJP) oraz słowniki frazeologiczne (PSF, SFJP, SFWP, WSF, WSFJP). Wykorzystywaliśmy głównie te opracowania, które konsekwentnie posługują się zbliżonymi do przyjętych przez nas kryteriami wyodrębniania jednostek nieciągłych, czyli ISJP i WSJP. Wiele zwrotów zawartych w słownikach stosujących inne definicje i odbiegające od przyjętych przez nas metod wyróżniania frazeologizmów okazywało się, z naszej perspektywy, połączeniem czasowników i różnego typu fraz niewerbalnych. Spośród słowników frazeologicznych w największym stopniu korzystaliśmy z tzw. słowników poznańskich, tj. PSF i SFWP. Natomiast słownik Stanisława Skorupki (SFJP), ze względu na bardzo szerokie rozumienie frazeologii – jako nauki o połączeniach wyrazów, traktowany był przez nas jako źródło uzupełniające. Sięgaliśmy również do sond słownikowych (Bogusławski, Garnysz-Kozłowska 1979; Bogusławski, Wawrzyńczyk 1993; Bogusławski, Danielewiczowa 2005). To właśnie z sond zaczerpnęliśmy takie zwroty,

⁶ Ścisłe mówiąc, Werbosław jest rozbudowaną wersją aplikacji Toposław, powstałej przy tworzeniu SAWY – *Słownika elektronicznego nazewnictwa Warszawy* (Marciniak i in. 2009), i wykorzystanej także do budowy SEJF-u – *Słownika elektronicznego jednostek frazeologicznych* (Czerepowicka 2014). Trzeba jednak zaznaczyć, że program przeszedł znaczne zmiany, uwzględniające specyfikę i poziom skomplikowania polskiej koniugacji, które w konsekwencji umożliwiły kodowanie frazeologizmów czasownikowych.

jak np.: {*ktoś / coś*} *bije rekordy {czegoś}*; {*coś*} *umiera śmiercią naturalną*, {*coś*} *daje głos* ('szczeka'). Dodatkowym źródłem frazeologizmów były opracowania naukowe. Stanowiły one źródło takich jednostek, jak np. {*coś*} *pada na podatny grunt* czy {*coś*} *przyprawia {kogoś} o zawrót głowy* (Liberek 2006). Ponadto własne kwerendy tekstowe pozwoliły znaleźć jednostki, które nie są dotychczas odnotowane w słownikach. To zwroty zasłyszane, z literatury, prasy czy Internetu. Należą do nich np. {*ktoś*} *nie urodził się wczoraj*, {*ktoś*} *spuści kogoś na drzewo / po brzytwie*, {*ktoś*} *nie ma wszystkich w domu*, {*ktoś*} *ma jakiś ból głowy {z czymś}*, {*ktoś*} *czarno {coś} widzi*, {*ktoś*} *ugotuje {kogoś} na twardo*, {*ktoś*} *umoczy dzioba*.

Korzystając na etapie gromadzenia jednostek z różnych źródeł leksykograficznych, stykamy się, jak łatwo przewidzieć, z różnicami co do postaci i liczby odnotowywanych frazeologizmów. Problem rozbieżności interpretacyjnych (bo tego obrazem są odmienne postaci związków w różnych słownikach) można uznać za typowy dla wielości źródeł, nie do uniknięcia nawet przy podobnym czy tożsamym spojrzeniu metodologicznym na dane zjawisko, w przypadku jednostek wielowyrazowych występujący z całą ostrością. Przykładowe rozbieżności co do postaci jednostek podano w poniższej tabeli. Uwzględniono w niej cztery słowniki ogólne i dla porównania słownik frazeologiczny wydawnictwa PWN – WSF:

SJPSz	USJP	ISJP	WSJP	WSF
<i>iść, pójść pod młotek</i> (licytacyjny)	<i>coś idzie, poszło pod (licytacyjny) młotek</i>	'Mówimy, że coś idzie lub poszło pod młotek albo że coś oddano pod młotek , jeśli jest to wystawione na licytacji lub zostało sprzedane na licytacji'; FPS + CZ	<i>coś idzie pod młotek; coś poszło pod młotek</i>	<i>coś idzie, poszło pod (licytacyjny) młotek</i>
<i>pleść trzy po trzy</i>	<i>bredzić, mówić, pot. klepać, pleść itp. trzy po trzy</i>	'Jeśli ktoś plecie, gada, mówi itp. trzy po trzy , to mówi bez sensu'; FPS + CZ	<i>trzy po trzy</i>	<i>bredzić, mówić, klepać, pleść itp. trzy po trzy</i>
<i>mówić na wiatr, rzucać słowa na wiatr</i>	<i>mówić na wiatr, rzucać słowa na wiatr</i>	'Jeśli ktoś nie mówi na wiatr lub nie rzuca słów na wiatr , to nie składa obietnic, których potem nie dotrzymuje'; FCZ NDK ZWYKLE NEG	<i>ktoś rzuci słowa na wiatr; ktoś rzuca słowa na wiatr</i> <i>rząd. ktoś mówi na wiatr</i>	<i>nie mówić na wiatr, nie rzucać słów na wiatr</i>

Tab. 1. Rozbieżności w zapisie frazeologizmów w wybranych słownikach

Przypadków podobnych różnic jest znacznie więcej. Spójrzmy jeszcze na dwa przykłady takich rozbieżności. W SJPSz, USJP i WSF ciąg *patrzeć na kogoś jak na raroga* jest odnotowany jako frazeologizm czasownikowy. Z kolei redaktorzy ISJP i WSJP wymienionego ciągu nie uznają za frazeologizm, lecz tylko za jedno z regularnych znaczeń leksemu *raróg*, por. definicję w ISJP: 'jeśli ktoś patrzy na kogoś jak na **raroga**, uważa kogoś za **raroga** itp., to patrzy na niego jak na jakąś osobliwość, podejrzliwie i niechętnie'. W WSJP na pięć przykładów ilustrujących omawiane zna-

czenie rzeczownika *raróg* tylko jeden to kolokacja z czasownikiem *patrzeć*. Inny typowy przykład opisywanych rozbieżności to ciąg *co koń wyskoczy*, który w SJPSz, USJP i WSF traktowany jest jako fraza czasownikowa z kilkoma wariantami werbalnymi (*gnać / jechać / pędzić co koń wyskoczy*), zaś w ISJP i WSJP opisywany ciąg odnotowuje się jako frazę przysłówkową.

Prezentowane powyżej przykłady ilustrują pewne ogólne tendencje w notacji nieciągłych jednostek werbalnych w poszczególnych słownikach. O ile ISJP oraz WSJP raczej „skracają” postać jednostki, redukując ją do postaci frazy przysłówkowej, o tyle słowniki tradycyjne przejawiają tendencje rozszerzające, jako frazeologizmy werbalne traktując frazy przysłówkowe czy rzeczownikowe. Takie rozbieżności można wytłumaczyć różnicami metodologicznymi (zwłaszcza uwzględnianiem pojęcia klasy otwartej / zamkniętej w ustalaniu granic jednostek w ISJP i WSJP), ale także celami i założeniami leksykografów. Słowniki tradycyjne oraz nowsze wpisujące się w tradycyjną metodologię leksykograficzną mniejszą wagę przywiązują do postaci jednostki nieciągłej i jej gramatycznej charakterystyki, a zasadniczym ich celem jest podanie użytkownikowi typowego sposobu użycia (stąd często pojawiające się w opisywanych przypadkach ciągi bliskoznaczników czasownikowych, np. *brać, czytać, kupować* itp. *jak leci*). Z kolei w ISJP i WSJP znacznie więcej uwagi poświęcono zdyscyplinowanemu wyodrębnieniu jednostek nieciągłych i ich gramatycznej charakterystyce, a typowe użycia zostały omówione w definicjach i / lub zilustrowane przykładami. Na marginesie warto odnotować, że również między ISJP i WSJP nie ma pełnej zgodności co do postaci frazeologizmów. Wątpliwości „zwrot czy wyrażenie?” ISJP wydaje się rozstrzygać, traktując formy czasowników jako zewnętrzne otoczenie jednostki, por. np. *ktoś nie powiedział ostatniego słowa* (WSJP) vs. *ostatnie słowo* (ISJP), *ktoś nie da złamanego grosza* (WSJP) vs. *złamany grosz* (ISJP), *ktoś nie pamięta o bożym świecie* (WSJP) vs. *boży świat* (ISJP). Podane przykłady rozbieżności potwierdzają, że jednostki nie są dane, trzeba je wydobywać z tekstów, oraz że ostateczna interpretacja należy do leksykografa, podobnie jak interpretacja co do użyteczności danego słownika jako źródła.

Na etapie gromadzenia jednostek pomocne były też dane korpusowe (zwykle z NKJP). Zgodnie z podejściem *corpus-based*, kwereńda korpusowa służyła weryfikacji wstępnych intuicji i hipotez co do postaci jednostki. I tak na przykład dane z NKJP spowodowały modyfikację pierwotnie zakładanej postaci jednostki *{ktoś} ma niezły ból głowy {z czymś}*. Ponieważ teksty wskazują na wymiennność elementu *niezły* w obrębie klasy otwartej (*spory, niemały, niezły, duży, wielki...*), a wymiennność ta nie narusza znaczenia i struktury frazeologizmu, ostatecznie zanotowaliśmy jednostkę *{ktoś} ma {jakiś} ból głowy {z czymś}*.

Sytuację przeciwną ilustruje odnajdywany w słownikach ciąg *{ktoś} koszulę / majtki w zębach nosił*. Korpus, poza postacią notowaną w słownikach, dostarcza przykładów ze składnikiem *pieluchę / pieluchy*:

(1) Nie ma się co oburzać. Po prostu Jagna wtedy pieluchę w zębach nosiła, to i nie pamięta. (NKJP – Usenet – pl.soc.dzieci)

(2) Niestety jako młodzienc, który w tychże latach (właściwie to za Gierka) stawiał dopiero pierwsze kroczki, albo jak wolą inni, pieluchy w zębach nosił, nie mogę ocenić na ile prawdziwe jest to przybliżenie. (NKJP – Usenet – pl.rec.ksiazki)

Nieuzasadnione byłoby jednak na podstawie zaledwie kilku przykładów z forów internetowych, nieznajdujących potwierdzenia w intuicji leksykografa, uznać dany ciąg za kolejny wariant nominalny. Modyfikację zwrotu można w wielu przypadkach zinterpretować jako spontaniczną, jednokrotną innowację nadawcy, mającą podkreślić znaczenie zwrotu, bądź jako nieświadomą kontaminację czy inne błędne przytoczenie zwrotu. W tym wypadku mamy do czynienia ze stosunkowo częstą i kłopotliwą dla leksykografa sytuacją, mianowicie z niedoborem tekstów, ale też z problemem interpretacji źródeł. Problem ten rysuje się ze szczególną ostrością w przykładach znajdujących z pomocą wyszukiwarek internetowych (np. Google, Yahoo).

Niewątpliwą zaletą dużego korpusu tekstów jest to, że daje on leksykografowi pewien zbiór poświadczeń, mogący zawierać takie warianty lub formy analizowanych jednostek, których badacz – odwołując się do swojej intuicji – nie brał wcześniej pod uwagę (Hanks 1990: 40; Biber 2015: 163). W tym sensie korpus wspomaga introspekcję, zapobiegając popadnięciu badacza w pewien schematyzm myślenia. W omawianych wyżej wypadkach dane korpusowe nie przesądzały o decyzjach dotyczących kształtu frazeologizmu, ale były ważnym argumentem, który należało wziąć pod uwagę (Hanks 1990: 32; Bańko 2012: 289). Ostatecznym weryfikatorem postaci jednostki nieciągłej pozostawała kompetencja językowa samych leksykografów.

2.2. Ustalanie paradygmatów frazeologizmów werbalnych

Drugi etap pracy nad słownikiem to ustalanie paradygmatów wyodrębnionych nieciągłych jednostek werbalnych. Tu przydatność obu analizowanych typów źródeł jest ograniczona kilkoma względami: korpusów – przede wszystkim przyjmowanym przez nas stanowiskiem nietekstowym, słowników – ubogą na ogół informacją fleksyjną o związkach frazeologicznych. Leksykografia starsza informacji gramatycznej o frazeologizmach nie podaje, co wynika m.in. z myślenia o nich w kategoriach wtórnych elementów systemu lub w kategoriach łączliwości wyrazów, ale też z koncentracji na opisie znaczenia związków. Słowniki nowsze, z których przynajmniej część traktuje frazeologizmy jako jednostki leksykalne i wobec tego opisuje je analogicznie do leksemów, dają pewne wskazówki co do odmiany. Konsekwentnie czynią to ISJP i WSJP (a także sondy słownikowe), choć w różny sposób. WSJP, podając informacje o odmianie zwrotu, odsyła zazwyczaj do jego głównego komponentu, przy którym z kolei znaleźć można paradygmat odtworzony za SGJP. Przy każdym zwrocie jest też informacja o przechodniości oraz o ewentualnym odpowiedniku aspektowym. Ponadto niekiedy dodatkowo zamieszczono informację o ograniczeniach w odmianie lub użyciu form, zob. np. *ktoś nie jest z cukru* – „zwykle cz. ter.”. ISJP z kolei nie odsyła wprost do paradygmatu czasownika, ale przy części związków podaje obok

informacji gramatycznych wskazówki co do ograniczeń fleksyjnych. Opis tych samych zwrotów może się w obu słownikach różnić. I tak na przykład jednostka *{coś} piechotą nie chodzi* w ISJP została opatrzona komentarzem: FCZ NDK TYLKO TER, zaś w WSJP podano: „fraza czasownikowa ndk, npch; odmienny: *nie chodzić*”. Z kolei zwrot *{ktoś} utopi {kogoś} w łyżce wody* w ISJP skomentowano następująco: FCZ DK ZWYKLE WAR / BEZOK, zaś w WSJP: „fraza czasownikowa dk, npch; odmienny: *utopiłby*. Spotykane również użycia formy czasu przyszłego: *ktoś utopi kogoś w łyżce wody*”. Tego typu informacje są dla autorów słownika fleksyjnego przydatne, jednocześnie jednak niektóre z nich dalekie od celu, jaki sobie wyznaczają. Słownik paradygmatów musi bowiem podawać formy wprost, innymi słowy – przynajmniej przy określonej koncepcji takiego słownika – powinna być w nim wyraźnie postawiona granica między ‘jest’ a ‘nie ma’, explicite powinien on stwierdzać, czy dana forma jest dopuszczalna w paradygmacie czy też – pomimo że dałoby się ją utworzyć (na przykład komponent czasownikowy daje taką możliwość) – jest ona wykluczona, jej sensowne użycie nie jest możliwe (np. z powodu semantycznych własności zwrotu). Dla autorów i użytkowników to informacja zasadnicza.

Określając paradygmat, posługujemy się zatem głównie własną kompetencją językową (introspekcją), porównując ewentualnie przykłady, w których dostrzegamy ograniczenia fleksyjne, z materiałem słownikowym i korpusowym. Korpus jest użyteczny na tym etapie prac, jak wspomniano we wstępie, do weryfikacji hipotez, także pod względem ich pełności (*corpus-based approach*)⁷. Na pewno, przeszukując bazy tekstów, nie można oczekiwać potwierdzenia użycia każdej formy, dla frazeologii werbalnej sytuacja typowa jest bowiem taka, że korpusy (także Internet) nie poświadczają wielu (żeby nie powiedzieć – większości) form z paradygmatu danego zwrotu. I tak na przykład dla zwrotu *{ktoś} kupuje kota w worku* w słowniku przyjmujemy pełny paradygmat (zgodny z paradygmatem czasownika *kupować*). Kwerenda w NKJP dostarcza w podkorpusie zrównoważonym poświadczeń dla jedynie 11 form (w tym dla kilku po zaledwie jednym przykładzie). Nie znajdujemy w nim zatem między innymi przykładów użycia omawianego zwrotu w czasie przyszłym i trybie przypuszczającym, 2. os. lm. trybu rozkazującego, zaś czas przeszły jest reprezentowany jedynie przez formę 3 os. lm. rodz. mos. Z kolei cały korpus NKJP zawiera przykłady użycia zwrotu we wszystkich formach czasu teraźniejszego i trybu rozkazującego oraz po jednej formie (3 os. lp. rodz. męs.) dla czasu przyszłego i trybu warunkowego. Takie wyniki w żaden sposób nie przesądzają o pełności / defektywności paradygmatu.

Jako przykład sytuacji, w której korpus jest pomocny w weryfikowaniu hipotez i rozstrzygnięciu wątpliwości, „podpowiadając” rozwiązanie przynajmniej co do niektórych form paradygmatu, może posłużyć przywoływany już wyżej zwrot *{ktoś} ko-*

⁷ Oczywiście posłużyć może on także do wskazania form i kontekstów typowych, częstych, a więc do uzyskania pewnych informacji pragmatycznych. Informacje takie są jednak w tworzonej słowniku zdecydowanie drugoplanowe, np. danych o frekwencji nie podajemy w sposób systematyczny, niekiedy zamieszczamy jedynie uwagę w komentarzach.

szulę w zębach nosił ('ktoś był młody w czasach, o których mowa'), który według WSJP ma paradygmat defektywny – nie jest używany w czasie teraźniejszym i przyszłym. Kwerenda korpusowa dostarcza przekonujących przykładów użycia opisywanego frazeologizmu w czasie teraźniejszym, których zgodnie z informacjami zawartymi w WSJP i ISJP nie należałoby się spodziewać:

(3) Przeżegnajże się lewą ręką, dziewucho! A cóż jabym u Pazderskich robiła? Stara užera się od rana do nocy, czeladzie porozpędzali, same chłopaki, sama koprowina, co jeszcze w zębach koszule nosi. (https://pl.wikisource.org/wiki/Pod_prawem/I)

(4) Spać, smyki! – zawołał Pantera ze swego legowiska. – Koszulę to jeszcze w zębach nosi, a już spiskuje. (NKJP; M. Rodziewiczówna, *Lato leśnych ludzi*)

Formy te nie wchodzą w konflikt z semantyką jednostki, podobnie jak formy bezokolicznika:

(5) O, widzicie ją jaka śmiała! A zdawać by się mogło, że powinna jeszcze koszulę w zębach nosić? (jaswily.pl/gmina/pliki/wesele.doc)

(6) Przestałem już nosić koszulę w zębach i dostałem swój pierwszy spinning. (<https://www.scribd.com/document/357862607/Ota-Pavel-S-mierc-pie-knych-saren>)

A zatem paradygmat interesującej nas jednostki nie ogranicza się do form czasu przeszłego, chociaż one zdecydowanie przeważają. Akceptowalne są, choć przez użytkowników używane sporadycznie, formy czasu teraźniejszego i bezokolicznik. Brak natomiast czasu przyszłego, a także form trybu rozkazującego, przypuszczającego, imiesłowu i bezosobnika (tych informacji słowniki już nie podają).

Inny przykład tej samej sytuacji to frazeologizm *{coś} spaliło na panewce*, który ma według Andrzeja Markowskiego tylko czas przeszły (Markowski 2005: 235). W opracowaniu *Kultura języka polskiego* błędną formę zilustrowano następującym cytatem:

(7) *Tym samym jeden z celów wprowadzenia systemu, czyli usprawnienie pracy wydziałów komunikacji, spali na panewce.

Mamy tu więc sytuację użycia czasu przyszłego dla stwierdzenia, że coś z pewnością się nie uda. Jednakże kwerenda korpusowa ujawnia również takie przykłady, w których gramatyczna forma *futurum* jest użyta w syntaktycznym i pragmatycznym kontekście np. do wyrażenia wątpliwości, warunku, por.

(8) I diabli wiedzą, co robić, bo na gwałt muszę jeszcze skądś wytrzasnąć ze sto tysięcy. Inaczej wszystko spali na panewce. (NKJP; T. Dołęga-Mostowicz, *Drugie życie doktora Murka*)

(9) Teraz wszystko zależy od rodzaju zamków. Jeśli będą skomplikowane, pewnie cały plan spali na panewce. (NKJP; J. Głębski, *Kuracja*)

Same formy gramatyczne istnieją zatem w paradygmacie zwrotu, choć ich użycie ograniczone jest do niepodstawowej funkcji *futurum*. Tworząc słownik paradygmatów, uświadamialiśmy sobie niejednokrotnie, że skłonni jesteśmy, tak jak w wyżej przy-

wołanym opracowaniu, wykluczyć jakieś formy, bo nie znajdowaliśmy dla nich uzasadnienia w ciągach separowanych lub kongruentnych. Przykłady w rodzaju *spalić na panewce* ujawniają pewien dodatkowy problem przy tworzeniu słownika fleksyjnego: podaje on formy, które izolowane mogą budzić sprzeciw użytkownika, ale użyte w określonym kontekście są całkowicie zgodne z normą. Korpus takie właśnie konteksty niewątpliwie pomaga znaleźć. Podane przykłady potwierdzają zatem, z jednej strony, przydatność *corpus-based approach* w tworzeniu omawianego typu słownika, jego użyteczność w sprawdzaniu hipotez, z drugiej – pokazują konieczność weryfikacji uzyskanych danych na podstawie własnej kompetencji (analizy własności semantyczno-strukturalnych zwrotu). Takie refleksje na temat źródeł korpusowych w pracy nad opisywanym słownikiem nie są zaskakujące. Potwierdzają niewątpliwie stanowisko tych językoznawców, którzy twierdzą, że badacz, nie korpus, jest instancją rozstrzygającą, interpretatorem źródeł (zob. też uwagi Bańki (2001) na temat ćwiczenia z wyrażeniem *nie bez kozery*).

3. Podsumowanie

W artykule staraliśmy się opisać stopień przydatności różnych źródeł w tworzeniu słownika paradygmatów zwrotów frazeologicznych. Pierwsze z wymienionych na wstępie – metaźródła słownikowe – mają niewątpliwie zastosowanie jako istotny element w ustalaniu / weryfikowaniu listy jednostek. Drugie – źródła korpusowe – poza funkcją egzemplifikacyjną, były przez nas wykorzystywane w weryfikowaniu hipotez, choć w ograniczony sposób. Ograniczenia te (także przy preferowanym tu podejściu *corpus-based approach*) wynikają w sposób oczywisty z charakteru słownika. Bańko (2001: 43) w odniesieniu do słownika ogólnego zadał pytanie, czy w zakresie odmiany słowniki powinny odzwierciedlać system czy uzus. Odpowiedział na nie następująco: „słowniki mają odzwierciedlać system, ale nie powinny eksponować form nie poświadczonych w korpusie”. W przypadku słownika gramatycznego nie da się zastosować tej zasady, a więc różnicować statusu form, skoro paradygmaty w założeniu mają być podane wprost. Po drugie, jeśli przyjmujemy stanowisko systemowe, o włączeniu lub nie danej formy do paradygmatu rozstrzyga analiza właściwości znaczeniowo-strukturalnych zwrotu. Korpus natomiast może posłużyć do podania pewnych elementów charakterystyki pragmatycznej i znalezienia kontekstów nieprzewidzianych przez badacza (czy nawet z jego perspektywy zaskakujących). Jak zauważa Patrick Hanks (1990: 40), języki naturalne pełne są nieprzewidzianych faktów, które korpusy pomagają zidentyfikować. Jednocześnie stwierdzić trzeba, że korpusy lepiej niż w opisie gramatycznym sprawdzają się w badaniach sposobu użycia i frekwencji danej jednostki. Względem korpusu nie zajmujemy zatem żadnej z dwóch postaw skrajnych: tj. z jednej strony nie uznajemy introspekcji za jedyne wiarygodne narzędzie językoznawcy, a kwerendy korpusowej za „zbieranie motyli” nie posuwające wiedzy językowej badacza (jak twierdzi N. Chomsky), a z drugiej – nie kierujemy się absolutnym zaufaniem do korpusu i traktowaniem introspekcji jako najbardziej

ułomnego narzędzia lingwistycznego (jak twierdzą np. przedstawiciele kognitywizmu, por. np. Tummers, Heylen, Geeraerts 2005).

Na zakończenie można zadać sobie pytanie: czy sam korpus (w obecnej postaci) i źródła słownikowe wzbogacone o opracowania naukowe z dziedziny frazeologii są wystarczające, żeby stworzyć (opracować) słownik fleksyjny frazeologizmów? Odpowiedź musi brzmieć: nie. Z jednej strony wynika to z braku poświadczeń stanowiącej większości form w korpusach i milczenia słowników co do ich statusu. Z drugiej strony, o ile metaźródła w postaci słowników dostarczają nam „opracowanych” informacji na temat niektórych form fleksyjnych, o tyle sam korpus – tylko informacji o wystąpieniu form w tekstach. W żaden sposób nie rozstrzyga, czy było to użycie przewidywane przez system, czy innowacja językowa funkcjonalnie nieuzasadniona, czy wreszcie chociażby metajęzykowe użycie danego ciągu.

Podsumowując, w tworzeniu fleksyjnego słownika zwrotów większość decyzji co do statusu danej formy jest oparta właściwie wyłącznie na kompetencji językowej leksykografów, a nawet tam, gdzie źródła słownikowe czy opracowania naukowe wskazują status danej formy, każdorazowo powinien on być weryfikowany w oparciu o inne źródła (w tym korpusowe) i kompetencję językową słownikarzy.

Bibliografia

- BAŃKO Mirosław (2001): *Z pogranicza leksykografii i językoznawstwa. Studia o słowniku jednojęzycznym*. – Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.
- BAŃKO Mirosław (2012): NKJP w oczach leksykografa. – [w:] Adam PRZEPÍORKOWSKI, Mirosław BAŃKO, Rafał L. GÓRSKI, Barbara LEWANDOWSKA-TOMASZCZYK (red.): *Narodowy Korpus Języka Polskiego*. – Warszawa: PWN, 283–290.
- BĄBA Stanisław (1998): Uwagi o odmianie polskich zwrotów przez czasy. – *Poznańskie Spotkania Językoznawcze* III, 7–14.
- BIBER Douglas (2015): Corpus-based and corpus-driven analyses of language variation and use. – [w:] Bernd HEINE, Heiko NARROG (red.): *The Oxford Handbook of Linguistic Analysis*. – Oxford, New York: Oxford University Press, 159–191.
- BOGUSŁAWSKI Andrzej (1976): O zasadach rejestracji jednostek języka. – *Poradnik Językowy* 8, 356–364.
- BOGUSŁAWSKI Andrzej, GARNYSZ-KOZŁOWSKA Teresa (1979): *Addenda do frazeologii polskiej. Zeszyt wstępny*. – Edmonton: Carbondale.
- BOGUSŁAWSKI Andrzej, WAWRZYŃCZYK Jan (1993): *Polszczyzna jaką znamy. Nowa sonda słownikowa*. – Warszawa: Uniwersytet Warszawski.
- BOGUSŁAWSKI Andrzej, DANIELEWICZOWA Magdalena (2005): *Verba polona abscondita. Sonda słownikowa III*. – Warszawa: Elma Books.
- CONRAD Susan (2010): What can a corpus tell us about grammar? – [w:] Anne O'KEEFFE, Michael MCCARTHY (red.): *The Routledge Handbook of Corpus Linguistics*. – London–New York, NY: Routledge, 227–240.
- CZEREPOWICKA Monika (2014): SEJF – Słownik elektroniczny jednostek frazeologicznych. – *Język Polski* XCIV, 1, 116–129.
- CZEREPOWICKA Monika, KOSEK Iwona, PRZYBYSEWSKI Sebastian (2014): O projekcie elektronicznego słownika odmiany frazeologizmów czasownikowych. – *Polonica* XXXIV, 115–123.

- ČERMÁK František (2003): Source materials for dictionaries. – [w:] Piet VAN STERKENBURG (red.): *A practical guide to lexicography*. – Amsterdam–Philadelphia: John Benjamins Publishing, 18–25.
- GÓRSKI Rafał L. (2012): Zastosowanie korpusów w badaniu gramatyki. – [w:] Adam PRZEPIÓRKOWSKI, Mirosław BAŃKO, Rafał L. GÓRSKI, Barbara LEWANDOWSKA-TOMASZCZYK (red.): *Narodowy Korpus Języka Polskiego*. – Warszawa: PWN, 291–299.
- GRABOWSKI Łukasz (2015): O frazeologii z perspektywy językoznawstwa korpusowego. Przegląd głównych nurtów badawczych ostatniego dwudziestolecia w Wielkiej Brytanii i USA. – *Problemy frazeologii europejskiej* X, Łask: Oficyna Wydawnicza LEKSEM, 23–48, online: <http://pfe.leksem.pl/wp-content/uploads/Grabowski-1.pdf>.
- HANKS Patrick (1990): Evidence and intuition in lexicography. – [w:] Jerzy TOMASZCZYK, Barbara LEWANDOWSKA-TOMASZCZYK (red.): *Meaning and Lexicography*. – Amsterdam–Philadelphia: John Benjamins Publishing Company, 31–41.
- KILGARIFF Adam (2013): Using corpora as data source for dictionaries. – [w:] Howard JACKSON (red.): *The Bloomsbury companion to lexicography*. – London–New York: Bloomsbury Academic, 77–96.
- KOSEK Iwona (2013): Paradygmaty zwrotów frazeologicznych – problemy opisu leksykograficznego. – [w:] Gabriela DZIAMSKA-LENART, Jarosław LIBEREK (red.): *Perspektywy współczesnej frazeologii polskiej. Między teorią a praktyką frazeograficzną*. – Poznań: Wydawnictwo Naukowe UAM, 51–61.
- LEWICKI Andrzej Maria (1976): *Wprowadzenie do frazeologii syntaktycznej. Teoria zwrotu frazeologicznego*. – Katowice: Uniwersytet Śląski.
- LEWICKI Andrzej Maria (1986): Składnia związków frazeologicznych. – *Biuletyn Polskiego Towarzystwa Językoznawczego* XL, 75–82.
- LEWICKI Andrzej Maria, PAJDZIŃSKA Anna (2001): Frazeologia – [w:] Jerzy BARTMIŃSKI (red.): *Współczesny język polski*. – Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, 315–333.
- LEWICKI Andrzej Maria, PAJDZIŃSKA Anna, REJAKOWA Bożena (1987): *Z zagadnień frazeologii. Problemy leksykograficzne*. – Warszawa: PWN.
- LIBEREK Jarosław (2006): Różnicowanie aspektowe komponentów werbalnych w stałych związkach frazeologicznych. Uwagi wstępne. – [w:] Stanisław BĄBA, Piotr FLICIŃSKI, Jarosław LIBEREK (red.): *Z zagadnień frazeologii, stylistyki i kultury języka*. – Poznań: Wydawnictwo Poznańskie.
- LUBAŚ Władysław, red. (1988): *Wokół słownika współczesnego języka polskiego. Materiały konferencji w Paszówce 26–28 XI 1986 r.* – Wrocław: Ossolineum, PAN.
- MARCINIAK Małgorzata, RABIEGA-WIŚNIEWSKA Joanna, SAVARY Agata, WOLIŃSKI Marcin, HELIASZ Celina (2009): Constructing an Electronic Dictionary of Polish Urban Proper Names. – [w:] *Recent Advances in Intelligent Information Systems (Proceedings of the Balto-Slavonic Natural Language Processing Workshop, Kraków)*. – Warszawa: Oficyna Wydawnicza EXIT, 743–749.
- MARKOWSKI Andrzej (2005): *Kultura języka polskiego. Zagadnienia leksykalne*. – Warszawa: PWN.
- MOON Rosamund (1998): *Fixed Expressions and Idioms in English. A Corpus-Based Approach*. – Oxford: Oxford University Press.
- PRZYBYSZEWSKI Sebastian (2015): Some problems with the description of paradigms of Polish verbal multi-word units. – [w:] Enrique GUTIÉRREZ RUBIO, Marta FALKOWSKA, Ekaterina KISLOVA, Marzena STĘPIEŃ (red.): *Beiträge der Europäischen Slavistischen Linguistik (POLYSLAV)*, t. 18, Wiesbaden: Harrassowitz Verlag, 213–223.
- SALONI Zygmunt (2001): *Czasownik polski. Odmiana, słownik*. – Warszawa: PWN.
- SINCLAIR John (1995): *Corpus, concordance, collocation*. – Oxford: Oxford University Press.
- TUMMERS Jose, HEYLEN Kris, GEERAERTS Dirk (2005): Usage-based approaches in Cognitive Linguistics: A technical state of the art. – *Corpus Linguistics and Linguistic Theory* 1 (2), 225–261.
- WALTER Elizabeth (2010): Using corpora to write dictionaries. – [w:] Anne O'KEEFE i Michael MCCARTHY (red.): *The Routledge Handbook of Corpus Linguistics*. – London–New York, NY: Routledge, 428–443.
- ŻMIGRODZKI Piotr, opr. (2015): *Wielki słownik języka polskiego PAN. Zasady opracowania*, wyd. II. – Kraków.

Słowniki i korpusy

- ISJP – Mirosław BAŃKO, red. (2000): *Inny słownik języka polskiego*, I–II. – Warszawa: PWN.
- NKJP – Adam PRZEPióRKOWSKI, Mirosław BAŃKO, Rafał. L. GóRSKI, Barbara LEWANDOWSKA-TOMASZ-CZYK (red.): *Narodowy Korpus Języka Polskiego*. – Warszawa: PWN; <http://www.nkjp.pl>.
- PSF – Stanisław BĄBA, Jarosław LIBEREK, Gabriela DZIAMSKA-LENART, red. (2001): *Podręczny słownik frazeologiczny języka polskiego*. – Warszawa: PWN.
- SFJP – Stanisław SKORUPKA (1985): *Słownik frazeologiczny języka polskiego*. – Warszawa: Wiedza Powszechna, wydanie IV.
- SFWP – Stanisław BĄBA, Jarosław LIBEREK, red. (2001): *Słownik frazeologiczny współczesnej polszczyzny*, Warszawa: PWN.
- SGJP – Zygmunt SALONI, Marcin WOLIŃSKI, Robert WOŁOSZ, Włodzimierz GRUSZCZYŃSKI, Danuta SKOWROŃSKA, red. (2015): *Słownik gramatyczny języka polskiego*. Wyd. III on-line, <http://sgjp.pl>.
- SJPD – Witold DOROSZEWSKI, red. (1958-1969): *Słownik języka polskiego*, I–XI. – Warszawa: PWN.
- SJPSz – Mieczysław SZYMCZAK, red. (1996): *Słownik języka polskiego*, I–III. – Warszawa: PWN.
- SWJP – Bogusław DUNAJ, red. (1996): *Słownik współczesnego języka polskiego*. – Warszawa: Wydawnictwo Wilga.
- USJP – Stanisław DUBISZ, red. (2003): *Uniwersalny słownik języka polskiego*, I–IV. – Warszawa: PWN.
- WSF – Anna KŁOSIŃSKA, Elżbieta SOBÓL, Anna STANKIEWICZ, red. (2005): *Wielki słownik frazeologiczny z przysłowiami*. – Warszawa: PWN.
- WSFJP – Piotr MÜLDNER-NIECKOWSKI (2003): *Wielki słownik frazeologiczny języka polskiego*. – Warszawa: Świat Książki.
- WSJP – Piotr ŻMIGRODZKI, red. nauk.: *Wielki słownik języka polskiego*, <http://www.wsjp.pl>

doi: <http://dx.doi.org/10.17651/bptj.74.4>

MICHAIL L. KOTIN

Uniwersytet Zielonogórski, Zielona Góra

Möglichkeiten und Grenzen der korpusbasierten Grammatikforschung

Możliwości i ograniczenia w zastosowaniu korpusów w badaniach nad gramatyką.

Streszczenie

W centrum uwagi artykułu znajdują się korelacje pomiędzy introspekcją, ankietowaniem oraz analizą opierającą się na korpusach tekstowych. Metodologia „types and tokens” jest dziś uważana za najbardziej współczesną i powszechnie stosowaną. Jednocześnie brak uwzględnienia współzależności między częstotliwością wystąpień (tokenów) a profilem poszczególnych tekstów tworzących całościowy korpus, który bezpośrednio wpływa na specyfikę użycia badanych form gramatycznych, a z drugiej strony, brak uwzględnienia czynników „pozakorpusowych” (przede wszystkim introspekcji, jako wymiennego kryterium kompetencji językowej), wywołuje niekiedy poważne kolizje między wynikami analizy korpusowej a rzeczywistym stanem rzeczy w sferze funkcjonowania kategorii gramatycznych. Kolejnym współczynnikiem, który zazwyczaj nie jest brany pod uwagę w ilościowych badaniach korpusowych, są specyficzne zastosowania analizowanych form (implikatury konwencjonalne), niekonieczne uwarunkowane kontekstem, który może zostać rzeczywiście wyłoniiony nawet z dobrze opracowanego korpusu, jako relewantny w stosunku do określenia funkcji odpowiedniej jednostki bądź formy gramatycznej.

Słowa kluczowe: lingwistyka korpusowa, introspekcja, kompetencja, typy, tokeny

Potentials and limitations of corpus-based grammar research

Abstract

The article deals with selected aspects of a corpus-based investigation of the categorial functions of language entities in comparison with other research methods, i.e. introspection and native speaker's questioning. The types-and-tokens methodology is nowadays considered most modern and thus it is commonly used. At the same time, the ignorance of the interdependence between the frequency of tokens and the individual profiles of texts which constitute the total corpus often leads to essential collisions between the

results of the corpus-based analyses and the real categorial functioning of the forms in question. On the other hand, the ignorance of corpus-independent factors (in particular, introspection as an essential criterion of language competence) sometimes triggers serious research discrepancies as well. A further factor which is often ignored in empirical corpus-based research are the specific uses of the analysed forms in an utterance (conventional implications) which do not always depend on the contexts which form the entity- and utterance-relevant corpus in question.

Keywords: corpus linguistics, introspection, competence, types, tokens

Die zunehmende Favourisierung von Untersuchungen, welche auf immer umfangreicheren Korpora basieren, hat eine nahezu unveräußerliche Anwendung von Methoden und Techniken der Korpusanalyse in einem jeden neuen Forschungsprojekt herbeigeführt. Zu den unmittelbaren Forschungsaufgaben werden heute immer häufiger nicht nur analytische korpusbasierte Abhandlungen gezählt, sondern auch die Aufstellung annotierter elektronischer Textkorpora als solche (vgl. Lüdeling, Kytö 2008-2009, Kratochvilová, Wolf 2010). Die erstellten Korpora sollen dann sowohl Linguisten als auch Spezialisten anderer Forschungsdisziplinen zur Verfügung gestellt werden, wobei die Korpora dermaßen aufgestellt und annotiert werden müssen, dass sie möglichst breiten Forscherkreisen als Textgrundlage dienen könnten, d.h. keine zu eng geschnittene bzw. für die Lösung von nur bestimmten Aufgaben konzipierte Form haben dürfen. Daher wird diese Arbeit verständlicherweise in Kooperation von Sprach- und Informationswissenschaftlern geleistet. Letztere erhalten von den Linguisten nicht selten gewisse Auflagen, auf deren Basis sie dann sowohl Textkorpora als auch deren mannigfaltige Verlinkungen bzw. Verweise ausarbeiten, welche einem schnelleren Zugang zu den für weitere Forschungszwecke erforderlichen Daten bewerkstelligen. Da statistische Daten bei dieser Arbeit äußerst wichtig sind, werden die Korpora dermaßen aufgestellt, dass der Umfang des Untersuchungsstoffs grundsätzlich keine Zweifel bezüglich ihrer quantitativen Suffizienz zulässt. Daher ist das übergeordnete Ziel stets eine Erstellung von möglichst umfangreichen Korpora, welche mehrere Millionen Entitäten zählen. Dies ist nun ein Grund dafür, dass einer der Hauptbegriffe der Korpuslinguistik der Begriff der Tokens ist, d.h. der Gesamtzahl von Okkurrenzen der zu analysierenden Entitäten, welche im gesamten Korpusumfang vorkommen. Die Vorkommenshäufigkeit sprachlicher Einheiten und Konstruktionen in einem möglichst umfangreichen „Korpusganzen“ gilt bei diesem Herangehen als das sicherste empirische Kriterium. Dagegen ist der Begriff der Typen, d.h. der typischen, abstrakten Modelle, welche den Inbegriff der Tokens ausmachen, immer öfter als eine im Vergleich zu den Tokens untergeordnete und sekundäre Kategorie behandelt.

Jede Interpretation des Funktionierens der zu untersuchenden grammatischen Sprachform oder -konstruktion erfordert somit eine angemessene solide empirische Belegbasis in Form statistisch erfasster Gebrauchsfrequenzen von Tokens, welche die jeweiligen sprachlichen Phänomene repräsentieren. Eine dermaßen rigoros formulierte Forderung bringt dabei allerdings nicht nur evidente Vorteile mit sich, sondern

kann auch in vielen Fällen Nebenwirkungen auslösen, die paradoxerweise die Glaubwürdigkeit bzw. Zuverlässigkeit der Forschungsergebnisse nicht erhöhen, sondern vielmehr vermindern. Ungewollte Folgen der Forschungsphilosophie, die die Korpora für die glaubwürdigste und zuverlässigste Quelle einer jeden Analyse erklärt, sind durch eine ganze Reihe von Faktoren bedingt, die den Gegenstand der nachfolgenden Überlegungen ausmachen. Es handelt sich hierbei vor allem um die Korrelation zwischen drei Methoden der empirischen Analyse, die sich aus der jeweiligen Quelle ergeben, der der empirische Stoff entspringt, nämlich: a) Introspektion, die auf individueller Sprachkompetenz der Sprachträger beruht; b) Informantenbefragung, oder Fragebogenmethode und c) Analyse, die sich auf Textkorpora stützt (vgl. Kratochvilová, Wolf 2010: 3-4).

Der letzteren Methode liegen entsprechend vorbereitete und für weitere Forschung angemessene Korpora zu Grunde, die heutzutage vorwiegend in elektronischer Form vorliegen und in aller Regel sehr umfangreich sind. Die Analyse als solche besteht in den meisten Fällen in der Auswertung der Gebrauchsfrequenz der zu untersuchenden Sprachformen (so genannter *Tokens*) im Vergleich mit deren Eigenschaften von allgemein-kategorialem Charakter (so genannter *types*, also Typen). Die *types- and tokens*-Methodologie gilt gegenwärtig als „modern“ und wird so gut wie überall verwendet. Zugleich löst die fehlende Berücksichtigung der Wechselbeziehungen zwischen der Tokenfrequenz und dem Profil der ein Korpus bildenden Texte, das einen unmittelbaren Einfluss auf die Spezifik der Verwendung der zu untersuchenden grammatischen Formen hat, nicht selten schwerwiegende Kollisionen zwischen den Ergebnissen der Korpusanalyse und dem „Ist-Stand“ im Kategorisierungsbereich der grammatischen Formen aus. Diese Diskrepanzen werden zusätzlich durch eine fehlende Berücksichtigung „außerkorporaler“ Faktoren (vor allem der Introspektion als eines essentiellen Kriteriums sprachlicher Kompetenz) verstärkt.

Fangen wir mit Introspektion an, d.h. mit der Bewertung der Zulässigkeit (Grammatikalität) einer gegebenen Sprachform vom Standpunkt sowohl des Sprachsystems als auch der Sprachnorm, welche durch den Sprachforscher selbst vorgenommen wird, wobei es sich um einen Muttersprachler der zu untersuchenden Sprache handelt oder zumindest um eine nahezu muttersprachlich kompetente Person. Noam Chomsky spricht hier von angeborenen grundlegenden Sprachkompetenzen, allen voran der syntaktischen Kompetenz, d.h. der Fähigkeit, eine Nominalphrase (NP) mit einer Verbalphrase (VP) zu verbinden (vgl. Chomsky 1996: 30). Wie sicher bzw. wie glaubwürdig ist nun Introspektion? Die Antwort auf diese Frage ist bei weitem nicht selbstverständlich. Die Korpuslinguistik setzt per definitionem ihrer Existenz voraus, dass Korpusdaten viel „objektiver“ als individuelle Introspektion sind und daher die Rolle dieser letzteren, welche stets subjektiv und somit relativ ist, naturgemäß abschwächen. Nicht selten hört man von Verfechtern des Alleinvertretungsanspruchs der Korpuslinguistik, dass das, was die Benutzer der Introspektionsmethode für falsch halten, in Wirklichkeit im Korpus durchaus vorkommt. Die Frage bleibt dennoch dieselbe, nämlich: Was folgt aus dieser Tatsache bzw. wie soll sie von den Sprachforschern interpretiert werden? Darf das Korpus Beispiele mit sog. Performationsfehlern haben?

Freilich muss diese Frage affirmativ beantwortet werden. Insbesondere gilt dies für immer häufiger aufgestellte Internetkorpora oder z.B. für Transkriptionen gesprochener Texte. Die nächste Frage betrifft den „Status“ dieser Fehler. Sind das System- oder Normfehler? Wovon zeugen sie? Geht es um Fehler, die wegen Vernachlässigung der Sprachnorm entstanden sind, also nicht normgerechte Aussprache oder Schreibweise etc., welche der Sprechende bzw. Schreibende einfach nicht bemerkt hat, oder aber um Fehler, die vom Sprachbenutzer nicht als fehlerhafte Verwendungen eingestuft werden, da er sie für richtig hält? Im letzten Fall muss wiederum zwischen dem Fehlen angemessener Sprachkenntnisse und der Tendenz in der Sprachentwicklung differenziert werden, in deren Folge die Fehler von heute zu den Normen von morgen werden. Empirische Korpusstudien allein geben keine Antwort auf die oben gestellten Fragen. Sie können jedoch durchaus beantwortet werden, wenn die Beobachtung der Dynamik sprachlicher Prozesse bzw. ihrer Entwicklungstendenzen mit individueller Introspektion verschränkt wird, da diese letztere festzustellen erlaubt, inwiefern diese oder jene Form vom System der Muttersprache toleriert und somit als Norm- bzw. Ususvariante behandelt werden kann oder aber als fehlerhaft bzw. unkorrekt aus der Sicht der in der jeweiligen Sprache existierenden Basisgrundlagen verworfen werden muss. Ein weiteres wichtiges Problem ist der Status von Beispielen, die vom jeweiligen Sprachbenutzer als potentielle Korpusbelege eingestuft werden. Eine gute Illustration hierfür ist die Antwort eines Referenten auf einer Fachtagung auf die Kritik seitens seines Opponenten, seine Beispiele seien lediglich selbst konstruiert und könnten daher kaum als echte empirische Objektbeispiele gelten, da sie nicht aus einem Korpus exzerpiert seien: „Als Muttersprachler bitte ich Sie nun, meine Beispiele in das Korpus aufzunehmen!“ Die Antwort war im Grunde richtig, da jeder Muttersprachler natürlich mit seiner Sprachkompetenz für die Korrektheit seiner eigenen „Belege“ bürgt. Das einzige, was hier noch bezweifelt werden könnte, wäre die Verwendung der Einzelbelege ohne Kontext und eben lediglich als Demonstrationsbeispiele und nicht etwa als Ergebnis einer nicht fachlich intendierten, freien *parole*. Dennoch bleibt die Frage offen, wieso fremde Belege grundsätzlich anders als Eigenbelege gewichtet werden müssen. Insgesamt sind alle Versuche, durch eine künstliche „Objektverfremdung“ die „Objektivität“ der Ergebnisse zu erhöhen bzw. die Subjektivität zu vermindern, gerade in der Sprachforschung stets fraglich. Die Spezifik des Forschungsobjekts Sprache liegt nämlich darin, dass Forschungssubjekt und Forschungsobjekt maximal nahe sind und daher nicht künstlich getrennt werden können. Es geht letztendlich um eine spezifische, symbolische Form der Bewusstseinstätigkeit, bei deren Beobachtung und Deskription wiederum Eigenreflexionen des Forschers nie völlig ausgeschlossen werden können. Anders kann über Sprache als symbolische Bewusstseinsform gar nicht gesprochen werden (vgl. Mamardašvili, Pjatiorskij 2011: 24-39). Chomsky (2002: 98) stellt fest, die Datensammlung via große Textkorpora habe erst dann Sinn, wenn dahinter eine bereits formulierte Theorie stehe, welche mit Hilfe empirischer Daten überprüft werden solle. Wenn jedoch, wie oben bereits erwähnt wurde, die Korpusaufstellung und -annotation als eine selbständige, wenngleich prätheoretische, Forschungsaufgabe behandelt werden, sei dieser metho-

dologische Anspruch laut Chomsky generell abwegig. Die Aufstellung von Hypothesen sollte demnach der Korpusaufstellung vorausgehen, und dies bedeutet seinerseits, dass Introspektion – welche ja offenkundig notgedrungen der erste Schritt dabei sein müsse – logischerweise im Vorgriff auf weitere Schritte, darunter auf die Sammlung empirischer Daten, vorzunehmen sei.

Außer der individuellen Introspektion, die gegen die herrschende Skepsis ein durchaus sicheres, wissenschaftlich fundiertes und unersetzbares Instrument der Analyse von sprachlichen Entitäten und Strukturen ist, wird in der sprachwissenschaftlichen Analyse die Befragung der Muttersprachler benutzt. Abhängig von einer Reihe von Faktoren kann auch diese Methode wesentliche Ergebnisse hinsichtlich der Spezifik des Funktionierens und der Entwicklungstendenzen der zu untersuchenden sprachlichen Einheiten ausweisen. Das dabei vom Sprachforscher gesetzte Ziel kann allerdings erst dann in vollem Ausmaß erreicht werden, wenn der Fragebogen korrekte und methodologisch einwandfreie Formulierungen enthält, die nicht nur mit der gestellten Forschungsaufgabe einher gehen, sondern auch die Möglichkeit einer angemessenen und methodologisch korrekten Interpretation der Antworten von Respondenten garantieren. Ferner sind in diesem Fall (ebenso wie bei Korpusstudien) die Anzahl der Befragten, ihre entsprechende regionale, soziale und altersbezogene Stratifikation sowie die Berücksichtigung der Differenzen in ihrer Ausbildung, ihren Wohnorten, ihrer Mobilität, der Teilnahme an Sprachkontakten etc. erforderlich. Im Grunde genommen ist die Informantenbefragung eines der Verfahren einer Objektivierung individueller Introspektion des Forschers durch deren Vergleich mit der Introspektion anderer Muttersprachler. Der epistemologische Wert der Ergebnisse der infolge der Befragung gesammelten Fragebogen bezüglich unterschiedlicher Forschungsziele und anhängig von der jeweiligen Eigenart der zu untersuchenden Entitäten und Strukturen kann u.U. sehr hoch sein. Es handelt sich nämlich um eine spezifische Form einer „gemeinschaftlichen Introspektion“ durch Kontrastierung diverser individueller Sprachkompetenzen, d.h. in der Terminologie der anthropozentrischen Linguistik der Warschauer Schule, der jeweiligen Idiolekte (vgl. F. Grucza 1997: 11, S. Grucza 2011: 151) von Personen, die sich derselben „Sprache“, d.h. desselben Polylekts bedienen.

Zurückkehrend zur Korpusanalyse, diesmal jedoch in ihrem Vergleich mit Introspektion und Befragung von Muttersprachlern, kann man feststellen, dass sie im Gegensatz zu den sonstigen Methoden einen anderen Ausgangspunkt für empirische Forschungen hat, sodass sich weder andere Analysemethoden durch Korpusanalyse ersetzen lassen noch vice versa. Vom Standpunkt des Sprachwesens haben wir hier hierbei mit unterschiedlichen Erscheinungsformen der Sprache zu tun. Bei der Untersuchung umfangreicher Korpora beschäftigen wir uns vor allem mit mehr oder minder ausschöpfender Fixierung der analysierten Entitäten in deren tatsächlicher textueller Umgebung. Bei der Anwendung der Introspektionsmethode oder der Informantenbefragung werden dagegen potentielle Kontexte festgelegt, die um die zu untersuchende Entität herum entstehen, bzw. sämtliche möglichen Erscheinungsvarianten dieser Entität. Dies ist die Methode der Ermittlung des Potentials einer Sprachform

(darunter einer grammatischen Form), wohingegen die Korpusanalyse ausschließlich auf einer realen, „schon stattgefundenen“ Verwendung der jeweiligen Einheit auf der Basis der Ermittlung der Gesamtzahl der Tokens beruht. Bei der Anwendung der Introspektion oder der Informantenbefragung werden daher *potentielle* Kontexte untersucht, während die Korpusanalyse *reale* Kontexte liefert, in denen die Entität *bereits verwendet* worden ist. Dabei gibt es keine Garantie, dass (i) alle registrierten Verwendungsfälle „korrekt“, d.h. systemkonform und normgerecht sind, und (ii) andere Verwendungen nicht zu kurz kommen, welche in einem sogar sehr umfangreichen Korpus nicht registriert wurden, dabei aber vom Standpunkt eines Muttersprachlers (Introspektion) oder mehrerer Muttersprachler (Informantenbefragung) durchaus zulässig sind.

Ein weiterer Aspekt, der gewöhnlich in quantitativen Korpusstudien missachtet wird, sind spezifische Verwendungen der analysierten Sprachformen (konventionelle Implikaturen), die in einem unterschiedlichen Grad kontextbedingt sind. Bezüglich dieses Problems spricht V. Ágel (2006) von der Opposition von Korpus- und Textprofilanalyse. Die letztere dient nämlich einer Ausfüllung der Lakunen in der quantitativen korpusbasierten Tokenfrequenzanalyse. Das Korpus ist nämlich normalerweise nicht imstande, einige Grundfragen zu beantworten, welche unumstritten für eine jede Untersuchung von Bedeutung sind, die mit der Funktionsgeltung grammatischer Formen im Zusammenhang steht. Zwar ist es möglich sich vorzustellen, dass eine detaillierte, außerordentlich minutiöse Annotation eines Belegkorpus die Ermittlung und Berücksichtigung der Verschränkungen zwischen dem Texttyp bzw. der Textsorte und der Eigenart der Verwendung grammatischer Strukturen darin gewährleisten kann, aber eine vollständige Erfassung eines jeden ins Korpus aufgenommenen partiellen Textprofils ist absolut ausgeschlossen.

Hier ein Beispiel, das diese These partiell illustrieren kann. Der Satz *Die Pommes müssen schon auf dem Tablett liegen* kann je nach Situation folgende Lesarten haben: (i) die Anordnung des Küchenchefs für den Kellner, die Pommes sofort auf das Tablett zu legen, (ii) die Reaktion eines unzufriedenen Kunden auf das Fehlen der bestellten Pommes auf dem Tablett oder (iii) Vermutung des Sprechers, dass die Pommes auf dem Tablett liegen. In beiden ersten Fällen geht es um verschiedene konversationelle Implikaturen beim Beibehalten der Grundbedeutung des Modalverbs, d.h. um so genannte deontische Modalität bzw. „Wurzelmodalität“ (root modality) (vgl. u.a. Palmer 1986: 102, de Haan 1997: 7). Diese Modalitätsart beruht auf der Semantik der Notwendigkeit der Ausführung einer bestimmten Tätigkeit oder ihrer Unterlassung durch das agentive Satzsubjekt. Daher ist die Wurzelmodalität stets subjektbezogen und somit in das „primäre Satzschema“ integriert, wodurch die allgemeine Faktizität der Aussage per se gewährleistet ist. Im konkreten Beleg ist aber das Satzsubjekt metaphorisch verwendet und deckt sich daher nicht mit dem Agens, welches kovert ist: Es handelt sich nämlich um die Notwendigkeit der Ausführung einer Tätigkeit durch das ungenannte, aber gemeinte Agens. Aus diesem Grund ist auch die Aktionsartsemantik des verbalen Prädikats auf Satz- und auf Kontextebene unterschiedlich. Das Verb *liegen* gehört ja zu der Aktionsartklasse des sog. *states* im Schema von Z. Vendler

(1957). Die kontextbedingte Lesart des Prädikats in der Gesamtaussage ist dagegen eine dynamische, da die uneigentliche Verwendung des Satzsubjekts die ebenso „uneigentliche“ Semantik des Verbs zur Folge hat. Der Sinn der Aussage ist daher [der Kellner soll die Pommes schnellstmöglich auf das Tablett *legen*], d.h. *gemeint* ist kein *state*, sondern (nach Vendlers Terminologie) ein *accomplishment*. Dies verändert seinerseits die temporale Perspektive, da die Aussage präsupponiert, dass die Pommes noch nicht auf dem Tablett liegen und erst in der nahen Zukunft darauf gelegt werden müssen. Insgesamt liegt ein direkter Sprechakt vor, welcher nun die kategoriale Lesart des Prädikats determiniert.

Im letzteren Fall (iii) handelt es sich dagegen um eine *starke Vermutung* des Sprechenden, dass die Pommes bereits auf dem Tablett liegen. Er sieht es zwar nicht, nimmt es aber – aus welchen Gründen auch immer – stark an. Hier liegt keine Feststellung der Sachlage vor, sodass dem Satz im Ganzen kein Wahrheitswert zugewiesen wird, sondern lediglich ein Verweis des Sprechers auf seinen Wissenstand kodiert wird (vgl. Fritz 2000: 3-4; 157-162; 201; 209). Derartige Sätze haben keine Bindung des Modalverbs an das Satzsubjekt, da die Modalität lediglich sprecherbezogen ist und somit über den gesamten Satzinhalt operiert. Dies ist ein klassischer Fall der so genannten *epistemischen* Modalverbbedeutung (vgl. u.v.a. Hundt 2003). Bei derartigen Verwendung von Modalverben ist das Subjekt – egal, ob es das reale oder das virtuelle Agens impliziert – von der Subjekt-Modalverb-Achse grundsätzlich ausgeschlossen und koordiniert mit dem finiten Modalverb lediglich formal-morphologisch. Es liegt dabei weder eine semantische noch eine tiefenstrukturell-syntaktische Korrelation vor. Daraus folgt nun, dass keine weiteren temporalen oder aktionsartbezogenen Implikationen vorliegen. Der Satz kodiert uneingeschränkt gegenwartsbezogene Temporalität und nicht-transformative Aktionsart. Das Verb *liegen* ist ein reales *state* in der Terminologie von Vendler, und das Präsens kennzeichnet Gegenwartsbezug. Diese Verwendung entspricht völlig der Generalisierung von W. Abraham und E. Leiss (2008: XII-XII), epistemische Modalität sei im Gegensatz zur deontischen (Wurzelmoralität) zur Imperfektivität affin. Die epistemische Lesart blockiert daher die genuine Semantik der Notwendigkeit beim Modalverb, welche ihrerseits die präferente zukunftsbezogene Lesart der Aussage triggert. Deswegen wird bei (i) und (ii) unabhängig von unterschiedlicher Interpretation der Proposition die erwartete Realisierung der Tätigkeit in der Zukunft ausgedrückt, während in der Lesart (iii) vorausgesetzt wird, dass diese Tätigkeit *wahrscheinlich* schon realisiert worden ist, sodass die Semantik der Wahrscheinlichkeit den Ausdruck der Zukunft gewissermaßen „ersetzt“.

Hier entsteht die Frage, ob es Instrumente gibt, welche eine technisch mögliche Feststellung dermaßen minutiöser Unterschiede bei der Interpretation semantischer Differenzen dieser Art auf der Grundlage der Korpus-Methodik zulassen. Diese Frage ist weder einfach noch trivial. Freilich lässt sich eine angemessene Interpretation der Semantik von Modalverben in den Kontexten des hier besprochenen Typs ausschließlich aufgrund der Korpusanalyse relativ problemlos ableiten. Allerdings enthält diese Prozedur keine Feststellung der Wechselbeziehungen zwischen der Verwendung des Modalverbs in dieser oder jener Bedeutung und dem Profil des jeweiligen Textes.

Folglich ist keine Ermittlung der Verbindungen zwischen dem Textprofil und der Tokenfrequenz von Modalverben in wurzelmodaler respektive epistemischer Lesart allein aufgrund der Korpusanalyse möglich. Textprofile, welche per definitionem ein Bindeglied zwischen der Semantik der Entität im Satz und ihrer Verwendung im Gesamtkorpus bilden müssten, können paradoxerweise nicht in quantitativen Korpusstudien als autonome Gegenstände mit wissenschaftlichem Eigenwert benutzt werden. Daher ist eine Integration der Textprofilanalyse in die Korpusanalyse nicht möglich. Dafür kann die Korpusanalyse jederzeit um Textprofilanalyse ergänzt werden. Ansonsten stellen beide Methoden alternative Untersuchungsverfahren dar.

Aus dem Gesagten folgt nun die Notwendigkeit der wechselseitigen Ergänzung von vier oben beschriebenen autonomen Typen empirischer Forschungsmethoden in der Linguistik, und zwar: (i) Introspektion, (ii) Informantenbefragung, (iii) Korpusanalyse und (iv) Textprofilanalyse. Jede der genannten Methoden ist zugleich relativ autonom und unzureichend für die komplexe Zielsetzung der sprachwissenschaftlichen Forschung. Zwei erste Methoden gestatten dabei die Sprache auf der Ebene potentieller Äußerungen dank den im Gehirn vorhandenen Sprachmodulen zu untersuchen, während zwei letztere Methoden die Sprache auf der Ebene der realen Verwendung entsprechender Sprachformen in konkreten Texten und Textkorpora beschreiben.

Abschließend noch eine Anmerkung bezüglich der Nutzung der Ergebnisse der oben dargestellten Methoden zum fachdidaktischen Zweck. Die Korpusanalyse liefert hier einerseits umfangreichen empirischen Stoff, dessen größter Vorteil die Authentizität der Objektbeispiele aus dem Quellenkorpus ist; andererseits sind aber auch der Korpusanalyse deutliche Grenzen gesetzt. Sie ergeben sich vor allem aus dem „Modellcharakter“ der Sprachvermittlung im Sprachunterricht. Die Verwendung einer Entität in einem gewissen konkreten Kontext ausschließlich auf der Grundlage von realen, aus dem Quellenkorpus stammenden Belegen führt unumgänglich zur Entstehung einer Situation, in der der Schüler bzw. Student mit einem konkreten *Exemplar* zu tun hat, das nicht auf ein abstraktes Modell zurückgeht. Die Sprachvermittlung beruht jedoch vor allem auf einer Reihe von Verallgemeinerungen, und zwar sowohl im Bereich grammatischer Formen als auch Situationen oder Kontexte, in denen sie vorkommen. Das Korpus gibt uns lediglich konkrete Beispiele, die zwar wegen ihrer Authentizität wertvoll sind, aber die Prozedur der Generierung analoger Konstruktionen in anderen Kommunikationssituationen keinesfalls erleichtern – schon wegen ihrer strikten Anbindung an eine gewisse, nicht immer wiederholbare Situation. Mehr noch: Es gibt keine ausführliche Darstellung bzw. Beschreibung dieser Situation, sodass sogar ein Modell fehlt, welches der Simulierung einer vergleichbaren Situation oder eines vergleichbaren Kontextes dienen könnte, in denen die jeweilige Sprachform verwendbar wäre. Was wir in diesem Fall haben, ist einzig und allein ein gewisser Korpusbeleg, von dem gesagt werden kann, dass er ein authentisches Beispiel der Verwendung dieser Entität ist. Seine weitere Bewertung muss auf Introspektion beruhen, d.h. größtenteils gerade auf eigener Sprachkompetenz und im Falle von Fremdsprachen auf Aneignung analoger Formen oder Konstruktionen in vergleichbaren Kom-

munikationssituationen, welche nolens volens mehr oder minder simuliert und somit von Textkorpora unabhängig sind.

Literaturverzeichnis

- ABRAHAM Werner, LEISS Elisabeth (2008): Introduction. – [w:] Werner ABRAHAM, Elisabeth LEISS (red.): *Modality-Aspect Interfaces: Implications and typological solutions*. – Amsterdam: Benjamins, XI–XXIV.
- ÁGEL Vilmos, HENNIG Mathilde, red. (2006): *Grammatik aus Nähe und Distanz. Theorie und Praxis am Beispiel von Nähetexten 1650–2000*. – Tübingen: Niemeyer.
- CHOMSKY Noam (1996): *Powers and Prospects. Reflections on human nature and the social order*. – London: Pluto Press.
- CHOMSKY Noam (2002): *On Nature and Language*. – Cambridge: Cambridge University Press.
- FRITZ Thomas (2000): *Wahr-Sagen. Futur, Modalität und Sprecherbezug im Deutschen*. – Hamburg: Buske.
- GRUZA Franciszek (1997): Języki ludzkie a wyrażenia językowe, wiedza a informacja, mózg a umysł ludzki. – [w:] Franciszek GRUZA, Maria DAKOWSKA (red.): *Podejścia kognitywne w lingwistyce, translatoryce i glottodydaktyce*. – Warszawa: Wydawnictwo Uniwersytetu Warszawskiego, 7–21.
- GRUZA Sambor (2011): Lingwistyka antropocentryczna a badania okulograficzne. – *Lingwistyka Stosowana* 4, 149–162.
- HAAN Ferdinand de (1997): *The Interaction of Modality and Negation. A Typological Study*. – New York–London: Garland.
- HUNDT Markus (2003): Zum Verhältnis von epistemischer und nicht-epistemischer Modalität im Deutschen. Forschungspositionen und Vorschlag zur Neuorientierung. – *Zeitschrift für Germanistische Linguistik* 31, 343–381.
- KRATOCHVILOVÁ Iva, WOLF Norbert Richard (2010): *Kompendium Korpuslinguistik. Eine Bestandsaufnahme aus deutsch-tschechischer Perspektive*. – Heidelberg: Winter.
- LÜDELING Anke, KYTŮ Merja (2008–2009): *Corpus Linguistics An International Handbook*. – Berlin–New York: de Gruyter.
- MAMARDAŠVILI Merab K., PIATOGORSKIJ Aleksandr M. (2011): *Simvol i soznaniye*. – Sankt-Petersburg: Azbuka.
- PALMER Frank R. (1986): *Mood and Modality*. – Cambridge: Cambridge University Press.
- VENDLER Zeno (1957): Verbs and Times. – *The Philosophical Review* 66 (2), 143–160.

doi: <http://dx.doi.org/10.17651/bptj.74.5>

ALEKSANDRA JANOWSKA

Uniwersytet Śląski, Katowice

Źródła a typy badań historycznojęzykowych

Streszczenie

Artykuł poświęcony został zagadnieniom źródeł językowych, wykorzystywanych przez badaczy historii języka. Zachowana do naszych czasów dokumentacja historycznojęzykowa jest bardzo różnorodna. Pierwszy narzucający się podział dotyczy wiarygodności dokumentów pisanych i pewności faktu językowego. Drugi związany jest z typem tekstów: chodzi o tekst jako źródło materiału językowego i tekst o funkcji metajęzykowej. Językoznawca, przystępując do analizy historycznej, zawsze dokonuje wyboru, jeśli chodzi o ekscerpcję materiału. To kwestia nie tylko liczby źródeł poddanych obserwacji, ale także ich rodzaju. Za każdym razem wybór ten rzutuje na sposób przeprowadzanej analizy i na wysuwane wnioski.

Słowa kluczowe: źródła językowe, słowniki historyczne, historia języka

Data sources in historical linguistic research

Abstract

The article discusses the diversity of language data sources which are used in the research work of historical linguists. The extant historical linguistic documentation varies significantly. The first division that seems particularly salient concerns the reliability of written documents and the certainty of a linguistic fact. The second division is based on the different types of text: text as a source of language material and text about language. A researcher who engages in a historical linguistic analysis always makes a choice as far as the excerption of the material is concerned. Decisions concerning the number of sources which are examined and their type influence the manner in which the analysis is conducted as well as the conclusions that are drawn.

Keywords: language data sources, historical dictionaries, language history

Jerzy Topolski w *Metodologii historii* pisze: „Źródłem historycznym w sensie ogólnym jest wszystko to, skąd czerpać możemy (źródło w sensie potencjalnym) lub skąd czerpiemy (źródło w sensie efektywnym) informacje o faktach historycznych”. (Topolski 1968: 274).

Tak szeroko zakrojone pole obserwacji naukowej mimo zmiennych koncepcji historiograficznych tworzy fundament pracy badacza, nie dziwi zatem, że specyfice i klasyfikacji źródeł historycznych poświęca się tak wiele uwagi teoretycznej.

Dla historyka języka źródło jest pojęciem, wydawać by się mogło, prostszym do opisu, niewymagającym tak głębokiej refleksji naukowej, jego zakres bowiem jest węższy. Można powiedzieć, że lingwistę interesuje to, co zostało zapisane (rzadko i jedynie w wypadku najnowszej historii będą to też nagrania czy zebrane ustne wspomnienia, ankiety). To ów materialny fakt językowy, który stanowi podłoże interpretacji, wnioskowania, ale też tworzenia (należałoby powiedzieć raczej: rekonstruowania) innych, domniemyanych bytów językowych.

A jednak „prostota” tych źródeł jest pozorna. Rzecz nie tylko w ich różnorodności i wielości, ale także w uwikłaniach w kontekst pozajęzykowy, który w wypadku wieków przeszłych nie sposób lekceważyć. Dawny tekst bez zaplecza historycznego niewiele jest wart. Jego badania bez odwołania do wiedzy ogólnej, kulturowej – posłużmy się słowami Marii Wojtyły Świerzowskiej – „siłą rzeczy muszą przekształcić się w czczą zabawę, żonglerkę formą, układanie i dopasowywanie za wszelką cenę puzzli, w których poszczególne elementy pochodzą z całkiem różnych układanek” (Wojtyła-Świerzowska 2009: 119).

Toteż tekstologię historyczną, mocno osadzoną w przestrzeni historyczno-kulturowej, uznać należałoby za jeden z najważniejszych nurtów współczesnych badań. Jeżeli podstawą obserwacji dawnego języka jest źródło pisane, o źródle tym powinniśmy wiedzieć jak najwięcej. Informacje te bowiem umożliwiają właściwą ocenę faktów językowych. Doskonale świadectwo potrzeby ciągłej weryfikacji naszej wiedzy w tym względzie stanowią nowe odczytania, interpretacje zabytków średnio-wiecznych¹.

1. Dla lingwistyki diachronicznej istotne są pytania o to, jak umiejętnie wykorzystać tak opracowane źródła. Poszukiwanie najbardziej adekwatnych metod opisu języka oznacza poszukiwanie również najlepszej bazy materiałowej. Nic więc dziwnego, że od lat toczy się swoisty spór dotyczący przydatności różnych źródeł i jednocześnie wzajemnej przystawalności, jeśli chodzi o materiał w nich zawarty. Warto zatem uporządkować dotychczasowe sądy.

Pierwszy narzucający się podział tekstów dotyczy ich wiarygodności jako świadectwa języka danej epoki, pisarza itd. Chodzi o różne poziomy niepewności zapisanego znaku językowego. W wypadku rękopisów, zwłaszcza manuskryptów średnio-wiecznych, to przede wszystkim trudności w odczytaniu, zrozumieniu zapisu. Sama transliteracja to, jak powszechnie wiadomo, interpretacja, nie mówiąc już o transkrypcji (Twardzik 1997, Mika 2015).

Wspomniana niepewność wynika też często z typu tekstu-źródła. Czym innym jest tekst, który możemy jednoznacznie zlokalizować w czasie i przestrzeni, czym

¹ Zob. np. wydanie *Rozmyślenia przemyskiego*: Keller, Twardzik 1998–2004; Trawińska 2005; Mika 2012; Deptuchowa, Wanicowa 2010; Wanicowa 2009 itd.

innym dokument „zawieszony w próżni”. Czym innym wreszcie jest źródło oryginalne, a czym innym kopiowane (zmieniane przez kopistę czasem bezwiednie, a czasem z pełną świadomością wprowadzenia koniecznych według niego zmian), lub też wtórnie wydane, obciążone arbitralnymi wyborami i rozstrzygnięciami wydawców (por. np. Kuźmicki 2015). Sprawę komplikuje ponadto fakt, że zabytki językowe, zwłaszcza te najstarsze, średniowieczne, często nie są homogeniczne, mają charakter wielowarstwowy, a do tych warstw badacze docierają z dużą trudnością². Wszystko to zaciera rzeczywisty obraz języka pierwotnego dzieła.

Oczywiście powodów wymuszających szczególną ostrożność badawczą względem analizowanego zapisu jest znacznie więcej. Za przykład mogą posłużyć wątpliwości, jakie budzą hasła słownikowe i w leksykonach XVI-wiecznych, i znacznie późniejszych, nawet XX-wiecznych. Pojawia się pytanie, na ile ta dokumentacja materiałowa jest miarodajna. Chodzi chociażby o neologizmy tworzone przez leksykografów na użytek tych dzieł czy też zwykłe błędy.

To tylko wybrane problemy, pokazujące, jak dużego namysłu wymaga penetracja dawnych dokumentów.

A jakie konsekwencje dla badań ma wspomniana niepewność faktów? Wbrew pozorom niekoniecznie musi stanowić przeszkodę w analizach językowych. Wszystko zależy od typu przeprowadzanych obserwacji. W opisach syntetycznych, gdy bierze się pod uwagę większą liczbę źródeł, omawiana kwestia wiarygodności nie odgrywa zbyt dużej roli. Jeśli mamy sporą egzemplifikację danego problemu, niejasne fakty giną w natłoku innych i zwykle się ich nawet nie wydziela. Brak pewności ma jednak znaczenie w analizach szczegółowych.

Dodać warto, że paradoksalnie właśnie trudne do opisu przypadki stymulują rozwój myśli językoznawczej, zmuszają do tworzenia nowych narzędzi badawczych, stają się też dobrym probierzem dotychczasowych ustaleń.

2. Zachowane źródła nie zawsze są, jak zauważył znawca historii Henri Marrou, takie, jakie chcielibyśmy, żeby były. I nie chodzi tylko o kwestię ich wiarygodności. Albo nie ma ich na tyle, ile potrzebujemy, doświadczamy tego najczęściej, badając najstarsze wieki, albo jest ich zbyt dużo i badacz czuje, że jest przytłoczony ich ciężarem (Marrou 2011: 79–80). historyk języka znajduje się w bardzo podobnej sytuacji. Stąd też i inna uwaga wspomnianego naukowca warta jest uwagi: „Nie można stać się historykiem tak sobie [...], należy nauczyć się rozumienia istnienia, natury, warunków i sposobów użytkowania różnych kategorii źródeł historycznych” (Marrou 2011: 83).

Zbyt mała liczba poświadczeń faktu językowego bardzo utrudnia analizę, ale i wielość źródeł to, wbrew pozorom, spory kłopot dla badaczy. Selekcja materiału w badaniach historycznojęzykowych ma znaczenie fundamentalne, nie jest działaniem

² O wielowartościowości dawnych zabytków zob. Mika 2015. Tam też obszerna bibliografia prac poświęconych analizom poszczególnych zabytków językowych.

wyłącznie mechanicznym, ale wiąże się ze ściśle wypracowanymi zasadami, a w każdym razie – powinna się z nimi wiązać. Wybór, jaki jest dokonywany, to kwestia i liczby źródeł poddanych obserwacji, i ich rodzaju. I za każdym razem decyzja badacza zdradza jego sposób patrzenia na język.

W tym kontekście należałoby wspomnieć o drugim ważnym podziale źródeł, który wiąże się z wydzieleniem dzieł o funkcji metajęzykowej, a zatem poradników, podręczników, słowników itd.³ Chodzi przede wszystkim o te ostatnie: dawne dzieła leksykograficzne i współczesne słowniki historyczne. Zwykle wykorzystuje się je jako podstawę materiałową, czasem – jeśli chodzi o stare leksykony – jako świadectwo dawnej kompetencji językowej, a wreszcie także jako opracowania naukowe przedstawiające własną propozycję opisu. Jako baza materiałowa wzbudza jednak wiele kontrowersji. Toczono dyskusje o tym, „jaka dokumentacja językowa uprawomocnia czynione konstatacje i uwiarygodnia wnioski: słownik czy tekst” (Pastuchowa 2016, 142), pokazują i braki, i zalety tego typu źródeł w badaniach językowych (por. np. Wanicowa 2017, Kleszczowa 2015, Danielewiczowa 2014, Masłej 2015).

Dla historyka języka, szukającego odpowiedniej egzemplifikacji, kłopotliwa może być przede wszystkim kwestia ingerencji w materię językową, z jaką mamy do czynienia w tym typie źródeł: słowniki z natury swojej krępują własnymi interpretacjami, dokonanymi wyborami, także źródłowymi, obarczone są sposobem myślenia czasów, w których powstawały. Dają opis z konieczności uproszczony. Z jakimi problemami trzeba się zmierzyć w takim wypadku, pokazuje m.in. Magdalena Danielewiczowa, analizując szczegółowo jedno z haseł SXVI (Danielewiczowa 2014) i podkreślając konieczność oderwania się we własnej pracy od leksykograficznych propozycji.

Dlaczego zatem sięgamy po słowniki jako źródło dokumentacyjne? Z dwóch powodów. Po pierwsze, decyduje o tym przede wszystkim obszerny materiał, zwłaszcza w wypadku współczesnych słowników historycznych (np. Sstp, SXVI, SXVII/XVIII). Jest to fakt nie do przecenienia, mimo że czasem trudniej o pewną segregację materiału (typ źródeł, czas ich powstania itd.). To prawda, że badacz ekscerpujący teksty ma poczucie, iż panuje nad materiałem w tym sensie, że ma możliwość świadomego wyboru odpowiednich dzieł, poznania ich specyfiki (była o tym mowa wyżej), a przede wszystkim ma swobodę oceny opisywanych zjawisk⁴. Niemniej praca z tekstami wiąże się zawsze ze stosunkowo niewielkim zbiorem przykładów. Nawet i digitalizacje zabytków, tworzenie internetowych korpusów nie zlikwidują w pełni ograniczeń związanych z fizycznymi możliwościami badacza, gdy chodzi np. o analizy obszernych zagadnień (np. procesów obejmujących kilka wieków). To kwestia nie tylko samej ekscepcji, ale również porządkowania zbioru. Pole obserwacji jest więc z oczywistych względów zawężone. Warto też przytoczyć w tym miejscu uwagi Krystyny Kleszczowej, zestawiającej materiał czerpany z tekstu i słownika:

³ O podziale tym pisała m.in. Kleszczowa (2009).

⁴ Pod tym względem wyróżnia się SXVII/XVIII, który od początku tworzony jest jako słownik internetowy, dający znacznie więcej możliwości wyszukiwania potrzebnych danych.

W moim przekonaniu nie da spodziewanych efektów sumowanie bardzo dokładnych nieraz opisów jednostek leksykalnych, opracowanych na podstawie tekstów. Nie neguję zasadności takich prac, ale doświadczenie mnie uczy, że „wyłuskiwanie” znaczenia leksykalnego z opisów „przeziąkniętych” znaczeniami kontekstowymi to praca żmudna, nieefektywna, powiedzieć wręcz można – syzyfowa. (Kleszczowa 2015: 16)

Coraz częściej traktuje się słowniki historyczne przede wszystkim jako duże korpusy językowe. Ułatwiają to digitalizacje i odpowiednie wyszukiwarki.

Jest jednak drugi powód, dla którego omawiane współczesne leksykony stanowią ważne źródło. To, co uznaliśmy wyżej za ich defekt (jego charakter interpretacyjny, stosowane uproszczenia), w wielu wypadkach okazuje się użyteczną i ważną cechą. Artykuł hasłowy to zawsze wynik wielomiesięcznych żmudnych prac specjalistów (por. uwagi Zofii Wanicowej (2017)). Otrzymujemy zatem materiał (czasem bardzo obszerny lub trudno dostępny) już wstępnie zweryfikowany, „poukładany”, co znacznie ułatwia i przyspiesza pracę badawczą, czasem wręcz w ogóle ją umożliwia. To bezcenna pomoc. Nie zwalnia to oczywiście naukowca od uwagi i własnych analiz.

3. Wydaje się jednak, że niewłaściwe jest pytanie o to, co uznać należałoby za lepsze źródło materiałowe: teksty czy słownik? Takie przeciwstawienie stanowi znaczne uproszczenie z co najmniej dwóch powodów.

Po pierwsze – różne poziomy języka wymagają różnego opisu i także innej dokumentacji. Typ źródeł musi być dopasowany do typu przeprowadzanych badań. Innych tekstów poszukuje się, gdy omawia się zagadnienia systemowe (np. fleksyjne czy nawet słowotwórcze), innych wymagają np. studia semantyczne, tekstologiczne.

Zwróćmy uwagę na to, że w badaniach fleksyjnych, a przede wszystkim składniowych zwykle uwagę zwraca się na teksty ciągłe (por. np. Kępińska 2006, Siuciak 2008, Słoboda 2012). Badania słowotwórcze, ukierunkowane na poszukiwanie kierunków ewolucji, preferują materiał słownikowy (por. przykładowo Szczaus 2005, Zarębski 2012, także z ośrodka katowickiego: Kleszczowa 1998, Pastuchowa 2000, Janowska 1999). I taki wybór nie jest tylko kwestią pewnych schematów, przyzwyczajęń, ale raczej szukaniem optymalnych rozwiązań. Analizy np. składniowe, stylistyczne z konieczności wymagają dłuższego kontekstu, czasem kontekstu całego dzieła. W przekrojowych opracowaniach słowotwórczych natomiast, gdy liczba badanych jednostek jest rzędu kilku tysięcy, nie ma możliwości własnej, samodzielnej ekscerpacji przykładów. Dla derywatologa nieistotne są też niuanse znaczeniowe, gdyż poszukuje ogólnego schematu semantycznego. Rola słowników jest więc nie do przecenienia, nawet jeśli weźmie się pod uwagę zawarte tam możliwe błędy (w słowotwórstwie liczą się serie, a nie pojedyncze fakty). Pamiętajmy też, że słowniki nakierowane są z natury swojej na zagadnienia leksykalne, toteż i w analizach tego typu najczęściej sięga się właśnie po nie.

Druga przyczyna to kwestia stosowanej metodologii i celu opracowania. Każdy bowiem wybór, także dotyczący ekscerpacji, jest już obciążony pewnymi założeniami wstępnymi. Trudno byłoby nie zgodzić się z uwagami cytowanego już wcześniej Jerzego Topolskiego, że informacje uzyskane ze źródeł nie są neutralne:

[...] nie można byłoby przeprowadzić większości analiz dotyczących postępowania historyka bez odwołania się do zespołu uznanych i przyjętych przez niego twierdzeń i dyrektyw, na gruncie których dokonuje on poznania i badania przeszłości. (Topolski 1968: 275)

Tak samo w wypadku językoznawcy – jego koncepcja języka, metodologia, którą stosuje, wyznaczają i sposób zbierania materiału, i jego opis. W toczonych dyskusjach na temat źródeł czasem zapomina się o tym, wydawałoby się dość oczywistym, fakcie. Prawie każdy wybór wiąże się jednocześnie z zawężaniem pola obserwacji. Barwnie scharakteryzował ów problem Henri Marrou: „Teoria jest jak projektor, oświetlający silnym światłem rzeczywistość i obiekty, które są w polu jego światła, zostawiając jednak w całkowitej ciemności wszystko pozostałe” (Marrou 2011: 186).

Przypatrzmy się dwóm najbardziej charakterystycznym przykładom.

Językoznawcy skupieni na bardzo szczegółowych badaniach diachronicznych, dla których istotne jest umiejscowienie faktu językowego w przestrzeni i czasie, pokazanie ewolucji wraz z jej etapami, wypracowali bardzo precyzyjne procedury zbierania materiału, związane z koncepcją płaszczyzn synchronicznych. Popularyzatorką tej metody w językoznawstwie polskim była Irena Bajerowa (por. np. Bajerowa 1964), a stosowali ją tak znakomici językoznawcy, jak np. Alina Kowalska (1976), Ryszard Rzepka (1985).

Koncepcja opisu ewolucji języka za pomocą przekrojów czasowych zdominowała analizy historyczne przez pewien czas i do dziś jest wykorzystywana⁵. Podstawę metody, co dla nas jest ważne, stanowi bardzo dokładny, przemyślany wybór materiału. Kanon tekstów dobiera się według ściśle wyznaczonych kryteriów. Pod uwagę brane są: rodzaj tekstu, jego tematyka, miejsce i czas wydania, autor itd. Każdy wycinek czasowy winien być reprezentowany podobnym typem źródeł i podobną długością tekstów. Miało to dać gwarancję reprezentatywności materiału. Nie chodzi zatem o to, aby tych tekstów było jak najwięcej, ale żeby tworzyły ściśle określony zbiór. Gdzieś w tle funkcjonowało założenie, że tylko różnorodność źródeł daje podstawy do mówienia o języku ogólnym i że taki zrównoważony korpus stwarza obiektywny obraz ewolucji. W ten sposób bada się głównie zmiany gramatyczne, ale nie tylko⁶. Nie w każdym jednak wypadku tę metodę można stosować, trudności spotykamy chociażby wtedy, gdy sięgamy po materiał staropolski niekompatybilny⁷ z materiałem z późniejszych wieków.

W badaniach poświęconych szerszym zjawiskom w dłuższych okresach czasu lub gdy mowa jest o procesach panchronicznych, można pozwolić sobie na inny wybór źródeł. I z czymś takim mamy do czynienia np. w badaniach ewolucji perspekty-

⁵ Ślady tej koncepcji odnajdujemy też w pracach odległych tematycznie. Widoczne są choćby w staranności i precyzji doboru źródeł, doboru uwzględniającego zasadę proporcjonalności, zob. np. Niewiara 2009.

⁶ Ten sposób zbierania materiału zastosowała np. Aleksandra Krupianka w swojej słowotwórczej pracy (1979).

⁷ Zachowany do naszych czasów materiał staropolski jest znacznie uboższy w porównaniu choćby z dokumentacją z XVI w. Chodzi nie tylko o liczbę zabytków, ale także o ich sprofilowaną tematykę.

wicznej (termin ten zaproponowała Krystyna Kleszczowa (1998)), w których istotny jest schematyczny obraz ewolucji językowej, zakreslenie głównych kierunków zmian bez całej gamy procesów ubocznych, nieistotnych w końcowym efekcie. W grę wchodzi w związku z tym przede wszystkim źródła, dające różnorodny materiał bez jego szczegółowej charakterystyki, a zatem głównie słowniki, korpusy. W takim ujęciu zestawia się i porównuje dwie płaszczyzny czasowe najbardziej oddalone od siebie, w tym wypadku: średniowiecze i współczesność, prymat jednak przyznając w opisie tej ostatniej (zob. też Kleszczowa 2003; Janowska, Pastuchowa 2005).

Jak ważne są cele i sposoby badań, widać na przykładzie oceny możliwości wykorzystania w analizach *Słownika staropolskiego*. W książce *Wyrażenia funkcyjne w średniowiecznej polszczyźnie z perspektywy składniowej* Tomasa Miki i Agnieszki Słobody zwraca uwagę stwierdzenie, że na hasło słownikowe, w tym wypadku związane z wyrażeniem funkcyjnym, „należy patrzeć jak na zbiór danych pogrupowanych w funkcjonalno-semantyczne podzbiory w sposób synchroniczny nieuwzględniający ani czynników chronologicznych, ani odmianowych” (Mika, Słoboda 2015: 96).

Trudno byłoby nie zgodzić się z powyższym sądem⁸. Dla badań dynamiki zjawisk zachodzących w średniowieczu, a takie zadanie stawiają sobie autorzy tych słów, to znaczne utrudnienie, ale już dla analiz proponowanych przez Kleszczową nie ma znaczenia, a wręcz jest swoistym ułatwieniem, odpowiada bowiem proponowanej koncepcji ogólnego opisu i przyjętego założenia, że okres ten traktować można jako jedną płaszczyznę czasową.

Te dwa różne spojrzenia na słownik, a z każdym z nich należałoby się zgodzić, chyba najwyraźniej pokazują, że dyskusja nad źródłami nie może być prowadzona w próżni metodologicznej (por. też Woźniak 2011).

Oczywiście należy pamiętać, że stosunkowo rzadko badacze decydują się na rygorystyczny wybór jednego z typów źródeł. Często łączy się je, traktując jako wzajemnie uzupełnienie. Taką kompilację odnajdziemy w wielu pracach diachronicznych (zob. przykładowo: Krążyńska 2000–2012, Stąpor 2008). Albo są to badania oparte na tekstach, a słownik traktowany jest „jako cenne narzędzie uzupełniające” (Mika, Słoboda 2015: 100), albo podstawę stanowi słownik z dodatkową ekscerpcją tekstów lub, co częstsze, kartotek leksykograficznych, typem źródeł pośrednim między tekstami a słownikami. To, co jest bardziej eksponowane jako źródło, zależy od stawianych przed sobą zadań⁹.

Co ciekawe, w niektórych opracowaniach możemy odnaleźć analizy różnic, jakie zachodzą we wnioskowaniu w wyniku wykorzystania odmiennych typów baz źródłowych. Warto zwrócić na nie uwagę, gdyż są to ważne uwagi, uzmysławiające od-

⁸ Komplikacji związanych z wykorzystaniem materiału słownikowego jest znacznie więcej, zob. Mika, Słoboda 2015: 91–103.

⁹ Oczywiście wybór źródeł może mieć jeszcze inny charakter. Przykładem jest praca W. Książek-Bryłowej (1992), w której autorka bierze pod uwagę dane słownikowe i dane zawarte w różnych opracowaniach poświęconych omawianemu zagadnieniu.

mienności opisywanej materii badawczej i jednocześnie zmuszające do zastanowienia się nad konsekwencjami dokonanych wyborów¹⁰.

Zapewne można byłoby zadać pytanie, czy w obecnych czasach, gdy pojawiają się nowe możliwości ekskserpcji materiału, powyższe rozważania są jeszcze aktualne. Mamy przecież korpusy internetowe, digitalizacje zabytków, wydania internetowe słowników itd.¹¹ Stanowi to znaczne ułatwienie pracy, a jednak podstawowe pytania, wątpliwości pozostają w dalszym ciągu te same. I, jak sędzę, upłynie jeszcze sporo czasu, zanim w pełni wypracowane zostaną nowe narzędzia badawcze.

Bibliografia

- BAJEROWA Irena (1964): *Kształtowanie się systemu polskiego języka literackiego XVIII wieku*. – Wrocław: Zakład Narodowy im. Ossolińskich.
- DANIELEWICZOWA Magdalena (2014): Opisać aby – uwagi na marginesie ujęcia w Słowniku polszczyzny XVI wieku. – [w:] Krystyna KLESZCZOWA, Anna SZCZEPANEK (red.): *Wyrażenia funkcyjne w perspektywie synchronicznej i porównawczej*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego, 181–193.
- DEPTUCHOWA Ewa, WANICOWA Zofia (2010): Tajemnice staropolskich glos. – *LingVaria* V, 2, 105–111.
- GRUSZCZYŃSKI Włodzimierz, ADAMIEC Dorota, OGRODNICZUK Maciej (2013): Elektroniczny korpus tekstów polskich z XVII i XVIII wieku (do 1772 roku) – prezentacja projektu badawczego. – *Polonica* XXXIII, 309–316.
- JANOWSKA Aleksandra (1999): *Funkcje przestrzenne przedrostków czasownikowych w polszczyźnie*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- JANOWSKA Aleksandra, PASTUCHOWA Magdalena (2005): *Słowotwórstwo czasowników staropolskich. Stan i tendencje rozwojowe*. – Kraków: Wydawnictwo Universitas.
- KĘPIŃSKA Alina (2006): *Kształtowanie się polskiej kategorii męsko- i niemęskoosobowości : język wobec płci*. – Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.
- KELLER Felix, TWARDZIK Waław, wyd. (1998–2004): *Rozmyślanie przemyskie. Transliteracja, transkrypcja, podstawa łacińska, niemiecki przekład*, t. 1–3. – Freiburg: Weiher Verlag.
- KLESZCZOWA Krystyna (1998): *Staropolskie kategorie słowotwórcze i ich perspektywiczna ewolucja. Rzeczownik*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- KLESZCZOWA Krystyna (2003): *Staropolskie derywaty przymiotnikowe i ich perspektywiczna ewolucja*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- KLESZCZOWA Krystyna (2009): Językoznawstwo historyczne w świetle współczesnej historiografii. – [w:] Jolanta GWIÓDZK, Krystyna KLESZCZOWA (red.): *Faktografia w badaniach historycznych*. – Katowice: Biblioteka Śląska, 155–161.
- KLESZCZOWA Krystyna (2015), *U źródeł polskich partykuł. Derywacja funkcjonalna, przemiany, zaniki*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- KOWALSKA Alina (1976): *Ewolucja analitycznych form czasownikowych z imiesłowem na -l w języku polskim*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- KRĄŻYŃSKA Zdzisława (2000–2012): *Staropolskie konstrukcje z przyimkami*, cz. 1–5. – Poznań: Wydawnictwo „Poznańskie studia Polonistyczne”.

¹⁰ Zob. np. uwagi zawarte w pracach: Pawelec 2003, Woźniak 2011.

¹¹ Przewodnikiem po zasobach internetowych jest praca: Pałka, Kwaśnicka-Janowicz 2017. Por. też bardziej szczegółowe rozważania o korpusach słownikowych: Gruszczyński, Adamiec, Ogrodniczuk 2013.

- KRUPIANKA Aleksandra (1979): *Czasowniki z przedrostkami przestrzennymi w polszczyźnie XVIII wieku*. – Warszawa–Poznań–Toruń: Państwowe Wydawnictwo Naukowe.
- KSIĄŻEK-BRYŁOWA Władysława (1992): Warianywność fleksji w historii i gwarach języka polskiego. – [w:] Teresa SKUBALANKA, Władysława KSIĄŻEK-BRYŁOWA: *Wariantywność polskiej fleksji*. – Wrocław: Zakład Narodowy im. Ossolińskich, 117–191.
- KUŹMICKI Marcin (2015): *Wydanie wielkopolskich rot sądowych w świetle najnowszych ustaleń badawczych*. – *LingVaria* 2, 205–219.
- MARROU Henri (2011): *O poznaniu historycznym*. Tłum. Hubert ŁASZKIEWICZ. – Kęty: Antyk Marek Derewiecki.
- MASŁEJ Dorota (2015): Słownik a tekst. Problem znaczenia odpowiedników leksykalnych w staropolskich przekładach Psalmu 50. – [w:] Tomasz MIKA, Dorota ROJSZCZAK-ROBIŃSKA, Olga STRAMCZEWSKA (red.): *Staropolskie Spotkania Językoznawcze. Jak badać teksty staropolskie?* – Poznań: Wydawnictwo Rys, 103–117.
- MIKA Tomasz (2012): *Kazania świętokrzyskie* – od rękopisu do zrozumienia tekstu. – Poznań: Wydawnictwo Rys.
- MIKA Tomasz (2015): *Tekst staropolski jako odmienny obiekt badań? W poszukiwaniu narzędzi opisu*. – *LingVaria* 2, 235–250.
- MIKA Tomasz, SŁOBODA Agnieszka (2015): *Wyrażenia funkcyjne w średniowiecznej polszczyźnie z perspektywy składniowej. Wybrane problemy badawcze*. – Poznań: Uniwersytet Śląski, Wydawnictwo Rys.
- NIEWIARA Aleksandra (2009): *Kształty polskiej tożsamości. Potoczny dyskurs narodowy w perspektywie etnolingwistycznej (XVI–XX wiek)*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- PAŁKA Patrycja, KWAŚNICKA-JANOWICZ Agata (2017): *Przewodnik po elektronicznych zasobach językowych dla polonistów (słowniki, kartoteki, korpusy, kompendia)*. – Kraków: Towarzystwo Miłośników Języka Polskiego.
- PASTUCHOWA Magdalena (2000): *Zmiany semantyczne i strukturalne czasowników odrzeczownikowych w polszczyźnie*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- PASTUCHOWA Magdalena (2016): Wielonurtowość polonistycznych badań językoznawczych. Próba typologii. – *Postscriptum Polonistyczne* 2 (18), 135–145.
- PAWELEC Radosław (2003): *Dzieje „sztuki”. Leksemy i pojęcia*. – Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.
- RZEPKA Wojciech (1985): *Demorfologizacja rodzaju w liczbie mnogiej w polszczyźnie XVI–XVII wieku*. – Poznań: Wydawnictwo Naukowe UAM.
- SIUCIAK Mirosława (2008): *Kształtowanie się kategorii gramatycznej liczebnika w języku polskim*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- SŁOBODA Agnieszka (2012): *Liczebnik w grupie nominalnej średniowiecznej polszczyzny. Semantyka i składnia*. – Poznań: Wydawnictwo Rys.
- STĄPOR Izabela (2008): *Kształtowanie się normy dotyczącej fleksji liczebników polskich od XVI–XIX wieku*. – Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.
- SZCZUAŚ Agnieszka (2005): *Rzeczownikowe synonimy słowotwórcze w polszczyźnie XVI wieku*. – Szczecin: Wydawnictwo Naukowe Uniwersytetu Szczecińskiego.
- TOPOLSKI Jerzy (1968): *Metodologia historii*. – Warszawa: Państwowe Wydawnictwo Naukowe.
- TRAWIŃSKA Maria (2005): *Fonetyka wielkopolskich rot sądowych*. – Warszawa: Slawistyczny Ośrodek Wydawniczy przy Instytucie Slawistyki PAN.
- TWARDZIK Wacław (1997): *O ważniejszym aniżeli dotychczas tekstu staropolskiego czytaniu i jakie z niego pożytki płyną rozprawa śliczna i podziwienią godna*. – Kraków: Instytut Języka Polskiego Polskiej Akademii Nauk.
- WANICOWA Zofia (2009): *Ignota, dubia, reperta. Czytać i rozumieć staropolszczyznę*. – Kraków: Wydawnictwo Lexis.

- WANICOWA Zofia (2017): Czy Słownik staropolski potrzebuje obrony? – *Studia Linguistica* 12, 308–315.
- WOJTYŁA-ŚWIERZOWSKA Maria (2009): Faktografia w badaniach historycznojęzykowych. – [w:] Jolanta GWIOŹDZK, Krystyna KLESZCZOWA (red.): *Faktografia w badaniach historycznych*. – Katowice: Biblioteka Śląska, 115–122.
- WOŹNIAK Ewa (2011): Z warsztatu badacza semantyki historycznej – uwagi o metodzie i źródłach. – [w:] Bogusław DUNAJ, Maciej RAK (red.): *Badania historycznojęzykowe. Stan, metodologia, perspektywy*. – Kraków: Księgarnia Akademicka, 71–79.
- ZARĘBSKI Rafał (2012): *Rzeczownikowe prefiksy obcego pochodzenia w historii języka polskiego*. – Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Wykaz skrótów

- SXVII/XVIII – *Słownik języka polskiego XVII i 1. połowy XVIII wieku*, t. 1, red. Krystyna SIEKIERSKA. – Kraków 1999.
- Elektroniczny słownik języka polskiego XVII i XVIII wieku*, red. Włodzimierz GRUSZCZYŃSKI, http://xvii-wiek.ijp-pan.krakow.pl/pan_klient/ lub <http://sxvii.pl>.
- SXVI – *Słownik polszczyzny XVI wieku*, red. Maria Renata MAYENOWA, t. 1–36. – Wrocław–Warszawa–Kraków 1966–2010.
- Sstp – *Słownik staropolski*, red. Stanisław URBAŃCZYK, t. 1–11. – Wrocław–Kraków 1953–2002.

doi: <http://dx.doi.org/10.17651/bptj.74.6>

DOROTA ROJSZCZAK-ROBIŃSKA

Uniwersytet im. Adama Mickiewicza, Poznań

Łacińskie źródła jako kontekst badań historycznojęzykowych¹

Streszczenie

Pytania o łacińskie źródła średniowiecznych polskich tekstów od dawna były obecne w badaniach nad początkami polskiego języka religijnego. W ostatnich latach, m.in. dzięki nowemu wydaniu *Rozmyślenia przemyskiego* (tzw. wydanie fryburskie), w którym równolegle do tekstu zestawiono podstawę źródłową, problem źródeł zaczęli podejmować coraz częściej także historycy języka, najczęściej poprzestając na istniejących spisach źródeł. W artykule wskazano kilka problemów, które dzięki własnym poszukiwaniom źródeł łacińskich mogły uzyskać nowe rozwiązania. Autorka jest przekonana, że choć poszukiwania źródeł nie są zadaniem historii języka, to powinny stale towarzyszyć badaniom najstarszych tekstów religijnych.

Wyrazy kluczowe: łacina, źródło, apokryf, średniowiecze, edycja tekstu, stylistyka

Latin sources in historical linguistic research

Abstract

Questions about the Latin sources of medieval Polish texts have been asked from the outset of research in the origins of Polish religious language. The new edition of *Rozmyślenie przemyskie* (Przemysł Meditation), in which Latin sources are juxtaposed with Old-Polish text, has triggered interest in sources among Polish historical linguists, most of whom have accepted the existing lists of sources. The article points to a number of problems for which new solutions have been found thanks to a new search for Latin sources. The author is convinced that while searching for sources is not the task of the history of language, it should continually accompany studies on the oldest religious texts.

Keywords: Latin, source, apocrypha, Middle Ages, text edition, stylistics

¹ Badania realizowane w ramach grantu Narodowego Centrum Nauki (projekt nr 2017/26/E/HS2/000083) pt. „Początki języka polskiego i kultury religijnej w świetle średniowiecznych apokryfów Nowego Testamentu”.

Po łacińskie źródła historycy języka sięgają prawie od zawsze. Nie sposób sobie wyobrazić np. diachronicznych badań składniowych bez kontekstu dzieł łacińskich. Znacząca część staropolskich tekstów to utwory tłumaczone lub częściowo oparte na źródłach obcych. Bez porównania z podstawą nie jest możliwa analiza rozwoju polskiej składni. Szczególnie widoczne jest to przy analizowaniu łacińskich konstrukcji składniowych i ich adaptacji na gruncie języka polskiego czy obcego. Np. kalka łacińskiej konstrukcji ACI, słabo obecna w *Rozmyślaniu przemyskim*² (dalej – RP), jest dużo bardziej popularna w XVI w.³ Również na kartach XVI-wiecznej *Sprawy chędogiej* (Sch) jest znacznie więcej kalk tej konstrukcji. Zenon Klemensiewicz zauważył tę prawidłowość: „szerokie jej [tj. konstrukcji ACI – DRR] użycie przypada na wiek XVI” (Klemensiewicz 1999: 314).

Podobnie jest z szykiem przydawki zaimkowej w grupie imiennej. Zauważono, że w tekstach starszych prawie zawsze stosuje się szyk naturalny dla języków słowiańskich (zob. Koziara 1998; Masłej 2013; Rojszczak-Robińska 2016: 283–289). Natomiast w tekstach XVI-wiecznych w języku religijnym widoczna jest tendencja do stosowania szyku łacińskiego z przydawką po rzeczowniku (RP 531/18–19: *i słożył swe odzienie*; Sch 24v/13–14: *i położył odzienie swoje*⁴).

Do źródeł łacińskich sięgają też badacze leksyki dawnej. Nie sposób wyobrazić sobie badań staropolskiej synonimii bez odniesienia do wspólnej podstawy źródłowej (tylko w takich sytuacjach, gdy jednej jednostce łacińskiej odpowiada kilka polskich słów, można w miarę bezpiecznie mówić o synonimii; zob. np. Białoskórska 1994; Kleszczowa 2002). Obecne są zatem w świadomości badaczy glos (np. Belcarzowa 1981, 1983, 1997, 2001), przekładów psalterzy (np. Kamińska 1981), modlitw podstawowych (np. Masłej 2016) czy kanonów mszy św. (np. Leńczuk 2013).

Także stylistyka, szczególnie staropolska, sięga do źródeł łacińskich, by analizować sposób przekładu, wierność tłumaczenia, dobór tekstów źródłowych w wypadku dzieł wieloźródłowych, sposób wprowadzania cytatów, korzystanie z rozwiązań retorycznych wypracowanych przez źródło lub samodzielność w tym zakresie. Zagadnienia można by mnożyć (zob. np. Ostrowska 1978; Mika 2002; Kwilecka 2003; Rojszczak-Robińska 2012).

Sięganie do źródeł łacińskich jest też częstą praktyką edytorów tekstów dawnych. Czasem jest to proste zestawianie wydawanego tekstu z jego łacińską podstawą (jak np. w wydaniu staropolskiego przekładu *Ewangelii Nikodema*, zob. Izydorczyk, Wydra 2007). Stosowaną praktyką jest też uzupełnianie fizycznych (wynikających z materialnego uszkodzenia rękopisu) luk w tekście za źródłem łacińskim. Tak np. postępują wydawcy *Kazań świętokrzyskich* w miejscach, w których znane jest źródło uszkodzonego fragmentu tekstu, jak tu uzupełniają za *Ewangelią wg św. Mateusza* 2,1: *Teć to, prawi, gdaż przydzie, zbawi l<ud swój ot> wielikich grzechow* (KŚw IV/2-3). Zdarza

² Olga Ziółkowska (2016) notuje zaledwie 5 przykładów na 847 kartach.

³ Wg badań Ziółkowskiej – 14 użycie na 122 kartach *Rozmyślań dominikańskich* (RD).

⁴ O ile nie zaznaczono inaczej, fragmenty RP pochodzą z wydania fryburskiego (Keller, Twardzik 1998–2000). Cytaty z pozostałych apokryfów podaję we własnej transkrypcji.

się też poprawianie przypuszczalnych błędów kopisty czy wręcz – uzupełnianie tekstu polskiego za podstawą źródłową, mimo że rękopis nie jest uszkodzony, jak np. w tym fragmencie RP: *Tedy odpowiedzieli jemu niektorzy mędrzy <a licemiernicy> rzekąc* (RP 330/5-6), dodając za źródłem (Mt 12,38⁵: *Tunc responderunt ei quidam de Scribis et Phariseis, dicentes*) *a licemiernicy* (szerzej na ten temat zob. Rojszczak-Robińska 2012, 2017; Mika 2017). Do problemu jeszcze wrócę.

W ostatnich latach widoczny jest rozwój badań nad źródłami staropolskich tekstów. Szczególnie widać to na przykładzie *Rozmyślenia przemyskiego*. Niektóre źródła wskazał już pierwszy wydawca tekstu, Aleksander Brückner (1903). Przełomowa okazała się jednak praca Tadeusza Dobrzeńieckiego (1969) *Łacińskie źródła Rozmyślenia przemyskiego*. Ponadstustronicowe zestawienie tekstu polskiego z łacińską podstawą (głównie: *Vita rhythmica*, *Historia scholastica* Piotra Comestora i Ewangelie) na wiele lat zamknęło poszukiwania źródeł dla tego najobszerniejszego staropolskiego tekstu. Temat powrócił wraz z nowym wydaniem RP. Niektóre fragmenty, dla których źródeł nie odnalazł Dobrzeńiecki, w wydaniu fryburskim (RP 1998, 2000) są już lokalizowane. Po tym wydaniu wskazano dodatkowo ok. 150 fragmentów źródłowych (Mazurkiewicz 2004, 2007; Rojszczak 2007, Rojszczak-Robińska 2015), a nawet dwa lata temu wskazano kolejne możliwe zależności i kierunki poszukiwań (Borowiec 2015a, 2015b, 2016).

Rozmyślenie przemyskie jest pod tym względem zabytkiem szczególnym. Nie wiele innych dzieł średniowiecznych może się poszczycić tak precyzyjnie zestawionymi spisami źródeł, są wśród nich m.in. *Kazania świętokrzyskie* (Stępień 2009) czy *Kazanie praskie* (Mazurkiewicz, Stępień 2007).

Co warto zauważyć, wśród badaczy, włączających się w uzupełnienia spisu źródeł dla polskich tekstów średniowiecznych, prawie nie ma językoznawców, co oczywiście nie znaczy, że z nich nie korzystają.

W niniejszym artykule zbieram efekty wieloletnich badań nad źródłami staropolskich apokryfów, a szczególnie nad podejściem staropolskich autorów do tychże. Celem jest pokazanie zalet własnych poszukiwań źródłowych, a także wskazanie kilku sytuacji, w których odkrycie źródła dla danego fragmentu lub dodanie nowego źródła do dotychczas odnalezionych zmieniło interpretację analizowanego fragmentu.

Znalezienie nowego źródła pozwala podtrzymać lub odrzucić decyzje wydawców o wprowadzeniu zmian do tekstu (szerzej o roli kontekstu w wydawaniu tekstów dawnych zob. Borowiec 2015a; Rojszczak-Robińska 2017), jak tutaj, w opisie uzdrowienia teściowej Szymona Piotra. Czytamy:

transliteracja: vzryącz ya myly Ies^{us} dotknął szye yey rzekąc a nyemocz ta ysta nathychmyast yą opvsczyła y byla vsdroyvona

⁵ Cytaty z Pisma Świętego pochodzą z wydania: *Biblia łacińsko-polska, czyli Pismo Święte Starego i Nowego Testamentu. Podług tekstu łacińskiego Wulgaty, i przekładu polskiego x. Jakóba Wujka T.J. z komentarzem Menochiusza T.J. przełożonym na język polski*. Przedruk: Komorów 1997.

transkrypcja fryburska: Użrąc ją miły Jesus dotknął sie jej r[z]ęki[cz] a niemoc ta ista natychmiast ją opuściła i była uzdrowiona (RP 288/7–10)

źródło wskazane w wydaniu: et tetigit manum ejus, et dimisit eam febris, et surrexit, et ministrabat eis (Mt 8,15).

Są dwa kanoniczne opisy tego wydarzenia. Jeden we wskazanej wyżej *Ewangelii wg św. Mateusza* (Jezus uzdrawia dotknięciem ręki). Drugie przedstawienie tej sceny znajduje się w *Ewangelii wg św. Łukasza*, wg którego do uzdrowienia konieczne było słowo – rozkaz: *et stans super illam imperavit febrī; et dimisit illam* (Łk 4,39). Staropolski autor znał zapewne obie wersje (o tworzeniu harmonii ewangelicznych w RP zob. Rojszczak-Robińska 2012: 57–68; 2014). Być może należałoby zatem – zgodnie z rękopisem – pozostawić w tekście zapisane w nim *rzekąc* (*dotknął szye yey rzekąc*) i transkrybować tekst następująco: *Użrąc ją miły Jesus dotknął sie jej rzekąc a niemoc ta ista natychmiast ją opuściła i była uzdrowiona*. Mam wrażenie, że pisarz RP połączył obie wersje – uzdrowienie dokonało się przez słowo i dotyk.

Pozostaje inny problem – jest to jedyne *rzekąc* w tekście RP, które nie wprowadza wypowiedzi cudzej. Skąd mogła się wziąć taka konstrukcja? Otóż w wypadku uzdrowień opisanych na kartach Pisma Świętego, a także w licznych – znanych z pewnością autorowi RP – hagiografiach ważne były rodzaj i kolejność działań. Niekiedy do przywrócenia zdrowia wystarczył sam gest (jak przy uzdrowieniu ślepego z Betsaidy – Mk 8,22–26 lub człowieka chorego na puchlinę – Łk 14,1–6) albo samo słowo (jak przy uzdrowieniu syna dworzanina – Mt 8,5–13 czy niewidomego pod Jerychem – Mk 10,46–52). Czasem potrzebne było jedno i drugie, ale wówczas też miało znaczenie to, czy pierwszy był gest (jak w wypadku uzdrowienia trędowatego – Mt 8,1–4 czy paralytyka – Mk 2,1–12), czy pierwsze było słowo (jak przy uzdrowieniu epileptyka – Mk 9,14–29 lub niewiasty w szabat – Łk 13,10–17). Tutaj, ponieważ nie ma żadnej wersji łączącej oba sposoby (św. Augustyn w *De consensu Evangelistarum* w ogóle nie porusza kwestii tej rozbieżności), autor nie wiedział. Takie skonstruowanie zdania daje możliwość, by w tej kwestii nie podejmować decyzji, by zachować wrażenie, że czynności są jednoczesne (*dotknął sie rzekąc* ‘dotknął jednocześnie mówiąc’). Ponadto, należy zauważyć, że w Ewangeliach, gdy uzdrowieniu towarzyszyło zarówno słowo, jak i gest, właśnie takiej konstrukcji używano, np.:

Et extendens Jesus manum, tetigit eum, dicens: Volo, mundare (Mt 8,3)

Tunc tetigit oculos eorum, dicens: Secundum fidem vestram fiat vobis (Mt 9,29)

Et extenens manum, tetigit eum dicens: Volo: mundare (Łk 5,13)

Ipsa autem tenens manum ejus clamavit, dicens: Puella, surge (Łk 8,54).

Oczywiście tam zawsze po *dicens* następowało przytoczenie słów Jezusa. Autor RP zna te rozwiązania:

Tako nasz miły zbawiciel dotknął sie ich oczu rzekąc: „Podług waszej wiary stań sie wam” (RP 297/21–298/2).

Tutaj jednak apokryfista nie miał za kim przytoczyć wypowiedzi Jezusa, bo żaden z Ewangelistów jej nie zanotował. To tłumaczy nietypowe dla RP użycie imiesłowu *rzekąc*.

Czasami bez wskazania dodatkowych źródeł językoznawca może jedynie zaobserwować pewne zjawiska, nie ma jednak podstaw do ich interpretacji. Tak jest np. gdy zauważy, że tłumaczenie danego wersetu jest niepełne, że staropolski pisarz nie przetłumaczył jakiegoś fragmentu. Na podstawie samej analizy wierności wobec źródła nie można odpowiedzieć na pytanie o przyczynę.

Opis rozmnożenia chleba w RP kończy się zdaniem: *Ostawiwszy tu miły Jesus tushcze, wsiadszy w łodzią i poszedł od nich* (RP 360/9–11), będące tłumaczeniem wersu 39. rozdziału 15. *Ewangelii wg św. Mateusza* (Mt 15,39): *Et dimissa turba, ascendit in naviculam, et venit in fines Magedan*. Autor nie przetłumaczył informacji o miejscu docelowym: *in fines Magedan*. Możliwą motywację takiej decyzji poznamy dopiero w kontekście innego tekstu źródłowego. Wskazywano już na to, że brak (omiłanie) może być sposobem na rozwiązanie kwestii rozbieżnych wersji ewangelistów (zob. Rojszczak-Robińska 2015). Także tutaj musimy uwzględnić dodatkowo *Ewangelię wg św. Marka*, w której podano inną nazwę miejscową: *Et statim ascendens navim cum discipulis suis, venit in partes Dalmanutha* (Mk 8,10). Magedan (inaczej Magdala) i Dalmanuta to dwa miasteczka starożytnej Galilei, oba położone nad Jeziorem Tyberiadzkim. Polski autor nie podjął decyzji, które z nich wskazać, więc ominął to miejsce.

Podobna sytuacja jest też w innym miejscu, w opisie chodzenia po wodzie⁶:

Jako już był wieczor, tako jego zwolenicy poszli ku morzu. A gdyż są wstąpili w łodzią, tako już mrok padł, iże było ciemno. Miły Jesus wtenczas k nim nie przyszedł, ale sam ostał na puszczy (RP 350/10–15; transkrypcja – DRR).

Ewangeliści różnili się między sobą m.in. w kwestii wskazania celu podróży. Źródłem bezpośrednim jest *Ewangelia wg św. Jana*: *Ut autem sero factum est, descenderunt discipuli ejus ad mare. Et cum ascendissent navim, venerunt trans mare in Capharnaum: et tenebrae jam factae erant, et non venerat ad eos Jesus* (J 6,16–17), który zapisał, że uczniowie udawali się do Kafarnaum. Św. Marek podał, że płynęli do Bet-saidy (*Et statim coegit discipulos suos ascendere navim, ut praecederent eum trans fretum ad Bethsaidam, dum ipse dimitterer populum* Mk 6,45). A św. Mateusz nie wskazał w ogóle celu podróży (*Et statim compulit Iesus discipulos ascendere in naviculam, et praecedere eum trans fretum, donec dimitteret turbas* Mt 14,22). W zachowanej wersji RP brak jest tłumaczenia fragmentu *in Capharnaum*. To nie błąd kopyisty, ani niewierność wobec źródła, lecz raczej świadoma decyzja o opuszczeniu tego, co w źródłach różne. To jeden z kilku wypracowanych przez pisarza RP sposo-

⁶ Przykład ten omawiałam już w: Rojszczak-Robińska 2014: 117.

bów na utrzymanie spójności fabularnej źródeł (zwłaszcza Ewangelii). Jeśli będziemy korzystać z wydania (które zawsze jest interpretacją tekstu), to często takich miejsc w ogóle nie zauważymy:

Jako już był wieczor, tako jego zwolenicy poszli ku morzu. A gdyż są wstąpili w łodzią, <poszli przez morze do Kafarnaum>. Tako już mrok padł, iże było ciemno a miły Jesus wtenczas k nim nie przyszedł, ale sam ostał na puszcy (RP 350/10–15; edycja fryburska).

Tak jest też w miejscu, będącym tłumaczeniem wersetu 35. rozdziału 22. *Ewangelii wg św. Łukasza*:

Et dixit eis: Quando misi vos sine sacculo, et pera et calceamentis: numquid aliquid defuit vobis? At illi dixerunt: Nihil (Łk 22,34–36).

Polscy autorzy przetłumaczyli ten fragment tak:

A potem, obrociwszy sie ku zwolnikom swym i rzekł jim: „Namili synowie, gdym was słał ku przepowiadaniu Ewangeliej przez workow i przez toboły, aza wam czego kiedy niedostało?”. A oni rzekli są jemu: „Nic niedostawało nam” (SCh 35v/16–36r/3).

A miły Jesus począł potym mówić k nim: „Kiedym was słał przez miecha i przez toboły, azali wam czego niedostawało?” A oni rzekli wszyscy: „Nic, wszegosmy dosyć mieli” (RP 581/12–18; transkrypcja – DRR).

Zatym miły Jesus począł lepak mówić k nim: „Gdym was słał bez miecha albo kalety, azali wam czego niedostawało?” A oni rzekli wszytcy: „Nic, wszegośmy dosyć mieli” (MPJ 46r/15–19).

W swojej wypowiedzi Jezus przypomina wcześniejsze rozesłanie uczniów, tzw. mowę misyjną. Każdy z Ewangelistów zachował ją w nieco innym kształcie, w różny sposób przekazując zalecenia Jezusa. Wg św. Mateusza Jezus posłał uczniów bez sandałów (*Non peram in via, neque duas tunicas, neque calceamenta, neque virgam* Mt 10,10), wg św. Marka w sandałach (*Et praecepit eis, ne quid tollerent in via, nisi virgam tantum; non peram, non panem, neque in zona aes, sed calceatos sandaliis, et ne induerentur duabus tunicis* Mk 6,8–9). Św. Łukasz samą mowę misyjną przytoczył dwa razy⁷, raz nie mówiąc w ogóle o butach (*Et ait ad illos: Nihil tuleritis in via, neque virgam, neque peram, neque panem, neque pecuniam, neque duas tunicas habetis* Łk 9,3), drugi raz stwierdzając, że Jezus posłał uczniów bez sandałów (*Nolite portare sacculum, neque peram, neque calceamenta, et neminem per viam salutaveritis* Łk 10,4). Trzej polscy autorzy pominęli tę informację, która była kłopotliwa.

Wydawcy RP i tu zdecydowali się uzupełnić ten fragment w oparciu o Pismo Święte, co sprawia, że decyzja staropolskiego autora o stworzeniu harmonii ewangelicznej pozostaje niezauważalna:

A miły Jesus począł potym mówić k nim: „Kiedym was słał przez miecha i przez toboły, <i przez botow>, azali wam czego niedostawało?” A oni rzekli wszyscy: „Nic, wszegosmy dosyć mieli” (RP 581/12–18; edycja fryburska).

⁷ Trzy, jeśli uwzględnić przypomnienie o niej w omawianym fragmencie.

Własne poszukiwania dodatkowych źródeł są pomocne przy ocenie prawidłowości niektórych interpretacji. Np. ten fragment RP:

Po małej chwili, jakoby po godzinie, przystępując ci, którzy około ognia stali i rzekli Piotrowi: „Zaprawdę ty jeś z jego zwolenników, bo cię twa mowa zjawia, bo jeś galilejski człowiek” (RP 690/17–23),

Ewa Ostrowska analizowała, porównując jego początek (*po małej chwili, jakoby po godzinie*) do *Lamentu świętokrzyskiego* (*ciężka moja chwila, krwawa godzina*). Badaczka twierdziła:

Chwila – godzina. Szeregi ciężka moja *chwila*, krwawa *godzina* oraz występujące w dodatkach do „Bogurodzicy”: „Już nam czas, *godzina* Grzechowa się kajaci” nie są, jakby się mogło zdawać, parami synonimów charakterystycznymi dla stylu tego okresu. [...] Szereg występujący w „Żalach”, *chwila – godzina*, był stosowany rzadko, podobnie jest w „Rozmyśłaniu przemyskim” (690): *po małej chwili, jakoby po godzinie*; to „jakoby” mogłoby być próbą unowocześnienia zwrotu, który wcześniej mógł brzmieć: *po chwili, po godzinie*. – *Godzina* i tu, i w „Żalach” w znaczeniu ‘moment’; por. jeszcze współczesne: *przyjść w złą godzinę* ‘zły czas, moment’ (Ostrowska 1978: 166).

W rzeczywistości nie jest to kwestia stylistyczna, nie jest to też unowocześnienie zwrotu. To konsekwencja wyborów źródłowych. Trzecie zaparcie się św. Piotra według św. Mateusza (a i Marka – Mk 14,70) dokonało się krótko po drugim: *Et post pusillum accesserunt, qui stabant, et dixerunt Petro* (Mt 26,73), według świętego Łukasza dopiero po godzinie: *et intervallo facto quasi horae unius* (Łk 22,59).

Dopiero uwzględnienie obu potencjalnych źródeł pozwala dostrzec, że w wypadku RP zestawienie słów *chwila – godzina* to efekt łączenia rozbieżnych wersji Ewangelistów. Staropolski pisarz przejął za św. Mateuszem niedokładne określenie czasu *po małej chwili* i następnie doprecyzował za św. Łukaszem: *jakoby po godzinie*. *Jakoby* jest tłumaczeniem łacińskiego *quasi*, nie „próbą unowocześnienia zwrotu”, jak chce Ostrowska.

Czasem potrzebne jest nie tyle nowe źródło łacińskie, ile nowe spojrzenie na toż źródło i rozważenie, co mógł widzieć w nim staropolski autor lub jak mógł je rozumieć. Przykładowo w scenie przeprowadzania Jezusa przez rzekę Cedron w *Sprawie chędogiej* czytamy:

Potem pan Jesukryst z wielkiego strachu swej męki i przykrego jęcia, a też z okrutnego i ciężkiego targania, pragnienie cierpiąc, pić pożytał, jako mowi w żołtarzu: „Z bystrości albo z utrudzenia na drodze pił” (SCh 53v/17–54r/5).

Autor SCh, stwierdzając, że Jezus był spragniony, przytoczył fragment *Psalmu 109*: *De torrente in via bibet* (Ps 109,7). Ten sam cytat pojawił się też w analogicznym miejscu *Rozmyśłania przemyskiego* i *Rozmyślań dominikańskich*:

Wtenczas nasz miły zbawiciel ilkoż z strachu trudnej męki, telko z onego okropnego siepania, telko od onego przykrego jęcstwa ciężko pragnąc picia a pożądać – bo o tem prorokował Dawid profeta rzekąc: „Z strumienia albo z rzeki na drodze będzie pić” (RP 661/6–13).

Pana Jesusa zepchnęli w rzeczkę barzo kamienistą, abociem widzieli, iż Jesus pożałował picia jako człowiek umęczony i umordowany, jakoż o tem Dawid prorokował rzekąc: „<Z> strumienia na drodze będzie pić” (RD 82/14–20).

Łacińskie *torrens* to ‘potok’, czyli ‘rzeczka mniejszych rozmiarów’. Autor RD *de torrente* przełożył, używając polskiego rzeczownika *strumień*, pisarz RP przy pomocy dwóch rzeczowników, tj. *strumień* i *rzeka*. W średniowieczu mogło to być już utarte połączenie. *Słownik staropolski* notuje wiele zestawień tego typu, np.:

- nyekthorey rzeki alybo strvmyenya – aliquo fluvio vel rivulo (*Sul* 21);
- myedzy dwiema dziedzinama strvmyeny abo rzeka (łac. *alvei fluviorum* *Dział* 7);
- ad fluminem seu torrentem, kv rzecze, kv strumyenyv, Cedron (*Kalużn* 272).

Autor SCh łącińskie *torrens* także oddał dwoma polskimi słowami, tj. *bystrość* i *utrudzenie*. Słowo *bystrość* mogło funkcjonować w znaczeniu ‘woda, rzeka’. Mogło zatem być odpowiednikiem leksykalnym łącińskiego *torrens*. *Słownik polszczyzny XVI wieku* notuje znaczenie ‘szybki, rwący prąd, nurt’. *Utrudzenie* z pewnością nie odpowiada łącińskiemu *torrens*. Szukając przyczyny takiego zestawienia tych dwóch polskich słów, postawiłam dwie hipotezy:

albo kopista miał przed sobą rękopis, w którym w sposób być może już zniekształcony lub niewyraźny zapisano *strumienia*, a w pierwotnym tekście był szereg z *bystrości* albo z *strumienia* (Rojszczak-Robińska 2016: 85).

To jednak wątpliwe, bo pisarz musiałby być wyjątkowo nieuważny, a zachowana kopia SCh jest dość staranna. Jako drugą przyczynę dopuściłam:

podejście autora SCh do prorocत्व starotestamentalnych. Widać bowiem, że rozumiał te, które pojawiały się w Ewangeliach (znał być może komentarze do Pisma, które te kwestie wyjaśniały). Nie rozumiał jednak teologii fragmentów pochodzących z innych, niż ewangeliczne, źródeł (Rojszczak-Robińska 2016: 85).

Jeśli spojrzymy na ten przykład z innej perspektywy, z perspektywy łącińskiego źródła, dojrzymy trzecią możliwość. Polski pisarz mógł nie być pewny, jakie słowo widzi w rękopisie, który być może miał przed sobą, czy *torrens*, czy *tormentum*. Jeśli spojrzymy na Ablatiw, to okazuje się, że podobieństwo jest duże: *de torrente* lub *de tormento*. Łacińskie *tormentum* to ‘męka, cierpienie, ból’. Pisarz SCh, nie będąc pewien, jaki rzeczownik widzi, być może zdecydował się na oddanie obu znaczeń: z *bystrości* albo z *utrudzenia* na drodze. To tylko potwierdza jego słabą znajomość prorocत्व starotestamentalnych, którą pokazały analizy wszystkich miejsc, w których się do nich odwołuje (szerzej zob. Rojszczak-Robińska 2016: 46–94).

Podobnie jest, gdy w źródłach łacińskich pojawia się czasownik *salvo* ('wyzwalać, zbawiać', także 'leczyć'). Polscy pisarze tłumaczyli go na różne sposoby, czasem prawdopodobnie myląc z czasownikiem *salveo* (o znaczeniu 'być pozdrowionym', używanym w formułach powitalnych i pożegnalnych). Np. w tym fragmencie RP tłumaczonym z *Ewangelii Nikodema*:

clamabant: "Hosanna filio David" [...] dicentes: Salva igitur, qui es in caelis, benedictus, qui venit in nomine Domini (EN 1.3);

a wołali: „Osanna synu Dawidowu, jakoby rzekł: Zdrow bądź, który *jesteś* na niebiesiech; błogosławiony, który wyszedł w imię boże” (RP 752/14-17).

Pisarz RP przełożył *salva* jako *zdrow bądź*. Być może widział (lub zapamiętał – jeśli cytował z pamięci) *salve*, nie *salva*. *Akta Piłata* funkcjonowały w wielu różnych redakcjach, a polscy pisarze z pewnością korzystali z innych odpisów. Autor SCh to samo zdanie przełożył inaczej: *A wołając mówili: „Zbaw nas, jenże jeś w niebiesiech i błogosławiony”* (SCh 75r/16–17).

W innym miejscu autor RP tłumaczył *salvum me fac* (dosłownie: 'uczyni mnie zbawionym/uleczonym') jako *uzdrowi mię*:

Salvum me fac, Deus, quoniam intraverunt aquae usque ad animam meam (Ps 68,2).

A tu mienią, iż miły Krystus siedząc w tej ciemnicy i piał ten psalm: „Salvum me fac, Deus. Uzdrowi mię, Gospodnie, iż przyszły wody aż do dusze mojej” (RP 733/1–6).

Pisarz SCh przełożył dokładnie:

I mówią mistrzewie Świętego Pisma, iż w ten czas mówił pan Jesus ten psalm: „Zbawionego mie uczyni Boże, bo weszli są wody aż ku duszy mojej” (SCh 63v/5–6).

W jeszcze innym miejscu autor RP oddał oba znaczenia:

Osanna, przez ktore sie słowo tako wiele rozumie, jakoby rzekł: O Gospodnie, zbaw mię albo zdrow bądź a slutuj sie nad nami rychlej i zbaw nas naszego zamęta (RP 754/1–5).

Świadomość tego, że zapis źródłowy mógł być niejasny (łatwo pomylić *e* z *a* – *salve* – *salva*), ułatwi interpretację takich miejsc.

Poszukiwanie źródeł nie jest zadaniem językoznawstwa, ale – w wypadku badania początków polszczyzny literackiej, zatem najstarszych tekstów religijnych – nie sposób poprzestać na istniejących spisach źródeł, których teksty funkcjonowały w wielu odmiennych redakcjach, różniących się od siebie nie tylko treścią, czasem fleksją, leksyką, konstrukcjami składniowymi, szykiem czy retoryką. Spisy zestawiane są najczęściej przez niejęzykoznawców, zatem osoby niewyczulone na subtelne różnice słów czy konstrukcji.

Z tego powodu takie poszukiwania źródłoznawcze powinny towarzyszyć interpretacji tekstów dawnych, w szczególności wieloźródłowych. Podobnie jak próba odwrócenia perspektywy – spojrzenie od strony łacińskiego źródła.

Wykaz skrótów

EN – *Ewangelia Nikodema*
 KŚw – *Kazania świętokrzyskie*
 MPJ – *Męka Pana Jezusowa*
 RD – *Rozmyślania dominikańskie*
 RP – *Rozmyślanie przemyskie*
 SCh – *Sprawa chędogo o męce Pana Chrystusowej*

Bibliografia

- BELCARZOWA Elżbieta (1981): *Głosy polskie w łacińskich kazaniach średniowiecznych. Część I.* – Wrocław: Ossolineum.
- BELCARZOWA Elżbieta (1983): *Głosy polskie w łacińskich kazaniach średniowiecznych. Część II.* – Wrocław: Ossolineum.
- BELCARZOWA Elżbieta (1997): *Głosy polskie w łacińskich kazaniach średniowiecznych. Część III.* – Wrocław: Wydawnictwo Naukowe DWN.
- BELCARZOWA Elżbieta (2001): *Głosy polskie w łacińskich kazaniach średniowiecznych. Część IV.* – Wrocław: Wydawnictwo Naukowe DWN.
- BIAŁOSKÓRSKA Mirosława (1994): Staropolskie semantyczne dublety słowotwórcze rzeczowników i ich kontynuacje w polszczyźnie XVI oraz XX wieku. – [w:] Marian KUČAŁA, Zdzisława KRAŻYŃSKA (red.): *Studia historycznojęzykowe I.* – Kraków: IJP PAN, 29–42.
- BOROWIEC Karolina (2015a): Filologiczne wydanie *Rozmyślania przemyskiego* a czeski *Život Krista Pána.* – *LingVaria* 2, 179–192.
- BOROWIEC Karolina (2015b): W stronę badań słowiańskich kontekstów staropolskich apokryfów (1). O czeskiej podstawie fragmentu *Rozmyślania przemyskiego.* – *Slavia Occidentalis. Linguistica* 1, 7–25.
- BOROWIEC Karolina (2016): W poszukiwaniu czeskiego kontekstu *Rozmyślania przemyskiego.* *Rozmyślanie przemyskie* a *Život Krista Pána.* Rekonesans. – [w:] Jerzy BARTMIŃSKI, Artur TIMOFIEJEW (red.): *Rozmyślanie przemyskie – świadectwo średniowiecznej kultury religijnej i językowej.* – Przemysł: Państwowa Wyższa Szkoła Wschodnioeuropejska, 95–110.
- BRÜCKNER Aleksander (1903): *Literatura religijna w Polsce średniowiecznej*, t. 2: *Pismo Święte i apokryfy.* – Warszawa: Księgarnia Gebethera i Wolffa.
- DOBRZENIECKI Tadeusz (1969): Łacińskie źródła „Rozmyślania przemyskiego”. – [w:] Jerzy LEWAŃSKI (red.): *Średniowiecze. Studia o kulturze* 4. – Wrocław: Ossolineum, 196–521.
- GÓRSKI Karol, KURASZKIEWICZ Władysław, wyd. (1965): *Rozmyślania dominikańskie.* – Wrocław: Ossolineum.
- IZYDORCZYK Zbigniew, WYDRA Wiesław (2007): *A Gospel of Nicodemus preserved in Poland.* – Turnhout: Brepols.
- KAMIŃSKA Maria (1981): *Psalterz floriański. Monografia językowa.* – cz. 1, Wrocław: Ossolineum.
- KELLER Felix, TWARDZIK Waław, wyd. (1998–2004): *Rozmyślanie przemyskie. Transliteracja, transkrypcja, podstawa łacińska, niemiecki przekład.* – Freiburg i. Br.: Weiher.
- KLEMENSIEWICZ Zenon (1999): *Historia języka polskiego.* – Warszawa: PWN.

- KLESZCZOWA Krystyna (2002): Bogactwo staropolskiej synonimii. – [w:] Ireneusz OPACKI (red.): *Dzieło literackie i książka w kulturze. Studia i szkice ofiarowane Profesor Renardzie Ocieczek w czterdziestolecie pracy naukowej i dydaktycznej*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego, 279–287.
- KOZIARA Stanisław (1998): O szyku zaimkowej przydawki dzierżawczej w polskich przekładach Psalterza (na przykładzie psalmu Miserere). – *Język Polski* LXXVIII, 62–69.
- KWILECKA Irena (2003): *Studia nad staropolskimi przekładami Biblii*. – Poznań: Wydawnictwo UAM.
- LEŃCZUK Mariusz (2013): *Staropolskie przekazy kanonu Mszy Świętej. Wariantywność leksykalna*. – Warszawa: Wydawnictwo IBL.
- MASŁEJ Dorota (2013): Problem szyku w staropolskich przekładach *Modlitwy Pańskiej*. – *Kwartalnik Językoznawczy* 4, 33–43.
- MASŁEJ Dorota (2016): *Modlitwa Pańska w polskim średniowieczu. Znad staropolskich rękopisów*. – Poznań: Wydawnictwo Rys.
- MAYENOWA Renata, POPŁAWSKI Feliks, red. (1966–2016): *Słownik polszczyzny XVI wieku*. – Wrocław: Ossolineum, Warszawa: IBL PAN.
- MAZURKIEWICZ Roman (2004): Nowo znalezione źródła łacińskie do tomu II. – [w:] Felix KELLER, Waclaw TWARDZIK (wyd.): *Rozmyślanie przemyskie. Transliteracja, transkrypcja, podstawa łacińska, niemiecki przekład*. – Freiburg i. Br.: Weiher, 544–546.
- MAZURKIEWICZ Roman (2007): Kilkanaście dalszych uzupełnień do źródeł „Rozmyślania przemyskiego”. – [w:] Roman LASKOWSKI, Roman MAZURKIEWICZ (red.): *Amoenitates vel lepores philologiae*. – Kraków: Wydawnictwo Lexis, 198–201.
- MAZURKIEWICZ Roman, STĘPIEŃ Paweł (2007): *Źródła i konteksty Kazania praskiego (rekonesans)*. – [w:] Roman LASKOWSKI, Roman MAZURKIEWICZ (red.): *Amoenitates vel lepores philologiae*. – Kraków: Wydawnictwo Lexis, 202–221.
- MIKA Tomasz (2002): *Maryja, Jezus, Bóg w „Rozmyślaniu przemyskim”. O nazywaniu osób*. – Poznań: Wydawnictwo Naukowe UAM.
- MIKA Tomasz (2017): O idealizacji i innych postawach wobec tekstu staropolskiego. – *Język Polski* XCVII, 7–23.
- OSIEWICZ Marek, WÓCIK Rafał, WYDRA Wiesław, wyd. (2015): *Męka Pana Jezusowa*. – Poznań: Wydawnictwo Naukowe UAM.
- OSTROWSKA Ewa (1978): *Z dziejów języka polskiego i jego piękna*. – Kraków: Wydawnictwo Literackie.
- ROJSZCZAK Dorota (2007): Do spisu źródeł „Rozmyślania przemyskiego”. – [w:] Roman LASKOWSKI, Roman MAZURKIEWICZ (red.): *Amoenitates vel lepores philologiae*. – Kraków: Wydawnictwo Lexis, 276–284.
- ROJSZCZAK-ROBIŃSKA Dorota (2012): *Jak pisano „Rozmyślanie przemyskie”*. – Poznań: Wydawnictwo Rys.
- ROJSZCZAK-ROBIŃSKA Dorota (2014): Teologia w służbie językoznawstwa. Nowe perspektywy w badaniach nad najstarszymi polskimi tekstami. – *LingVaria* 2, 115–122.
- ROJSZCZAK-ROBIŃSKA Dorota (2015): Łacińskie źródła *Rozmyślania przemyskiego*. Pytania, problemy perspektywy. – [w:] Tomasz MIKA, Dorota ROJSZCZAK-ROBIŃSKA, Olga STRAMCZEWSKA (red.): *Staropolskie Spotkanie Językoznawcze 1: Jak badać teksty staropolskie*. – Poznań: Wydawnictwo Rys, 135–172.
- ROJSZCZAK-ROBIŃSKA Dorota (2016): *Staropolskie pasje: Rozmyślanie przemyskie, Sprawa chędogo, Karta Rogawskiego. Źródła – język – fabuła*. – Poznań: Wydawnictwo Rys.
- ROJSZCZAK-ROBIŃSKA Dorota (2017): Wydawanie staropolskich narracji biblijno-apokryficznych. Rola kontekstu. – [w:] Karolina BOROWIEC i in. (red.): *Staropolskie Spotkania Językoznawcze 2: Jak wydawać teksty dawne*. – Poznań: Wydawnictwo Rys, 255–278.
- STĘPIEŃ Paweł (2009): Objasnienia źródłowe. – [w:] Paweł STĘPIEŃ (red.): *„Kazania świętokrzyskie”. Nowa edycja. Nowe propozycje badawcze*, Warszawa: Biblioteka Narodowa, 241–272.

URBAŃCZYK Stanisław, red. (1953–2002): *Słownik staropolski*. – Warszawa–Wrocław–Kraków: PAN.

VRTEL-WIERCZYŃSKI Stefan, wyd. (1933): *Sprawa chędogo o męce Pana Chrystusowej i Ewangelia Nikodema*. – Poznań: PTPN.

ZIÓŁKOWSKA Olga (2016): *Specyfika języka „Rozmyślań dominikańskich” w świetle składni i leksyki*. – Poznań: Wydawnictwo Rys.

doi: <http://dx.doi.org/10.17651/bptj.74.7>

JOANNA KULWICKA-KAMIŃSKA

Instytut Języka Polskiego UMK, Toruń

Piśmiennictwo religijne Tatarów Wielkiego Księstwa Litewskiego jako źródło badań historycznojęzykowych

Streszczenie

Przedmiotem opisu są pochodzące z okresu od XVI do XX w. zabytki piśmiennictwa religijnego Tatarów Wielkiego Księstwa Litewskiego. Należą do nich księgi podstawowe: rękopisy Koranu, tefsiry, kitaby, chamaili; księgi pomocnicze: sufry, tedźwidy, wokabularze; amulety: hramotki, daławary, nuski, tablice, muhiry. Zabytki te powstawały w wielokulturowej i wielojęzycznej Rzeczypospolitej Obojga Narodów, w środowisku muzułmanów, żyjących w warunkach diaspory w otoczeniu chrześcijańskim. Miało to wpływ na wielowarstwowość tatarskiej twórczości, w tym na jej multilingwizm, czyli obecność języków orientalnych oraz słowiańskich zapisanych alfabetem arabskim. Inną cechą wskazującą na wielosegmentowość tatarskich ksiąg jest łączenie różnych tradycji religijnych i przekładowych. Stąd wskazuje się na związek tych zabytków z literaturą biblijno-psalterzową europejskiego średniowiecza i renesansu oraz z tureckimi tefsirami. Na szczególną uwagę zasługuje fakt, że Tatarzy już w XVI w. dokonali trzeciego na świecie przekładu Koranu na język słowiański, tj. polszczyznę północnokresową.

Słowa kluczowe: dziedzictwo kulturowe, kitabistyka, historia języków polskiego i białoruskiego, Koran, tefsir, transliteracja

Religious writings of the Tatars of the Grand Duchy of Lithuania as a source of data in historical and linguistic research

Abstract

The object of description are the 16th—20th-century religious writings of the Tatars of the Grand Duchy of Lithuania, literary monuments, which include basic books: manuscripts from the Quran, tefsirs, kitabs, chamails; auxiliary books: sufirs, tajwids, vocabularies; and amulets: hramotkas, dalavars, nuskas, tables, muhirs. They were written in the multicultural and multilingual Polish-Lithuanian Commonwealth, in the community of Muslims, who formed a diaspora in a Christian environment. This

contributed to the Tatar works being multilayered and multilingual, which is reflected in the presence of Oriental and Slavic languages written in the Arabic script. The combining of various religious traditions and translation methods makes the Tatar books multisegmental and it is argued that they are related to the Biblical-psaltic literary tradition of the European Middle Ages and Renaissance, and to the Turkish tefsirs. Worthy of emphasis is the fact that already in the 16th century the Tatars translated the Quran into a Slavic language, a dialect of Polish used in the North Borderland, and it was the third translation of Quran in the world.

Keywords: cultural heritage, kitabistics, history of Polish and Belarusian, Quran, tefsir, transliteration

1. Tatarskie piśmiennictwo przekładowe¹

1.1. Geneza

Polska terminologia muzułmańska zaczęła kształtować się w XVI w.² W drugiej połowie tego stulecia powstawały prawdopodobnie pierwsze zabytki piśmiennictwa Tatarów Wielkiego Księstwa Litewskiego (WKL). W celu ustalenia chronologii i przyczyn ich genezy można posłużyć się m.in. metodą filologiczną w odniesieniu do tekstu najstarszego traktatu, autorstwa samych Tatarów, z 1558 r. (*Risāle-i Tātār-i Leh*) oraz poddając analizie zawartość treściową i cechy językowe tatarskich manuskryptów³, a także badając kontekst historyczny (m.in. wpływy reformacji i kontrreformacji)⁴.

Za bezpośrednią przyczynę powstania tatarskiej literatury przekładowej uznaje się utratę przez Tatarów ich ojczystego języka, tj. dialektów kipczackich, które przynieśli ze sobą osadnicy ze Złotej Ordy. Wyznawali oni islam sunnicki, tworząc diasporę w otoczeniu chrześcijańskim. Z czasem przyjęli języki autochtonów – polski lub białoruski. Proces ten następował od XV w., a moment jego zakończenia datuje się na drugą połowę XVI stulecia (por. Łapicz 1986: 33–60). Wówczas to zaczęła powstawać słowiańskojęzyczna, lecz zapisana alfabetem arabskim, tatarska literatura przekła-

¹ Charakterystykę podstawowych cech piśmiennictwa Tatarów Wielkiego Księstwa Litewskiego oparto na ustaleniach zawartych w pracy Kulwickiej-Kamińskiej (2013).

² Por. Kulwicka-Kamińska 2004: 14: „Z powodu braku, w polszczyźnie tego okresu, większej ilości leksemów związanych z kręgiem kultury i religii islamu (choć istniał już przekład Koranu na język łaciński) omówiono hasło *soltan* i wyrażenie *Mahometowy grob* (zanotowane w cytacie), jak również apelatywa nazywające orientalne realia życia, które w późniejszych dobach nabierają nowego znaczenia, związanego z kulturą i religią islamu”.

³ Problem może być jednak bardziej złożony nie tylko ze względu na właściwości polszczyzny północnokresowej, ale też na sakralny wymiar tekstów, a co za tym idzie sięganie do określonych źródeł (psalterze, przekłady Biblii) i konserwującą rolę języka religijnego. W grę może zatem wchodzić przełączanie się na inny kod – kod języka religijnego. W tatarskich zabytkach można bowiem wyróżnić cały szereg cech, charakterystycznych dla staropolskiej – średniowiecznej i renesansowej – prozy religijnej.

⁴ Zjawisko to jest szeroko opisywane w literaturze przedmiotu – por. m.in. Drozd 1999, Łapicz 2007: 99–117, Gądomski, Łapicz 2008–2009: 33–58, Łapicz 2009: 293–310, Suter 2004, Тарэлка, Сынькова 2009, Miškinienė 2001, Konopacki 2010: 129.

dowa, której głównym celem było zapewnienie Tatarom odrębności etnicznej i tożsamości kulturowej, identyfikowanej już w XVI w. wyłącznie poprzez islam (Konopacki 2010: 130).

W kitabistyce przyjmuje się, że podstawowymi cechami tatarskiego piśmiennictwa są: zapis rękopiśmienny, anonimowość autorów, wielojęzyczność – obecność warstw słowiańskiej (polskiej i białoruskiej) oraz orientalnej (tureckiej, arabskiej i perskiej), stosowanie alfabetu arabskiego do zapisu tekstów orientalnych i słowiańskich oraz dominację tematyki religijnej, związanej z islamem.

1.2. Autorzy i kopiści

Twórczość religijna Tatarów WKL – analogicznie do innych literatur rękopiśmiennych – jest anonimowa (Drozd 1995: 36). Powodem owej anonimowości mogła być motywacja religijna – dążenie do ideału pracy na chwałę Boga, przyświecające już twórcom epoki średniowiecza⁵. Ponadto wskazuje się, że przekaz rękopiśmienny czynił z kopisty współautora, a ze społeczności – właściciela praw autorskich (Drozd 1995: 36).

Nie znając konkretnych nazwisk autorów, można opierać się jedynie na przesłankach historycznojęzykowych i społecznych. Istotne byłoby więc sięgnięcie do archiwaliów i poszukanie tatarskich autorów przekładów wśród zamożnej elity rodowej stanu duchownego. Z elity tej rekrutowali się bowiem także posłowie, gońcy, tłumacze etc. w służbie dyplomatycznej Rzeczypospolitej (zwłaszcza do połowy XVII w.)⁶. Jest w tym analogia do twórców rękopiśmiennych zabytków doby staropolskiej, którzy należeli do ówczesnej (XIII–XV w.) elity intelektualnej, czyli do sfery wyższego duchowieństwa (za: Walczak 1993: 36–37).

Artur Konopacki (2010: 159) przytacza imiona dwóch pisarzy-kopistów, mianowicie Hodyny (kitab z 1645 r.) oraz imama mińskiego Urjasza syna Ismaila (tefsir datowany na 1686 r.). Jednakże inne opracowania, np. *Белорусские тексты, писанные арабским письмом, и их графико-орфографическая система* podają, że Hodyna był tłumaczem jednego z tekstów zawartych w kitabie (Антонович 1968: 125). Analogicznie przedstawia się rzecz z imamem Urjaszem, który według ustaleń Andrzeja Drozda i Michaiła Tarełki był tłumaczem tefsiru z 1686 r. (Drozd 2004: 237–248; Тарелка 2006: 34–43).

Inna sytuacja dotyczy datowania i autorstwa kopii rękopisu, albowiem kopiści z reguły informowali o dacie i miejscu ukończenia pracy, nazwisku przepisyującego,

⁵ Takie ideały głoszono także w renesansie, choć z reguły już podpisywano swe dzieła – por. *Słownik łacińsko-polski*, 1544: „Ja brat Bartłomiej z Bydgoszczy, Minorityta, nie jestem i nigdy nie byłem człowiekiem lubiącym się chęłpić swoją pracą, człowiekiem, który by podejmował trud pisarski dla zdobycia rozgłosu, czy próżnej chwały (jak mnie o to różni ludzie obwiniają), lecz pracowałem dla dobra Rzeczypospolitej i dla pomnożenia chwały Bożej, 1544 r.”.

⁶ Por. badania historyków i heraldyków w tym zakresie, m.in. Bairašauskaitė 2006: 27–40, Dumin 1991: 147–163; 2006, Konopacki 2010 oraz zagadnienie znajomości języków wschodnich w dawnej Polsce – Baranowski 1950, Reychman 1950, jak też uwagi na ten temat – Drozd 1995: 36–39.

a czasem i zamawiającego księgę (np. znany jest kopsta tefsiru z 1725 r. Jest nim Bogdan Assanowicz, syn Szabana, imama z Łowczyc – za: Тарэлка 2015: 183). Kopistami byli zazwyczaj mułlowie, muezini i znachorzy, a także „ludzie starsi, traktujący swą pracę jako jeden z pobożnych uczynków” (Borawski, Dubiński 1986: 250). Do głównych miejsc kopiowania tatarskiej literatury religijnej należały: Mińsk, Lachowicze, Śmiłowicze, Słonim, Nowogródek (Konopacki 2010: 160).

Historię losów danego zabytku można odtworzyć na podstawie informacji zawartych w kolofonie, umieszczanym przeważnie na końcu tekstu i sygnalizującym tym samym zakończenie czynności kopiowania. Kolofon obecny w środku rękopisu może świadczyć o powstawaniu manuskryptu w etapach lub kopiowaniu go przez więcej niż jedną osobę (Radziszewska 2010).

1.3. Rodzaje

Wśród zabytków piśmiennictwa tatarskiego zachowanych do czasów współczesnych wyróżnia się kilka typów (opracowano za: Drozd, Dziekan, Majda 2000: 12–16)⁷. Należą do nich **księgi podstawowe**:

- rękopisy Koranu – jedne z najpopularniejszych tatarskich ksiąg rękopiśmiennych, liczące około 200–300 kart. Na ich zawartość składają się: Koran w języku arabskim oraz zasady recytacji tej Księgi, a także modlitwy. Świętą Księgę islamu zwykle kopiowano i przekazywano z pokolenia na pokolenie jako jedno z najcenniejszych dóbr⁸;
- tefsiry – w islamie są to komentarze do Koranu. U Tatarów WKL mianem tefsirów określa się obszerne manuskrypty zawierające pełny tekst Koranu, z podwierszowym przekładem na polszczyznę północnokresową, uzupełnionym o warstwę egzegetyczną, a także o modlitwy i obrzędy towarzyszące recytacji Koranu. Liczą średnio 400–500 kart. Dla społeczności muzułmańskiej były niezwykle kosztowne, stąd zamawiano je jako *wakuf* (ar. waqf ‘w prawie muzułmańskim zapis pieniężny na cele religijne, charytatywne lub publiczne’) dla meczetu⁹;
- kitaby – zabytki o zróżnicowanej objętości i treści, mające charakter kodeksów. Cechuje je heterogeniczna forma. Zawierają teksty o różnorodnej tematyce, w tym: opowiadania o prorokach i „świętych” islamu, fragmenty apokryfów, Koranu, tefsirów i Biblii, a także tedźwidy, słowniczki do nauki języków orientalnych – arabskiego i tureckiego, ponadto utwory moralizatorskie, często związane z tematyką eschatologiczną, hadisy, utwory dewocyjne i modlitewne, ko-

⁷ Por. typologię i szczegółową charakterystykę tatarskiego piśmiennictwa – Szynkiewicz 1935: 138–139), Łapicz 1986: 65–69; 2008a: 37–38, Radziszewska 2008: 138–143, Suter 2004: 6–9, Konopacki 2010: 131–153 i szereg innych.

⁸ Panuje pogląd, że Koran przepisywano do początków XX w., jednakże znane są przykłady sporządzania późniejszych kopii, np. rękopis Koranu z lat 70. XX w., przepisany przez imama Sulejmana Rafałowicza z Iwia.

⁹ Najpełniej ten typ zabytku omawia Suter (2004), ponadto Łapicz (2008b: 70).

mentarze do niektórych sur Koranu, opisy obrzędów, nauki dotyczące rytuałów religijnych i prawa muzułmańskiego, polemiki religijne, prognostyki, teksty magiczne, rzadziej natomiast teksty o tematyce pozareligijnej, jak np. fragmenty literatury staropolskiej, głównie chrześcijańskiej proveniencji. Liczą zazwyczaj 150–300 kart¹⁰;

- *chamaile*¹¹ – tatarskie modlitewniki. Na ich zawartość składają się teksty o tematyce religijnej, wśród których są też opisy rytuałów muzułmańskich, związanych z filarami i dogmatami islamu. Księgi te są przykładem łączenia wiary monoteistycznej z obrzędowością ludową (tj. zwyczajami i wierzeniami koczowniczych ludów tureckich, a także ludności polskiej i białoruskiej) oraz magią. Stąd obecne są w nich, np. prognostyki, formuły magiczno-lecznicze, formuły egzorcyzmów ludowych itp. Wyróżnia się zazwyczaj dwa typy *chamaile*ów: *molliański* – służący *mołłom* (ar. *mawla* ‘pan’; stąd *mołna*, *molla*/*mułla* ‘duchowny tatarski’) do odprawiania obrzędów oraz *fałdżejski* – służący *fałdżej*om (ar. *fa’l* ‘(dobry) znak, dobra wróżba’; stąd tur. *fal* ‘przepowiadanie, wróżenie’¹² + tur. *-dži*), czyli osobom parającym się wróżbiarstwem. Rzadko występują one w czystej formie, zwykle łączą w sobie treści *molliańskie* i *fałdżejskie*. Liczą średnio od 100 do 300 kart¹³; **księgi pomocnicze**:
- *sufry* – niewielkiej objętości rękopisy, zawierające jedną trzydziestą część Koranu, zwaną *dżuz* (ar. *ğuz* ‘część, udział’; ‘część składowa, składnik’)¹⁴. Mają cel dydaktyczny – służą do nauczania dzieci, a także zawierają zestawy modlitw do odmawiania nad zmarłym w noc przed pogrzebem;
- *tedżwidy* – podręczniki napisane w języku tureckim. Służą do nauki recytacji Koranu z wyjaśnieniem zasad gramatycznych. Zawierają międzywierszowy przekład polsko-białoruski świętej Księgi islamu. Są sporadycznie spotykane w *kitabach* bądź *chamaile*ach;
- *wokabularze* – księgi rękopiśmienne, zawierające zazwyczaj turecką leksykę i frazeologię wraz z tłumaczeniami na języki polski i/lub białoruski; **amulety**:
- *hramotki* – zwoje z modlitwami noszone w celach ochronnych. W treści zawierają głównie tzw. *ajaty* ochronne i inne wybrane wersety koraniczne o takiej funkcji, ponadto zasady wiary muzułmańskiej, arabskie modlitwy, służące zapewnieniu życia wiecznego oraz chroniące przed „złymi przygodami”, formuły i figury magiczne (tzw. *talsimy* – por. ar. *ṭilasm* ‘talizman’) itp.;

¹⁰ Tematykę najstarszych znanych dziś *kitabów* przytacza Konopacki (2010: 139–140). Szczegółowo ten typ piśmiennictwa charakteryzują: Jankowski, Łapicz 2000: 13–20 oraz Łapicz 1986, a także Dufala 2008–2009: 205, jak również Miśkinienė 2000: 30–36.

¹¹ Różne etymologie tej nazwy zestawia Radziszewska (2010).

¹² Znaczenia terminów tureckich podano m.in. za: Stachowski 1975–1986 oraz Jankowski, Łapicz 2000.

¹³ Najpełniejszą, jak dotąd, charakterystykę rękopiśmiennych *chamaile*ów zawiera rozprawa Radziszewskiej (2010) oraz eadem (2008–2009: 221–227); por. także: Miśkinienė 2001.

¹⁴ Każda z trzydziestu części dzieli się z kolei na dwie partie – *hizb* (ar. *ḥizb* m.in. ‘sześćdziesiąta część Koranu’), zaś jeden *hizb* składa się z czterech ćwiartek – *rub* (ar. *rub* ‘ćwierć, ćwiartka, czwarta część, ułamek: jedna czwarta, kwarta’). Podział ten ułatwia recytację (np. w miesiącu ramadan można recytować jedną część każdego dnia). Ponadto jest pomocny przy uczeniu się Koranu na pamięć.

- daławary – zwoje modlitewne składane do grobu wraz ze zmarłym oraz rękopisy tworzone i wykorzystywane przez żyjących. Są pisane w języku arabskim bądź tureckim (rzadziej). Zawierają: wybrane sury i wersety z Koranu, zasady wiary (tzw. Pięć imanów), modlitwy pokutne i zapewniające zbawienie. Za życia służą więc ochronie przed wszystkim, co złe, a po śmierci pomagają w pozytywnym rozstrzygnięciu Sądu Ostatecznego (por. Radziszewska 2010: 221);
- nuski – niewielkich rozmiarów karteczki, zawierające wypisane z chamailu teksty arabskich modlitw, formuł lub figur magicznych o charakterze ochronnym lub leczniczym;
- tablice – papierowe lub kartonowe plansze, wieszane w meczetach i domach. Zawierają opisy obrzędów, zasady etyki, okolicznościowe formuły modlitewne, wybrane wersety koraniczne i ich tłumaczenia itp. Pełnią funkcję informacyjną;
- muhiry – ozdobne plansze lub tkaniny, zawieszane na ścianach domów i meczetów. Porównuje się je do tzw. świętych obrazów w Kościele katolickim. Przedstawiają budowle i miejsca o wartości sakralnej lub magicznej. Zawierają również liczne inskrypcje koraniczne. Pełnią funkcję dekoracyjną oraz modlitewną, stąd wyróżnia się muhiry inskrypcyjne i muhiry przedstawieniowe (por. ich obszerną charakterystykę: Łyszczarz 2013: 258–261).

Powyższy podział związany z kryterium formy i treści jest intuicyjnie przyjmowany przez samych Tatarów (por. Drozd, Dziekan, Majda 2000: 12).

Ponadto można wymienić:

- wyciągi z kitabów zawierające głównie zasady praktyki religijnej (najstarszy tego typu zabytek datuje się na połowę XVII w. – por. opis <http://www.islamicmanuscripts.info/reference/books/Drozd-2000-Tatarow/Drozd-2000-Tatarow-048-083.pdf>);
- krótkie podręczniki i słowniki języka tureckiego (por. Miśkinienė 1995; 2008a; 2008b).

W najnowszym katalogu zabytków Tatarów WKL (Тарэлка 2015: 6–8) proponuje się podział na: **księgi-kodeksy**, do których zalicza się rękopisy Koranu, sufry, tefsiry, kitaby, chamaily, tzw. leczebniki, czyli księgi lecznicze; **zwoje**, do których należą daławary i hramotki, oraz **jednoarkuszowe rękopisy**, wśród których wymienia się wyciągi z tedźwidów, wyciągi z chamailów, nuski, modlitwę *Lāhi*. Wskazuje się ponadto, że w katalogu z 2011 r. był obecny taki typ zabytku jak tedźwid. Podaje się też mało znane manuskrypty, wymykające się wyżej przyjętej klasyfikacji: rękopis z domu modlitwy w Słoniemiu zawierający tylko jedną modlitwę w języku tureckim oraz rękopis z Narodowej Biblioteki Białorusi, zawierający tylko przekład 78–114 sur Koranu i objaśnienie do nich w języku polskim – jest to zabytek o cechach tefsiru i kitab. Henryk Jankowski (2015: 140) postuluje uzupełnienie powyższej typologii o materiał epigraficzny (napisy nagrobne) oraz dokumenty, w tym pisma prawne i parafialne, a także korespondencję prywatną.

1. 4. *Wielojęzyczność*

Zabytki Tatarów WKL cechuje wielowarstwowość językowa, na którą składają się warstwa słowiańska (polska i białoruska) i warstwa orientalna (arabska, turecka – głównie osmańsko-turecka i perska). Istotne jest przy tym zwrócenie uwagi na różne typy zabytków tatarskiego piśmiennictwa. W świecie muzułmańskim język arabski był – i jest nadal – językiem sfery sacrum (por. Drozd, Dziekan, Majda 2000: 22). W tatarskiej diasporze waloru języka religijnego nabrał ponadto język turecki, chociaż uważano go za język niższej rangi. Dlatego też większe wpływy tego języka widoczne są w innego typu zabytkach niż tefsiry. Jeden z tłumaczy – Hodyna (Kitab z Kazania) – wprost napisał, że przełożył tę księgę z języków perskiego i tureckiego (Антонович 1968: 125).

Jeśli zaś chodzi o języki słowiańskie, na które przekładano muzułmańskie teksty religijne ze źródeł arabskich lub tureckich, to zabytki piśmiennictwa Tatarów WKL odzwierciedlają stosunki socjolingwistyczne panujące w Księstwie, gdzie na substrat białoruski nałożył się język polski (Drozd 1999: 27–29; Suter 2004: 126). Przeważają więc teksty mieszane językowo: białoruskie i białorusko-polskie, co ilustruje panującą wśród Tatarów dwujęzyczność. Wyjątek stanowią natomiast tatarskie tefsiry i teksty modlitewne, liturgiczne, które przekładano na język polski (polszczyznę północnokresową) ze względu na ich wysoką rangę – objawionego słowa Bożego (por. Suter 2004: 10–12, 31–33, 84).

1.5. *Cechy rodzajowe*

Tatarskie zabytki to teksty o zróżnicowanej objętości, klasyfikacji genologicznej (heterogeniczny charakter literatury religijnej – wykorzystanie innorodzajowych cech dokonuje się nawet w ramach jednego utworu), jak też tematyce. Zawierają one obok tekstów arabskich – białoruskie, polskie, rzadziej rosyjskie, a także tureckie. „Czynności przekładowe wymagały od podejmujących ten trud dość znacznej erudycji, a przede wszystkim znajomości języków orientalnych, w jakich pisane były teksty muzułmańskie, tzn. arabskiego, tureckiego i perskiego. Tłumacze musieli znać jednocześnie języki, na jakie przekładali, czyli białoruski i polski”¹⁵. Języki te Tatarzy wprowadzali do swego piśmiennictwa posługując się alfabetem arabskim. Osoby dokonujące pierwszych przekładów tekstów religijnych na języki polski i białoruski dobrze znały wymowę arabską i turecką, a może także wymowę innych języków z kręgu krajów kultury arabskiej. Celem ich pracy był fonetyczny zapis tłumaczeń słowiańskich za pomocą liter arabskich. W zapisie tekstów dominuje więc zasada fonetyczna, a nie reguły ortograficzne. Tłumacze i kopiści wykorzystywali jednakże swoje doświadczenia w zakresie ortografii polskiej i białoruskiej oraz arabskiej i tureckiej. Dlatego w tekstach, obok przeważających zapisów fonetycznych, pojawiają się nierzadko zapisy podporządkowane właściwym dla danego okresu zasadom pisowni polskiej i białoruskiej, a także arabskiej i tureckiej (za: Łapicz 1986: 88).

¹⁵ „A jeżeli nasi nie mogą rozmawiać po Arabsku lub Turecku, za to mają znajomość innych języków: każdy bowiem w naszym kraju posiada dwie mowy, tam powszechnie używane” (fragment z *Risāle-i Tātār-i Leh* za: Łapicz 1986: 61) – por. Suter 2004: 92.

1.6. Zawartość

Tatarzy tworzyli teksty samodzielnie bądź adaptowali fragmenty literatury staropolskiej i dostosowywali je do kanonów i dogmatów islamu bądź też przekładali teksty orientalne na języki polski i/lub białoruski. Ponadto zapożyczali ze źródeł wschodnich teksty arabskie lub tureckie i zapisywali je alfabetem arabskim. A. Drozd (1999: 42) podaje, że Tatarzy korzystali ze źródeł turkojęzycznych, do których zalicza piśmiennictwo Złotej Ordy bądź Turkiestanu w języku czagatajskim, a na przełomie XV/XVI w. piśmiennictwo z Turcji Osmańskiej i zabytki tureckie o cechach późniejszych oraz ze źródeł arabskich, którymi najczęściej były modlitwy w klasycznym języku literackim, powstałe w Anatolii, ale też na Powołżu, Krymie czy w Turkiestanie.

2. Słowiańskie aljamiado

Zabytki religijne Tatarów WKL należą do tzw. słowiańskiego aljamiado. Na świecie prowadzone są badania nad literaturą typu aljamiado (por. ar. *al-luġa al-‘aġamiyya* ‘język obcy’ – za: Dziekan 2015: 75–99). Obejmują one twórczość w języku miejscowym, zapisywaną odpowiednio przystosowanym alfabetem arabskim i rozwijającą się pod silnym wpływem tradycji i kultury islamu. Charakterystyce tej odpowiada np. literatura muzułmańska w zachodniej i południowej Afryce, a także w południowej i wschodniej Azji, literatura aljamiado na Półwyspie Iberyjskim, muzułmańskie teksty serbskie w Bośni i Hercegowinie, literatura albańskich bejtedżinów. Kryteria twórczości aljamiado spełnia także piśmiennictwo Tatarów WKL – jest realizowane w językach miejscowych: polskim i/lub białoruskim, zostało zapisane odpowiednio przystosowanym alfabetem arabskim, rozwijało się pod silnym wpływem tradycji i kultury islamu wyznawanego przez społeczność muzułmańską WKL (por. Łapicz 2014: 59–70). Charakteryzuje się ono również cechami tylko sobie właściwymi – inną niż wyżej wymieniona literatura genezą oraz przeznaczeniem. Tatarska twórczość rękopiśmienna nie miała bowiem celu misyjnego, a służyła zachowaniu przez Tatarów, żyjących w warunkach diaspory, własnej tożsamości. To istotnie różni ją od literatury pisanej grafiańską arabską na Bałkanach czy na Półwyspie Iberyjskim. Do Albanii lub innych krajów bałkańskich (Bośni, Hercegowiny) islam dotarł w wyniku podboju przez muzułmańską Turcję. Ich literatura miała charakter świecki i składała się głównie z utworów lirycznych. W Hiszpanii zaś katolicyzm narzucono moryskom, tj. hiszpańskim muzułmanom, którzy po kryjomu praktykowali islam. Ich literatura obejmowała z kolei modlitewniki, traktaty religijne itd. Twórczość litewsko-polskich muzułmanów była w znakomitej większości literaturą religijną, powstającą w warunkach dobrowolnej, niewymuszonej sławizacji i powoli postępującej chrystianizacji, co niewątpliwie przyczyniło się do dokonania przez Tatarów WKL pierwszego na świecie przekładu Koranu na język słowiański w formie tatarskiego tefsiru.

3. Znaczenie tatarskiego piśmiennictwa dla badań historycznojęzykowych, w tym dla rozwoju kitabistyki

Kitabistyka to oryginalna, naukowa subdyscyplina filologiczna (Gadomski, Łapicz 2008–2009: 247–268), łącząca filologię polską i słowiańską z orientalistyką (arabistyką i turkologią) oraz z kulturoznawstwem i religioznawstwem. Przedmiotem jej badań jest materialne i niematerialne dziedzictwo kulturowe Tatarów WKL, a w szczególności oryginalna, rękopiśmienna literatura, zwłaszcza religijna w postaci kitabów, choć odnosi się ona także do innych odmian rękopiśmiennych tekstów litewsko-polskich muzułmanów (por. Suter 2004: 3; Miškinienė 2009a: 235; Тарэлка, Сынькова 2009: 16. Najistotniejsze zadania i postulaty badawcze kitabistyki zostały zaprezentowane i szczegółowo omówione w artykule Łapicza (2008a: 38–44) oraz w opracowaniu: Gadomski, Łapicz 2008–2009: 46–51).

W 2015 r. powołano na Wydziale Filologicznym Uniwersytetu Mikołaja Kopernika w Toruniu specjalistyczną jednostkę naukowo-badawczą – Centrum Badań Kitabistycznych. Pracami Centrum kieruje Joanna Kulwicka-Kamińska. Do zadań statutowych jednostki należą: inicjowanie, prowadzenie i koordynowanie interdyscyplinarnych i międzynarodowych działań o charakterze naukowo-badawczym, edytor-skim oraz popularyzatorskim z zakresu kitabistyki, a w szczególności:

- 1) wprowadzenie do świadomości społecznej oraz do obiegu naukowego oryginalnego źródła filologicznego, jakim jest piśmiennictwo religijne muzułmanów WKL, wraz z pierwszym, dokonany prawdopodobnie już w XVI w. przekładem Koranu na polszczyznę północnokresową¹⁶;
- 2) prowadzenie długofalowych prac dokumentacyjnych, edytorskich i badawczych o fundamentalnym znaczeniu dla dziedzictwa i kultury nie tylko narodowej, ale też europejskiej;
- 3) gromadzenie, opracowywanie i udostępnianie tekstów źródłowych do interdyscyplinarnych badań nie tylko filologicznych, ale także historycznych, socjologicznych, religioznawczych, literaturoznawczych, kulturoznawczych, teologicznych etc., zwłaszcza że można wskazać podobieństwa systemowe między rękopisami tatarskimi i polskimi rękopiśmiennymi tekstami średnio-wiecznymi i renesansowymi;

¹⁶ Wieloletnie badania w tym zakresie wieńczą prace nad przygotowaniem krytycznej edycji tatarskiego tłumaczenia, podjęte przez międzynarodowy interdyscyplinarny zespół naukowców w ramach projektu Narodowego Programu Rozwoju Humanistyki, pt. *Tefsir – projekt filologiczno-historycznego opracowania oraz krytycznego wydania tzw. tefsiru Tatarów Wielkiego Księstwa Litewskiego z 2. połowy XVI w. (pierwszego przekładu Koranu na język polski)*. Międzynarodowy projekt „TEFSIR”, którego autorami są Joanna Kulwicka-Kamińska i Czesław Łapicz, był realizowany w latach 2013–2016 w ramach Narodowego Programu Rozwoju Humanistyki, moduł 1.2 nr 12 H 12 0041 81. Od 2017 r. prowadzone są prace składające się na drugą część projektu (nr 11 H 16 0319 84). Ich rezultatem ma być krytyczne wydanie tatarskiego tefsiru, zawierające pełny tekst zabytku, czyli warstwę słowiańską w transliteracji oraz warstwę orientálną – transkrypcję wybranych fragmentów tekstu arabskiego i/lub tureckiego oraz opatrzone komentarzem filologiczno-historycznym. Szczegółowy opis projektu „TEFSIR”: <http://www.tefsir.umk.pl/>.

- 4) udział na zasadach eksperckich w inicjatywach innych ośrodków krajowych i zagranicznych podejmujących problematykę mniejszości narodowych i kulturowych, w tym, np. identyfikacja i opis zabytkowych rękopisów tatarskich;
- 5) edycja z komentarzem filologiczno-historycznym tłumaczenia Koranu na język polski z lat dwudziestych XIX w. autorstwa wileńskich filomatów – ks. Dionizego Chlewińskiego i Ignacego Domeyki (por. Kulwicka-Kamińska, Lewicka 2015: 65–79).

Badania nad literaturą religijną Tatarów WKL sytuują się na pograniczu językoznawstwa (historia języka polskiego, teolingwistyka, lingwistyka islamska, kitabisyka) i translatoologii.

Znaczenie tatarskiego piśmiennictwa jest niebagatelne, albowiem z jednej strony ukazuje związki między wielowątkową kulturą Rzeczypospolitej a światem islamu, z drugiej zaś – jest istotnym źródłem do prowadzenia badań filologicznych, gdyż zawiera niezbadane dotąd warstwy gramatyczne i leksykalne polszczyzny północno-kresowej z okresu od XVI do XX w. (por. Kulwicka-Kamińska 2017). Ponadto wszechstronnie ilustruje sposoby i metody sławizacji – polonizacji i białorutenizacji – orientalizmów (zwłaszcza arabizmów i turcyzmów) (por. Gadomski, Łapicz 2008–2009: 49; Kulwicka-Kamińska 2001: 2). Stanowi także bezcenny materiał do obserwacji procesów interferencyjnych i transferencyjnych w obrębie języków słowiańskich i kontaktów słowiańsko-orientalnych na wszystkich płaszczyznach językowych: graficzno-ortograficznej, fonetyczno-fonologicznej, leksykalno-semantycznej i gramatycznej (por. Łapicz 1986: 219–220), a zwłaszcza przenikania do polszczyzny – związanych semantycznie z islamem – zapożyczeń z języków arabskiego, tureckiego i perskiego (por. Kulwicka-Kamińska 2004).

Muzułmańskie piśmiennictwo religijne dostarcza również materiał do badań nad językiem religijnym, w tym leksykologią diachroniczną. Na jego podstawie można m.in. wskazać kształtowanie muzułmańskiej terminologii religijnej, włączając się tym samym w nurt badań nad leksyką i frazeologią polskiej reformacji. Dowiedziono bowiem związki literatury Tatarów WKL z biblistyką zachodnioeuropejską, w tym z renesansowymi tłumaczeniami Biblii na język polski, a zarazem z koraniczną tradycją przekładową w postaci m.in. tureckich tefsirów. Ponadto wskazano cechy wspólne tatarskiego piśmiennictwa i literatury biblijno-psalterzowej okresu średniowiecza (por. Drozd 1999; Kulwicka-Kamińska 2013). Stąd materiał źródłowy, jakim są tatarskie zabytki, dostarcza przykładów w zakresie sposobów oddawania arabskiej leksyki i frazeologii w językach słowiańskich, w tym także adaptacji form arabskich i tureckich do systemu gramatyczno-leksykalnego tych języków oraz w zakresie stosowanych przez tłumaczy metod przekładu, a zarazem metod transformacji pierwotnego wzoru, takich jak: symplifikacje, redukcje (np. elipsa, anakoluty), amplifikacje, uniwerbalizacje, deskrypcje, substytucje itp.

Prowadzone na podstawie tatarskiej literatury rękopiśmiennej badania do niedawna skupiały się na tekstach słowiańskich i fragmentarycznej analizie warstwy orientalnej. Badania orientalistyczne nad tatarskimi tekstami przekładowymi zostały podjęte głównie przez turkologów, którzy obszarem swoich poszukiwań badawczych

uczynili zapożyczenia wschodnie w języku tekstów i w mowie codziennej oraz imiennictwo Tatarów WKL. Przedmiotem ich analiz stał się również język etniczny Tatarów (por. ustalenia A. Drozda (1999: 15))¹⁷. Jeśli chodzi o arabskojęzyczną zawartość piśmiennictwa religijnego muzułmanów litewsko-polskich, to dotyczyły one wstępnego wskazania zawartości zabytków oraz grafii i ortografii tekstów arabskich obecnych w wybranych tekstach tatarskich (Kulwicka-Kamińska, Lewicka 2015: 70). Do końca lat 90. XX w. dorobek arabistyczny w zakresie kitabistyki stanowiły jedynie prace A. Drozda: studium poświęcone modlitwie za sułtanów wchodzącej w skład *chutb* pt. *Sułtan dua (święteczna modlitwa za sułtanów)* oraz monografia *Arabskie teksty liturgiczne w przekładzie na język polski XVII wieku*, której przedmiotem jest tłumaczenie na język polski trzech arabskich oracji świętecznych obecnych w cha-maile ze zbiorów Biblioteki Wydziału Orientalistycznego Uniwersytetu w Sankt Petersburgu (Kulwicka-Kamińska, Lewicka 2015: 70–71).

Przełom na gruncie obcym przyniosły publikacje kitabistów z Wilna (por. Miśkinienė 2009b) oraz prace Paula Sutura (2004, 2009), w których podjął on próbę oceny tłumaczenia Koranu w aspektach filologicznym i translatologicznym na podstawie sury Al-Baqary, zaś na gruncie polskim praca J. Kulwickiej-Kamińskiej (2013). Nowe perspektywy badawcze otwierają prace prowadzone w ramach projektu „TEFSIR” (zob. Kulwicka-Kamińska, Lewicka 2015: 71–72).

Postulowanym i wciąż aktualnym zagadnieniem badawczym jest zatem wielojęzyczna zawartość tatarskich rękopisów, której pełne i komplementarne opracowanie wymagałoby stałej współpracy slawistów i orientalistów. Badania stylistyczno-językowe i translatologiczne powinny objąć w pierwszej kolejności tatarskie tefsiry. Dla prowadzenia dalszych badań w zakresie przekładu religijnych tekstów muzułmańskich na języki spoza kręgu kultury i tradycji islamu ważne jest zatem krytyczne wydanie najstarszego polskiego tłumaczenia Koranu, czyli tzw. tefsiru, opatrzonego rzetelnym komentarzem filologicznym (por. Gadomski, Łapicz 2008–2009: 46)¹⁸.

¹⁷ Ze współczesnych turkologów można wymienić m.in. Jankowskiego (2008: 147–159; 2013: 149–163; 2015: 139–171) czy Majdę (1995: 139–158; 2013: 203–210).

¹⁸ O doniosłej roli tefsirów i ich unikatowości pisali m.in. Szyrkiewicz (1935: 138 i 140) – por. „Obecnie tefsiry stały się prawie unikatami i przechowują się w nielicznych rodzinach, jako relikwie. Trafiają się egzemplarze z w. XVIII. Ich tłumaczenie jest bardzo dobre, w duchu najlepszych komentatorów muzułmańskich” oraz „...kitabów pozostało bardzo mało, po jednym – po dwa na każdą parafję, a co się tyczy tefsirów, to na całą Rzeczpospolitą znajdzie się nie więcej dziesięciu” i Suter (2004: 6) – por. „Es gibt bis heute kein vollständiges Inventar der litauisch-tatarischen Handschriften. Die bekannten Manuskripte liegen verstreut in Museen und Bibliotheken in Wilna, Minsk, St. Petersburg, Kazan, Warschau, Leipzig oder London. Eine beträchtliche Anzahl von Handschriften befindet sich in Privatbesitz tatarischer Familien, wo sie von Generation zu Generation weitervererbt werden. Diese Manuskripte sind in der Regel der wissenschaftlichen Forschung nur beschränkt zugänglich. Immer wieder gehen Handschriften unwiederbringlich verloren, oft aus Unachtsamkeit oder Gleichgültigkeit dem islamisch-tatarischen Erbe gegenüber”. Przystępując w 2013 r. do prac nad projektem „TEFSIR” badacze z Polski, Litwy i Białorusi mieli świadomość istnienia zaledwie kilkunastu tefsirów. Obecnie wiadomo, że jest ich ponad dwadzieścia (por. Таpэлка 2015).

Bibliografia

- BAIRAŠAUSKAITĖ Tamara (2006): Rodzina tatarska w świetle ksiąg metrykalnych (wiek XIX). – *Biuletyn Historii Pogranicza* 7, 27–40.
- BARANOWSKI Bohdan (1950): *Znajomość Wschodu w dawnej Polsce do XVII wieku*. – Łódź: Łódzkie Towarzystwo Naukowe.
- BORAWSKI Piotr, DUBIŃSKI Aleksander (1986): *Tatarzy polscy. Dzieje, obrzędy, legendy, tradycje*. – Warszawa: Wydawnictwo Iskry.
- DROZD Andrzej (1995): O twórczości literackiej Tatarów w dobie staropolskiej. – [w:] *Tatarzy w Europie i na świecie*. – Poznań: Materiały z ogólnopolskiej sesji naukowej. 10–11 grudnia 1994 r., 33–47.
- DROZD Andrzej (1999): *Arabskie teksty liturgiczne w przekładzie na język polski XVII wieku. Zagadnienia gramatyczne na materiale chutb świętecznych*. – Warszawa: Wydawnictwo Akademickie DIALOG.
- DROZD Andrzej (2004): Koran staropolski. Rozważania w związku z odkryciem tefsiru mińskiego z 1686 roku. – *Rocznik Biblioteki Narodowej* XXXVI, 237–248.
- DROZD Andrzej, DZIEKAN Marek, MAJDA Tadeusz (2000): Piśmiennictwo Tatarów polsko-litewskich (XVI–XX w.). Zarys problematyki. – [w:] *Piśmiennictwo i muhiry Tatarów polsko-litewskich. Katalog zabytków tatarskich*, t. III. – Warszawa: Res Publica Multietnica.
- DUFALA Kristina (2008–2009): Legenda o św. Grzegorzu w kitabie Tatarów – muzułmanów Wielkiego Księstwa Litewskiego. – [w:] Aleksander GADOMSKI, Czesław ŁAPICZ (red.): *Chrestomatia teolingwistyki*. – Symferopol: Universum, 205–220.
- DUMIN Stanisław (1991): Szlachta tatarska w Wielkim Księstwie Litewskim i zmiany jej sytuacji prawnej w XVI–XVIII w. – *Roczniki Historyczne* LVII, 147–163.
- DUMIN Stanisław (2006): *Herbarz rodzin tatarskich Wielkiego Księstwa Litewskiego*. – Gdańsk: Związek Tatarów Polskich.
- DZIEKAN Marek (2015): Zastosowanie pisma arabskiego do zapisu wybranych języków indoeuropejskich. Perspektywa historyczno-porównawcza. – [w:] Joanna KULWICKA-KAMIŃSKA, Czesław ŁAPICZ (red.): *Tefsir Tatarów Wielkiego Księstwa Litewskiego. Teoria i praktyka badawcza*. – Toruń, http://www.tefsir.umk.pl/pliki/Tefsir_Tatarow_WKL.pdf, 75–99.
- GADOMSKI Aleksander (2008–2009): Опыт составления словаря русско-польской терминологии теолингвистики. – [w:] Aleksander GADOMSKI, Czesław ŁAPICZ (red.): *Chrestomatia teolingwistyki*. – Symferopol: Universum, 247–268.
- GADOMSKI Aleksander, ŁAPICZ Czesław (2008–2009): Teolingwistyka: historia, stan współczesny, perspektywy... – [w:] Aleksander GADOMSKI, Czesław ŁAPICZ (red.): *Chrestomatia teolingwistyki*. – Symferopol: Universum, 33–58.
- JANKOWSKI Henryk (2008): The Tatar Name of Sorok Tatary. Keturiasdešimt Totorių Discovered. – [w:] Tamara BAIRAŠAUSKAITĖ, Halina KOBECKAITĖ, Galina MIŠKINIENĖ (red.): *Orientas Lietuvos Didžiosios Kunigaikštijos Visuomenės Tradicijoje: Totoriai ir Karaimai*. – Vilnius: VU Leidykla, 147–159.
- JANKOWSKI Henryk (2013): Imiennictwo Tatarów litewsko-polsko-białoruskich w dawnych dokumentach. – [w:] Joanna KULWICKA-KAMIŃSKA, Czesław ŁAPICZ (red.): *Tatarzy Wielkiego Księstwa Litewskiego w historii, języku i kulturze*. – Toruń: TNT, 149–163.
- JANKOWSKI Henryk (2015): Cechy graficzne i językowe tekstów turkijskich w zapisie kopistów polsko-tatarskich. – [w:] Joanna KULWICKA-KAMIŃSKA, Czesław ŁAPICZ (red.): *Tefsir Tatarów Wielkiego Księstwa Litewskiego. Teoria i praktyka badawcza*. – Toruń, http://www.tefsir.umk.pl/pliki/Tefsir_Tatarow_WKL.pdf, 139–171.
- JANKOWSKI Henryk, ŁAPICZ Czesław (2000): *Klucz do raju. Księga Tatarów litewsko-polskich z XVIII wieku w przekładzie i opracowaniu Henryka Jankowskiego i Czesława Łapicza*. – Warszawa: Wydawnictwo Akademickie DIALOG.

- KONOPACKI Artur (2010): *Życie religijne Tatarów na ziemiach Wielkiego Księstwa Litewskiego w XVI–XIX wieku*. – Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.
- KULWICKA-KAMIŃSKA Joanna (2001): *Kształtowanie się polskiej terminologii islamistycznej*. – Toruń: praca doktorska, Uniwersytet Mikołaja Kopernika (komputeropis).
- KULWICKA-KAMIŃSKA Joanna (2004): *Kształtowanie się polskiej terminologii muzułmańskiej*. – Toruń: TNT.
- KULWICKA-KAMIŃSKA Joanna (2013): *Przekład terminologii religijnej islamu w polskich tłumaczeniach Koranu na tle biblijnej tradycji translatorskiej*. – Toruń: Wydawnictwo Naukowe UMK.
- KULWICKA-KAMIŃSKA Joanna (2017): Piśmiennictwo religijne Tatarów Wielkiego Księstwa Litewskiego jako oryginalne źródło do badań polszczyzny północnokresowej. – *Poznańskie Studia Polonistyczne* 24, 2, 85–110.
- KULWICKA-KAMIŃSKA Joanna, LEWICKA Magdalena (2015): Instytucjonalne formy badań nad dziedzictwem kulturowym Tatarów Wielkiego Księstwa Litewskiego. Centrum Badań Kitabistycznych (UMK, Toruń). – [w:] Artur KONOPACKI (red.): *Tatarszczyzna w badaniach. Konteksty interdyscyplinarne*. – Białystok: Związek Tatarów Rzeczypospolitej Polskiej Oddział Podlaski, 65–79.
- KURZOWA Zofia (1993): *Język polski Wileńszczyzny i kresów północno-wschodnich XVI–XX w.* – Warszawa–Kraków: Wydawnictwo Naukowe PWN.
- ŁAPICZ Czesław (1986): *Kitab Tatarów litewsko-polskich (Paleografia. Grafia. Język)*. – Toruń: Wydawnictwo UMK.
- ŁAPICZ Czesław (2007): Z zagadnień przekładu muzułmańskiej terminologii modlitewnej na język polski i białoruski. – [w:] Joanna KAMPER-WAREJKO, Joanna KULWICKA-KAMIŃSKA, Katarzyna NOWAKOWSKA (red.): *Z przeszłości i teraźniejszości języka polskiego. Księga pamiątkowa dedykowana Teresie Friedelównie*. – Toruń: Wydawnictwo Naukowe UMK, 99–117.
- ŁAPICZ Czesław (2008a): Kitabistyka a historia języka polskiego i białoruskiego. Wybrane zagadnienia. – *Rocznik Slawistyczny* LVII, 31–49.
- ŁAPICZ Czesław (2008b): Glosy, komentarze, objaśnienia etc., czyli o pozakoranicznych dopiskach w rękopiśmiennych tefsirach muzułmanów Wielkiego Księstwa Litewskiego. – [w:] Tamara BAIRASKAUSKAITĖ, Halina KOBECKAITĖ, Galina MIŠKINIENĖ (red.): *Orientas Lietuvos Didžiosios Kunigaikštijos Visuomenės Tradicijoje: Totoriai ir Karaimai*. – Vilnius: VU Leidykla, 69–80.
- ŁAPICZ Czesław (2009): Chrześcijańsko-muzułmańska interferencja religijna w rękopisach Tatarów Wielkiego Księstwa Litewskiego. – [w:] *Lietuvos Didžiosios Kunigaikštystės kalbos, kultūros ir rašties tradicijos*. – Vilnius: VU Leidykla, 293–310.
- ŁAPICZ Czesław (2014): Czy piśmiennictwo Tatarów – muzułmanów Wielkiego Księstwa Litewskiego jest słowiańskim aljamiado? – [w:] Mateusz WARCHAŁ (red.): *W podróży za słowem: księga pamiątkowa z okazji jubileuszu 70-lecia urodzin profesora Emila Tokarza*. – Bielsko-Biała: Akademia Techniczno-Humanistyczna w Bielsku-Białej, 59–70.
- ŁYSZCZARZ Michał (2013): *Młode pokolenie polskich Tatarów*. – Olsztyn–Białystok: UWM.
- MAJDA Tadeusz (1995): Turkish-Byelorussian-Polish Handbook. – *Rocznik Orientalistyczny* XLIX (2), 139–158.
- MAJDA Tadeusz (2013): Osmanizacja pisanego języka Tatarów polsko-litewskich. – [w:] Joanna KULWICKA-KAMIŃSKA, Czesław ŁAPICZ (red.): *Tatarzy Wielkiego Księstwa Litewskiego w historii, języku i kulturze*. – Toruń: TNT, 203–210.
- MIŠKINIENĖ Galina (1995): *Турэчка-беларускі размоўнік 1836 году з збораў Нацыянальнага Музею Літоўскай Рэспублікі ў Вільні*. – New York: Беларускі Інстытут Навукі й Мастацтва.
- MIŠKINIENĖ Galina (2000): O zawartości treściowej najstarszych rękopisów Tatarów litewskich. – *Rocznik Tatarów Polskich* VI, 30–36.
- MIŠKINIENĖ Galina (2001): *Sieniaus lietuvis totorių rankraščiai. Grafika. Transliteracija. Vertimas. Tekstų struktūra ir turinys*. – Vilnius: VU Leidykla.

- MIŠKINIENĖ Galina (2008a): Турецко-польский словарь из китаба Якуба Хасеневича (1840). – [w:] Tamara BAIRAŠAUSKAITĖ, Halina KOBECKAITĖ, Galina MIŠKINIENĖ (red.): *Orientas Lietuvos Didžiosios Kunigaikštijos Visuomenės Tradicijoje: Totoriai ir Karaimai*. – Vilnius: VU Leidykla, 105–121.
- MIŠKINIENĖ Galina (2008b): *Turkų-lenkų kalbų žodynėlis iš lietuvos totorių rankraščio (1840)*. – Vilnius: Instytut Języka Litewskiego.
- MIŠKINIENĖ Galina (2009a): Развитие китабистики в Вильнюсском университете. – *Kalbotyra* LIV (2), 231–237.
- MIŠKINIENĖ Galina (2009b): *Ivano Luckevičiaus kitabas – Lietuvos totorių kultūros paminklas*. – Vilnius: Instytut Języka Litewskiego.
- RADZISZEWSKA Iwona (2008): Rękopisy tatarskie na Podlasiu. – [w:] Tamara BAIRAŠAUSKAITĖ, Halina KOBECKAITĖ, Galina MIŠKINIENĖ (red.): *Orientas Lietuvos Didžiosios Kunigaikštijos Visuomenės Tradicijoje: Totoriai ir Karaimai*. – Vilnius: VU Leidykla, 138–143.
- RADZISZEWSKA Iwona (2008–2009): Chamaity Tatarów Wielkiego Księstwa Litewskiego jako teksty religijne i źródła filologiczne. – [w:] Aleksander GADOMSKI, Czesław ŁAPICZ (red.): *Chrestomatia teolingwistyki*. – Symferopol: Universum, 221–230.
- RADZISZEWSKA Iwona (2010): *Chamaity jako typ piśmiennictwa religijnego muzułmanów Wielkiego Księstwa Litewskiego (na podstawie słowiańskiej warstwy językowej)*. – Toruń: praca doktorska, Uniwersytet Mikołaja Kopernika (komputeropis).
- REYCHMAN Jan (1950): *Znajomość i nauczanie języków orientalnych w Polsce w XVIII w.* – Wrocław: Prace Wrocławskiego Towarzystwa Naukowego.
- STACHOWSKI Stanisław (1975–1986): *Studien über die arabischen Lehnwörter im osmanisch-türkischen*. – Wrocław–Warszawa–Kraków–Gdańsk: Prace Komisji Orientalistycznej PAN.
- SUTER Paul (2004): *Alfurkan Tatarski. Der litauisch-tatarische Koran-Tefsir*. – Köln–Weimar–Wien: Böhlau Verlag [przekł. na br. 2009: *Альфуркан татарскі. Каран-тэфсір татараў Вялікага Княства Літоўскага*. – Mińsk].
- SZYNKIEWICZ Jakób (1935): Literatura religijna Tatarów litewskich i jej pochodzenie. – *Rocznik Tatarski* II, 138–143.
- WALCZAK Bogdan (1993): Komu zawdzięczamy polski język literacki? – [w:] Irena BAJEROWA, Maria KARPLUK, Zenon LESZCZYŃSKI (red.): *Język a chrześcijaństwo*. – Lublin: Tow. Nauk. Katolickiego Uniwersytetu Lubelskiego, 23–42.
- АНТОНОВИЧ Антон (1968): *Белорусские тексты, писанные арабским письмом, и их графико-орфографическая система*. – Вильнюс: Министерство Высшего и Среднего Специального Образования Лит. ССР Вильнюсский Государственный Университет им. В. Капсукаса Кафедра Русского Языка.
- ТАРЭЛКА Міхась (2006): Калафоны Мінскага тэфсіра. – *Здабыткі* 8, 34–43.
- ТАРЭЛКА Міхась (2015): *Рукапісы татараў Беларусі XVIII – пачатку XXI стагоддзя з дзяржаўных і грамадскіх кнігазбораў краіны. Каталог*. – Мінск: Беларуская навука.
- ТАРЭЛКА Міхась, СЫНКОВА Ірына (2009): *Адкуль пайшлі ідалы*. – Мінск: Тэхналогія.

Źródła internetowe:

<http://www.islamicmanuscripts.info/reference/books/Drozdz-2000-Tatarow/Drozdz-2000-Tatarow-048-083.pdf>

doi: <http://dx.doi.org/10.17651/bptj.74.8>

ANGELIKA PAWLACZYK

Uniwersytet Mikołaja Kopernika, Toruń

Znaczenie metody biograficznej w badaniach gwary staroobrzędowców mieszkających w Polsce

Streszczenie

Celem niniejszego artykułu jest analiza przydatności i ukazanie sposobu wykorzystania metody biograficznej w badaniach zjawisk zachodzących w gwarze staroobrzędowców mieszkających w Polsce na podstawie materiałów zebranych podczas ekspedycji dialektologicznych w latach 2014–2017. Udostępnione badaczom prywatne dokumenty i nagrania rodzinne stanowią ważny materiał językowy i socjolingwistyczny. W przekazanym materiale biograficznym odnaleźć można nie tylko osobiste refleksje staroobrzędowców, ale również informacje na temat życia codziennego całej społeczności i/lub poszczególnych osób oraz relacje z określonych zdarzeń. Wspomnienia i relacje osób badanych są niezwykle ważne dla badaczy rosyjskiej gwary staroobrzędowców, ponieważ poszerzają korpus tekstów i nagrań o nowe sfery tematyczne (swobodne wypowiedzi na tematy osobiste, domowe i rodzinne), stylistyczne (artyistyczna funkcja gwary w utworach poetyckich), a także stanowią przykłady spontanicznych wypowiedzi (uroczystości wspólnotowe i rodzinne).

Słowa kluczowe: metoda biograficzna, dokumenty osobiste, staroobrzędowcy, rosyjska gwar

The biographical method in studies on the dialect of Polish Old Believers

Abstract

This article discusses the usefulness of the biographical method and describes the way it is employed in studying the dialect of the Old Believers living in Poland. Private family documents and recordings, which were made available to the researchers and collected during dialectological expeditions in the years 2014–2017, constitute a significant language and sociolinguistic material. In the biographical material, we can find not only personal reflections of the Old Believers, but also information about everyday life of the whole community and/or of specific people, as well as reports from particular events. Memories and reports of the respondents are especially important for the researchers of the Russian dialect of the Old Believers because they enrich the corpus of texts and recordings with new thematic (casual utterances con-

cerning personal life, household, family) and stylistic areas (artistic function of the dialect in poetic works), as well as with examples of spontaneous utterances (community and family celebrations).

Keywords: biographical method, private documents, Old Believers, Russian dialect

Istotny wpływ na rozwój socjologii miały badania współczesnego społeczeństwa w Stanach Zjednoczonych prowadzone przed ponad stu laty przez Williama Isaaca Thomasa. Zgromadzone listy polskich imigrantów po uprzednim opracowaniu miały zostać opublikowane jako materiał o charakterze źródłowym, co odpowiadałoby sformułowanemu później przez badacza postulatowi dotyczącemu „wykorzystania do celów badawczych takich materiałów jak dzienniki, księgi sądowe i kościelne, kazania i programy szkolne” (Leoński 1995: 124). Jednak rozpoczęta współpraca z Florianem Znanieckim nieco zmodyfikowała plany wydania wielkiej publikacji na rzecz pięciotomowego dzieła opatrzonego komentarzami oraz tekstami o charakterze metodologicznym. W ten sposób *Chłop polski w Europie i Ameryce* (Thomas, Znaniecki 1976) dał początek metodzie biograficznej i rozpoczął historię wykorzystania dokumentów osobistych jako nowego typu źródeł w socjologii (Leoński 1995: 124).

Metoda biograficzna w zależności od podejścia danego badacza może być rozumiana w różny sposób. Według Jana Szczepańskiego:

metodą dokumentów biograficznych nazywamy takie prowadzenie badań socjologicznych, w których do rozwiązania postawionego zagadnienia zbiera się tylko materiały zawierające relacje ludzi o ich uczestnictwie w zdarzeniach i procesach stanowiących przedmiot badań, a na podstawie tych relacji dokonuje się opisu danych procesów i stawia hipotezy wyjaśniające (Szczepański 1971: 577–578).

Czesław Gryko twierdzi, że:

Metodę biograficzną należy rozpatrywać z dwu punktów widzenia. Po pierwsze oparta jest ona na określonych założeniach ontologicznych. Na pierwszy plan wysuwa się tu człowiek jak aktywny podmiot kultury. Stanowi to podmiotowy (biograficzny) element kultury. Po drugie jest to pewna dyrektywa badawcza nakazująca uwzględnienie biografii w opisie kultury i uznająca za niedopuszczalną jej depersonalizację. Jest to biograficzny aspekt badań nad kulturą (Gryko 1989: 150–151).

Natomiast Jacek Leoński przyjmuje, że:

metoda biograficzna, metoda autobiograficzna, metoda dokumentów osobistych, metoda pamiętnikarska to określenia, które w socjologii występują zamiennie. Często używa się tych określeń na oznaczenie pewnego postępowania badawczego (Leoński 1993: 25).

Metoda biograficzna wykorzystywana jest przede wszystkim w naukach społecznych. Jednak jak zauważyła Kaja Kaźmierska:

od kilku dekad posługują się nią nie tylko socjologowie czy pedagodzy powołujący się na ugruntowaną w obrębie tych dyscyplin tradycję badawczą [...], ale też pedagodzy, historycy czy – coraz częściej – artyści (Kaźmierska 2013: 6).

Jej zastosowanie można odnaleźć również w językoznawstwie. Bardzo ważne jest, by w procesie badawczym nie pomijać jednak relacji i osobistych przeżyć członków badanej grupy społecznej oraz sposobu ich wypowiedzania się. Pozyskane w ten sposób materiały posłużyć mogą m.in. do zbadania biografii indywidualnej czy biografii całego społeczeństwa i wyodrębnienia dokonujących się we wspólnocie przemian. W ten sposób dla socjologa stanowią one będą istotne źródło analizy socjologicznej, a dla językoznawcy – lingwistycznej i socjolingwistycznej. Pisząc o sposobach ujmowania metody biograficznej, warto przyjąć za Jackiem Leońskim (1993) rozróżnienie na trzy płaszczyzny odniesienia kontekstu metodologicznego, w jakim ona występuje, tj. pewną orientację metodologiczną, przedmiot analizy, gdzie za temat przyjmuje się całą biografię lub jakiś jej fragment oraz sposób gromadzenia danych i ich analizy (Leoński 1993: 25). W przypadku tej pierwszej badacze z reguły odnoszą się do źródeł mających charakter relacji i osobistej refleksji dotyczącej życia jednostki i/lub całego społeczeństwa, które jeśli ujmowane są całościowo, mogą stanowić wartościowy materiał socjologiczny (Leoński 1993: 26). Do tego typu źródeł J. Leoński zalicza „pamiętniki, listy, wspomnienia, relacje, dzienniki, wypowiedzi na określone tematy czy historie mówione lub inaczej wywiady narracyjne” (Leoński 1993: 31–32). Wynika z tego, że w opisywanej metodzie wykorzystywane są przede wszystkim dokumenty osobiste, które w sposób subiektywny (bo z punktu widzenia członków badanej wspólnoty) przedstawiają historie, postawy i dokonujące się przemiany w społeczeństwie (Lutyński 1990: 84). Te z kolei, według Jana Szczepańskiego, mogą być wykorzystywane i opracowywane metodą konstruktywną, egzemplifikacji, analizy treści, statystyczną i typologiczną (Szczepański 1971: 597–599).

Celem niniejszego artykułu jest analiza przydatności i ukazanie sposobu wykorzystania metody biograficznej w badaniach zjawisk zachodzących w rosyjskiej gwarze staroobrzędowców¹ mieszkających w Polsce na podstawie materiałów zebranych podczas ekspedycji dialektologicznych² w latach 2014–2017. Dzięki regularnym wizytom w ośrodkach zamieszkiwanych przez staroobrzędowców możliwe jest dotarcie i opracowanie danych nie tylko z zakresu socjologii, ale także językoznawstwa. Gromadzony od 1999 r. materiał obejmuje przede wszystkim wywiady prowadzone z informatorami w różnym wieku na tematy związane z życiem codziennym, historią i sytuacją językową wspólnoty. Co ważne, mają one charakter swobodnych rozmów i prowadzone są w taki sposób, by informatorzy w opowieściach dzielili się swoimi

¹ Staroobrzędowcy to grupa społeczno-wyznaniowa, której przodkowie nie zaakceptowali reform wprowadzonych w połowie XVII w. przez patriarchę Nikona w Rosyjskiej Cerkwi Prawosławnej (Iwaniec 1977: 22). W wyniku migracji część rosyjskiej społeczności staroobrzędowej pod koniec XVII w. osiedliła się również na terenach Rzeczypospolitej, gdzie obecnie zamieszkuje ośrodek mazurski, suwaliński i augustowski (Grek-Pabisowa 1968: 36–38). Dla staroobrzędowców mieszkających w Polsce charakterystyczne jest posługiwanie się gwarą środkowowielkoruską typu pskowskiego, którą ich przodkowie wynieśli z okolic Pskowa, Nowogrodu i Wielkich Łuk.

² Od 1999 r. badacze z Katedry Filologii Słowiańskiej Uniwersytetu Mikołaja Kopernika w Toruniu uczestniczą w ekspedycjach dialektologicznych do miejsc zamieszkiwanych przez polskich staroobrzędowców, podczas których dokumentują nowy i uzupełniają zgromadzony już materiał językowy.

myślami, emocjami, refleksjami i doświadczeniami. W latach 2014–2017 materiał został uzupełniony o przekazane badaczom utwory poetyckie, wspomnienia i nagrania wideo autorstwa dwóch przedstawicieli badanej mniejszości. Wywiady oraz otrzymane dokumenty mogą być wykorzystywane i analizowane w różny sposób. Jednak ze względu na osobisty charakter części z nich, czyli wspomnianych nagrań audio i wideo oraz utworów poetyckich, zasadne wydaje się wykorzystanie w badaniach jakościowych metody biograficznej. W tym przypadku przedmiotem rozważań jest więc przede wszystkim ewolucja języka dwujęzycznej jednostki i/lub grupy na tle dokonujących się w ich życiu zmian.

Wykorzystanie metody biograficznej w badaniach gwary staroobrzędowców mieszkających w Polsce stało się możliwe dopiero po zakończeniu II wojny światowej. Wynika to przede wszystkim z długiej izolacji, a także doświadczanych podczas wojny represji i deportacji. Rozpoczęte przez polskich językoznawców w latach pięćdziesiątych badania potwierdzają jednak, że mimo początkowej nieufności starowierców stopniowo zanikała w nich obawa przed osobami spoza ich wspólnoty, co pozwoliło na zgromadzenie oraz eksplorację obszernego materiału językowego (por. Grek-Pabisowa 1999: 309). Obecnie obserwuje się, że staroobrzędowcy z ośrodka augustowskiego, szczególnie ci zaliczani do starszego pokolenia, starają się w każdy możliwy sposób propagować starowierską kulturę. Dlatego też chętnie dzielą się oni swoją wiedzą na temat historii, tradycji i obrzędowości również z osobami spoza swojej wspólnoty. W związku z tym przeprowadzone i udokumentowane do tej pory wywiady są pełne osobistych refleksji, opowieści z życia prywatnego oraz relacji z wydarzeń, które miały miejsce w przeszłości. Opowieści informatorów, zarówno te w języku polskim, jak i rosyjskiej gwarze, można uznać za istotne źródło danych socjolingwistycznych. Materiały te, ze względu na swoją treść, mogą zostać wykorzystane np. w badaniach lokalnego folkloru lub pamięci społecznej wspólnoty, ale z punktu widzenia analizy językowej istotna jest przede wszystkim ich forma. Ponieważ są to wypowiedzi w rosyjskim dialekcie i stanowią podstawę do badań stanu gwary i dokonujących się w niej zmian pod wpływem polskiego otoczenia językowego. Gromadzenie i analiza materiału językowego (niezależnie od tematu rozmowy) jest istotne ze względu na to, że gwara staroobrzędowców funkcjonuje praktycznie tylko w formie ustnej i jest to jedyny sposób jej utrwalenia oraz utworzenia korpusu do badań, a ponadto coraz częściej uznawana jest ona (także przez samych staroobrzędowców) za język zagrożony wymarciem. Obecnie coraz mniej jest użytkowników, którzy opanowali rosyjski dialekt na poziomie zbliżonym do znajomości ich podstawowego języka, czyli języka polskiego. Wśród nich można wyróżnić osoby, które posługują się gwarą w zakresie kodu rozwiniętego (wg definicji Basila Bernsteina) nie tylko w mowie, ale również w piśmie, co oznacza, że na tle pozostałych członków społeczności wyróżniają się oni m.in. bogatymi zasobami zachowanej leksyki gwarowej, rozbudowaną frazeologią, złożonością konstrukcji składniowych, różnorodnością stylistyczną (por. Boszkański, Piotrowski, Ziółkowski 1977: 106–131; Głuszkowski 2013: 125–130). Analiza niedawno odkrytych dokumentów osobistych w postaci amatorskich utworów poetyckich i tekstów pieśni dwóch przedstawicieli starszego

pokolenia staroobrzędowców z ośrodka augustowskiego pozwala na stwierdzenie, że mimo funkcjonowania gwary prawie wyłącznie w formie ustnej możliwe jest również wypracowanie własnego systemu zapisu gwary przez jej użytkowników. Dzięki zdobytemu w latach 2014–2017 materiałowi możliwe jest zatem wykorzystanie metody biograficznej nie tylko w procesie gromadzenia materiałów, czyli wywiadów z informatorami, ale także podczas późniejszej analizy utworów poetyckich, nagrań audio i wideo oraz tekstów o charakterze wspomnieniowym.

W czasie prowadzonych wywiadów badacze rosyjskiej gwary muszą przede wszystkim zdobyć zaufanie potencjalnych informatorów i nakłonić ich do dzielenia się opowieściami na tematy związane z ich życiem prywatnym oraz życiem całej wspólnoty. Co ważne, badacze nie ingerują w opowieści i nie przerywają rozmówcom, a kolejne pytania zadawane są w celu rozwinięcia lub przedłużenia wypowiedzi. Rozmowa kierowana jest więc w taki sposób, by informatorzy dzielili się swoją subiektywną oceną i własnym punktem widzenia pewnych wydarzeń, których byli uczestnikami. Najczęściej opowieści te związane są z pamięcią o różnych wydarzeniach historycznych, gdzie niezwykle ważnym źródłem informacji są przedstawiciele starszego pokolenia, których pamięć sięga nawet do czasów I wojny światowej. Podczas jednego wywiadu możliwe jest zatem zarejestrowanie swobodnych wypowiedzi staroobrzędowców na tematy osobiste, domowe i rodzinne, a to z kolei pozwala poszerzyć korpus tekstów i nagrań o materiały możliwie różnorodne pod względem leksykalnym i zawierające także rozbudowane struktury składniowe.

W 2016 r. toruńskim badaczom udało się dodatkowo zdobyć nagrania audio i wideo autorstwa jednej z nieżyjących już starowierek. Materiał ten stanowi niezwykle ważne źródło wiedzy o dialekcie staroobrzędowców, ponieważ udokumentowane zostały spontaniczne wypowiedzi członków społeczności w czasie odbywających się uroczystości wspólnotowych i rodzinnych, a także swoisty wideo-pamiętnik, swoją formą przypominający dzisiejsze internetowe video-blogi. Oprócz możliwości zapoznania się z samym przebiegiem tego typu imprez okolicznościowych prywatne dokumenty pozwalają na sprawdzenie, czy rozmowy prowadzone przez badaczy w czasie spotkań z informatorami różnią się w jakiś sposób od tych, w których nie uczestniczą osoby spoza ich wspólnoty. W związku z tym, że dokumenty powstały spontanicznie i bez udziału badaczy, stanowią możliwie naturalne źródło do badań językowych i dlatego są niezwykle cennym uzupełnieniem korpusu tworzonego do tej pory na bazie wywiadów z informatorami.

Innym ważnym materiałem, którego uzyskanie nie byłoby możliwe w czasie wywiadu, jest twórczość własna staroobrzędowców. Utwory poetyckie, mimo że nie zostały dotąd opublikowane, są znane i popularne wśród mieszkańców wsi, w której urodzili się ich autorzy. Początkowo korpus tekstów był niewielki, ponieważ ze względu na brak wykształcenia i związane z nim obawy przed krytyką staroobrzędowcy nie chcieli dzielić się swoją twórczością z osobami spoza ich wspólnoty. Jednak podczas kolejnych ekspedycji dotarcie do tekstów było już łatwiejsze. Teksty w większości poświęcone są tematom dotyczącym życia codziennego, historii i tradycji staroobrzędowców. Oprócz tego wśród zapisków w zeszytach można odnaleźć także relacje

z przebiegu różnych uroczystości, spisane historie życia poszczególnych mieszkańców wsi czy szczegółowe opisy miejsc i wydarzeń, w których uczestniczyły miejscowe poetki. Tego typu materiał pozwala przede wszystkim na poznanie myśli, emocji, poglądów i wartości wypowiadających się osób i świadczy o artystycznej funkcji gwary. Według Jacka Kubery tworzenie materiałów autobiograficznych czy tekstów literackich ma również duże znaczenie dla ich autorów ze względu na „właściwości terapeutyczne, porządkujące i rozwiązujące problemy życiowe, racjonalizujące osobiste wybory, dostarczające im intelektualnych i moralnych uzasadnień” (Kubera 2015: 51).

Wykorzystanie metody biograficznej w badaniach zjawisk zachodzących w gwarze staroobrzędowców mieszkających w Polsce jest przydatne z wielu powodów. Wywiady pozwalają na zarejestrowanie sposobu wypowiadania się danej dwujęzycznej jednostki, a w późniejszym postępowaniu – zbadanie zachowanych w jej idiolektcie cech gwarowych i wpływów języka polskiego. Dzięki regularnie prowadzonym badaniom oraz gromadzonemu od 1999 r. materiałowi językowemu można także sprawdzić, jak zmieniał się język konkretnej osoby i/lub całej grupy na tle dokonujących się w ciągu lat zmian w ich życiu. Wykorzystanie w badaniach metody biograficznej pozwala zatem nie tylko na zapoznanie się z refleksjami i emocjami informatorów czy dostarczonymi przez nich informacjami na temat życia codziennego społeczności, ale z punktu widzenia językoznawstwa pozyskany nowy rodzaj dokumentów osobistych, czyli m.in. wiersze, można uznać za ważny i unikalny materiał językowy, ponieważ (tak jak nagrania) umożliwiają one zbadanie zjawisk językowych i zachodzących w gwarze zmian oraz poprzez zróżnicowanie zastosowanych źródeł zapewniają większą wiarygodność wyników.

Metoda biograficzna wykorzystywana jest przede wszystkim do dokumentowania rosyjskiej gwary i, co się z tym wiąże, do poszerzania korpusu nagrań i tekstów o nowe sfery tematyczne. Zgromadzone materiały można uznać jednak za wartościowe również z wielu innych powodów. Dzięki zastosowaniu w badaniach metody biograficznej możliwe jest m.in. wyodrębnienie czynników socjolingwistycznych, które w większym lub mniejszym stopniu wpływają na życie staroobrzędowców i obecny stan ich gwary (zob. Гжибовский, Глушковский 2008: 200–214).

Dostęp do tego typu materiałów pozwala także zapoznać się z samodzielnie wypracowanym systemem zapisu rosyjskiej gwary, a to z kolei – na stworzenie przez badaczy charakterystyki języka pisanego staroobrzędowców i udzielenie odpowiedzi na pytanie, w jaki sposób zapisywana jest rosyjska gwara funkcjonująca dotąd tylko w formie ustnej (zob. Павлячик 2017: 269–279). Przeprowadzona analiza twórczości jednej z miejscowych poetek pokazała, że w zdecydowanej większości starowierka zapisuje wyrazy tak, jak je słyszy i jak podpowiada jej intuicja, nie kierując się ortografią rosyjskiego języka literackiego, ponieważ jej znajomość tego języka w mowie i w piśmie jest ograniczona. Najważniejszym kryterium w systemie zapisu jest fonetyczna postać rosyjskiej gwary staroobrzędowców w Polsce. Ponieważ informatorka nie posiada wykształcenia filologicznego, nie stosuje żadnego ze znanych systemów transkrypcji, a jedynie na podstawie intuicji dobiera znaki zarówno z polskiego, jak i rosyjskiego alfabetu, np.:

baćjan (pol. bocian, ros. аист, ros. st. bać'an)
 radzicieli (pol. rodzice, ros. родители, ros. st. raǰ'icël'i)
 oczeń (pol. bardzo, ros. очень, ros. st. 'očëń)
 iszczo (pol. jeszcze, ros. ещё, ros. st. išč'o)
 wsio (pol. wszystko, ros. всё, ros. st. vs'o)
 соньце (pol. słońce, ros. солнце, ros. st. s'ońcè)
 песьни (pol. pieśni, ros. песни, ros. st. p'e'sni)
 ушол (pol. odszedł, ros. ушёл, ros. st. uš'oł)

Utwory poetyckie stanowią wartościowy materiał również ze względu na to, że możliwe jest dzięki nim zbadanie oddziaływania polszczyzny na rosyjską gwarę. Dokumentowanie nagrań i tekstów pisanych pozwala na wyodrębnienie takich zjawisk jak przełączanie kodu i wstawki językowe, zapożyczenia, kalki i hybrydy językowe (przejawy interferencji wg klasyfikacji Doroty Paśko-Konecniak (2011: 54–59)). Oprócz tego możliwe jest również określenie, w jakich sytuacjach dochodzi do przemowienia elementów z języka polskiego oraz w których sferach tematycznych gwara jest najlepiej zachowana, a w których polskie wpływy są bardziej intensywnie (zob. Głuszkowski 2009: 115–131). Obecnie można zaobserwować, że na gwarę staroobrzędowców w coraz większym stopniu oddziałuje język polski. Przykładem może być występujące u staroobrzędowców zjawisko przełączania kodów w formie cytowania. Dokładne przytoczenie polskiego zdania lub jego fragmentu w wypowiedzi gwarowej można odnaleźć:

– zarówno w nagranych wywiadach, np. *Ja jem'u gavar'u pa r'usk'i to 'on gavar'it: ja nie rozumiem*,

– jak i w tekstach pisanych, np. *Это фашисты деревню окружили | Задумали злое дело | Еще темно – на рассвете ломот раздался в каждые двери | Ubierajcie się wszyscy – Будzie z wami jak z Żydami*.

Dla dwóch wspomnianych wcześniej przedstawicieli społeczności charakterystyczne jest posługiwanie się gwarą w zakresie kodu rozwiniętego. Przekazane dokumenty pozwalają na sprawdzenie, czy znajduje to swoje odzwierciedlenie również w powstałych w rosyjskiej gwarze tekstach pisanych. Zgromadzony materiał można zbadać pod kątem zachowanych w ich dialekcie cech fonetycznych, fleksyjnych i składniowych oraz leksyki gwarowej. W zakresie fonetyki w tekstach zauważalne jest odzwierciedlenie m.in. takich zjawisk, jak:

- silne akanie np. *pabieda, pasabiraj, cnacубa мебе, axoma*,
- regularnie występujący niepalatalny wariant spółgłoski č, np. *czem, czto, cwietoczki*,
- zamiana spółgłosek palatalizowanych [t'], [d'], [s'] polskimi średniojęzykowymi [ć], [ź], [ś] w wyniku zmiany korelacji miękkości, np. *budziet, prichadzi, ciebie, raściot*.

Wcześniej źródłem do przeprowadzenia tego typu badania były przede wszystkim wywiady i prowadzone rozmowy. Dzięki zdobytym w ostatnich latach nowym materiałom można rozszerzyć zakres badań i określić stopień, kierunek i dynamikę za-

chodzących pod wpływem polszczyzny zmian nie tylko w mowie, ale również w tekstach gwarowych. Forma tekstów pokazuje, w jaki sposób użytkownicy badanego języka postrzegają występujące w nim dźwięki poprzez dobór znaków służących do ich zapisu. Występujące w badanym materiale słowa *cwietoczki* i *budziet*, nie tylko stanowią potwierdzenie dotychczasowych obserwacji całkowitego wyparcia [č'] przez [č]³ oraz zastąpienia rosyjskiego palatalizowanego [d'] polskim środkowojęzykowym [ʒ], ale również pozwalają stwierdzić, że zmiany te dokonały się także w świadomości fonologicznej danej społeczności językowej. Oprócz tego dostęp do osobistych dokumentów staroobrzędowców umożliwia również zbadanie zależności między ich mową a pismem.

Analiza zgromadzonego materiału pozwala również na wyodrębnienie słów, które charakterystyczne są dla konkretnej jednostki, np. *matylioczki*, *gniezdzieczka*, *недельки*, *раненька*. Choć leksyka danej osoby nieco różni się od języka całej społeczności, to badając gwarę nie rezygnujemy z niej, ponieważ ze względu na zakres oddziaływania informatorki zajmującej wysoką pozycję w hierarchii społecznej (artystka, śpiewaczka, poetka) ważna jest obserwacja ewentualnego przenikania takiego słownictwa do języka innych osób. Przykładem może być tutaj jedna z przedstawicielek starszego pokolenia, która tworząc utwory poetyckie wzbogaca swoje teksty różnymi komentarzami, dopowiedzeniami, wyjaśnieniami znaczenia gwarowej leksyki czy rysunkami, które odzwierciedlają rzeczywistość przedstawioną w wierszach. W jednym z nich odnaleźć można np. fragment tekstu, w którym autorka pisze: [...], *бо демок было многа* (AP: „bo dzieci było dużo”) i chcąc doprecyzować oraz rozwinąć swoją myśl zamieszcza obok wyjaśnienie: *11 демок + Пана u Мама* (AP: „11 dzieci + Tata i Mama”).

Wszystkie wspomniane sposoby wykorzystania materiału potwierdzają, że metoda biograficzna jest niezwykle przydatna w badaniach zjawisk zachodzących w gwarze staroobrzędowców mieszkających w Polsce. Udostępnione badaczom prywatne dokumenty i nagrania rodzinne stanowią unikalny materiał socjolingwistyczny i językowy, ponieważ:

- swobodne wypowiedzi na tematy osobiste, domowe i rodzinne poszerzają korpus tekstów i nagrań o nowe sfery tematyczne,
- wciąż powstające wiersze potwierdzają, że rosyjska gwara staroobrzędowców może funkcjonować także w zakresie stylu artystycznego,
- materiał zarejestrowany podczas różnych uroczystości pozwala zapoznać się z przebiegiem rozmów staroobrzędowców, w których nie uczestniczą osoby spoza ich wspólnoty,
- odnaleźć w nich można osobiste refleksje staroobrzędowców, informacje na temat życia codziennego poszczególnych osób i całej społeczności oraz relacje z określonych zdarzeń.

³ O wahaniach wymowy [t'], [d'], [s'] / [č], [ʒ], [ʃ] oraz fakultatywnym użyciu [č] i [č'] w staro-wierskiej gwarze z lat 50. i 60. XX w. pisała w swojej pracy Iryda Grek-Pabisowa (1968). Obecnie we wszystkich pokoleniach używane jest jedynie [č].

Wszystko to pozwala na szczegółową i rzetelną analizę ewolucji języka dwujęzycznej jednostki i/lub całej grupy na tle dokonujących się w ich życiu zmian. W tym miejscu warto przywołać koncepcję Kwiryny Handke, która w swojej monografii zwraca uwagę na potrzebę prowadzenia podobnych badań. Przedmiotem jej rozważań jest język familijny analizowany na tle polszczyzny ogólnej/literackiej (Handke 1995: 10). Kategoria ta może być wykorzystywana m.in. do badania języka rodziny jako osobnej całości i/lub języka rodziny na tle całej społeczności. Kluczowy jest sposób patrzenia na przedmiot badań. W związku z tym K. Handke proponuje, by język familijny rozumiany był nie tylko jako zjawisko lingwistyczne, ale również jako zjawisko „na pograniczu różnych płaszczyzn aktywności człowieka: językowej, psychicznej, społecznej, prawnej i kulturowej” (Handke 1995: 35), co odnieść można także do badań gwary staroobrzędowców w Polsce i prowadzonych analiz (np. analizy języka jednostki/rodziny na tle gwary całej społeczności).

Ze względu na dokonujący się proces zamierania rosyjskiej gwary gromadzone przez toruńskich badaczy dokumenty, tj. nagrania, teksty pisane i wspomnienia, stanowią materiał unikalny dla językoznawców, ponieważ pozwalają na opis zachodzącego na oczach badaczy procesu śmierci języka, a także dla przedstawicieli nauk społecznych, ponieważ jednocześnie jest to forma rejestracji procesów związanych z zanikiem określonej kultury mniejszościowej. Docieranie do tego typu źródeł oraz ich analiza już teraz wzbudza zainteresowanie przedstawicieli młodszego pokolenia staroobrzędowców, szczególnie tych powracających do korzeni językowych i kulturowych, co dodatkowo przemawia za potrzebą rejestracji gwary.

Bibliografia

- BOSZKAŃSKI Zbigniew, PIOTROWSKI Andrzej, ZIÓŁKOWSKI Marek (1977): *Socjologia języka*. – Warszawa: Wiedza Powszechna.
- GLUSZKOWSKI Michał (2009): Dyglosja w społeczności staroobrzędowców regionu suwalsko-augustowskiego. – *Socjolingwistyka* 22–23, 115–131.
- GLUSZKOWSKI Michał (2013): *Socjologia w badaniach dwujęzyczności*. – Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.
- GREK-PABISOWA Iryda (1968): *Rosyjska gwara starowierców w województwach olsztyńskim i białostockim*. – Wrocław–Warszawa–Kraków: Zakład Narodowy im. Ossolińskich.
- GREK-PABISOWA Iryda (1999): Kierunki i stan badań nad staroobrzędowcami w Polsce. – [w:] Iryda GREK-PABISOWA (red.): *Staroobrzędowcy: szkice z historii, języka, obyczajów*. – Warszawa: Slawistyczny Ośrodek Wydawniczy, 303–315.
- GREK-PABISOWA Iryda, MARYNIAKOWA Irena (1980): *Słownik gwary starowierców mieszkających w Polsce*. – Wrocław–Warszawa–Kraków–Gdańsk: Zakład Narodowy im. Ossolińskich.
- GRYKO Czesław (1989): *Józef Chalasiński. Socjologiczna teoria kultury*. – Lublin: Wydawnictwo Lubelskie.
- HANDKE Kwiryna (1995): *Polski język familijny*. – Warszawa: Slawistyczny Ośrodek Wydawniczy.
- IWANIEC Eugeniusz (1977): *Z dziejów staroobrzędowców na ziemiach polskich XVII–XX w.* – Warszawa: Państwowe Wydawnictwo Naukowe.
- KAŹMIERSKA Kaja (2013): Badania biograficzne w naukach społecznych. – [w:] Kaja KAŹMIERSKA (red.): *Przegląd Socjologii Jakościowej: Metoda biograficzna w naukach społecznych*. – Łódź, 6–10.

- KUBERA Jacek (2015): Powieść autobiograficzna jako dokument osobisty i podobny do pamiętników konkursowych materiał badań socjologicznych. – *Studia Humanistyczne AGH* 14, 45–61.
- LEOŃSKI Jacek (1993): Różne sposoby ujmowania metody biograficznej w socjologii. – [w:] Teresa RZĘPA, Jacek LEOŃSKI (red.): *O biografii i metodzie biograficznej*. – Poznań: Wydawnictwo Nakom, 25–32.
- LEOŃSKI Jacek (1995): Historia wykorzystania dokumentów osobistych w socjologii. – *Ruch Prawniczy, Ekonomiczny i Socjologiczny* LVII, 123–128.
- LUTYŃSKI Jan (1990): Metoda biograficzna i sposoby otrzymywania materiałów w badaniach socjologicznych. – [w:] Jan WŁODAREK, Marek ZIÓLKOWSKI (red.): *Metoda biograficzna w socjologii*. – Warszawa–Poznań: Państwowe Wydawnictwo Naukowe, 83–90.
- PAŚKO-KONECZNAK Dorota (2011): *Wpływ polszczyzny na zasób leksykalny gwary staroobrzędowców na Suwalszczyźnie*. – Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.
- SZCZEPAŃSKI Jan (1971): *Odmiany czasu teraźniejszego*. – Warszawa: Książka i Wiedza.
- THOMAS William, ZNANIECKI Florian (1976): *Chłop polski w Europie i Ameryce*. – Warszawa: Ludowa Spółdzielnia Wydawnicza.
- ГЖИБОВСКИЙ Стефан, ГЛУШКОВСКИЙ Михал (2008): Социолингвистическая ситуация старообрядцев в деревнях Габове Гронды и Бур. – [w:] Леонид КАСАТКИН (red.): *Язык, культура, история. Сборник статей к XIV Международному съезду славистов*. – Москва: Языки славянских культур, 200–214.
- ПАВЛЯЧИК Анджелика (2017): Как пишут старoverы (на примере поэтических произведений старoverки старшего поколения августовского региона). – [w:] Katarzyna DEMBSKA, Dorota PAŚKO-KONECZNAK (red.): *Współczesne badania nad językiem rosyjskim i jego odmianami*. – Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, 269–279.

doi: <http://dx.doi.org/10.17651/bptj.74.9>

MACIEJ RAK

Uniwersytet Jagielloński, Kraków

Przydatność testamentów chłopskich w badaniach językoznawczych

Streszczenie

W artykule omówiono testamenty polskich chłopów (Kazimierz Dobrowolski, *Włościańskie rozporządzenie ostatniej woli na Podhalu w XVII i XVIII w.*, Kraków 1933; Janusz Łosowski, *Testamenty chłopów polskich od drugiej połowy XVI do XVIII wieku*, Lublin 2015) pod kątem ich przydatności w językoznawstwie. Wykazano, że teksty te mogą być wykorzystywane w badaniach antroponimii i mikrotoponimii, a także zmienności leksyki gwarowej i regionalnej. W pewnym zakresie omawiane dokumenty mogą być również przedmiotem badań etnolingwistycznych, gdyż ujawniają ludowe wierzenia, przesady i zwyczaje.

Słowa kluczowe: testamenty chłopów polskich, dialektologia, etnolingwistyka, historia leksyki gwarowej

Usefulness of peasants' testaments in ethnolinguistic research

Abstract

The article discusses the usefulness of testaments (published by Kazimierz Dobrowolski, *Włościańskie rozporządzenie ostatniej woli na Podhalu w XVII i XVIII w.*, Kraków 1933 and Janusz Łosowski, *Testamenty chłopów polskich od drugiej połowy XVI do XVIII wieku*, Lublin 2015) in linguistic research. It is argued that last wills can be used in studies on anthroponymy and microtoponymy, as well as on the variability in dialectal and regional lexis. To some extent, the discussed documents can also be the subject of ethnolinguistic research, as they reveal folk beliefs, superstitions and customs.

Keywords: Polish peasants' testaments, dialectology, ethnolinguistics, history of dialectal lexis

Uwagi wstępne

Testamenty – zwłaszcza w ostatnim trzydziestoleciu – stały się popularnym przedmiotem prac edytorskich. Najczęściej wydawano te dokumenty, których w archiwach jest najwięcej, czyli szlacheckie akty ostatniej woli (np.: Borkowska 1984; Klint 2008, 2011; Kowalkowski, Nowosad 2013). Mniej uwagi poświęcono rozporządzeniom mieszczan (np.: Dymmel 1997; Justyniarska-Chojak 2010; Danowska 2011; Lubczyński, Pielas, Suchojad 2005), a najmniej duchownych (pojedyncze teksty znajdziemy w edycji Jacka Kowalkowskiego i Wiesława Nowosada (2013)) i chłopów. W konsekwencji mamy tylko dwa obszerniejsze opracowania źródłowe chłopskich testamentów (pomijam edycje pojedynczych dokumentów): *Włościańskie rozporządzenie ostatniej woli na Podhalu w XVII i XVIII w.* Kazimierza Dobrowolskiego (dalej: Dobr) oraz (wydane przed trzema laty) *Testamenty chłopów polskich od drugiej połowy XVI do XVIII wieku* w opracowaniu Janusza Łosowskiego (dalej: Łos). Pierwszy zbiór obejmuje 120 dokumentów podhalańskich pochodzących z Archiwum Sądu Powiatowego w Nowym Targu, w drugim mamy 326 testamentów z Korony przechowywanych w 17 instytucjach w Polsce i za granicą (m.in. archiwa we Lwowie i Mińsku). W jednym i drugim przypadku testującymi były osoby lepiej sytuowane, bowiem najbiedniejsi chłopci po prostu nie mieli co przekazywać w spadku.

Jeśli idzie o ścisłość z zakresu terminologii prawniczej, mamy tu właściwie do czynienia nie tylko z testamentami, lecz także (i to częściej) z kodycyłami. Różnice między tymi dokumentami zostały określone w prawodawstwie rzymskim. Testament wskazywał dziedzica lub dziedziców poprzez wyznaczenie im części z całego majątku przeważnie wyrażonej ułamkami, a kodycył był wykazem rzeczy przypadających w spadku poszczególnym osobom (Dobr: 10–11).

Testamenty były już przedmiotem badań kulturoznawczych (np. Popiołek 2009; Dunin-Dudkowska 2013, 2014), językoznawczych (np. Zajda 2016) i tekstologicznych (np. Żmigrodzka 1997; Kość 2001, 2010, także wymienione wyżej prace Dunin-Dudkowskiej 2013, 2014). Skupiano się jednak przede wszystkim na dokumentach szlacheckich i mieszczańskich. Jeśli idzie o testamenty chłopskie, poświęcono im zaledwie 3 artykuły (Rak 2017, 2018a, 2018b), w których na materiale Dobr zostały omówione ludowa terminologia prawnicza (Rak 2018b), nazwy stopni pokrewieństwa i powinowactwa oraz nazwy płaciwa (Rak 2018a), a także przejawy życia religijnego, wierzenia i zwyczaje góralskie (Rak 2017).

Wiarygodność źródeł

Analizę dokumentów umieszczonych w Dobr i Łos należy poprzedzić uwagami dotyczącymi wiarygodności tych źródeł. Składają się na to trzy aspekty.

1. Przyjmuje się, że dopiero pod koniec XIX w. z inicjatywy Lucjana Malinowskiego rozpoczęto systematyczne zapisywanie ludowych wypowiedzi. Dzięki testamentom chłopskim można przesunąć tę datację o trzy wieki wstecz. Oczywiście nie

mamy w tych tekstach do czynienia ze stylem potocznym, który przede wszystkim interesuje dialektologów, gdyż ustalona konwencja rozporządzenia ostatniej woli wprowadza wyraźne ograniczenia, niewątpliwie jednak są to dokumenty, w których przemawiają chłopi. Należy nadmienić, że – podobnie jak w przypadku zabytków językowych staropolskich, na podstawie których można jedynie z przybliżeniem (bardziej czy też mniej dokładnym, to już kwestia interpretacji) orzekać o wycinkach dawnej polszczyzny – także testamenty chłopskie są zabytkami charakteryzującymi się wybiórczością. Dotyczy to zarówno płaszczyzny językowej (celowe unikanie fonetycznych cech gwarowych, w tym przede wszystkim mazureńskich i w drugiej kolejności samogłosek pochylonych), jak i kulturowej (pomijanie ludowych wierzeń dotyczących śmierci i pogrzebu, które – jak można wnioskować na podstawie prac etnologicznych (np. Fischer 1921) – były liczne). Za tę wybiórczość odpowiada, po pierwsze, konwencja testamentu, bowiem dokument ten miał określoną budowę i składniki, a także zawierał pewne ustalone formuły. Po drugie, pisarz (testatorzy najczęściej byli niepiśmienni, a nawet jeśli umieli pisać, woleli, żeby testament sporządził pisarz-urzędnik, dzięki czemu dokument miał większą moc prawną, często też znajdowali się w takim stanie zdrowia, który uniemożliwiał samodzielne pisanie) nakładał na wypowiedź filtr językowy i kulturowy. Po trzecie, testator, znajdując się w specyficznej sytuacji komunikacyjnej, która zyskiwała charakter oficjalności, choćby dlatego, że dokonywała się w obecności pisarza (osoby obcej) i dotyczyła bardzo poważnych, ostatecznych spraw, sam unikał gwary.

2. Testamenty jako dokumenty rękopiśmienne, nierzadko uszkodzone i nieczytelne, przysparzają wydawcom pewnych problemów, zwłaszcza w zakresie słownictwa gwarowego. Jeśli zdają sobie oni sprawę z możliwych błędnych odczytań, wówczas edycja ma większą wartość dla korzystającego z niej filologa. Na przykład zdanie:

Mam zagrodę swoją własną, pańszczyzna konna do lezki dwudniowa, tę legował na Katarzynę córkę swoją własną... (Łos: 315)

zostało opatrzone komentarzem „fragment niezrozumiały”. Jest to sygnał, by nie traktować formy *lezki* jako pewnej. W innym miejscu jako „przedmiot niezidentyfikowany” zostały opisane *litry* – *litry do wozu jedne* (Łos: 410). Ujawnia się w ten sposób nieznamość gwar, gdyż chodzi tu po prostu o latry ‘boczne deski kładzione wzdłuż wozu’, gwarowe: *lātry*, *lytry*, *letry* i *litry*.

Inaczej sprawa wygląda w przypadku zupełnie błędnego odczytania i interpretacji. Tak się stało z zestawieniem *niepile ręce*, pojawiającym się w testamentach i księgach sądowych z Małopolski (por. Zajda 2001b: 120). Ponieważ pierwszy jego człon jest oparty na wyrazie *pili* (tu w postaci zaprzeczonej *niepili*) ‘swój, z rodziny’ – znanym obecnie z gwar południowej, górskiej Małopolski (Kąs 2003: 485; Hodorowicz 2013: 252), dla niedialektologa przez swoją dyferencyjność jest niezrozumiały, ale graficznie podobny do imiesłowu przymiotnikowego biernego *pity* (*niepity*). Tak też więc został odczytany, a ponieważ zestawienie *niepite ręce* jest dość osobliwe, aby je uprawdopodobnić, dodano komentarz ‘sprzedać szybko bez litkupu, czyli umocnienia transakcji wypiciem alkoholu przez kontrahentów’ (Łos: 136, 260). Zarówno od-

czytanie, jak i komentarz są całkowicie błędne, jednak jako umieszczone w naukowej edycji mogą ze szkodą np. dla badaczy ludowego prawodawstwa upowszechnić się w pracach naukowych.

Czasem sytuacja z odczytaniem jest bardziej zawiła, co obrazuje przykład tym razem spoza testamentów. Kazimierz Nitsch, krytykując *Słownik gwar polskich* Jana Karłowicza (Karłowicz 1900–1911), podważył umieszczone tam hasło *kirkak* ‘na plecach’ z przykładem użycia *na kirkak nieść* ‘na plecach nieść’ (t. II: 356), pochodzące ze *Spisu wyrazów podhalskich* Augustyna Wrześniowskiego (1885), wskazując na źródło tego (jak sądził) nieporozumienia *Dziennik podróży do Tatrów* Seweryna Goszczyńskiego (1853). Nitsch wspólnie z Juliuszem Zborowskim (1952) objaśnili to błędnym zinterpretowaniem zasłyszanego *nieść na kirkak* ‘nieść trumnę na cmentarz żydowski’ (z niem. *Kirchhoff*) jako ogólnego ‘nieść cokolwiek’. Do dyskusji 21 lat po komunikacie Nitscha i Zborowskiego włączyła się Zofia Kurzowa (1973), która udowodniła, że rację miał jednak Goszczyński, gdyż omawiany wyraz w gwarze Skalnego Podhala jest realizacją słowackiego *krk* ‘kark’.

3. *Testamenty chłopów polskich od drugiej połowy XVI do XVIII wieku* zostały wydane w zmodernizowanej pisowni, która – jak czytamy we wstępie (Łos: LIII) – miała m.in. na celu „usunięcie błędów składniowych i fleksyjnych”, czyli w praktyce interesujących dialektologa cech gwarowych. Zachowały się tu jedynie nieliczne właściwości, jak zleksykalizowane w imionach przykłady prejotacji, np.: *Jagnieszka*, *Jewa* (Łos: 78).

Inaczej sprawa wygląda w przypadku *Włościańskiego rozporządzenia...*, w którym ingerencji było znacznie mniej. Jak zaznaczył K. Dobrowolski (Dobr: 79), „zachowano wszystkie właściwości językowe”, w tym np. mazurzenie, por.: *wnucka*, *wieczny*, *przysły* (Dobr: 198), *dziewcę* (Dobr: 104), *lemies* (Dobr: 107), *ceski* (Dobr: 125), *wnucek* (Dobr: 126), *cego* (Dobr: 170), i znacznie od niego rzadsze oraz bardziej ograniczone geograficznie formy Msc. lp. rz. męskich miękko tematowych, w których po zmazurzonych *š*, *ž* przez analogię do rz. twar dotematowych weszła końcówka - *’e*, np.: *po męzie* (Dobr: 104, 169), *przy Łukasie* (Dobr: 185), *na Odrowązie* (Dobr: 117), *na Ostrysie* (Dobr: 191). Należy jednak stwierdzić, że rozporządzenia ostatniej woli nie mogą być wiarygodnym źródłem do badań fonetyki gwarowej i jej ewolucji.

Wykorzystanie testamentów w badaniach językoznawczych

Podczas analizy językoznawczej testamentów chłopskich najpoważniejszym – jak sądzę – problemem jest oddzielenie słownictwa ogólnopolskiego XVI–XVIII w. od ówczesnej leksyki gwarowej i regionalnej. Wynika to nie tylko z niedostatku w znajomości stanu gwar sprzed kilkuset lat, ale też ze specyfiki rozwoju leksyki. Jak wiadomo, gwary przechowują dawniejsze wyrazy, dlatego charakterystyka testamentów zebranych w Dobr i Łos w kontekście słowników współczesnej polszczyzny może doprowadzić do błędnego wniosku, że jeśli dany leksem dziś jest gwarowy, taki status miał też wcześniej. Oczywiście dzięki *Słownikowi polszczyzny XVI wieku* (Mayenowa, Pełowski 1966–) i *Słownikowi polszczyzny XVII i I. połowy XVIII wieku* (Sie-

kierska 1999–) możemy uniknąć większości błędnych interpretacji, ale nie wszystkich, gdyż dzieła te nie zostały do tej pory ukończone.

Czułości wymaga również interpretacja takich zjawisk, jak obniżenie lub podwyższenie artykulacji samogłosek przed spółgłoskami sonornymi (por. *kupiela* (Łos: 25), *wykupieli* (Łos: 74), *kupieli* (Dobr: 134, 170), *robieli* (Dobr: 180), *czynieli* (Dobr: 110), *sprawieli* (Dobr: 117), *pieli* (Dobr: 198), *ożyniony*, *cyl* (Dobr: 179)) oraz zamienne używanie przyimków *bez* i *przez* (Dobr: 151). Podane przykłady dziś są charakterystyczne dla gwar, ale w XVI–XVIII w. należały też do ówczesnej polszczyzny ogólnej. Z kolei z zakresu fleksji pojawiająca się w Dobr końcówka *-e* w D. lp. rz. żeńskich miękкотematowych, np.: *z Olce* (Dobr: 111), *od siostrzenice* (Dobr: 140), może być traktowana zarówno jako cecha gwarowa, jak i dziedzictwo staropolskie. Zastąpienie *-e* < *-ě₃ w omawianym przypadku przez *-y* jest datowane na XVII w. (Klęmśiewicz, Lehr-Splawiński, Urbańczyk 1965: 288–289). Wymienione dwa przykłady pochodzą z dokumentów datowanych na początek 2. połowy XVIII w. Rozstrzygnięcie więc nie jest całkowicie pewne.

Kwestią jeszcze trudniejszą niż oddzielenie leksyki gwarowej od ogólnopolskiej jest rozgraniczenie regionalizmów i dialektyzmów. W odniesieniu do dawniejszych epok jest to obarczone tak dużą możliwością błędu, że lepiej za Małgorzatą Witaszek-Samborską (1985: 92) przyjąć szerokie rozumienie regionalizmu, w obrębie którego mieszczą się: 1) regionalizmy w tradycyjnym ujęciu (czyli właściwości językowe dopuszczalne w mowie potocznej warstw wykształconych danego obszaru), 2) dialektyzmy (cechy gwarowe, z punktu widzenia normy niedopuszczalne w mowie warstw wykształconych, będące przedmiotem badań dialektologii) i 3) zapożyczenia o ograniczonym zasięgu terytorialnym.

Mimo ograniczeń, które wymieniłem powyżej, testamenty mogą być wykorzystywane w dialektologii historycznej w celu pokazania zmian, jakie zaszły w słownictwie gwarowym w ciągu ostatnich kilkuset lat (oczywiście, nie chodzi tu o zanik wyrazów spowodowany zanikiem desygnatów). Przyjrzyjmy się więc wycinkowi pola leksykalnego nazw stopni pokrewieństwa i powinowactwa.

Zasięg germanizmu *kuzyn* i derywatów (*kuzynka*, *kuzynostwo*) pokrywa się z dawnymi granicami zaborów austriackiego i pruskiego (Szymczak 1966: 121), z kolei w dawnym zaborze rosyjskim mamy zestawienia *brat cioteczny*, *siostra cioteczna* i rzadziej *brat stryjeczny* i *siostra stryjeczna*. W gwarze podhalańskiej notujemy *kuzyna* (Hodorowicz 2013: 187) i *kuzyniaka* (Hodorowicz 2013: 187), czyli znajdujemy w ten sposób potwierdzenie zasięgu geograficznego omawianego leksemu. W Dobr (zbiore, który – jak już to zostało zaznaczone na wstępie – obejmuje jedynie podhalańskie dokumenty) pojawiają się jednak tylko zestawienia *brat cioteczny* (Dobr: 170) i *brat stryjeczny* (Dobr: 96, 98, 141, 183, 204). Testamenty są więc świadectwem tego, że podhalańska terminologia rodzinna uległa zmianie. Przy okazji można postawić pytanie, czy w tym przypadku pod wpływem polszczyzny galicyjskiej czy też niemieckiej. Odpowiedź, jak sądzę, nie jest dyskusyjna – pod wpływem pierwszego z kodów.

Jak się okazuje, testamenty chłopskie przydają się także przy ustalaniu lub weryfikowaniu geografii słownictwa gwarowego. I tak, w testamencie Agnieszki Łasociny

z Krowodrzy (dziś dzielnica Krakowa), który jest datowany na 25 I 1766 r., pojawia się wyraz *maselnik* ‘naczynie na masło’ (Łos: 293). Kartoteka *Słownika gwar polskich* PAN wskazuje, że ma on zasięg kaszubsko-słowiński: notuje go Bernard Sychta w *Słowniku gwar kaszubskich na tle kultury ludowej* (Sychta 1969: 55), a także Friedrich Lorentz w *Pomoranisches Wörterbuch* (1958 I: 491) i *Slovinisches Wörterbuch* (1908 I: 607). Testament Łasociny rozszerza jednak jego geografii znacznie poza Kaszuby i tereny przyległe.

Chłopskie rozporządzenia ostatniej woli dostarczają też informacji na temat wyrazów, form i znaczeń dotychczas niezanotowanych w słownikach. Przykładami „nowych” słów są:

– *ruchomizna* ‘mienie ruchome przekazywane w spadku’ (Dobr: 84, 109, 154, 129; Łos: 219, 229); wyraz poza testamentami spotykany w księgach sądowych, wcześniej omówił go Aleksander Zajda (2001a: 422–423);

– *szrajtarz* ‘bliżej nieokreślone narzędzie kowalskie’ (Łos: 296);

– *szparacz* ‘bliżej nieokreślone narzędzie kowalskie’ (Łos: 296);

– *koślina* ‘skoszone zboże’ (Łos: 43);

– *mlito* ‘zapłata od mlewa’ (Łos: 72), w słownikach znajdziemy *mliwo* / *mlewo*.

W grupie nieznanych słownikom form można wymienić następujące:

– *chomuc* (Łos: 374) odpowiednik ogpol. i oggw. *chomąta*;

– *madzelonka* (Łos: 266). SW (II: 906) notuje hasło *mazalanka* i formy *maza-lónka*, *mazelanka*, *mazelonka* i *melskonka*, ale pośród nich nie ma zarejestrowanej w Łos.

Z kolei jeśli idzie o dodatkowe, nieuchwycone w słownikach znaczenia, warto zwrócić uwagę na:

– *czworaki* ‘dwie pary bron’ (Łos: 310);

– *podwale* ‘wzmocnienie stołu’: „stół jeden wielki, drugi mniejszy laworowny z podwalem” (Łos: 290);

– *roczynę* ‘rodzaj podatku płaconego raz do roku’ (Łos: 19, 22 302).

Osobnym zagadnieniem jest wykorzystanie testamentów w badaniach onomastycznych, bowiem teksty te są nasycone antroponimami i mikrotoponimami, por.:

1. Stało się przy ludziach wiarygodnych i uczciwych, osobliwie przy **Stanisławie Rafaczu**, wójcie międzyczerwieńskim, przy **Janie Sabale**, przy **Janie Skubisu**, przy **Zaryckim Wojciechu**, przy **Gąsienicu Wawrzyńcu**, przy **Janie Stopku**, przy tak wielu ludziach wiarygodnych (Dobr: 99).

2. **Wierzch Nowego Bystrego** od **wierzchu Butorowa** swoje granice, na szerz ciągnie się do **wierzchu pod Ślodyczkę** i tam się na źródle, które jest tamże, pod zwyz położonym kończy, na dłuż zaś między potokiem **Bystrem** idą z obu stron wierzchy i swoje miedze albo granice prowadzą do granice wsi Ratułowa, dopieroż się zowie *ex antiquo* **Wierzch Nowego Bystrego**, pod którym potok **Bystry**, tak nazwany, swoje wejścia ma (Dobr: 83–84).

Do ciekawszych zwyczajów poświadczonych w rozporządzeniach ostatniej woli należało nadawanie takiego samego imienia rodzeństwu¹, por.:

¹ Nie było to typowe tylko dla chłopstwa, gdyż – jak utrzymuje tradycja – np. Jan Długosz nosił takie samo imię jak jego jedenastu braci rodzonych i przyrodnych.

1. **Najstarszy Tomasz, drugiemu także Tomasz**, trzeciemu Jan; córek dwie, Anna i Rejna, a teraz jest brzemienną (Dobr: 60).

2. Którą małżonkę [Annę – MR] z dziećmi zostawiam w roli w półtorej ośminie, w polanie czwartej części, Staselówka zwanej, z którego pola i polany obliguje synom, aby pomagali matce swojej córki wyposażyć według możliwości swojej, to jest **Reginę**, Zofję, Annę, Marjanę, **Reginę** i Agnieszkę (Dobr: 137).

3. A że z pierwszego małżeństwa dzieci pozostałe po śp. Magdalenie żonie, **Walenty, syn najstarszy** [...] aby kłótni ani prawa żadnego z dziećmi drugiego małżeństwa robić nie ważył się, to jest **Walentemu**, Janowi, Józefowi, Jakóbowi synom, córkom **Annie**, Katarzynie, Rozalji, Marjannie. Pozostałym córkom z pierwszego małżeństwa **Annie**, Reginie i Zofji, w służbie będącym, po talaru jednym naznaczam (Dobr: 193).

Zwrócił na to uwagę Sebastian Flizak w artykule *Moda w nadawaniu imion chrzestnych u górali pod Gorcami* (1937). Tytułową *modę* objaśnił popularnością pewnych imion, wpływem rodziców chrzestnych (którzy mogli wymagać, żeby dziecko nosiło takie imię jak ojciec lub matka chrzestna) lub tradycją parafialną (dzieciom nadawano imiona patronów parafii).

Według Flizaka (1937: 56) po północnej stronie Gorców nie nadawano synowi (córcie) imienia ojca (matki). Obawiano się, że sprowadzi to śmierć na rodzica, gdyż dwie osoby o takich samych personaliach nie mogą żyć pod jednym dachem. Na Podhalu ten przesąd nie był znany, na co wskazują testamenty – w kilkunastu dokumentach (np. Dobr: 92–93, 95, 99–100, 113) znajdziemy informację o tym, że syn pierwotny nosił imię ojca.

W społeczności wiejskiej ze względu na jej rolniczy charakter szczególnie ceniło krowę i konia. Nie dziwi więc, że w rozporządzeniach ostatniej woli wielokrotnie jest mowa o tych zwierzętach. Jednak tylko krowom nadawano imiona (są zapisane wielką literą), co może świadczyć o pewnej więzi emocjonalnej, potwierdzonej też przysłowiem *Krowiczka to nasza mamiczka, a koniczek to taticzek*.

Testamenty zebrane przez Łosowskiego dostarczają bogatego materiału zoonimicznego z różnych części Polski, który – jeśli idzie o aspekt motywacyjny – eksponuje:

1) umaszczenie: *Bielawa* (biała), *Płowucha* (płowa), *Kropicha* (w kropki), *Kwiatula* (w łaty), *Rościcha* 'krowa o maści czerwonej lub brązowej z białymi plamkami', *Sikora* 'krowa o umaszczeniu żółtawym, oliwkowym, tak jak ptak – sikorka', *Cyrana* 'krowa prawdopodobnie o umaszczeniu brązowym, tak jak ptak – cyranka';

2) inne cechy wyglądu zwierzęcia: *Głowucha* 'krowa o dużym łbie', *Rozena* 'krowa z dużymi rogami'. Jak wiadomo, w polszczyźnie nazw części ciała są zróżnicowane na ludzkie i zwierzęce, np. *noga* – *łapa*, *głowa* – *łeb*, *plecy* – *grzbiet*, co jako pierwszy opisał Zdzisław Kempf (1985). Leksem *głowucha* stanowi pozorne zaprzeczenie tej reguły. Pozorne, gdyż wskazuje na wyjątkowe wartościowanie krowy, w wyniku czego w gwarach w odniesieniu do tego zwierzęcia często używano leksyki zwyczajowo zarezerwowanej do opisu ludzi (por. Krawczyk-Tyrpa 2001: 110).

3) dzień, w którym zwierzę się urodziło: *Pędzicha* 'krowa, która urodziła się w poniedziałek'.

Z dokumentów zebranych w Dobr i Łos mogą skorzystać także badacze zależności języka i kultury. Jeśli idzie o wierzenia przedchrześcijańskie, w testamencie z 3 XI 1775 r. Jakuba Magiery z Bańskiej znajdujemy informację o tym, że budowanie domu na pogorzelisku przynosi nieszczęście (Dobr: 149). Uważano bowiem, że pożar (zwłaszcza wywołany piorunem) nawiedza te zabudowania, w których przebywa diabeł (por. Fischer 1927: 8).

Kolejne wierzenie dotyczy wiatru, który jest sprawcą chorób i nieszczęść: „Przy tym uprasza W. M. Pana, ażeby ani zły wiatr nie zawinął na jego małżonkę i na jego dziatki, jako na niego samego nie zawinął” (Dobr: 149). Powszechnie w Europie uważano, że wiatr to dusza opuszczająca ciało (Moszyński 1967: 480). Barbara Bazińska (1967: 88) zanotowała na Olczy (dziś dzielnica Zakopanego) przesąd mówiący o tym, że wiatr to dusza człowieka, która zbiegła z piekła. Według innych wierzeń silny wiatr był powodowany śmiercią samobójczą zwłaszcza przez powieszenie (stąd porównanie *wieje jakby się ktoś powiesił*) bądź śmiercią czarownicy lub czarownika, czyli ludzi bliskich diabłu.

W większości testamentów znajdujemy ujęte w kanwę katolicką bezpośrednie odwołania do pradawnego kultu Matki Ziemi, która daje życie i przysparza po śmierci:

1. [...] oddaję tedy Panu Bogu duszę, a ciało ziemi [...] (Dobr: 81).
2. Najprzód ciało oddaję ziemi, skąd poszło [...] (Dobr: 82).
3. Ciało, jako z ziemi wyszło, ziemi przy kościele czarnodunajekim podług obrządku chrześcijańskiego proszę oddać i przystojnie pogrzebić [...] (Dobr: 83).
4. Ciało zaś mizerne jako jest z ziemie i z gliny, oddaję ziemi (Dobr: 109, 124).
5. Ciało ziemi, skąd swój początek wzięło (Dobr: 135, 136, 139).
6. Ciało zaś, ręką boską z ziemi uformowane, na pożarcie robactwu ziemi oddaję [...] (Dobr: 122, 131, 138).

W Dobr i Łos leksem *ziemia* pojawia się w połączeniu z *matką* i zaimkiem *nasza*, otrzymujemy więc częste we wstępnej części testamentów zestawienie *nasza matka ziemia*:

1. Najprzód duszę swoją w ręce Pana Boga polecam, ciało zaś ziemi matce naszej, z której swój początek wzięło, oddaję (Dobr: 90).
2. Najprzód duszę Panu Bogu w ręce polecam, ciało ziemi, matce naszej oddaję (Dobr: 116).
3. Najprzód dusze moję grzeszną Bogu Wszechmogącemu, w Trójcy świętej Jedynemu oddaję, a ciało swoje ziemi, jako matce swojej, skąd początek wzięło (Dobr: 121).
4. Ciało ziemi matce naszej oddaję (Dobr: 127, 128 x 2, 137, 139, 140).

Niemal obowiązkowy w rozporządzeniach ostatniej woli nakaz pochowania ciała na cmentarzu wydaje się – z naszej perspektyw – niedorzecznością, przecież jest to miejsce przeznaczone na pochówek. Nakaz ten staje się jednak zrozumiały, gdy weźmiemy pod uwagę to, że na cmentarzu chowano tylko osoby, które na to zasłużyły, takie, które pojednane z Bogiem zmarły śmiercią naturalną. Jak wiadomo, samobójców chowano poza cmentarzem.

W kilku dokumentach (np. Szymona Sprocha z Lasku z 10 I 1750 r. (Dobr: 115); Andrzeja Chodura z Czarnej koło Łańcuta z 17 IX 1739 (Łos: 203–204)) jest mowa

o bożym obiedzie, czyli rodzaju stypy dla ubogich, którzy w zamian za pożywienie mieli się modlić za duszę zmarłego (Moszyński 1967: 260–261). Jak się okazuje, dziś na Podhalu i szerzej w południowej Małopolsce boży obiad nie jest znany, testamenty są więc historycznym świadectwem tego zwyczaju.

Podsumowanie

Chłopskie rozporządzenia ostatniej woli należą do najstarszych i prawdopodobnie najwiarygodniejszych świadectw słownictwa gwarowego i regionalnego oraz mentalności mieszkańców wsi. Dokumenty te są przede wszystkim przydatne do opisu zmienności leksyki, w ten sposób pozwalają jeszcze dokładniej poznać zasób wyrazowy polszczyzny w jej różnych odmianach. Stanowią też istotne źródło do prac onomastycznych, zwłaszcza z antroponimii i toponimii. Na ich podstawie możemy w pewnym zakresie prowadzić również badania etnolingwistyczne. Przydatność testamentów chłopskich jest więc bezdyskusyjna, tym bardziej dziwi, że do tej pory teksty te nie było przedmiotem większych opracowań językoznawczych.

Bibliografia

- BAZIŃSKA Barbara (1967): Wierzenia i praktyki magiczne pasterzy w Tatrach Polskich. – [w:] Włodzimierz ANTONIEWICZ (oprac.): *Pasterstwo Tatr Polskich i Podhala*, t. 7. – Wrocław: Zakład Narodowy im. Ossolińskich, 65–229.
- BORKOWSKA Małgorzata (1984): *Dekret w niebieskim ferowany parlamencie. Wybór testamentów XVII–XVIII wieku*. – Kraków: Znak.
- DANOWSKA Ewa, oprac. (2011): *Dług śmiertelności wypłacić potrzeba. Wybór testamentów mieszczan krakowskich XVII–XVIII wieku*. – Kraków: Wydawnictwo PAU.
- DOBR – Kazimierz DOBROWOLSKI (1933): *Włościańskie rozporządzenie ostatniej woli na Podhalu w XVII i XVIII wieku. Studia i materiały*. – Kraków: Nakładem Polskiej Akademii Umiejętności.
- DUNIN-DUDKOWSKA Anna (2013): Gatunki testamentowe jako zwierciadło kultur. Rozważania polsko-amerykańskie. – *Acta Universitatis Lodzensis. Kształcenie polonistyczne cudzoziemców* 20, 131–140.
- DUNIN-DUDKOWSKA Anna (2014): *Testament jako zwierciadło kultur. Polsko-amerykańskie studium komparatystyczne*. – Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.
- DYMMEL Piotr, oprac. (1997): *Testamenty mieszczan wojnickich 1599–1809*. – Wojnicz: Towarzystwo Przyjaciół Ziemi Wojnickiej.
- FISCHER Adam (1921): *Zwyczajne pogrzebowe ludu polskiego*. – Lwów: Wydawnictwo Biblioteki Zakładu Nar. im. Ossolińskich.
- FISCHER Adam (1927): *Czarownice w dolinie nowotarskiej*. – Lwów: Towarzystwo Ludoznawcze.
- FLIZAK Sebastian (1937): Moda w nadawaniu imion chrzestnych u górali pod Gorcami. – *Lud* XXXV, 55–59.
- GOSZCZYŃSKI Seweryn (1853): *Dziennik podróży do Tatrów*. – Petersburg: Nakładem B.M. Wolffa.
- HODOROWICZ Stanisław Andrzej (2013): *Słownik dawnych i współczesnych wyrażań górali Skalnego Podhala*. – Nowy Targ: Podhalańska Państwowa Wyższa Szkoła Zawodowa w Nowym Targu.

- JUSTYNIARSKA-CHOJAK Katarzyna (2010): *Testamenty i inwentarze pośmiertne z ksiąg miejskich województwa sandomierskiego (XVI–XVII wiek)*. – Kielce: Wydawnictwo Uniwersytetu Humanistyczno-Przyrodniczego Jana Kochanowskiego.
- KARŁOWICZ Jan (1900–1911): *Słownik gwar polskich*, t. 1–6. – Kraków: Nakładem Akademii Umiejętności.
- KAŚ Józef (2003): *Słownik gwary orawskiej*. – Kraków: Księgarnia Akademicka.
- KEMPF Zdzisław (1985): Wyrazy „gorsze” dotyczące zwierząt. – *Język Polski* LXV, 125–144.
- KLEMENSIEWICZ Zenon, LEHR-SPLAWIŃSKI Tadeusz, URBAŃCZYK Stanisław (1965): *Gramatyka historyczna języka polskiego*. – Warszawa: Państwowe Wydawnictwo Naukowe.
- KLINT Piotr, wyd. (2008): *Testamenty szlacheckie z ksiąg grodzkich wielkopolskich z lat 1631–1655*, „Źródła Dziejowe”, t. 27, red. M. GÓRNY. – Poznań–Wrocław: Wydawnictwo Historyczne.
- KLINT Piotr, wyd. (2011): *Testamenty szlacheckie z ksiąg grodzkich wielkopolskich z lat 1657–1680*, „Źródła Dziejowe”, t. 28, red. M. GÓRNY. – Poznań–Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- KOŚC Józef (2001): Komentarz tekstologiczny do siedemnastowiecznego testamentu Jakuba Suszy – biskupa chełmskiego. – [w:] Czesław KOSYL (red.): *Studia z historii języka polskiego i stylistyki historycznej ofiarowane profesor Halinie Wiśniewskiej na 50-lecie jej pracy naukowo-dydaktycznej*. – Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, 101–109.
- KOŚC Józef (2010): Konwencje tekstu testamentowego w średniowiecznych relacjach sądowych. – *Poznańskie Studia Polonistyczne. Seria Językoznawcza* XVI, 99–109.
- KOWALKOWSKI Jacek, NOWOSAD Wiesław, oprac. i wyd. (2013): *Testamenty szlachty Prus Królewskich z XVII wieku*. – Warszawa: Wydawnictwo DIG.
- KRAWCZYK-TYRPA Anna (2001): *Tabu w dialektach polskich*. – Bydgoszcz: Wydawnictwo Akademii Bydgoskiej im. Kazimierza Wielkiego.
- KURZOWA Zofia (1973): Podhalańskie na kirkach nieść, na kirkki wziąć. – *Język Polski* LIII, 369–371 [przedruk: Zofia KURZOWA (2007): *Z przeszłości i teraźniejszości języka polskiego* 4, wyb. i oprac. Monika SZPICZAKOWSKA, Mirosław SKARŻYŃSKI. – Kraków: Universitas, 389–392].
- LORENTZ Friedrich (1908–1912): *Slovinzisches Wörterbuch*, t. I–II. – St. Petersburg: Kaiserliche Akademieder Wissenschaften.
- LORENTZ Friedrich (1958): *Pomoranisches Wörterbuch*, t. I–IV. – Berlin: Akademie-Verlag.
- LUBCZYŃSKI Mariusz, PIELAS Jacek, SUCHOJAD Henryk, oprac. (2005): *Cui contingit nasci, restat mori. Wybór testamentów staropolskich z województwa sandomierskiego*. – Warszawa: Semper.
- ŁOS – Janusz ŁOSOWSKI, oprac. (2015): *Testamenty chłopów polskich od drugiej połowy XVI do XVIII wieku*. – Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.
- MAYENOWA Maria Renata, PEPIŃSKI Franciszek (od t. XVIII), red. (1966–): *Słownik polszczyzny XVI wieku*. – Wrocław–Warszawa–Kraków: Ossolineum (1966–1994), Instytut Badań Literackich PAN (1995–).
- MOSZYŃSKI Kazimierz (1967): *Kultura ludowa Słowian*, t. 2, cz. 1. – Warszawa: Książka i Wiedza.
- NITSCH Kazimierz, ZBOROWSKI Juliusz (1952), Kirkak ‘kark’. – *Język Polski* XXXII, 81–84 (Nitsch, 81–82, Zborowski, 82–84).
- POPIOLEK Bożena (2009): „Woli mojej ostatniej testament ten...”. *Testamenty staropolskie jako źródło do historii mentalności XVII i XVIII wieku*. – Kraków: Wydawnictwo Naukowe Uniwersytetu Pedagogicznego.
- RAK Maciej (2017): Życie religijne, wierzenia i zwyczaje Podhalan w świetle testamentów z XVII i XVIII wieku. – [w:] Iwona STECZKO, Renata DŻWIGOŁ (red.), *Dialog z Tradycją*, t. VI: *Dawna i współczesna kultura funeralna*. – Kraków: Collegium Columbinum, 245–257.
- RAK Maciej (2018a): Między historią języka a dialektologią – słownictwo testamentów góralskich z XVII i XVIII wieku. – [w:] Magdalena PASTUCHOWA, Mirosława SIUCIAK (red.), *Historia języka w XXI wieku. Stan i perspektywy*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego, 402–414.

- Rak Maciej (2018b): *Niepile ręce i inne przykłady ludowego słownictwa prawniczego* (na materiale chłopskich rozporządzeń ostatniej woli). – [w:] Renata PRZYBYLSKA, Maciej RAK, Agata KWAŚNICKA-JANOWICZ (red.), *Historia języka, dialektologia i onomastyka w nowych kontekstach interpretacyjnych*. – Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego, 231–242.
- SIEKIERSKA Krystyna, red. (1999–): *Słownik języka polskiego XVII i I. połowy XVIII wieku*. – Kraków: Lexis (wersja elektroniczna pod kier. Włodzimierza GRUSZCZYŃSKIEGO: http://xvii-wiek.ijp-pan.krakow.pl/pan_klient/index.php?wstep_zakl=biogr).
- SYCHTA Bernard (1967–1976): *Słownik gwar kaszubskich na tle kultury ludowej*, t. I–VII. – Wrocław–Warszawa–Kraków–Gdańsk: Zakład Narodowy im. Ossolińskich.
- SZYMCZAK Mieczysław (1966): *Nazwy stopni pokrewieństwa i powinowactwa rodzinnego w historii i dialektach języka polskiego*. – Warszawa: Państwowe Wydawnictwo Naukowe.
- WITASZEK-SAMBORSKA Małgorzata (1985): Regionalizmy fonetyczne w mowie inteligencji poznańskiej. – *Slavia Occidentalis* 42, 91–104.
- WRZEŚNIEWSKI Augustyn (1885): Spis wyrazów podhalskich. – *Pamiętnik Towarzystwa Tatrzańskiego* X, 1–26 [przedruk: Joanna OKONIOWA, red. (2006): *Studia dialektologiczne* III. – Kraków: Lexis, 185–232].
- ZAJDA Aleksander (2001a): Dwadzieścia dwa rzadkie wyrazy prawnicze z ksiąg sądowych wiejskich. – [w:] Zofia CYGAL-KRUPA (red.): *Studia językoznawcze. Dar przyjaciół i uczniów dla Zofii Kurzowej*. – Kraków: Universitas, 421–430.
- ZAJDA Aleksander (2001b): *Studia z historii polskiego słownictwa prawniczego i frazeologii*. – Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- ZAJDA Aleksander (2016): Testamenty jako źródło do dziejów słownictwa z dziedziny kultury materialnej. – [w:] Ewa MŁYNARCZYK, Ewa HORYŃ (red.): *Dialog z Tradycją*, t. V: *Językowe dziedzictwo kultury materialnej*. – Kraków: Collegium Columbinum, 203–227.
- ŻMIGRODZKA Bożena (1997): *Testament jako gatunek tekstu*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego.

doi: <http://dx.doi.org/10.17651/bptj.74.10>

KATARZYNA WOJAN
Uniwersytet Gdański, Gdańsk

Bibliografie słowników przekładowych jako narzędzie rekonstrukcji dziejów oraz oceny stanu i dorobku słownictwa narodowego

Streszczenie

Celem artykułu jest zaprezentowanie różnorodnych zastosowań tworzonych baz faktograficznych w postaci usystematyzowanych, zaopatrzonych w latentne informacje, bibliografii polskich słowników przekładowych obejmujących 136 języków obcych. Bazy bibliograficzne są fundamentem całościowego bądź specjalizowanego opisu analitycznego dziejów słownictwa narodowego (ich rekonstrukcji), w tym jego poszczególnych działów, których obiektem są języki obce (klasyczne, nowożytnie, wymarłe, skonstruowane, czyli tzw. „sztuczne”). Analiza zebranego w nich materiału umożliwia leksykografom redakcje początków danego działu słownictwa. Bibliografie stanowią punkt wyjścia do tworzenia podstaw naukowych komparatystyki leksykograficznej. Usystematyzowane językowo, chronologicznie i/lub dziedzinowo bibliografie stanowią ważkie, a zarazem miarodajne narzędzie opisu rozwoju historycznego leksykografii rodzimej, a także obiektywne narzędzie przeprowadzenia wieloaspektowych ocen jej obecnego stanu i dorobku zarówno w aspekcie kwantytatywnym, jak i kwalitatywnym.

Słowa kluczowe: bibliografia, leksykografia polska, słowniki przekładowe, historia, metodologia

Bibliographies of multilingual dictionaries as a tool for reconstruction and assessment of national lexicographic achievements

Abstract

The aim of the article is to present various applications of factographic databases in the form of systematized bibliographies of Polish translation dictionaries, covering 136 foreign languages. Bibliographic databases are the foundation of a comprehensive or specialized analytical description of the national history of lexicography (its reconstruction), including its individual branches focusing on foreign languages (classical, modern, extinct, constructed, i.e. so-called “artificial”). Analyzing the material collected in bib-

liographic databases enables researchers to redate the beginnings of a given branch of lexicography. Bibliographies constitute a starting point for creating the basis of lexicographic comparative studies. Bibliographies arranged according to linguistic, chronological and/or domain-related criteria are an important and reliable tool for describing the historical development of native lexicography, as well as an objective tool for conducting multi-faceted assessments of its current state and achievements in both qualitative and quantitative aspects.

Keywords: bibliography, Polish lexicography, bilingual and monolingual dictionaries, history, methodology

Wprowadzenie

Przed niespełną stuleciem Stefan Vrtel-Wierczyński (1886–1963), wybitny poznański bibliograf i bibliotekoznawca, historyk literatury polskiej i słowiańskiej, a także inicjator Polskiej Bibliografii Literackiej, zawarł w „Przeglądzie Bibliotecznym” znaczącą dla rozwoju polskiej bibliografistyki i naukoznawstwa konstatację, że każda dziedzina nauki „powinna mieć swoją bibliografię, opracowaną wyczerpująco i systematycznie, przystosowaną do jej potrzeb, obejmującą jej całokształt” (Vrtel-Wierczyński 1928: 7). Ów „Przegląd Biblioteczny”¹ był pierwszym polskim naukowym, a jednocześnie zawodowym, czasopismem bibliotekarskim², stanowił ponadto istotną „siłę sprawczą w usamodzielnieniu się i rozwoju dyscypliny bibliotekoznawstwa i informacji naukowej” (zob. Gruszka 2012³). Spisy bibliograficzne, jak powszechnie wiadomo, powstają od zarania dyscyplin naukowych (por. Matczuk 2014: 9), zaś w wiekach średnich „bibliografia była jednym z najważniejszych narzędzi przekazywania informacji, bez którego nie byłoby postępu naukowego” (Gmiterek 2015: 397). Współcześnie młodzi badacze specjalizujący się w metaleksykografii i metaterminografii obficie korzystają z dostępnych publikacji bibliografii słowników; do ich grona należy m.in. Marek Łukasik (2017: 51), który stwierdza, że „Bibliografie przedmiotu są istotnym narzędziem rejestrowania dorobku dziedziny. Potrzeba ich sporządzania z jednej strony wynika z chęci podsumowania pewnego okresu działalności naukowej/praktycznej, z drugiej zaś świadczy o «dojrzałości» dyscypliny i jest przejawem podjęcia własnych metabadań”.

Jedną z form piśmienniczych bibliografii jest monografia bibliograficzna, będąca „wyjątkowym opracowaniem” o nieschematycznej budowie, powstała „w wyniku badań jednego tylko tekstu – w jednym lub wielu wydaniach oraz jednego rodzaju i gatunku literackiego lub piśmienniczego” (Bajor 2005: 9). Monografie bibliogra-

¹ Pierwszą próbą polskiego czasopisma poświęconego teorii i praktyce bibliotekarstwa był kwartalnik „Przegląd Biblioteczny. Czasopismo ilustrowane poświęcone bibliotekoznawstwu, bibliotekarstwu i bibliografii” powstałe w 1908 roku.

² Dodatkiem ciągłym do „Przeglądu Bibliotecznego” jest *Bibliografia Analityczna Bibliotekoznawstwa i Informacji Naukowej* opracowywana w Instytucie Bibliograficznym Biblioteki Narodowej.

³ Cytat pochodzi z opisu książki: *Przegląd Biblioteczny: monografia* Zbigniewa Gruszki zawartego w: „Informator Sekcji Bibliotek Niepaństwowych Szkół Wyższych Zarządu Głównego Stowarzyszenia Bibliotekarzy Polskich” R. 8, z. 5 (210): poniedziałek, 4 lutego 2013 roku, s. 3.

ficzne⁴ mają wymiar heurystyczny (Wojan 2017b). Szeroko zakrojone konkretne działania naukowe cechuje się bowiem badawczością i odkrywczością, certyfikuje dynamikę rozwoju danej dziedziny wiedzy oraz stopień jej progresji. Tego typu publikacje stanowią bezsprzecznie miarodajne i w pełni użyteczne narzędzie badawczo-naukowe, nieocenioną pomoc w dydaktyce akademickiej, ponadto pełnią rolę praktycznego, nieodzownego informatorium dla specjalistów w danej dziedzinie (np. Grzegorzczak 1967; Czachowska, Loth 1989; prace Wojan). Zrównoważony rozwój bibliografistyki jest sprzężony z powstaniem i rozwojem wysokiej kultury naukowej. Metody bibliografów – „doskonalone przez stulecia – wciąż oddają usługi dyscyplinom naukowym, co więcej, wydatny jest udział bibliografii w powstawaniu nowych dziedzin” (Matczuk 2014: 9). Bibliografie stanowią fundamentalne systematyzacje poznawcze, wpisują się do „nauk o nauce” (Wojan 2017b).

Celem niniejszego artykułu jest, po pierwsze: krótkie zaprezentowanie różnorodnych zastosowań tworzonych baz faktograficznych w postaci kompletnych, usystematyzowanych, zaopatrzonych w latentne informacje, bibliografii polskich słowników przekładowych, które z założenia stanowić mają przydatne narzędzie w rekonstrukcji dziejów słownictwa oraz wieloaspektowej ocenie synchronicznej i diachronicznej stanu leksykografii polskiej; po drugie: zapoznanie z rezultatami bieżących prac badawczych w zakresie bibliografistyki leksykograficznej.

Dorobek naukowy w zakresie polskiej bibliografistyki leksykograficznej

W centralnym katalogu zbiorów polskich bibliotek naukowych NUKAT dla zapytania: „bibliografia” widoczne są 31 722 rekordy⁵ opisów dokumentów, dla „monografia bibliograficzna” – zaledwie 23 rekordy, zaś dla „bibliografia słowniki” – 290 rekordów⁶, przy czym – należy zauważyć – nie wszystkie tytuły odpowiadają tej klasyfikacji. Obecnie najstarszą udokumentowaną w katalogu bibliotecznym polską (załącznikową) bibliografią słowników jest *Księdza Stanisława Zaborowskiego Ortografia polska: z łacińskiego na polski język przełożona z przydaniem uwag tłómacza, tudzież Ortografii Seklucyana i spisu bibliograficznego grammatyki i słowników polskich* ułożona przez Andrzeja F. Kucharskiego z 1825 r. Z najnowszych zaś katalog rejestruje, różne pod względem genologii i przedmiotowego ukierunkowania, dwie pozycje: obszerny *Słownik biograficzno-bibliograficzny polskich franciszkanów konwentualnych zamordowanych i tragicznie zmarłych w latach 1939–1945* Łukasza Janneckiego oraz *Z dziejów leksykografii polsko-rosyjskiej. Tom 1: Słowniki lingwistyczne (bibliografia za lata 1700–2015)* Katarzyny Wojan – obie prace powstały w 2016 r. W zbiorze pozycji w kategorii „słowniki bibliograficzne” katalog NUKAT uwzględ-

⁴ Jak stwierdza Bajor (2005: 9), „termin bibliografia monograficzna w praktyce jest rzadko używany i wzbudza wiele kontrowersji wśród teoretyków bibliografii ze środowisk filologów i bibliologów”.

⁵ Stan z dn. 16.12.2017.

⁶ Stan z dn. 16.12.2017.

nia zaledwie 37 tytułów⁷, przy czym przeważają w nim opracowania Jana Wawrzyńcy, a zatem współczesne i w większości zorientowane na językoznawstwo sławistyczne i rusycystyczne.

Przytoczone powyżej dane statystyczne zaczerpnięte z systemu informacyjno-wyszukiwawczego nie są jednak pełne. W katalogu brakuje bowiem wielu ważnych książkowych pozycji z gatunku monografii bibliograficznych, jak też słowników bibliograficznych – nie zostały one bowiem przyporządkowane do właściwej kategorii (bądź podkategorii) hasła przedmiotowego. Można zatem stwierdzić, że powyższe dane jedynie fragmentarycznie odzwierciedlają stan rozwoju polskiej bibliografistyki. Ich zbiór w istocie jest liczniejszy. Dokładna analiza zawartości polskich katalogów bibliotecznych⁸ ujawnia różnorodne luki zaistniałe w procesie bibliografowania.

Bibliografia najczęściej rozumiana jest jako „uporządkowany zbiór opisów bibliograficznych dokumentów dobranych według określonych kryteriów, którego celem jest informowanie o istnieniu tych dokumentów” (Bajor 2005: 7). Bibliografie różnią się strukturą (formą) i przeznaczeniem; wyodrębnia się ich różne typy: adnotowane, przedmiotowe, bieżące, narodowe, komentowane, zalecające, specjalne, retrospektywne, międzynarodowe, materiały bibliograficzno-eksplanacyjne, katalogi, spisy, wykazy publikacji, indeksy, słowniki biobibliograficzne, przewodniki bibliograficzne i leksykalno-bibliograficzne, informatory i in. Jedną z głównych, a tym samym najbardziej powszechnych, form bibliografii jest spis bibliograficzny o schematycznym porządku, zwykle chronologicznym lub tematycznym/przedmiotowym (ogólnym lub szczegółowym).

Bazy bibliograficzne⁹ na potrzeby niniejszych badań leksykograficznych tworzone są na podstawie kwerendy źródeł – dokumentów (w tym rękopiśmiennych), katalogów bibliotecznych i wydawniczych, przewodników bibliograficznych, różnego rodzaju opracowań naukowych, recenzji, projektów naukowych, korpusów online. Gromadzone bibliografie zasobowe jednostek piśmienniczych mają charakter inwentaryzacyjny i adnotowany, zawierają odpowiednie opisy bibliograficzne oraz komentarze. Bazy bibliograficzne leksykografii obejmują opracowania genologicznie różnorodne – lingwistyczne, terminologiczne oraz encyklopedyczne. Przybierają one również charakter prospektywny, gdyż uwzględniają informacje o pracach przygotowywanych w ośrodkach naukowo-badawczych do publikacji. Zgromadzone zasoby bibliografii adnotowanych polskich opracowań słownikarskich umożliwią dokonanie klasyfikacji słowników z uwzględnieniem różnorodnych kryteriów i cech strukturalnych tekstów słownikowych.

⁷ Stan z dn. 16.12.2017.

⁸ Bazy katalogów bibliotecznych różnią się pod wieloma względami, m.in. pod względem zawartości adresów bibliograficznych, opisów bibliotecznych (odmienność słów kluczowych, kategorii haseł i podhaseł przedmiotowych, języka informacyjno-wyszukiwawczego), nie są też ze sobą zintegrowane, np. katalog główny Biblioteki Narodowej nie jest niestety kompatybilny z centralnym katalogiem bazy NUKAT, co powoduje szereg różnego rodzaju rozbieżności analitycznych.

⁹ Sposób przechowania gromadzonych danych źródłowych – własna baza danych na prywatnym nośniku pamięci.

W polskiej literaturze naukoznawczej gruntowne, obszerne indeksy słowników polskich (w tym translatorycznych) sporządzili m.in. Kazimierz Lewański (1959; 1963; 1964; 1972; 1973a, b, c) (słowniki z językami słowiańskimi), Piotr Grzegorzczak (1967) (słowniki narodowe, przekładowe, terminografia), Hanna Kwiatkowska (1977) (terminografia), Edward Stankiewicz (1984) (słowniki z językami słowiańskimi), Elżbieta Kędelska (1986; 1995) (słowniki łacińsko-polskie), Agnieszka Frączek oraz Ryszard Lipczuk (słowniki z językiem niemieckim) (Frączek i Lipczuk 2004; Frączek 2010), Tadeusz Piotrowski (2001a, b) (słowniki z językiem angielskim), Marek Łukasik (2007; 2014) (terminografia polsko-angielska), Danuta Szuper-Jakubiuk (2010) (słowniki handlowe z polskim, rosyjskim i ukraińskim) oraz Katarzyna Wojan (leksykografia polsko-rosyjska (Wojan 2006; 2016b), polsko-angielska (Wojan 2013b; 2014a), polsko-białoruska (Wojan 2013c), polsko-ukraińska (Wojan 2017a), polsko-łacińska (Wojan 2013d; 2014b), polsko-esperancka (Wojan 2013a), polsko-fińska (Wojan 2016c)). Ogromne dokonania na polu bibliografistyki leksykograficznej posiada Jan Wawrzyńczyk, który zainicjował i rozwinął odrębny kierunek, koncentrujący się na tworzeniu oryginalnych słowników leksykalno-bibliograficznych w zakresie językoznawstwa polskiego i rosyjskiego (zob. m.in. Wawrzyńczyk 1994; 1998; 2001; 2002; 2005; Wawrzyńczyk i Małek 2004; 2014–2015). W tym samym nurcie sytuuje się praca *Język potoczny w polskiej lingwistyce. Materiały do Przewodnika leksykalno-bibliograficznego* Wojan (2015).

Opracowania bibliograficzne przybierają częstokroć charakter monografii bibliograficznych (zob. np. prace: Wojan 2006; 2013b; 2014a; 2016b; Konefał 2016; za granicą zaś klasyczne opracowania: Collins 1955; Jones 2000 i in.), są ważną formą piśmienniczą, pełniącą rolę informatorium i kompendium wiedzy z danego obszaru (korzyści z nich płynące dla translatoryków, tłumaczy praktyków, naukoznawców, historyków danego obszaru nauki). Bibliografie słowników stanowią istotny wskaźnik częstości zaistnienia kontaktów językowo-kulturowych. Obecnie najlepiej opracowane pod względem analityczno-bibliograficznym jest słownictwo polsko-angielskie (zob. Piotrowski 2001b; Wojan 2013b; 2014a; Łukasik 2007; 2014), co z kolei jest kompatybilne z polską praktyką leksykograficzną dotyczącą obopólnych relacji tych języków. Z kolei, jak stwierdza Łukasik (2017: 47), „opracowania monograficznego nie doczekała się historia polskiej terminografii”. Jest to ważny postulat, bowiem leksykografia polska posiada pewną tradycję w praktyce i teorii tej dyscypliny. Skonkretyzowane zasoby bibliograficzne mogą z powodzeniem stać się dobrym narzędziem wieloaspektowego opisu jakościowego i ilościowego polskiej leksykografii terminologicznej zarówno jedno-, jak i wielojęzycznej.

Polska bibliografistyka leksykograficzna – w aspekcie genologicznym

Do tej pory drogą kwerend bibliotecznych (analizy zawartości katalogów) oraz tekstowych utworzono¹⁰ specjalistyczne bazy bibliograficzne dla dziesięciu głównych

¹⁰ Własna baza danych. Stan z dnia 5.02.2017.

typów strukturalnych leksykografii: 1) narodowej, 2) przekładowej, 3) terminografii jedno- i wielojęzycznej, 4) mimicznej, 5) brajlowskiej, 6) bio- i bibliografistyki, 7) encyklopedystyki i leksykografii inwentaryzacyjnej, 8) leksykografii onomastycznej, 9) kartografii lingwistycznej, 10) korpusologii i konkordancji.

Powyższa typologia strukturalna leksykografii może wzbudzać pewne kontrowersje, bowiem w polskim piśmiennictwie naukowym nie zaproponowano, jak dotąd, obiektywnej typologii ani też genologii tekstów słowników (zespołu cech w postaci wzorca przypisywanych pewnym typom tekstów słownikowych), a przede wszystkim nie opisano wyznaczników struktury gatunkowej słowników. Istniejące w polskiej literaturze językoznawczej klasyfikacje leksykografii cechuje duży stopień uogólnienia i uproszczenia (zob. np. w pracy Żmigrodzkiego (2005)). Szybko zmieniające się realia kulturowo-społeczno-gospodarcze implikują ewolucję i transformację różnych gałęzi słownikarstwa wychodzącego naprzeciw różnorodnym zapotrzebowaniom odbiorcy oraz wymogom rynku wydawniczego, integrowanie różnych gatunków tekstów (fuzję tradycyjnej formy słownika z innymi formami tekstowymi), interdyscyplinarność leksykografii (w praktyce i teorii). Leksykografowie nie postulują nawet zbudowania dynamicznego modelu gatunków słownikowych, skorelowanego z rytmem konwersji czytelniczej i odpowiadającego klimatowi intelektualnemu epoki, uwzględniającego różnorakie transgresje. Trzeba też pamiętać, że zasadniczo zmienia się instrumentarium naukowe, eksplozywnie rozwija się metodologia badawcza, powstają nowe możliwości repozytaryjne (cyfryzacja nauki), co z kolei otwiera słownikarzom pole do szerszej zakrojonych działań, zwiększając ich potencjał. Wymusza to innowacyjne aplikacje leksykograficzne, skoordynowane z różnymi obszarami działalności intelektualnej człowieka. W rzeczywistości sprowadza się to do zmiany paradygmatu leksykograficznego, poszerzenia go o nowe formy tekstowe i przypisanie mu nowych funkcji. Przywodzi to niejako konieczność rozszerzenia definicji słownika jako tekstu użytkowego, definiowanego dotąd najczęściej jako „zbiór uporządkowany materiału leksykograficznego” (Grzegorzczuk 1967: 10).

Stworzona autorska¹¹ baza bibliograficzna słowników polskich obejmuje różne typy strukturalne słownikarstwa; w większości mają one charakter translacyjny. Zasoby słownikarstwa narodowego są tu konstytuowane przez:

a) słowniki narodowe (ogólnodefinicyjne języka polskiego, ideograficzne, słowniki dla cudzoziemców, obrazkowe, historyczne, ortograficzne, ortoepiczne, wyrazów obcych, zabytki leksykograficzne);

b) słowniki odmian językowych (dialektów, gwar terytorialnych, gwar środowiskowych, leksyki bachicznej, polszczyzny potocznej, polszczyzny politycznej, wulgaryzmów);

c) słowniki frazeologiczne, paremiologiczne (przysłów, powiedzeń), skrzydlatografia (słowniki skrzydlatych słów i cytatów), aforystykony;

d) słowniki specjalne (słowniki a tergo, afektonimów, anagramów, antonimów, archaizmów, eponimów, erotyzmów, etykiety językowej, etymologiczne, eufemiz-

¹¹ Badania prowadzone w ramach naukowej Pracowni Leksykograficzno-Bibliograficznej afiliowanej na Wydziale Filologicznym Uniwersytetu Gdańskiego.

mów, frekwencyjne, gramatyczne, homonimów, języka pisarzy, neologizmów, onomatopei, paronimów, peryfraz, pleonazmów, skrótów, strukturalne, stylistyczne, słowotwórcze, synonimów).

Polskie słownictwo przekładowe na obecnym etapie rozwoju uwzględnia 136 języków obcych, będących obiektem rejestracji lub komparacji. Zbiór¹² ten (liczący obecnie 17 781 prac słownikarskich) podzielono na następujące działy¹³:

a) słowniki z żywymi językami tzw. dużymi, takimi jak: angielski (5419 pozycji), arabski (119), chiński (54), francuski (1736), hiszpański (461), niderlandzki (holenderski, flamandzki) (121), niemiecki (2390), portugalski (91), rosyjski (2449), ukraiński (225), włoski (542) – łącznie 13 607 pozycji leksykograficznych;

b) słowniki z żywymi językami tzw. mniejszymi używanymi w Europie, takimi jak: albański (11), baskijski (2), białoruski (104), bośniacki (17), bułgarski (72), celtyckie (1), chorwacki (74), czeski (208), duński (49), estoński (18), fiński (45), flamandzki (9), grecki (185), irlandzki (11), islandzki (18), jidisz (33), karaimski (10), kataloński (5), ladino (judeo-hiszpański) (1), litewski (116), luksemburski (1), łotewski (25), łużyckie: górno- i dolnołużycki (36), macedoński (14), maltański (2), norweski (63), retoromański (sursylwański) (2), romski (romani, cygański) (13), rumuński (54), serbski (39), serbskochorwacki (34), słowacki (103), słoweński (41), szkocki (1), szwedzki (131), tatarski (4), waloński (1), węgierski (143) – łącznie 1 696 pozycji leksykograficznych;

c) słowniki z żywymi językami używanymi poza Europą, takimi jak: afrikaans (5), ajnusi (~ ainu) (9), amharski (2), armeński (ormiański) (18)¹⁴, azerbejdżański (~ azerski) (1), chewa (~ chichewa ~ nyanja ~ chinyanja) (1), buriacki (2), dari (5), dagurski (2), dołgański (3), ewenkijski (5), gruziński (5), hawajski (2), hebrajski z hebrajskim biblijnym (116), hindustani (1) z odmianami: hindi (8) oraz urdu (1), indonezyjski (8), inuicki (1), jakucki (25), japoński (79), kamczadański (itelmeński) (6), kampa (1), karaczajsko-bałkarski (1), kaukaskie (2), kazachski (4), kilivila (~ kiriwina) (1), koreański (9), koriański (2), kurdyjski (1), laotański (1), lingala (1), malgański (1), mambwe (1), mandżurski (5), mołdawski (2), mongolski (25), nanajski (goldyjski, goldzki) (7), neomelanezyjski (~ tok pisin) (1), niwchijski (giliacki) (8), orocki (~ uilta ~ uiltyński) (5), oroczański (nani) (1), paszto (~ pusztu) (afgański) (3), perski (14), suahili (8), tadżycki (1), tajski (2), tunguskie (5), turecki (67), turkmeński (2), tybetański (2), ujugurski (1), udegejski (~ udihe) (3), ulczyjski (~ olcza ~ ulcza) (5), wietnamski (32) – łącznie 529 pozycji leksykograficznych;

d) słowniki z językami dawnymi, wymarłymi i rewitalizowanymi, takimi jak: aramejski (10), awestyjski (4), dalmatyński (16), egipski (4), etruski (1), fenicki (1), iliryjski (1), kipczacki (kumański, połowiecki) (10), koptyjski (2), łacina (1310), połab-

¹² W nawiasach podaje się liczbę pozycji słowników. Statystyką objęto wszystkie wydania, a zatem wznowienia, aktualizacje itp. – dane własne, stan badań z dn. 20.12.2017.

¹³ Przy podziale zastosowano kryterium liczby użytkowników języków oraz arealne (geograficzne).

¹⁴ Należy tu jednak zwrócić uwagę na fakt, że armeński, azerski i gruziński są oficjalnymi językami państw będących członkami Rady Europy.

ski (9), prasłowiański (3), pruski (8), ruski (4), sanskryt (23), staro-cerkiewno-słowiański (51), staroruski (13) – łącznie 1470 pozycji leksykograficznych;

e) słowniki z językami skonstruowanymi, takimi jak: ido (3), esperanto (324), latino (~ latina) sine flexione (1), paraglot (1), wspólny język słowiański (2) – łącznie 331 pozycji leksykograficznych;

f) słowniki z językami regionalnymi, takimi jak: arumuński (macedo-rumuński) (1), bojkowski (9), huculski (9), kaszubski (84), łemkowski (21), rezjański (1), xibe (~ xibo ~ sibe ~ sibo) (2), śląski (13), wilamowicki (4) – łącznie 144 pozycji leksykograficznych;

g) słowniki rejestrujące odmiany danego języka, takie jak: austriacka odmiana niemieckiego (1), egipski arabski (2), marokański arabski (1) – łącznie 4 pozycje.

Terminografia rodziła jedno-, dwu- i wielojęzyczna, a także dokumentacyjna, stanowi potężny dział leksykografii, opisującej co najmniej 475 dziedzin i przedmiotów. Obecnie obserwowana jest jej znaczna aktywizacja, jej zasoby słownikarskie są ogromne. Ten dział wymaga odrębnych, szczegółowych opisów. Prace badawcze w tym obszarze zapoczątkował już Łukasik (2007; 2014; 2017).

Aplikatywne gałęzie leksykografii stanowią słowniki mimiczne (języka migowego/miganego) (30 pozycji), a także brajlowskie (6).

Nader obszernym i wieloobszarowym działem słownictwa jest encyklopedystyka; dotychczas udokumentowano opracowania tematycznie odnoszące się do co najmniej 123 obszarów, dziedzin i tematów. Część słowników encyklopedycznych ma charakter przekładowy. W nauce polskiej, jak dotąd, nie zajmowano się całościowo historią encyklopedystyki.

Bio- i bibliografistykę tworzą – prócz słowników stricte bibliograficznych i bio-bibliograficznych – m.in. takie opracowania, jak słowniki pisarzy, hagiograficzne, genealogiczne i heraldyczne (herbarze). Pokażną gałąź leksykografii inwentaryzacyjnej stanowią spisy (geograficzne, osobowe i in.), inwentarze różnego typu; ten typ leksykografii jest ważny dla archiwizowania wiedzy i upowszechniania polskiego dorobku intelektualnego.

Do zadań leksykografii onomastycznej należy opracowywanie słowników i wykazów nazw. W stworzonej bazie wyodrębniono działy słowników nazw barw, archiwów, instytucji, ergonimów, etnonimicznych, glottonimicznych, biblijnych, osobowych, przezwisk, numizmatycznych, ojkonimów, miast, szkół, Warszawy, geograficznych, fizjograficznych, hydronimicznych, fitonimicznych, zoonimicznych i in. Zasoby polskiej leksykografii onomastycznej są ogromne, często też onomastykone konstruowane są za pomocą metody innojęzycznych komparacji.

Kartografię lingwistyczną, dotychczas w ogóle niewiązaną z leksykografią, prezentują naukowe atlasy lingwistyczne (48). Jest to niezwykle interesujący dział lingwistyki aplikatywnej.

Rozwijającą się w Polsce dziedziną leksykografii jest korpusologia i konkordancja.

Obecnie cenną aplikacją bibliografistyki stała się netografia.

Bibliografistyka jako dział leksykografii aplikatywnej

Bazy bibliograficzne stanowią fundament całościowego bądź też specjalizowanego opisu analitycznego dziejów – a ściślej: ich rekonstrukcji – słownikarstwa narodowego, w tym jego poszczególnych działów, których obiektem są języki obce (klasyczne, nowożytne, regionalne, wymarłe, rewitalizowane, skonstruowane, czyli tzw. „sztuczne”). Zgromadzone i uporządkowane bibliografie służą procesom weryfikacji określonych też związanych z kwestiami słownikarstwa polskiego, upowszechnionych w piśmiennictwie rodzimym i obcojęzycznym. Często dotyczą one chronologizacji słownikarstwa. Analiza zebranego materiału dokumentacyjnego w postaci usystematyzowanych chronologicznie bibliografii słowników umożliwia leksykografom dokonanie redatacji początków danego działu słownikarstwa (polsko-innojęzycznego). W ostatnich latach w polskim piśmiennictwie naukowym pewne redatacje już zostały poczynione (Wojan 2013b; 2013c; 2016b). W syntetycznym opracowaniu historii słownikarstwa polsko-angielskiego wyniki analizy zebranego materiału bibliograficznego słowników umożliwiły zmianę w datowaniu jego początków (Wojan 2013b), odmiennie podawanego przez badaczy w literaturze przedmiotu. W oparciu o zebrane i zweryfikowane adresy bibliograficzne ustalono, że pierwszą pracą leksykograficzną z językiem angielskim był *Dykcyonarzyk Geograficzny Czyli Opisanie Krolestw, Prowincyi, Miast, Biskupstw, Xięstw, Hrabstw, Margrabstw, Portow, Fortecy i innych mieysc znaczniejszych w czterech częściach świata* [...] Echarda Laurence’a w tłumaczeniu Franciszka Siarczyńskiego z 1782 r. Wcześniej zaś angielsi utrzymywali, że pierwszym opracowaniem zawierającym pierwiastek słownikarski była *Xiązka wypisów angielskich z słownikiem* Krystyna Lacha Szyrmy z 1828 r. (zob. Piotrowski 2001b: 186), bądź też (już nieco później) *Słownik potrzebniejszych wyrazów* zamieszczony w *Gramatyce dla Polaków chcących się uczyć angielskiego języka* Juliana Antonowicza¹⁵ z 1788 r. (zob. Podhajecka 2016). Pomijano rolę fundamentalnego *Słownika język polskiego* w sześciu tomach Samuela Bogumiła Lindego (pierwsze wydanie w Warszawie w latach 1807–1814), który w artykułach hasłowych notował liczne ekwiwalenty angielskie i staroangielskie. Na podstawie empirycznych analiz bibliograficznych stwierdzono też, że pierwszą pracą leksykograficzną na rynku amerykańskim, w której miało miejsce spotkanie języków: polskiego i angielskiego, był multilingwalny podręcznik *An Attempt Towards An International Language* Ludwika Zamenhofa z 1889 r., wyposażony w słowniczkę z polskim elementem leksykalnym: *International-English. English-International* (Zamenhof 1889). Uwzględnienie działalności wydawniczej Eliasza Kopijewicza (Kopijewskiego) wpłynęło na przesunięcie ram czasowych początków słownikarstwa z językiem rosyjskim w centrum. Kopijewicz, polsko-białoruski pedagog i wydawca, jest autorem *Номенклаторъ, на латинскомъ, рускомъ и немѣцкомъ іазыкъ* (*Nomenclator trium linguarum, Latine, Russice et Germanice*) wydanej w Amsterdamie w 1700 r. Za pierwszą pracę

¹⁵ Dziełko dostępne jest w repozytorium UJ – <https://jbc.bj.uj.edu.pl/dlibra/doccontent?id=302006> (dostęp: 12.12.2017).

posiadającą słowniczek polsko-rosyjski uznawano zaś *Łatwy sposób nauczania się po rosyjsku i po polsku czytać i pisać... z przydaniem... słów pod alfabetem co potrzebniejszych a używanych najczęściej* Franciszka Jaxycy Makulskiego z 1795 r. (Wojan 2006). Z kolei Tadeusz Piotrowski (2001b) wskazywał na – powstały w Petersburgu w Rosji – *Польский общій словарь и библейный, с польскою, латинскою и руссийскою новоисправленною библиями смечиван...* Cyriaka Kondratowicza z 1775.

Poniżej przytacza się wybrane datacje ważniejszych działów polskiego słownikarstwa przekładowego¹⁶. Istotną część wykazu stanowią redatacje dokonane dzięki zgromadzonym danym bibliograficznym. Listę opracowań inicjujących polski dział słownikarstwa przekładowego konstituują języki (wybór):

- **greka, niemiecki, arabski**: Jan Stanko, *Antibolomenum (Antidotarium, Antibolarium) Benedicti Parthi* (rps) 1472;
- **węgierski**: Silvestris Joannes Pannonius, *Dictionarius Latine, Germanice, Polonice et Ungarice*, Kraków 1531;
- **włoski**: [Anonim] (*Słownik*) *Dictionarius seu Nomenclatura quatuor linguarum Latine Italice Polonice & Theutonice, a prime cuius vtilissimus cum peregrinantibus tum domiresidentibus Adiecto vocabulorum indice. Vokabularz nowy czterech iezikow: Łaczińskiego, Włoskiego, Polskiego, Niemieckiego wssem w tey sławney koronie y inszym narodom barzo wzyteczny*, Kraków 1532;
- **turecki**: Bartholemej Georgijewiç, słownik polsko-turecki [dołączony do *Rozmowa z Turczynem o wierze krześcijańskiej*], Kraków 1548;
- **francuski**: [Anonim] *Dictionarius seu Nomenclatura Quattuor Linguarum, Latine, Gallice, Polonice et Teutonice...*, Kraków 1574;
- **litewski**: Konstanty Szyrwid, *Dictionarium trium linguarum in usum studiosae juventutis*, Wilno 1629;
- **języki sztuczne**: wspólny język słowiański: Marian z Jaślisk, *Dictionarium Sclavo-Polonicum*, 1641;
- **hiszpański**: [Anonim] *Dictionarium (Hexaglosson) cum Multis Colloquiis pro Diversitate Status Hominum, Quotidie Occurrentibus. Dykcjonarz sześci języków... teraz nowo polskim językiem objaśniony*, Warszawa 1646;
- **perski**: Franciszek (Mesgnien) Meniński, *Thesaurus linguarum orientalium Turcicae-Arabicae-Persicae... nimirum lexicon Turcico-Arabico-Persicum...*, Wiedeń 1680;
- **ormiański**: Deodatus Nierszesovicz (Deodat Nersesowicz), *Dictionarium Latino-Armenum Super Sacram Scripturam, et libros Divini Officii Ecclesiae Armenae, compositum per D. N. E. T.*, Rzym 1695;
- **egipski**: Karol Gotfryd Wojda (Wojde), *Lexicon Aegyptiaco-Latinum ex veteribus illius linguae monumentis summo studio collectum...*, Ozonii e Typographo Clarendoniano 1775;
- **ukraiński**: *Dykcjonarz, czyli Słownik Polsko-Ruski, Niemiecki, Łacinski. Słowiański, po czesci, Francuski etc z układem słow na trojakim stylu ku wygodzie*

¹⁶ Badania własne.

- i pożytkowi uczacych się takowych języków z różnych przewodów różnych autorów zebrany i napisany roku 1794 w Mohylewie, Mohylew 1794 (rps);*
- **japoński**: Golovnin Vasilij M., *O Japonii uwagi B. Gołownina, do których jest przydany Słowniczek japoński układu K.P. Thunberga*, przeł. Grzegorz Buczyński, Warszawa 1823;
 - **romski, hindi**: Ignacy Daniłowicz, *O Cyganach. Wiadomość historyczna, czytana na posiedzeniu publicznym Cesarskiego Uniwersytetu Wileńskiego, dnia 30 czerwca 1824 roku. Z dodaniem wyrazów i wzorów odmian gramatycznych języka tego narodu [wzór cygańskich wyrazów z indyjskimi porównanymi]*, Wilno 1824;
 - **mongolski**: Józef Szczepan Kowalewski, *Монгольско-русско-французский словарь...*, Kazań 1844–1849;
 - **sanskryt**: Kajetan Kossowicz, *Санскрито-русский словарь*, Petersburg 1854;
 - **awestyjski**: Ignacy Pietraszewski, *Zend-Avesta ou plutôt Zen-Daschta*, Vol. 1, *Chapitre I–VIII du Vendidad = Miano slavianskie w ręku jednej rodziny od trzech tysięcy lat zostające, czyli Nie Zendawesta, a Zędaszta, to jest życie dawcza książeczka Zoroastra*, Berlin 1858;
 - **kaszubski**: Pobłocki, *Słowniczek kaszubski*, Poznań 1875;
 - **niderlandzki, duński**: Erazm Lucjan Kasprowicz, *Tornister-Dolmetsche des deutschen Reichssoldaten im täglichen Werkehr mit den Grenzvölkern des Deutschen Reiches. (Dänisch, Holländisch, Französisch, Russisch, Polnisch und Böhmisch). Sammlung der nothwendigsten Wörter und Ausdrücke mit Wiedergabe ihrer Aussprache in deutscher Schrift*, Lipsk 1878;
 - **kurdyjski**: August Kościeszka-Żaba, *Dictionnaire kurde-français*, St.-Petersbourg 1879;
 - **portugalski**: Ludwik Włodek, *Ilustrowany przewodnik po Brazylii wraz ze słowniczkiem polsko-portugalskim*, Kraków 1909;
 - **jidis**: M. Wolfson, *Słownik polsko-żydowski*, Warszawa ca 1920;
 - **bantu**: Stanisław Hankiewicz, *Słownik polsko-cinyanja [cziniańdża, język używany przez Murzynów w Rodezji]*, (rps) ca 1928;
 - **karaimski**: Aleksander Mardkowicz, *Słownictwo karaimskie. Karaimsko-polsko-niemiecki słownik*, Łuck 1933;
 - **kipczacki**: Ananiasz Zajączkowski, *Słownik arabsko-kipczacki z okresu Państwa Mameluckiego*, Warszawa 1954;
 - **fiński**: Józef Trypućko, *Mały słownik polsko-fiński*, Warszawa 1962;
 - **baskijski**: Alfred F. Majewicz, *Słownik baskijsko-polski*, Poznań 1969 (nieukończony);
 - **irlandzki**: Aidan Doyle, Edmund Gussmann, *An ghaeilge: podręcznik do nauki języka irlandzkiego*, Lublin 1991 [zawiera słownik irlandzko-polski i polsko-irlandzki].

Usystematyzowane językowo, chronologicznie i/lub dziedzinowo (według zagadnień/tematów) bibliografie rejestrujące różnego typu opracowania słownikarskie, indeksy, glosaria itp., wydane w Polsce lub za granicą (których (współ)autorami i/lub

(współ)redaktorami są Polacy), stanowią ważne, a zarazem miarodajne narzędzie opisu rozwoju historycznego leksykografii rodzimej, a także obiektywne narzędzie przeprowadzenia wieloaspektowych ocen jej obecnego stanu i dorobku zarówno w aspekcie kwantytatywnym (np. dynamiki opracowania słowników, stopnia intensywności ruchu wydawniczego, częstotliwości wznowień i aktualizacji konkretnych pozycji słownikarskich itd.), jak i kwalitatywnym (m.in. ewaluacji tendencji dominujących w językoznawstwie aplikatywnym, zapotrzebowania na dane typy słowników, ukierunkowania na języki wyjściowe i docelowe w słowniku, analiz materiałowych opracowań leksykograficznych itd.). Bibliografie słowników stanowią ważne źródło danych statystycznych oraz źródło danych faktograficznych do analiz językoznawczych (w ramach np. kontaktologii językowej). Uzyskane dane kwantyfikacyjne dotyczące liczby poszczególnych typów opracowań leksykograficznych są wykorzystywane do oceny ogólnego i szczegółowego rozwoju słownikarstwa zarówno współczesnego, jak i historycznego (Wojan 2013a,b,c,d; 2014a,b; 2016a; 2017a). Na podstawie przeprowadzonych empirycznych analiz danych bibliometrycznych opisano syntetycznie m.in. słownikarstwo polskie z językiem angielskim okresu lat 1782–2012 (Wojan 2013b), zmiany w najnowszej leksykografii polskiej (Piotrowski 2016), jak też terminografię polską (Łukasik 2007; 2014; 2017).

Bibliografie stanowią również punkt wyjścia do tworzenia podstaw naukowych komparatystyki leksykograficznej, która pozostaje jeszcze w bardzo wczesnym stadium rozwoju.

Podsumowanie

Bibliografie słowników są użytecznym, miarodajnym i efektywnym narzędziem badawczym służącym empirycznym analizom bibliometrycznym prowadzonym na dużym materiale faktograficznym. Stanowią dobrą bazę materiałową, dokumentacyjną, źródłową, źródłoznaczącą. Wymagają one dalszych pogłębionych opracowań naukowych, aktualizacji i weryfikacji. Zastosowana metodyka uporządkowania i deskrypcji materiału faktograficznego daje możliwość zaprezentowania, po pierwsze: bogactwa dziedzictwa kulturowego i dorobku intelektualnego polskiego narodu, po drugie: wkładu do dorobku światowego uczonych i działaczy polonijnych oraz mających polskie korzenie.

Bibliografia

- BAJOR Agnieszka (2003): *Monografia bibliograficzna: analiza modelowa*. Częstochowa: –Wydawnictwo „E Media”.
- BAJOR Agnieszka (2005): *Z zagadnień form bibliografii: monografia bibliograficzna*. –Warszawa: Wydawnictwo SBP.
- COLLISON Robert Lewis (1955): *Dictionaries of foreign languages: A bibliographical guide to the general and technical dictionaries of the chief foreign languages, with historical and explanatory notes and references*. – London: The Hafner Publishing Company.

- COLLISON Robert Lewis (1971): *Dictionaries of English and Foreign Languages: A Bibliographical Guide to Both General and Technical Dictionaries with Historical and Explanatory Notes and References*. 2nd ed. [rev. and enlarged]. – New York: The Hafner Publishing Company.
- CZACHOWSKA Jadwiga, LOTH Roman (1989): *Przewodnik polonisty. Bibliografie, słowniki, biblioteki, muzea literackie*. Wyd. 3 zm. Seria Vademecum Polonisty. – Wrocław: Zakład Narodowy im. Ossolińskich.
- FRĄCZEK Agnieszka (2010): *Słowniki polsko-niemiecki i niemiecko-polskie z przełomu XVII i XVIII wieku: analiza leksykograficzna*. – Warszawa: Universitas Varsoviensis.
- FRĄCZEK Agnieszka, LIPCZUK Ryszard (2004): *Słowniki polsko-niemieckie i niemiecko-polskie: historia i teraźniejszość*. – Wołczkowo: In Plus.
- GMITEREK Grzegorz (2015): III Ogólnopolska Konferencja Naukowa „Bibliografia – teoria, praktyka, dydaktyka”, Warszawa, 17–18 XI 2015 r. – *Z badań nad książką i księgozbiorami historycznymi* 9: 395–399. – URL: <http://www.lis.uw.edu.pl/badan/?p=252> (dostęp: 31.07.2017).
- GRUSZKA Zbigniew (2012): *Przegląd Biblioteczny: monografia*. Seria: Nauka – Dydaktyka – Praktyka. – Warszawa: Stowarzyszenie Bibliotekarzy Polskich.
- GRZEGORCZYK Piotr (1967): *Index lexicorum Poloniae: bibliografia słowników polskich*. – Warszawa: Państwowe Wydawnictwo Naukowe.
- Informator Sekcji Bibliotek Niepaństwowych Szkół Wyższych Zarządu Głównego Stowarzyszenia Bibliotekarzy Polskich* 8, z. 5 (210): poniedziałek, 4 lutego 2013 roku. – URL: http://biblioteka.wsb-nlu.edu.pl/uploadedFiles/file/ISBNik_2013_02_04.pdf (dostęp: 10.11.2017).
- JANECKI Łukasz (2016): *Słownik biograficzno-bibliograficzny polskich franciszkanów konwentualnych zamordowanych i tragicznie zmarłych w latach 1939–1945*. – Niepokalanów: Wydawnictwo Ojców Franciszkanów.
- JONES William Jervis (2000): *German Lexicography in the European Context: A descriptive Bibliography of Printed Dictionaries and Word Lists Containing German language 1600–1700*. – Berlin–New York: Walter de Gruyter.
- KĘDELSKA Elżbieta (1986): *Łacińsko-polskie słowniki drukowane pierwszej połowy XVI wieku i ich stosunek do źródeł czeskich*. – Wrocław: Zakład Narodowy im. Ossolińskich.
- KĘDELSKA Elżbieta (1995): *Studia nad łacińsko-polską leksykografią drugiej połowy XVI wieku*. – Warszawa: Slawistyczny Ośrodek Wydawniczy przy Instytucie Slawistyki PAN.
- KIETLICZ-WOJNACKI Waław (1980): *Polskie osiągnięcia naukowe na obczyźnie: od średniowiecza do II wojny światowej*. – Lublin: Wydawnictwo Lubelskie.
- KOŁODZIEJCZYK Edmund (1911): *Bibliografia sławianoznawstwa polskiego*. – Kraków: Akademia Umiejętności.
- KONDRATOWICZ Cyriak (1775): *Польскій общій словарь и библейный, с польскою, латинскою и русскою новоисправленною библиями смечиван...* – Petersburg.
- KONEFAŁ Ewa (2016): *Przeładoznawstwo rosyjskie*, t. 1: *Autoreferaty dysertacji 1937–2015*. – Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.
- KORPIEWICZ (KORPIEWSKI) Eliasz (1700): *Номенклаторъ, на латинскомъ, рускомъ и немѣцкомъ языкъ (Nomenclator trium linguarum, Latine, Russice et Germanice)*. – Amstelodami.
- KUCHARSKI Andrzej F. (1825): *Księżda Stanisława Zaborowskiego Ortografija polska: z łacińskiego na polski język przełożona z przydaniem uwag tłomacza, tudzież Ortografii Seklucjana i spisu bibliograficznego grammatyki i słowników polskich*. – Warszawa: Drukarnia Józefa Węckiego.
- KWIATKOWSKA Hanna (1977): *Bibliografie, encyklopedie i słowniki polskie: [wybór bibliograficzny]*. – Warszawa: Biblioteka Narodowa.
- LEWANSKI Richard Casimir (1959): *A Bibliography of Polish Dictionaries with a Supplement of Lusatian and Polabian Dictionaries*. – New York: The New York Public Library.
- LEWANSKI Richard Casimir (1963): *A Bibliography of Slavic Dictionaries. Vol. 3: Russian*. – New York: New York Public Library.

- LEWANSKI Richard Casimir (1964): *Bibliography of Dictionaries in the Field of Library Science and Related Subjects*. – Paris: [s.n.].
- LEWANSKI Richard Casimir (1972): *A Bibliography of Slavic Dictionaries*. Vol. 3: *Russian*. The Johns Hopkins University Bologna Center Library Publications, 7. World Bibliography of Dictionaries. 2nd rev. and enl. ed. – Bologna: Editrice Compositori.
- LEWANSKI Richard Casimir (1973a): *A Bibliography of Slavic Dictionaries*. Vol. 1: *Polish*. The Johns Hopkins University Bologna Center Library Publications, 4. World Bibliography of Dictionaries. 2nd rev. and enl. ed. – Bologna: Editrice Compositori.
- LEWANSKI Richard Casimir (1973b): *A Bibliography of Slavic Dictionaries*. Vol. 2: *Belorussian, Bulgarian, Czech, Kashubian, Lusatian, Old Church Slavic, Macedonian, Polabian, Serbocroatian, Slovak, Slovenian, Ukrainian*. The Johns Hopkins University Bologna Center Library Publications, 5. World Bibliography of Dictionaries. 2nd rev. and enl. ed. – Bologna: Editrice Compositori.
- LEWANSKI Richard Casimir (1973c): *A Bibliography of Slavic Dictionaries*. Vol. 4: *Supplement*. The Johns Hopkins University Bologna Center Library Publications, 7. World Bibliography of Dictionaries. – Bologna: Editrice Compositori.
- ŁUKASIK Marek (2007): *Angielsko-polskie i polsko-angielskie słowniki specjalistyczne (1990–2006): analiza terminograficzna*. – Warszawa: Katedra Języków Specjalnych Uniwersytetu Warszawskiego.
- ŁUKASIK Marek (2014): *Najnowsze słowniki specjalistyczne z terminologią angielską i polską. Przegląd bibliograficzny w ramach analizy terminograficznej za lata 2007–2013*. – [w:] Marek ŁUKASIK, Beata MIKOŁAJEWSKA (red.): *Języki specjalistyczne: wczoraj, dziś i jutro*. Studi@ Naukowe 17. – Warszawa: Wydawnictwo Naukowe Instytutu Kulturologii i Lingwistyki Antropocentrycznej Uniwersytetu Warszawskiego, 261–277. – URL: <http://www.sn.iksi.uw.edu.pl/documents/7732735/0/SN+17+Marek+%C5%81lukasik%2C%20Beata+Miko%C5%82ajewska+%28red.%29%2C%20J%C4%99zyki+specjalistyczne+wczoraj%2C%20dzi%C5%9B%20i+jutro.pdf> (dostęp: 3.10.2017).
- ŁUKASIK Marek (2017): *Polska szkoła terminograficzna*. – [w:] Sambor GRUCZA, Marta MAŁACHOWICZ (red.): *Polskie i europejskie nurty terminologiczne*. – Warszawa: Wydawnictwo Naukowe Instytutu Komunikacji Specjalistycznej i Interkulturowej Uniwersytetu Warszawskiego, 26–69. – URL: http://www.academia.edu/31262986/Pre-Print_Polska_szko%C5%82a_terminograficzna_Polish_school_of_terminography (dostęp: 3.12.2017).
- MAKUŁSKI Franciszek Jaxyc (1795): *Łatwy sposób nauczania się po rosyjsku i po polsku czytać i pisać... z przydaniem... słów pod alfabetem co potrzebniejszych a używanych najczęściej*. – Warszawa: w Drukarni P. Zawadzkiego.
- MATCZUK Alicja (2014): *Polskie bibliografie nauk humanistycznych i społecznych do roku 1989: historia i metodyka*. – Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.
- PIOTROWSKI Tadeusz (2001a): *Słowniki angielsko-polskie i polsko-angielskie w dobie najnowszej. Próba analizy bibliograficznej*. – [w:] Włodzimierz GRUSZCZYŃSKI et al. (red.): *Nie bez znaczenia... Prace ofiarowane Profesorowi Zygmuntowi Saloniemu z okazji jubileuszu 15 000 dni pracy naukowej*. – Białystok: Wydawnictwo Uniwersytetu w Białymstoku, 203–210.
- PIOTROWSKI Tadeusz (2001b): *Zrozumieć leksykografię*. – Warszawa: Wydawnictwo Naukowe PWN.
- PIOTROWSKI Tadeusz (2016): *Zmiany w najnowszej leksykografii polskiej*. – [w:] Monika ZAŚKO-ZIELIŃSKA, Małgorzata MISIAK, Jan KAMIENIECKI, Tomasz PIEKOT (red.): *Lingua et gaudium. Księga jubileuszowa ofiarowana Profesorowi Janowi Miodkowi*. – Wrocław: Oficyna Wydawnicza Atut, 431–438.
- PODHAJECKA Mirosława (2016): *A history of Polish-English, English-Polish bilingual lexicography (1788–1947)*. – Opole: Wydawnictwo Uniwersytetu Opolskiego.
- STANKIEWICZ Edward (1984): *Grammars and Dictionaries of the Slavic Languages from the Middle Ages up to 1850: an Annotated Bibliography*. – Berlin: Mouton Publishers.

- SZUPER-JAKUBIUK Danuta (2010): Bibliografia słowników przekładowych handlowych i okołohandlowych: (język polski, rosyjski, ukraiński): za lata 1990–2010 w układzie chronologicznym. – *Linguo-didactica* 14, 243–250.
- VRTEL-WIERCZYŃSKI Stefan (1928): O najpilniejszych zadaniach naszej bibliografii. – *Przegląd Biblioteczny* 3: 191–209.
- WAWRZYŃCZYK Jan (1994): *Przewodnik bibliograficzno-leksykalny po współczesnym językoznawstwie rosyjskim*. – Toruń: Uniwersytet Mikołaja Kopernika.
- WAWRZYŃCZYK Jan (1998): *Mały przewodnik leksykalno-bibliograficzny po współczesnym językoznawstwie ogólnym i polonistycznym*. – Warszawa: Katedra Lingwistyki Formalnej Uniwersytetu Warszawskiego.
- WAWRZYŃCZYK Jan (2001): *Słownik bibliograficzny językoznawstwa ogólnego i polonistycznego*. – Warszawa: Dialog.
- WAWRZYŃCZYK Jan (2002): *Współczesne językoznawstwo rosyjskie, polskie, ogólne. Słownik bibliograficzny*, z. 1–2. – Łódź–Warszawa: Katedra Literatury i Kultury Rosyjskiej Uniwersytetu Łódzkiego.
- WAWRZYŃCZYK Jan (2005): *Słownik bibliograficzny językoznawstwa ogólnego i polonistycznego*. – Warszawa: Wydawnictwo Akademickie Dialog.
- WAWRZYŃCZYK Jan, MAŁEK Eliza (2004): Z materiałów do słownika bibliograficznego języka rosyjskiego. Terminologia lingwistyczna, wybrane terminy wiedzy o kulturze i literaturze, neologizmy, hapaks legomena. – Warszawa: Semiosis Lexicographica 16.
- WAWRZYŃCZYK Jan, MAŁEK Eliza (2014–2015): *Русская лексика. Библиографические характеристики* 1–10. – Warszawa: Mila Hoshi.
- WOJAN Katarzyna (2006): *Język rosyjski w polskiej leksykografii przekładowej. Bibliografia słowników za lata 1795–2005*. – Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.
- WOJAN Katarzyna (2013a): Esperanto w leksykografii polskiej. Bibliografia słowników (1888–2012). – *Język. Komunikacja. Informacja* 8, 133–149.
- WOJAN Katarzyna (2013b): *Język angielski w polskiej leksykografii*, t. 1: *Słowniki przekładowe lingwistyczne i encyklopedyczne wydane w latach 1782–2012*. – Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.
- WOJAN Katarzyna (2013c): *Język białoruski w polskiej leksykografii polskiej. Bibliografia słowników (za lata 1807–2012)*. – *Acta Albaruthenica* 13, 219–233.
- WOJAN Katarzyna (2013d): *Z polsko-łacińskiej leksykografii przekładowej, cz. I: Bibliografia słowników z językiem łacińskim za lata 1964–1989 w układzie chronologicznym*. – *Symbolae Philologorum Posnaniensium Graecae et Latinae* 1 (23), 161–191.
- WOJAN Katarzyna (2014a): *Język angielski w polskiej leksykografii*, t. 2: *Słowniki przekładowe terminologiczne wydane w latach 1782–2012*. – Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.
- WOJAN Katarzyna (2014b): *Z polsko-łacińskiej leksykografii przekładowej, cz. II: Bibliografia słowników z językiem łacińskim za lata 1990–2011 w układzie chronologicznym*. – *Symbolae Philologorum Posnaniensium Graecae et Latinae* 1 (24), 227–268.
- WOJAN Katarzyna (2015): *Język potoczny w polskiej lingwistyce. Materiały do Przewodnika leksykalno-bibliograficznego*. – [w:] Marcelina GRABSKA, Żanna ŚLADKIEWICZ (red.): „*Mówimy jak mówimy... Gdzie ukryta jest potoczność?*” – Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego, 43–83.
- WOJAN Katarzyna (2016a): Wkład Polaków do dorobku światowego słownictwa (z językiem rosyjskim w tle). – *Studia Rossica Gedanensia* 3, 167–200.
- WOJAN Katarzyna (2016b): *Z dziejów leksykografii polsko-rosyjskiej*, t. 1: *Słowniki lingwistyczne (bibliografia za lata 1700–2015)*. – Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.
- WOJAN Katarzyna (2016c): *Leksykografia polsko-fińska*. – [w:] Katarzyna WOJAN: *Język fiński w teorii i praktyce*. – Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego, 250–257.
- WOJAN Katarzyna (2017a): *Leksykografia polska z językiem ukraińskim w opisie statystycznym: z bibliografią za lata 1794–2016*. – *Slavia Orientalis* 1 (LXVI), 171–194.

- WOJAN Katarzyna (2017b): Rosyjskie piśmiennictwo naukowe z zakresu przekładoznawstwa. Rzecz o monografii bibliograficznej Ewy Konefał pt. „Przekładoznawstwo rosyjskie, t 1: Autoreferaty dysertacji 1937–2015” (Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk 2016). – *Acta Polono-Ruthenica* 3, 147–160.
- ZAMENHOF Ludwik (1889): *An Attempt Towards An International Language*. By Dr. Esperanto (Warsaw, Russia). Translated by PHILLIPS, Henry, Jr. A Secretary of the American Philosophical Society. *Together with an International-English Vocabulary* compiled by The Translator. – New York: Holt Henry & Co.
- ŻMIGRODZKI Piotr (2005): *Wprowadzenie do leksykografii polskiej*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego.

doi: <http://dx.doi.org/10.17651/bptj.74.11>

MONIKA BIESAGA

Instytut Języka Polskiego PAN, Kraków

Poprawność polityczno-kulturowa w słowniku. Pola tematyczne a typy mikrostrukturalnych informacji leksykograficznych

Streszczenie

W artykule omówiony został problem poprawności polityczno-kulturowej w słowniku. Istotność zagadnienia wzrasta wraz z rozwojem leksykografii internetowej i komputerowej, umożliwiającej łatwe prezentowanie onomazjologicznego porządku haseł, w tym również grup znaczeń potencjalnie wrażliwych kulturowo. Stąd w niniejszym artykule wskazane zostaną te pola tematyczne leksyki, które są w szczególności związane z poprawnością polityczno-kulturową. Następnie omówionych zostanie kilka wybranych przykładów (odnoszących się do różnych elementów mikrostruktury słownika: definicji, cytatów, kolokacji, ilustracji), które pozwolą praktycznie zanalizować ten problem.

Słowa kluczowe: poprawność polityczna, leksykografia, onomazjologia, pole tematyczne, pole semantyczne, pole leksykalne, leksykografia internetowa

Political and cultural correctness in dictionaries. Thematic fields and types of microstructural lexicographic information

Abstract

The aim of this paper is to address the problem of political and cultural correctness in dictionaries. Its significance increases in parallel to the development of electronic lexicography which facilitates the presentation of the onomasiological ordering of dictionary entries, including meanings that are potentially culture-sensitive. The article focuses on those thematic fields that are related to political and cultural correctness. A discussion of selected examples referring to various elements of the dictionary microstructure, such as definitions, citations, collocations, and illustrations, will serve as an illustration of the analysed issues.

Keywords: political correctness, lexicography, onomasiology, thematic field, semantic field, lexical field, Internet lexicography

Rewolucja internetowa a makrostruktura słownika

Od kilkunastu lat, za sprawą rozwoju językoznawstwa komputerowego oraz zwrotu ku pragmatycznym aspektom poszczególnych nauk, także lingwistyki, sztuka tworzenia oraz kompilowania słowników przeżywa swój prawdziwy renesans. Jednym z bodaj najbardziej kluczowych udogodnień współczesnej leksykografii, internetowej i elektronicznej, jest przewyższenie ograniczeń związanych z ilością informacji. Na naszych oczach kruszy się klasyczna typologia słowników (opracowania ogólne, frazeologiczne, gramatyczne czy etymologiczne). Czytelnik internetowy oczekuje bowiem od autorów dzieła jak najbardziej wszechstronnego. Choć nadmiar informacji może być utrudnieniem dla odbiorców (por. Gouws, Tarp 2016), zaś umiejscowienie poszczególnych typów danych na stronie jest i przez dłuższy czas będzie przedmiotem licznych badań eksperymentalnych, samą tendencję uznać należy za pozytywną, a nawet rewolucjonizującą praktykę słownikową. Co więcej, językoznawstwo komputerowe pozwala autorom dowolnie żonglować poszczególnymi rodzajami danych. Dzięki temu odbiorca może zapoznać się nie tylko z interesującym go hasłem, lecz przejść również do haseł w nietrywialny sposób z nim powiązanych (słowotwórczo, pragmatycznie czy tematycznie, por. Biesaga 2017).

Porządek onomazjologiczny w semantyce i leksykografii

Właśnie nad całościowym podziałem tematycznym znaczeń w odniesieniu do kulturowej zawartości haseł, szczególnie w świetle poprawności politycznej, zamierzam w niniejszym tekście się skupić.

Od wielu stuleci leksykografia rozróżnia dwa makrostrukturalne typy uporządkowania artykułów hasłowych. Pierwszy typ, alfabetyczny, opiera się na następstwie haseł zgodnym z narastającymi literami alfabetu. Jest to porządek najczęstszy, intuicyjny i najbardziej w leksykografii skonwencjonalizowany. Przeciwstawia się mu typ drugi, semantyczny, zwany również onomazjologicznym, konceptualnym, ideograficznym (por. Burkhanov 2010: 20, 103–105; Hartmann, James 2001: 5, 101–102, Piotrowski 2001: 37). Tutaj hasła prezentowane są użytkownikowi zgodnie z ich przynależnością do grupy powiązanych ze sobą desygnatów, tworzących całości pojęciowe, tym samym umożliwiając kształtowanie się w umyśle schematów kognitywnych, w praktyce komunikacyjnej zaś – realizację i rozumienie wypowiedzi na dany temat. Choć z psychologicznego punktu widzenia porządek niealfabetyczny wydaje się bardziej naturalny (por. Mikołajczak-Matyja 2005: 17–18), ze względu na rozmycie granic całości pojęciowych i szereg innych problemów semantycznych (polisemia, podwójna przynależność pojęciowa rozlicznych sensów itd.), był on stosunkowo rzadko stosowany. W wielu słownikach o przynależności tematycznej informowały pośrednio kwalifikatory specjalistyczne (Batko-Tokarz 2008: 32). Jak natomiast wnioskować można na podstawie rozwoju słowników internetowych i elektronicznych, płynne przechodzenie użytkownika pomiędzy porządkami alfabetycz-

nym i onomazjologicznym zyskuje i będzie zyskiwało na popularności i znaczeniu (por. Piotrowski 2001: 56). Czytelnicy mogą dzięki temu udogodnieniu poznać znaczenia słów z tej samej grupy semantycznej, wzbogacać słownictwo i wiedzę o świecie, leksykografowie zaś – utrzymać uwagę odbiorców przez czas dłuższy niż związany z czytaniem wyłącznie jednego hasła.

Porządku semantycznego nie da się zatem ignorować dłużej w leksykografii skierowanej do szerokiego grona czytelników. Rodzi się pytanie, jaki model uporządkowania leksykonu wybrać jako podstawę grupowania znaczeń, mających następnie w określonej wiązce ideograficznej przyciągnąć na dłużej uwagę czytelnika.

Co ciekawe, na rynku leksykograficznym znajdziemy sporo słowników semantycznych (por. Rembiszewska 2005). Niezwykle rzadko jednak ambicją ich autorów jest globalny i wyczerpujący podział sensów leksyki na poszczególne grupy. Zazwyczaj oddają one onomazjologiczne uporządkowanie słownictwa tylko w odniesieniu do wybranych grup semantycznych, co wynika z celu danego słownika (np. opracowania nieogólne, specjalistyczne mają wyłącznie charakter dyferencjalny; pomoce glottodydaktyczne zogniskowane są wokół nauczania leksyki podstawowej itd.). Tymczasem stworzenie, choćby w przybliżeniu, pełnego podziału pojęciowego słownictwa umożliwi pełniejszą prezentację leksykonu (korzyści leksykologiczne, semantyczne, kognitywne, socjolingwistyczne itd.).

Prób całościowego opisu ideograficznego słownictwa podejmowano dotychczas relatywnie niewiele¹. W lingwistyce europejskiej i amerykańskiej najbardziej znany jest podział stworzony przez Petera Marka Rogeta w 1852 r. (stosowany do dzisiaj w zaktualizowanej wersji w słownikach onomazjologicznych² oraz na przykład w Wikipedii do porządkowania sensów i ilustracji graficznych). W 1933 r. Franz Dornseiff opublikował dla języka niemieckiego pracę *Der deutsche Wortschatz nach Sachgruppen* (aktualizowaną i wznawianą do dziś, por. Dornseiff 2004). W r. 1952 Julio Casares wydał po pięciu latach pracy nad leksyką hiszpańską *El Diccionario ideológico de la lengua española*. W tym samym roku Rudolf Hallig oraz Walter von Wartburg ogłosili drukiem swoją pracę *Begriffssystem als Grundlage für die Lexikographie*. Interesująco przedstawiają się również polskie propozycje. W 1992 r. Andrzej Markowski wraz ze współpracownikami opracował podział nieco okrojony, dotyczący leksyki wspólnoodmianowej. Został on przedstawiony w pracy *Leksyka wspólna różnym odmianom polszczyzny* (Markowski 1992). Z ostatnich lat pochodzi propozycja Barbary Batko-Tokarz, stworzona na potrzeby podziału tematycznego słownictwa w jednym ze słowników współczesnej polszczyzny (Batko-Tokarz 2008). Za każdym razem istotą semantycznego podziału całości znaczeń leksykonu jest wyznaczenie pewnych grup głównych, makrogrup ideograficznych. Makropola rozpadają się natomiast na mniejsze pola. Liczba poziomów waha się średnio w poszcze-

¹ Warto w tym miejscu wspomnieć o nieco odmiennym projekcie, jakim jest WordNet. Jego celem jest stworzenie, w oparciu o różnorakie metody informatyczne (analizy korpusów), sieci psycholingwistycznych powiązań między jednostkami leksykonu (por. Miller, Beckwith, Fellbaum, Gross, Miller 1993: 1).

² Spolszczoną wersją dzieła Rogeta jest pierwszy polski słownik onomazjologiczny: *Dobór wyrazów* Romana Zawilińskiego z 1926 r. (Zawiliński 1985).

gólnych opracowaniach od dwóch do pięciu, sześciu (porównanie w: Mikołajczak-Matya 2005).

Jak zaznaczają teoretycy leksykografii (por. Burkhanov 2010: 105–106), w przypadku słowników ideograficznych mówić można o trzech sposobach organizacji danych. Uporządkowanie tezaursowe opiera się na stworzeniu apriorycznej siatki pól pojęciowych. Typ drugi – tematyczny, związany z klasyfikacją desygnatów, służyć ma przede wszystkim nauczaniu języków obcych. Natomiast typ trzeci – systematyczny – zasadza się na semantycznym porządkowaniu znaczeń leksemów reprezentujących tę samą część mowy. Bliższe przyjrzenie się wymienionym wcześniej klasyfikacjom onomazjologicznym prowadzi do wniosku, że podział na tezaury, opracowania tematyczne i systematyczne jest pewnego rodzaju konstruktem teoretycznym. W istocie autorzy tworzą swego rodzaju hybrydy, oparte na, w mniejszym lub większym stopniu, apriorycznych kategoriach (w części opracowań kognitywnie intuicyjnych, zatem bliskich podziałom naturalnym), skonfrontowanych następnie z autentycznym materiałem leksykalnym w całym jego bogactwie.

Na potrzeby niniejszego tekstu zastosowana zostanie klasyfikacja tematyczna słownictwa Barbary Batko-Tokarz (2008). Jest to koncepcja, w której duże klasy tworzą kontinuum, uwzględniająca problem słownictwa atematycznego (wyodrębnionego w osobną grupę) oraz kwestię mieszania się perspektywy potocznej i naukowej. Ponadto całość po kilku latach stosowania w odniesieniu do leksyki współczesnej polszczyzny została lekko zmodyfikowana i obecnie zgodna jest w dość znacznym stopniu z charakterystyką leksykonu dzisiejszego języka polskiego.

Poprawność polityczna w kulturze i leksykografii

Zanim skonfrontujemy wspomnianą klasyfikację tematyczną z problemem poprawności polityczno-kulturowej w słowniku, warto przeanalizować, czym jest poprawność polityczna. W definicji pochodzącej z *Encyklopedii PWN* (2017) czytamy:

poprawność polityczna, ang. political correctness, zasada unikania określeń, które mogłyby być uznane za przejaw dyskryminacji w stosunku do osób ze względu na ich przynależność rasową, narodową, wyznaniową, pochodzenie społeczne lub orientację seksualną.

Wnioskując na podstawie debat toczących się w mediach oraz literatury politologicznej (por. Szahaj 2010), zaproponowaną wyżej definicję uznać należy za zbyt wąską. W dyskursie związanym z poprawnością polityczną mowa bowiem również o określeniach kobiet, osób niepełnosprawnych, starszych czy wyróżniających się wyglądem. Kluczem do właściwego postrzegania problemu poprawności polityczno-kulturowej w debacie publicznej jest zatem kwestia tożsamości człowieka i ogółu cech, które tę tożsamość budują. Właściwy sposób wypowiadania się byłby zatem związany z szacunkiem wyrażanym innym (Szahaj 2010: 174).

Jest więc poprawność polityczna określonego rodzaju filtrem nakładanym na język. Wydaje się, że tego rodzaju współczesne działania przypominają, w pewnym

stopniu, charakterystyczną dla dawniejszych słowników, postawę preskryptywną przy opisywaniu leksykonu (np. w SJPD, por. tamże: XVI; Piotrowski 2001: 90). Wtedy tendencje normatywistyczne uzasadniane były w pierwszym rzędzie kształtowaniem języka wpływającego na kulturę (propagowanie właściwej stylistyki wypowiedzi zaowocuje doskonaleniem systemu). U podstaw obu tendencji, działania w zgodzie z poprawnością polityczną oraz preskryptywizmu leksykograficznego, znajduje się to samo przekonanie. Pomiędzy językiem i stosunkiem człowiek–rzeczywistość zachodzi relacja wzajemności. Język kształtuje się w oparciu o rzeczywistość zantropologizowaną, może on jednak wpływać również na przemiany jej postrzegania (Rudnicka 2016: 51). Istotą poprawności politycznej jest zatem przyjęcie pewnego rodzaju perspektywy społeczeństwa idealnego, innego niż wynikałoby to z dotychczasowego doboru środków językowych, stojących w niektórych wypowiedziach w sprzeczności z szacunkiem okazywanym jednostce ze względu na jej tożsamość.

Pod względem semantycznym realizacja postulatów poprawności polityczno-kulturowej związana byłaby z ze świadomym podejściem leksykografa do problemu prezentacji cech denotacyjnych i, w szczególności, konotacyjnych. W perspektywie teorii rzemiosła leksykograficznego poprawność obyczajowa wymusza na autorze przyznanie, że granica pomiędzy słownikiem (opisującym język) i encyklopedią (opisującą świat) jest płynna. Słownik, opisując język, opisuje świat.

Treści kulturowe w słownikach

Ślady rozważań metaleksykograficznych, skupionych wokół tematu kultura–słownik, odnajdziemy w dość licznych pracach badawczych. Dotychczas w lingwistyce polskiej w szczególności opisywany był warsztat słownikarza jako niewielka część rzeczywistości kulturowej, otaczającej opracowanie referencyjne (Rudnicka 2016: 50). Napotkać można było również rozważania analogiczne do teoretyczno-literackich (jak rzeczywistość polityczna czy historyczna wpłynęła na dzieło, por. rozważania o SJPD, np. Piotrowski 2001: 90). Stosunkowo rzadko zajmowano się problemem wpływu czynników obyczajowych na ukształtowanie treści opracowań referencyjnych. Tymczasem słowniki są niezwykle nośnym świadectwem kultury czasów, w których powstały.

Gdyby zapytać, jaka część artykułu hasłowego w słowniku podlega wpływom rzeczywistości kulturowo-obyczajowej, z powodzeniem odpowiedzieć można, że w zasadzie każda (np. siatka haseł, relacje semantyczne). Im większy słownik, z bardziej różnorodnymi typami danych, tym świadomość, podejrzliwość i ostrożność kulturowa leksykografa powinny być większe. Wskazać można jednak typy danych szczególnie wrażliwe na treści związane z poprawnością polityczną, np. cytaty, definicje, dobór kolokacji czy ilustrację graficzną. To właśnie przykłady dotyczące tych czterech typów danych słownikowych staną się podstawą materiałową niniejszych rozważań.

Pola związane z poprawnością polityczno-kulturową

Po scharakteryzowaniu zagadnienia poprawności politycznej w ogóle, natomiast przed przejściem do omówienia wybranych przykładów, warto wskazać, które z grup onomazjologicznych są szczególnie wrażliwe na wymogi poprawności polityczno-kulturowej w słowniku. W świetle opracowań metaleksykograficznych (Lorentzen, Trap-Jensen 2016: 139–142) najszybciej dezaktualizują się hasła związane z prawem wpływającym na życie człowieka (np. edukacja, układy społeczne), systemami monetarnymi, technologią, nazwami własnymi, markami handlowymi, taksonomiami biologicznymi oraz właśnie poprawnością polityczną. Wśród obszarów semantycznych, zależnych od kultury (kategoria szersza od poprawności politycznej), wymienia się w teorii leksykografii: ekonomię, prawo, politykę, znaczenia związane z życiem codziennym (gotowanie, czynności codzienne, wychowywanie dzieci, prowadzenie dyskusji), por. Bergenholz, Nielsen 2013: 469. Wszystkie wymienione pola, dezaktualizujące się i zależne od kultury, wskazywane były przez autorów częściowo apriorycznie, częściowo zaś indukcyjnie – na podstawie uprzednich warsztatowych doświadczeń leksykograficznych. Spojrzenie na problem polityczno-kulturowej poprawności z uwzględnieniem całościowego ideograficznego podziału leksykonu umożliwi wskazanie większej liczby hipotetycznie wrażliwych oraz, z drugiej strony, neutralnych pól tematycznych, sprzężonych bądź nie z wartościowaniem tożsamości człowieka.

Jeżeli weźmiemy pod uwagę, wzmiankowany wyżej, podział tematyczny leksykonu autorstwa B. Batko-Tokarz (2008) oraz, wyliczone wcześniej, obszary poprawności polityczno-kulturowej (por. Szahaj 2010), dojdziemy do wniosku, że w zasadzie niemal całe makropole CZŁOWIEK JAKO ISTOTA FIZYCZNA wydaje się objęte tego rodzaju zastrzeżeniami (zwłaszcza podpola BUDOWA I FUNKCJONOWANIE CIAŁA LUDZKIEGO [w tym seksualność], BIEG ŻYCIA, OKREŚLENIA FIZYCZNOŚCI CZŁOWIEKA). W szczególności zwrócić należy uwagę na sensy wartościujące, odnoszące się do cech fizycznych. Nieco mniej drażliwe wydają się znaczenia związane z PIELĘGNACJĄ CIAŁA.

W makropolu CZŁOWIEK JAKO ISTOTA PSYCHICZNA zwrócić należy uwagę na następujące pola, powiązane z subiektywnością wartościowania, która powinna być na przykład w definicji lub za pomocą innych odpowiednich środków podkreślana: STANY PSYCHICZNE CZŁOWIEKA, USPOSOBIENIE CZŁOWIEKA, OCENA I WARTOŚCIOWANIE, RELACJE MIĘDZYLUDZKIE. W mniejszym stopniu zanurzone w problemach poprawności polityczno-kulturowej wydają się obszary: DZIAŁALNOŚĆ INTELEKTUALNA CZŁOWIEKA i DZIAŁALNOŚĆ ARTYSTYCZNA CZŁOWIEKA, choć i tu sensy związane z ocenianiem winny być odzwierciedlane w sposób ostrożny.

W makropolu CODZIENNE ŻYCIE CZŁOWIEKA uwaga powinna być zachowana w przypadku opisu niektórych znaczeń z podpola RODZINA, w pewnym stopniu PRACA. Dalece mniej wrażliwe wydają się sensy odnoszące się do: NAJBLIŻSZEGO ŚRODOWISKA ŻYCIA CZŁOWIEKA, UBRANIA, JEDZENIA, CZASU WOLNEGO oraz SPORTU (naturalnie poza sensami wartościującymi).

W obszarze CZŁOWIEK W SPOŁECZEŃSTWIE warto zwrócić uwagę na znaczenia związane z polami: ZASADY WSPÓŁŻYCIA SPOŁECZNEGO, PRZYNALEŻNOŚĆ I PODZIAŁ TERYTORIALNY, PRAWO I ŁAMANIE PRAWA, TRADYCJA ORAZ RELIGIA I KOŚCIÓŁ. O ile nie mamy do czynienia z określeniami wartościującymi, mniej drażliwe wydają się pola: FUNKCJONOWANIE PAŃSTWA, WOJSKO I WOJNA, JĘZYK, EDUKACJA I OŚWIATA, ŚRODKI MASYOWEGO PRZEKAZU, BIZNES I FINANSE.

W makropolu CZŁOWIEK I PRZYRODA wspomnieć należy o polu ŚWIAT ZWIERZĄT (zwłaszcza sensy związane z kulturowo nośną obecnie relacją człowiek–zwierzę). Podpola: NIEBO I WSZECHŚWIAT, POGODA, ZIEMIA, ŚWIAT ROŚLIN rysują się jako neutralne tożsamościowo. Makropola CZŁOWIEK I TECHNIKA oraz KATEGORIE FIZYCZNE prezentują się również jako niezwiązane bezpośrednio z analizowaną problematyką.

Jak zatem widać, podział ideograficzny słownictwa umożliwia bardziej precyzyjne wskazanie pól sprzężonych z problemami poprawności polityczno-kulturowej. Warto nadmienić, że obszary te nie muszą się ujawniać bezpośrednio w znaczeniach do nich się odnoszących. Pośrednio wątpliwości mogą rodzić się w znaczeniach z innych grup, przy których opisie na przykład odnotowywane są określonego rodzaju kolokacje czy cytaty.

Studia przypadku

– dane leksykograficzne a poprawność polityczno-kulturowa

A. Definicje

Przegląd innych kłopotliwych zagadnień, które związane są z wrażliwymi obyczajowo polami tematycznymi, pojawiły się natomiast w mojej praktyce leksykograficznej, rozpocznę od wątpliwości definicyjnych. Dość jaskrawy przykład wiąże się z podstawowym znaczeniem słowa *małżeństwo* (pole tematyczne RODZINA). Jak wiadomo, w polskim systemie prawnym nie ma instytucji małżeństwa między osobami tej samej płci. Tymczasem leksykograf, analizując choćby kolokacje i cytaty w korpusach internetowych, napotka dość częste użycia, dotyczące małżeństw jednopłciowych, odnoszące się jednak do realiów zagranicznych czy związane z dyskursem światopoglądowym. Jak w takim wypadku postąpić? Czy dostosować definicję do polskiego systemu prawnego, ignorując dane językowe (korpusowe) i narażając się na krytykę? Może lepiej sformułować definicję na tyle szeroko, aby obejmowała również inne systemy prawne, oddawała natomiast perspektywę szerszą niż polskocentryczna³? Z pewnością nie uda się stworzyć konstruktu o charakterze globalnym (zbyt duże zróżnicowanie prawidłowości pozajęzykowych), trzeba więc wybrać jakiegoś rodzaju optykę oddawania zjawisk obyczajowych.

Innym przykładem, związanym z polami OKREŚLENIA RELACJI MIĘDZYLUDZKICH i SEKSUALNOŚĆ, może być seria wyrazów *awanse*, *zalatoty* czy *konkury*. Oczywiście

³ Więcej o różnych perspektywach kulturowych w leksykografii: Haß 2010.

wiążą się one z pewnego rodzaju odcieniem książkowym czy żartobliwym wypowiedzi. Jak jednak sformułować definicję? Czy powinna ona zawierać informację o działaniach skierowanych ku osobie płci przeciwnej, która komuś się podoba? Czy, nawet nie znajdując odpowiednich poświadczeń, powinno się poszerzyć definicję do formuły uwzględniającej również ewentualne konteksty związków jedno-płciowych?

B. Kolokacje

Problemy pojawiają się również w przypadku tak intersubiektywnych danych jak kolokacje (pozyskiwane z zasobów NKJP). Jeżeli spojrzymy na wyniki przyporządkowane do słów *konkubent* czy *macocha* (pola RODZINA i ZWIĄZKI MIĘDZY-LUDZKIE), okaże się, że mamy tam bardzo wiele częstych połączeń, które mówią sporo o cechach konotacyjnych znaczenia w polskim kręgu kulturowym, nie służą natomiast w żadnym stopniu oddawaniu cech denotacyjnych (np. *pijany konkubent*, *zabić konkubenta*, *konkubent znęcał się*). Powszechnie wiadomo, że wyraz używany jest w tekstach urzędowych, na przykład policyjnych, zazwyczaj w wypowiedziach przedstawiających różnego rodzaju sytuacje patologiczne. Problem polega na tym, że wspomniane kolokacje dominują, niejako zatem zafałszowują rozpoznanie znaczenia i mogą być obraźliwe dla pewnych osób. Czy należy je usunąć, czy też lepiej je zostawić, równoważąc przykładami wskazującymi na pozytywne wartościowanie?

Podobny przypadek wiąże się z kolokacjami jednoznaczeniowego leksemu *lichwiarz*. Słowo przyporządkować można do stosunkowo neutralnych pól PRACA i FINANSE, jednak jedno z zaledwie kilku powtarzalnych i dosłownych połączeń (obok *starego lichwiarza* oraz szeregu *kupiec i lichwiarz*) to *żydowski lichwiarz*. Tutaj natomiast wkraczamy w trudne kulturowo pola RELIGIA-KOŚCIÓŁ oraz PRZYNALEŻNOŚĆ I PODZIAŁ TERYTORIALNY. Czy, bazując na prawidłowościach społecznych z przeszłości, odzwierciedlonych następnie w korpusach językowych, powinno się tę ostatnią kolokację zachować? Czy też, ponieważ *lichwa* jest zjawiskiem aksjologicznie wątpliwym, zaś kolokacja utrwałać będzie niekorzystne stereotypy narodowościowe, niemające odbicia we współczesnym języku czy współczesnej rzeczywistości, połączenie należy usunąć ze spisu?

C. Cytaty

Problemy rodzą się też przy doborze właściwych cytatów. Weźmy pod uwagę rzeczownik *lanie* w jego podstawowym znaczeniu. Definicja będzie oczywiście mówiła o przemocy fizycznej polegającej na wielokrotnym uderzeniu kogoś. Wiadomo, że najbardziej prototypowe kulturowo przykłady, aktualne czy też nie, będą dotyczyły kar fizycznych wymierzanych przez rodziców czy opiekunów dzieciom. Wkraczamy zatem na wrażliwe pole RODZINA oraz pole EDUKACJA ORAZ WYCHOWANIE. Czy takie przykłady, w świetle dzisiejszej poprawności kulturowej, powinny się pojawić w słow-

niku? Czy, jeśli tak, cytat odnoszący się do przemocy rodzicielskiej jako prototypowy może być wzmiankowany na pierwszym miejscu, czy może lepiej umieścić go na dalszych pozycjach, aby mógł zostać przez czytelnika przeoczony?

Wątpliwości co do doboru cytatów rodzą również znaczenia sprzężone ze współczesnym dyskusjami światopoglądowymi. Weźmy pod uwagę na przykład serię wyrazów *aborcja*, *aborcjonista*, *aborcjonistka* (makropola CZŁOWIEK JAKO ISTOTA FIZYCZNA I PRAWO I ŁAMANIE PRAWA). Czy autor słownika powinien odnotować jedynie cytaty neutralne, eksplikujące znaczenie, niezwiązane natomiast z żadną postawą ideową? Czy też lepiej byłoby umieścić po jednym cytacie wartościującym, przy czym każdy odnosiłby się do innej pozycji światopoglądowej?

D. Ilustracje graficzne

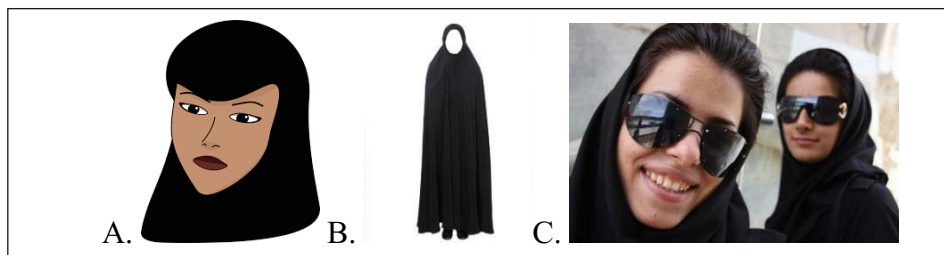
Szereg trudności rodzi również dobór ilustracji graficznej w słowniku, w szczególności w polach sprzężonych z poprawnością polityczno-kulturową. W polskich opracowaniach leksykograficznych *rodzinę* ujmuje się jako rodziców wraz dziećmi. Która z poniższych ilustracji byłaby trafniejsza?



A. PublicDomainPictures, CC0, <https://pixabay.com>; B. White 77, CC0, <https://pixabay.com>;
C. PublicDomainPictures, CC0, <https://pixabay.com>

Czy może najlepszym rozwiązaniem byłoby umieszczenie kilku ilustracji, wskazujących na różne możliwe scenariusze, sprzężone z polską rzeczywistością kulturową? Czy umieścić tam również ilustracje niezwiązane z polskim etnosem?

Inny przykład to różne zasłony, używane przez kobiety wyznające islam (*burka*, *czarczaf*, *czador*, *nikab*). Pod względem teorii leksykograficznej są to wręcz modelowe znaczenia, które należałby zilustrować graficznie, gdy słowa nie mają w przypadku tych sensów wystarczającej mocy przekazywania wiedzy (por. Al-Kasimi 1977: 100–101). Problem polega na silnej zależności znaczeń od pól bardzo wrażliwych kulturowo, RELIGIA-KOŚCIÓŁ oraz PRZYNALEŻNOŚĆ I PODZIAŁ TERYTORIALNY. W takiej sytuacji każdy element kontekstu zdjęcia (por. poniższe przykłady dla leksemu *czador*) może w zasadzie świadczyć o poglądach przekazywanych obocznie przez opracowanie leksykograficzne. Pierwsza ilustracja może być odczytana jako związana jest negatywnymi emocjami przedstawianej postaci, druga wydaje się sztucznie zdepersonalizowana, trzecia – zbyt śmiała.



A. OpenClipart-Vectors, CC0, <https://pixabay.com>; B. SamirA: Chador, flickr.com;
C. <http://nomad123.blogspot.com/2010/05/czador-obowiazkowe-nakrycie-glowy.html>.

Podsumowanie

Podsumowując, problem poprawności polityczno-kulturowej w słowniku jest ważny. Szczególnie jeśli wziąć pod uwagę możliwość łączenia podobnych znaczeń w grupy i przedstawiania ich czytelnikowi w postaci onomazjologicznej (zbiór powiązanych lematów czy ilustracji dodatkowych przy czytanim hasła). Taki sposób prezentacji z wielokrotni jeszcze zastosowaną strategię leksykograficzną lub może stać się dowodem na niekonsekwencję przy jej konstruowaniu.

W idealnej sytuacji leksykograf powinien w swojej pracy być osobą całkowicie neutralną światopoglądowo, postrzegać wykonywane zadania w sposób dyskursywny, wykazywać się ostrożnością i czujnością w odniesieniu do znaczeń i elementów opisu sprzężonych z pewnymi polami tematycznymi, także sensów wartościujących. Do decyzji autora lub autorów słownika należało będzie, czy w egzemplifikacji semantycznej dążyć będą do całkowitej obiektywności polityczno-kulturowej, czy postarają się uchwycić i zobrazować różnorakie odcienie wartościowania, obecne w materiale językowym. Warto wspomnieć również możliwość deklaracji *a priori* autorskiego punktu wyjścia opisu (np. scjentyzm czy pozytywizm w SJPD, ogląd świata zdroworozsądkowy, charakterystyczny dla przeciętnego użytkownika języka w ISJP). Z całą pewnością odrzucić należy natomiast niestaranność warsztatową, cechującą się bezrefleksyjnym umieszczaniem informacji potencjalnie wrażliwych. Z drugiej strony, zdecydowanie skrytykować należy także taki dobór materiału ilustracyjnego w opracowaniu, który promować ma określoną postawę światopoglądową, reprezentowaną przez autora czy autorów, kosztem innych poświadczonych elementów wartościowania językowego. Dzieło leksykograficzne, jak zostało to wyżej nadmienione, jest świadectwem swojego czasu, oddziałuje pod względem treści na czytelników i utrwała nie tylko systemowy, lecz także, pośrednio, (kon)tekstowy obraz świata. Warto również pamiętać, że krytyczny ogląd słownika, dokonywany przez osoby niezwiązane z językoznawstwem, bardzo często dokonywany jest w oparciu o opis właśnie kulturowo wrażliwych haseł, o czym świadczy korespondencja kierowana przez czytelników do autorów opracowań (por. Lorentzen, Trap-Jensen 2016).

Niniejszy artykuł ma charakter wstępny, ogólnoleksykograficzny. W dalszej kolejności warto pogłębić rozważania, uwzględniając możliwości i cele różnych słowników, ogólnych i specjalistycznych (np. dla osób uczących się danego języka, pochodzących z różnych kręgów kulturowych), deskryptywnych i preskryptywnych (np. stylistycznych), o rozbudowanej strukturze czy dostarczających pewien ściśle określony typ informacji.

Bibliografia

- AL-KASIMI Ali (1977): *Linguistics and Bilingual Dictionaries*. – Leiden: E.J. Brill.
- BATKO-TOKARZ Barbara (2008): Tematyczny podział słownictwa w Wielkim słowniku języka polskiego. – [w:] Piotr ŻMIGRODZKI, Renata PRZYBYLSKA (red.): *Nowe studia leksykograficzne 2*. – Kraków: Lexis, 31–48.
- BERGENHOLZ Henning, NIELSEN Sandro (2013): The Treatment of Culture-bound Items in Dictionaries. – [w:] Rufus H. GOUWS, Ulrich HEID, Walter SCHWEICKARD, Herbert Ernst WIEGAND (red.): *Dictionaries. An International Encyclopedia of Lexicography. Supplementary Volume*. – Berlin–Boston: de Gruyter, 469–481.
- BIESAGA Monika (2017): Stan współczesnej leksykografii internetowej (na przykładzie haseł termin i terminologia w jednojęzycznych słownikach ogólnych angielszczyzny. – [w:] Renata PRZYBYLSKA, Władysław ŚLIWIŃSKI (red.): *Terminologia specjalistyczna w teorii i praktyce językoznawców słowiańskich*. – Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego, 137–150.
- BURKHANOV Igor (2010): *Lexicography. A Dictionary of Basic Terminology*. – Rzeszów: Wydawnictwo Uniwersytetu Rzeszowskiego.
- CASARES Julio (1959): *Diccionario ideológico de la lengua Española*. – Barcelona: Editoria Gustavo Gili, S.A.
- DORNSIEFF Franz (2004): *Der deutsche Wortschatz nach Sachgruppen*. – Berlin–New York: Walter de Gruyter.
- Encyklopedia PWN (2017): <http://encyklopedia.pwn.pl/haslo/3960506/poprawnosc-polityczna.html> (dostęp: 27.11.2017).
- GOUWS Rufus H., TARP Sven (2016): Information Overload and Data Overload in Lexicography. – *International Journal of Lexicography*, 31 lipca (dostęp zaawansowany), doi:10.1093/ijl/ecw030.
- HARTMANN R.R.K., JAMES Gregory (2001): *Dictionary of Lexicography*. – London–New York: Routledge.
- HAß Ulrike, (2010): *In Search of the European Dimension of Lexicography*, tekst wystąpienia (The Fifth International Conference for Historical Lexicography and Lexicology, Oxford, St. Anns College, 16–18.06), <http://www.linse.unidue.de/publikationenliste/articles/in-search-of-the-european-dimension-of-lexicography.html?pdf=1855> (dostęp: 27.11.2017).
- LORENTZEN Henrik, TRAP-JENSEN Lars (2016): What, When and How? – the Art Of Updating an Online Dictionary. – [w:] Tinatin MARGALITADZE, GEORGE MELADZE (red.): *Proceedings of the XVII EURALEX International Congress: Lexicography and Linguistic Diversity*. – Tbilisi: Ivane Javakishvili Tbilisi State University, 138–145.
- MARKOWSKI Andrzej (1992): *Leksyka wspólna różnym odmianom polszczyzny*, t. 1. – Wrocław: Wiedza o Kulturze.
- MIKOŁAJCZAK-MATYJA Nawoja (2005): Hierarchiczne uporządkowanie słownictwa a tradycyjna leksykografia. – *Investigationes Linguisticae* XII, 17–32.
- MILLER George A., BECKWITH Richard, FELLBAUM Christiane, GROSS Derek, MILLER Katherine (1993): *Introduction to WordNet: An On-line Lexical Database*. <http://wordnetcode.princeton.edu/5papers.pdf> (dostęp: 27.11.2017).

- NKJP (2017): *Narodowy Korpus Języka Polskiego*, www.nkjp.pl (dostęp: 27.11.2017).
- PIOTROWSKI Tadeusz (2001): *Zrozumieć leksykografię*. – Warszawa: Wydawnictwo Naukowe PWN.
- Roget's International Thesaurus. Fourth edition*, red. Robert L. CHAPMAN. – New York 1977: Thomas Y. Crowell, cop.
- RUDNICKA Ewa (2016): Słownik jako tekst kultury. – *Język Polski* 4, 47–64.
- REMBISZEWSKA Dorota K. (2005): Słownictwo tematyczne w dwudziestowiecznych opracowaniach leksykograficznych. – *Poradnik Językowy* 8, 37–52.
- SJPD: *Słownik języka polskiego*, red. Witold DOROSZEWSKI, t. 1. – Warszawa 1958: Państwowe Wydawnictwo Wiedza Powszechna.
- SZAJHAJ Adam (2010): *E pluribus unum? Dylematy wielokulturowości i politycznej poprawności*. – Kraków: Universitas.
- ZAWILIŃSKI Roman (1985): *Dobór wyrazów. Słownik wyrazów bliskoznacznych i jednoznacznych*, przedruk wydania z 1926 r. – Warszawa: Wydawnictwa Artystyczne i Filmowe.

doi: <http://dx.doi.org/10.17651/bptj.74.12>

BARBARA LEWANDOWSKA-TOMASZCZYK

State University of Applied Sciences in Konin

PAUL A. WILSON¹

University of Lodz

Sources of data and methodological foundations of a contrastive linguistic analysis of emotion concepts

Abstract

The aim of the paper is a presentation of methodological foundations of semantics of emotion concepts in contrastive linguistic analyses and identifying the sources of data in Polish and English obtained by less frequently used research methods. The main assumption is the presence of not fully determinate concepts in linguistic semantics and the need to interpret emotion concepts in terms of prototypical meanings and cluster models. This requires the application of cognitive linguistics categories and cultural models of emotions and emotion events. The methodological part demonstrates the qualitative and quantitative nature of sources and research (GRID, sorting and ranking, cognitive linguistic analysis of corpus data, metaphoricity). The examples analyzed in fuller detail are Eng. *compassion* and Pol. *współczucie* in their respective emotion clusters. The conclusions present implications of this interdisciplinary and multi-method approach to develop a new methodology of contrastive research in emotion studies.

Keywords: cognitive linguistics, collocations, *compassion*, contrastive analysis, emotion, GRID, language corpora, semantic approximation, prosodic profiles, semantic clusters, sorting and ranking methodology, *współczucie*

Źródła danych i podstawy metodologiczne kontrastywnej analizy językowej pojęć wyrażających uczucia i emocje

Streszczenie

Celem artykułu jest prezentacja podstaw metodologii badań nad znaczeniem pojęć wyrażających emocje i uczucia w aspekcie kontrastywnym oraz wskazania na źródła danych do badań na przykładzie języka pol-

¹ The authors' contributions to the research and text of the paper were equal.

skiego i angielskiego w kontekście rzadziej stosowanych metod badawczych. Głównym założeniem pracy jest obecność znaczeń nie w pełni określonych w semantyce językowej oraz konieczność interpretacji znaczeń emotywnych w kognitywnych terminach *znaczeń prototypowych* z jednej strony i *klasterów znaczeniowych* z drugiej. Wymaga to sięgnięcia do kategorii *językoznawstwa kognitywnego* oraz *kulturowych modeli uczuć* i emocji oraz zdarzeń emotywnych. W części metodologicznej pracy zaprezentowano źródła i badania *danych ilościowych* oraz, czerpiące z inspiracji językoznawstwa kognitywnego, badania jakościowe pojęć emotywnych i ich metaforyki (kwestionariusze GRID, zadania online sortowania i rankingu oraz badania korpusowe). Jako przykład omówiono angielskie pojęcie *compassion* w szerszym klasterze emotywnym oraz polskie pojęcie *współczucie*, jak też wzorce ich występowania w autentycznym materiale językowym. We wnioskach przedstawione są implikacje tak pojętych analiz interdyscyplinarnych, czerpiących z połączenia metodologii i źródeł badań psychologicznych, socjologicznych, kulturowych i językowych dla nowego podejścia w językowych badaniach kontrastywnych.

Słowa kluczowe: językoznawstwo kognitywne, kolokacje, *compassion*, analiza kontrastywna, emocje i uczucia, GRID, korpusy językowe, aproksymacja semantyczna, profile prozodyczne, klastery semantyczne, metodologia sortowania i rankingu, *współczucie*

1. Introduction

Studies of emotions, as expressed in language, methods of their investigation and sources of data, are related to the phenomena we consider crucial for our discussion, namely the relationship between external reality language and cognition, language and culture, and the status of meaning in language and in verbal communication in general. In other words, in Continental linguistic tradition, what is addressed in research concerning linguistic meanings are both onomasiological as well as semasiological phenomena and relations (Wierzbowski 1980: 31; Geeraerts 2002), immersed in a broad cultural context. The former lead to qualitative internal *meaning distinctions* between concepts (e.g. polysemy), the latter involve *the naming* of lexical senses and are related with the quantitative methods, identifying structural weight measures, tendencies, cognitive salience and prototypicality (Geeraerts 2002: 29). Numerous Polish scholars (for example, Wierzbicka 1985, 1999, 2004; Nowakowska-Kempna 1995; Mikołajczuk 2012, in preparation; Krawczak 2014) have been studying emotion concepts in terms of models within a Cognitive Linguistics tradition, which will also be a point of reference in our study. The phenomena signalled in the introduction are *meaning approximation* in interaction (Lewandowska-Tomaszczyk 2012) and *semantic reconceptualization* (Lewandowska-Tomaszczyk 2010), which are part and parcel of any communicative act, verbal or non-verbal, carried out in one or more languages, spoken or written. These assumptions are discussed within the claim of the present study, concerning the emotion cluster equivalence thesis both in monolingual (synonymy) as well as in translational contexts. This discussion will also show the reasons for the use of a multidisciplinary approach, involving a number of different research methods in emotion studies, which results in considering numerous sources of data in the research endeavor (consult also Wilson 2012 and Wilson and Lewandowska-Tomaszczyk 2012).

2. Meaning approximation in interaction

The concepts people use in communication exchanges are usually only partially overlapping (see Lewandowska-Tomaszczyk 2012 and 2017b for details). Some of our thinking tends to be more *effable*, that is, possible to express in a language (Katz 1978), while a large part of it such as some image-verbal thoughts, visions, emotions or dreams remain more felt than expressed, more imagined than put in words. In other words, meanings are, as we learn from Cognitive Linguistic paradigms, notoriously underdetermined by their fuzziness and absence of rigid categorial boundaries. Another crucial parameter of approximation is connected with the language typological parameter – a linguistic type a given system represents, in which some ontological categories are verbally marked and can be expressed, while some others are absent or left non-verbalized in the system. Connected with this is the type of linguistic construction, in which the speaker has to or can (considering the typological differences) choose to convey a message. The selection of a particular construction also symbolizes a particular semantic perspective conveyed by these forms via different *profiling* and cognitive *construal* of a given scene or event (Langacker 1987, 1991).

Consequently, the semantic scope of an individual interactant's meanings leaves an uncovered space in the interpretation of a message, as a part of a respective organization of mental structures. Its exponent is a system (language, visual system), which is complemented by *inferential structures* – an ability to complement a message with further (unexpressed) thought(s) or image(s), for example, by the activation of an Idealized Cognitive Model (cf. Lakoff 1987) of a respective scene. Although language users apply numerous strategies (synonymy, paraphrase, polysemy, super-/sub-ordinate category members, etc.) to achieve their communicative goal, such strategies are always used at the expense of maintaining the exact *sameness* of meaning either in one language or in translation. What takes place in such contexts is communication via monolingual or cross-linguistic meaning approximation rather than identity (Lewandowska-Tomaszczyk 2012).

The meanings of emotion terms are not very much like those of concrete concepts, which are more effable in terms of explicit semantic characterization and cover the necessary and sufficient properties of a given sense. Our position in this respect is rather that parts of the meanings of emotion concepts are accessible via behavioural and experiential properties; however, they are conceptualized not as direct meanings but as networks of different kinds of meaning, that is, in terms of figurative forms. Moreover, cross-linguistic comparisons also show that what is a fairly crisp concept in one language may be considered a blended notion in another (see Fauconnier and Turner 1998; consult also Wilson, Lewandowska-Tomaszczyk and Niiya 2013).

3. Emotion concepts in contrast

Emotions can be defined as bodily and mental states of arousal in an experiencer which develop as a consequence of some prior or co-temporal or co-temporally en-

visaged stimulation (cause). The rise of an emotion is conditioned both by the character of the stimulus, which triggers the experiencer's positive or negative reactions, the cultural and social conventions of the community they have been brought up in, which regulate their external behavior including language, and the experiencer's past experiences and their individual mental and behavioural predispositions.

Each of these facets of emotions can be taken as a subject of research scrutiny and adequately investigated. They can also be considered points of reference in a contrastive analysis of emotion concepts. Bodily effects can be adequately measured by means of observational, experimental and medical methods. Since the publication of the influential book by Ekman (2004) on the universality of basic emotion facial expressions, the thesis has been supported or challenged by numerous studies. One of the most recent publications by Gendron et al. (2014) argues against the universal status of perceptions of emotion from facial expressions. Instead, supporting their evidence with the emotion perception of participants from the United States and the Himba ethnic group from the Keunene region of northwestern Namibia, the authors hypothesize that evidence of universality depends on the conceptual and cultural context.

4. Culture

A common thread in many of the definitions of culture proposed throughout years of research involves a few *constitutive properties* (Kluckhohn and Kelly 1945; Damen 1987), which center on *patterning*, that is human models for behaviour and living, the *sharing* of these patterns, their *structure* and *repetitiveness*, *learnability* and *acquisition*, *cultural transmission* in terms of signs and symbols, and *cultural imagery* and its *embodiment* in artefacts and in human achievement in thought and language. Culture is a broad and elaborate concept that involves real or fictitious heroes and events, concerns symbols and imagery, describes verbal and non-verbal rituals such as speech events and hand shaking, profile values to identify good and evil, and creates or imposes numerous social practices (Lewandowska-Tomaszczyk 2011). However, a particular culture primarily involves *patterns of emotions and thoughts* shared by a given community of people. In the present paper we aim to determine how people's shared ideas and cultural practices, in other words the *cultural models* shared by a certain society, shape the feelings and emotions of that society.

5. Culture and emotions

There are a number of elements that can influence variation in emotion representation across cultures (cf. also Langacker 2014). As Lewandowska-Tomaszczyk and Wilson (2017) explain with respect to shame and guilt, these include religion, honour, face, and individualism vs. collectivism. Although it is beyond the scope of the present paper to provide a detailed assessment of how specific emotions are influenced by

such cultural elements, it should be noted that some of them (e.g. religion) are linked to what Lakoff discusses in his *Moral Politics* (1996), and synthetically in *Philosophy in the Flesh* (1999).

The cultural dimensions identified by Hofstede (1980) have provided an in-depth understanding of such cross-cultural variation. The specific dimensions identified by Hofstede (1980) are as follows: *collectivism versus individualism* – whether we construe ourselves as autonomous, self-determined entities vis-à-vis interdependent members of a group; *power distance* – the extent to which one accepts the unequal distribution of power by subordinate members of institutions; *masculinity versus femininity* – emphasis on success and money versus caring for others and quality of life, respectively; *uncertainty avoidance* – how much individuals are adversely affected by ambiguity or unknown situations in their lives; *long-term orientation* – the degree to which individuals are prepared to decline immediate social success in favour of greater emphasis on long-term aims; *indulgence versus restraint* – emphasis on fun, enjoyment, natural human desires and self-gratification versus the restriction, regulation and control of the indulgence in human desires and needs, respectively.

5.1. Individualism and collectivism

The collectivism-individualism dimension is central in the present chapter as it is key to the understanding of the differences between British English *compassion* and its Polish equivalent, *współczucie*, which is foregrounded in our explication of the compatability of the three methodologies that we employ, namely GRID, online emotions sorting methodology, and the corpus linguistics methodology. The emphasis on individualism-collectivism in the present paper is not an attempt to present a skewed representation of the importance of this dimension above other cultural elements such as religion, honour and face, but is determined by its relevance to the choice of emotions, namely *compassion* and *współczucie*.

5.2. Individualism

Individualism is characterised by the following central features (e.g. Hofstede 1980; Oyserman, Coon and Kemmelmeier 2002; Triandis 2001; Diener and Diener 1995):

- The primacy of one's autonomy with personal aims, which places emphasis on the self as opposed to the in-group in terms of achieving success and fulfilling one's potential
- The foregrounding of individual emotions, values and thoughts compared with those shared by the in-group; satisfaction in one's life is more central to self-construal
- consistency with the main tenet of exchange theory, the weighting of beneficial and detrimental factors are central in social relationships
- relationships are relatively more temporary, but emphasis is placed on respect

- a greater emphasis on rights and more social distance from in-groups; individuals have greater individual responsibility, have fewer duties towards the in-group and have greater freedom to make individual choices
- relationships within the professional sphere are based more on contractual obligations
- judgements regarding the causes of actions of others are based on the individual as opposed to the situational

5.3. Collectivism

Collectivistic cultures are underscored by the following features (e.g. Hofstede 1980; Oyserman, Coon and Kimmelmeier 2002; Triandis 1995; Markus and Kitayama 1991):

- a greater perceived social distance between in-groups and out-groups, with more of an impermeable border between them
- as opposed to individual autonomy, there are closer interpersonal relations in the in-group that are based on equality and principles based on generosity
- duties and obligations in the in-group are determined by social standing
- satisfaction and contentment are based on the construal of the self that centres on the successful execution of social roles and duties that are in turn derived from common aims and harmonious relations in in-groups that have a greater permanence and are characterised by a greater focus on the benefit of others
- due to the differences between members of the in-group the range of attitudes, values and behavior is broader in scope
- the regulation of the outward expression of emotions to maintain harmony in in-group relationships
- attributing cause during interpersonal perception centres on situational and social factors as opposed to the individual

5.4. Individualism and Collectivism: Criticisms and Influences

More recent research has questioned the established view of the stable social constructs of individualism and collectivism presented above. The main focal point of this challenge has centred on both conceptual and methodological criticism. Other research has highlighted the possible dynamism in these constructs, calling in question whether the traditional distinction between the relatively more individualistic Britain and the more collectivistic Poland might need to be revised.

5.4.1. Conceptual and Methodological Criticism

One of the main criticisms is that individualism and collectivism are conceptually “fuzzy” with many definitions and assessments being too broad and diffuse (Brewer and Chen 2007). Bond (2002: 76) similarly calls for a more detailed understanding of these multifaceted constructs that permit “many different operationalizations”. Fiske

(2002: 87) questions the validity of the “characterization of cultures according to IND and COL” and proposes that instead of merely identifying how collectivistic cultures differ from individualistic cultures, we need to gain a more complete understanding of the intrinsic nature of the many collectivistic cultures around the world. It has also been argued that the individualistic and collectivistic construct has been applied too readily in attempts to understand a wide range of diverse cross-cultural issues (Bond 2002; Earley and Gibson 1998).

A number of methodological criticisms have been directed at empirical investigations into individualism and collectivism. Oyserman, Coon and Kimmelmeier (2002) highlight a number of methodological limitations in their comprehensive review of the literature. By using undergraduate participants, who may be more individualistic and less collectivistic (Fiske 2002), the generalizability of the results of many studies to the wider society is restricted. Other studies restrict the generalizability of the results to other nations, racial groups, or ethnic groups, as they focus narrowly on a comparison between European American undergraduates and either undergraduates from a single Asian country or a single ethnic minority group within America. Another problem concerns the different ways that researchers use and conceptualise individualism and collectivism. For example, when collectivism is assessed on the basis of belonging to in-groups and seeking advice from others, Americans see themselves as relatively collectivistic; however, they rate themselves as low in collectivism when researchers use duty to in-group in their measurement of collectivism. Finally, Oyserman, Coon and Kimmelmeier (2002) found that very few of the cross-cultural comparisons that they reviewed provided information on the ethnic background of their samples.

This criticism of individualism and collectivism has not received unanimous support. Schimmack, Oishi and Diener (2005) argue in favour of individualism as an important and valid element of intercultural differences and that the apparent inconsistencies in the results can be explained by cross-cultural differences in self-report response styles. As noted by Brewer and Chen (2007: 134), when these response styles are statistically controlled “horizontal individualism shows high convergent validity with Hofstede’s (1980) original rankings of emotions on the I-C dimension”. However, in contrast with individualism, which is deemed to have a valid definition and assessment, Schimmack, Oishi and Diener (2005) propose that the concept of collectivism may need to be reassessed. Such a distinction alludes to a rejection of the position that individualism and collectivism are conceptual opposites. Accumulating evidence supports the statement by Oyserman, Coon and Kimmelmeier (2002: 8) that individualism and collectivism “are better understood as domain-specific, orthogonal constructs differentially elicited by contextual and social cues”.

More research is required to address these conceptual and methodological concerns. The validity of these objections notwithstanding, the classification of cultures on the basis of individualism and collectivism has provided a great deal of insight in cross-cultural investigations.

6. Emotions: definition and types

As proposed by Cosmides and Tooby (2000), emotions are a set of superordinate programmes which coordinate or harmonize the activation of various componential subordinate programmes such as approach, avoidance, resting, etc. Distinct emotions are associated with a number of *Emotion Event-types* (Lewandowska-Tomaszczyk and Wilson 2013) that recur both in the course of evolutionary history and in child development and, when needed, the most successful programmes are activated. The *emotion clusters* that we refer to rely on material from close conceptual domains that are based on a form of family resemblance (e.g. joy, happiness) and exploit partial synonymy as well as hyponymy and metonymy links among emotion terms such as those between *sadness – sorrow – grief*. Emotion clusters can either be negative or positive (e.g. *anger and hate* or *compassion and kindness*). The values of such properties are not considered discrete as was argued above, but they are rather graded with blurred edges and boundaries².

In order to investigate a complex structure of emotions a number of methodological tools have been used in our study. They will be discussed in the sections to follow with the use of the concepts of English *compassion* and Polish *współczucie* and their cluster members as the subject of analysis in the present paper.

7. GRID method of emotion studies

The complex nature of emotions that are conceptually structured in clusters rather than single emotions underscores the need for empirical investigation to employ diverse methodologies in a cross-disciplinary approach. To meet these demands, the present study employed the GRID, online emotions sorting, and language corpus methodologies (see below). Dimensions and components are an integral part of the GRID instrument in its investigation of emotion prototype structure. Support for the dimensional structure of emotion concepts has come from a number of studies (e.g. Fontaine, Scherer and Soriano 2013). In addition to enabling analyses based on these emotion dimensions, the wide range of emotion features in the instrument allows relevant components to be isolated and analysed. The main advantage of the online emotions sorting methodology is the enriched meaning of emotion clusters that it provides, which extends horizontally, providing information on the relationships between clusters. Using this methodology we can not only establish that the British English version of *compassion* is more positive than its more negative Polish counterpart, *współczucie*, but additionally discover, by determining its prototypical location within clusters and how it relates to emotions in other clusters, the possible reasons for this cross-cultural difference. The corpus methodology provides more detailed materials

² Emotion clusters are differentiated from the experience of ambivalent feelings, so-called *mixed feelings*, considered a mutually exclusive option such as a *love/hate relationship*, which represents an *approach-avoidance* conflict, a consequence of activating conflicting emotion subprogrammes.

on the specificity of particular emotion terms, which also includes information on how they function in various contexts. The context is recovered by examining the distribution (i.e. occurrence) of a particular term in various structures and identifying the frequency of their co-occurrence with neighboring words. The type of meaning they represent and the structure in which they are used make it possible to identify both the VALENCE and intensity with which the emotion terms are used in discourses.

The three methodologies that were employed, GRID, online emotions sorting, and the cognitive corpus methodology, were chosen on the basis of the broad aims of the study, which were to compare and contrast the structures of the British English and Polish emotion domains. As outlined above, each of these methodologies focus on different but compatible aspects of emotion meaning and it is the composite picture that emerges from these that provides an understanding of how such meaning is conceptualized.

7.1. GRID

The GRID instrument is a methodological tool of major importance that has been developed recently and implemented with good effect to investigate emotions in the cross-linguistic and cross-cultural domain (Scherer 2005; Fontaine et al. 2002, 2007, 2013; Fontaine, Scherer and Soriano 2013; Fontaine 2013). It employs a system of dimensions and components, which bring about insight into the nature of emotion prototypical structures. The GRID project is coordinated by the Swiss Center for Affective Sciences at the University of Geneva in collaboration with Ghent University and is a worldwide study of emotional patterning across 23 languages and 27 countries. The GRID instrument comprises a Web-based questionnaire in which 24 prototypical emotion terms are evaluated on 144 emotion features.

7.2. GRID Methodology

Participants completed the GRID instrument in a controlled Web study (Reips 2002), in which each participant was presented with four emotion terms randomly chosen from the set of 24 and asked to rate each in terms of the 144 emotion features³. They rated the likelihood that each of the 144 emotion features can be inferred when a person from their cultural group uses the emotion term to describe an emotional experience. A 9-point scale was employed that ranged from *extremely unlikely* (1) to *extremely likely* (9) – the numbers 2 to 8 were placed at equidistant intervals between the two ends of the scale, with 5 *neither unlikely, nor likely* in the middle and participants typed their ratings on the keyboard. It was clearly stated that the participants needed to rate the likelihood of occurrence of each of the features when somebody who speaks

³ Each participant rated a random selection of four of the emotion terms on the basis of the 144 emotion features as the task would have been too long and demanding if each participant rated all 24 emotions on each of these features.

their language describes an emotional experience with the emotion terms presented. Participants rated all four emotion terms on the basis of a certain emotion feature sequentially on separate screens before rating these four emotion terms on the next emotion feature and so on until all four emotion terms had been rated on the basis of all 144 emotion features.

The British English sample comprised 201 British English participants (124 females) with a mean age of 21.5 years (all born and raised in Britain and spent most of adult lives in Britain), and the Polish sample comprised 124 Polish participants (95 females) with a mean age of 23.2 years (all born and raised in Poland and spent most of adult lives in Poland). As each participant rated four emotion terms, the participant details are idiosyncratic to each of the emotion terms. For example, the participant details for the compassion emotion term are as follows: 33 British English-speaking subjects (21 females; mean age 23.2 years) rated *compassion* and 29 Polish-speaking subjects (26 females; mean age 25.6 years) rated *współczucie*.

7.3. Components of Emotion

The 144 GRID emotion features represent activity in all six of the major components of emotion. Thirty-one features relate to *appraisals*, eighteen to *bodily reactions*, twenty-six to *facial, vocal or gestural expressions*, forty to *action*, twenty-two to *feelings*, and four to *emotion regulation*. Three further features refer to other qualities, such as *frequency* and *social acceptability* of the emotion. Participants are asked to rate the likelihood of these features for the various emotions. This methodology is comprehensive in its scope as it allows the multicultural comparison of emotion conceptualisations on all six of the emotion categories recognised by emotion theorists (Ellsworth and Scherer 2003; Niedenthal, Krauth-Gruber and Ric 2006; Scherer 2005).

The significance of the GRID components of emotion is made clear when one considers the Component Process model of emotion (CPM, Scherer 1984, 2005), which proposes that rather being emotion states, emotions comprise a number of componential processes that are activated in response to certain events that encompass one's goals, values and needs. The main emotion components in the GRID methodology are *appraisals*, *bodily reactions*, *expressions*, *action*, and *feelings*, all of which have been shown to have a substantial representation in the meaning of emotion terms (Fontaine and Scherer 2013). Emotion words bear information about these components across different languages and cultures. Fontaine and Scherer (2013) observed that the *feelings* component contributes the most to differentiating between the meanings of different emotion words, with *appraisals*, *bodily reactions*, *expressions*, and *action*, having a lesser, albeit similar, influence.

7.4. Emotion Dimensions

In addition to its componential element the GRID also offers the means to analyse the dimensional structure of emotions. Fontaine (2013) identifies the assessment of ex-

periences relating to affect and emotions as being central to such dimensional approaches. Support for the dimensional structure of emotion concepts has come from a number of studies. However, how the emotional domain is organised in terms of the number and nature of these dimensions has been a matter of recent debate. In contrast with an earlier proposal of three dimensions (EVALUATION, POTENCY and ACTIVATION) by Osgood, May and Miron (1975), later suggestions have tended to settle on a two-dimensional model comprising VALENCE and AROUSAL (e.g. Yik, Russell and Feldman-Barrett 1999). Fontaine and Scherer (2013: 107) argue that when emphasis is placed on the *meaning* of emotion words, research shows that a three dimensional model, comprising VALENCE, POWER and AROUSAL, offers the most comprehensive representation of dimensional space (Shaver et al. 1987; Shaver, Wu and Schwartz 1992; Fontaine et al. 2002).

Analyses performed on GRID data have consistently produced four dimensions. In an initial study of the dimensional structure of emotions using the GRID instrument, Fontaine et al. (2007) derived a four-dimensional structure for English, French and Dutch that added the dimension of NOVELTY to the previous three-dimensional model of VALENCE, POWER and AROUSAL. Analyses conducted on the data from all of the languages represented in the GRID project have reproduced this four-dimensional structure (Fontaine, Scherer and Soriano 2013). To determine the dimensional structure of the Polish and British English data in the present study, principle components analysis (PCA) with varimax rotation was performed on the combined dataset of these two languages (see above for details). The four-dimensional solution that was selected comprised the same dimensions as Fontaine et al. (2007) and Fontaine, Scherer and Soriano (2013) and accounted for 81.9% of the total variance. The first dimension (VALENCE) accounted for 52.9% of the variance, the second dimension (POWER) for 15.5%, the third dimension (AROUSAL) for 8.3%, and the last dimension (NOVELTY) for 5.1%. A sensory cue GRID feature was included in a dimension if it achieved a 0.6 loading on this dimension *and* did not have a higher loading on another dimension. The VALENCE dimension is characterised by appraisals of intrinsic pleasure and goal conduciveness. Other features include action tendencies of approach versus avoidance, and pleasant emotions versus unpleasant emotions. Specific examples of features associated with this factor include “felt positive”, “wanted to sing and dance”, “in itself unpleasant for the person”, “felt inhibited or blocked”, and “incongruent with own standards and ideals”. POWER includes appraisals of control, with the feelings of power and weakness being particularly salient. It is also characterised by appraisals of interpersonal dominance or submission, and by urges to either initiate action or refrain from this. This dimension includes features such as “assertive voice”, “felt submissive”, and “wanted to take the initiative her/himself”. The AROUSAL dimension is mainly characterised by sympathetic arousal (e.g. rapid heartbeat and readiness for action). The features associated with this dimension include “breathing getting faster”, “felt hot”, “sweat”, and “spoke faster”. The fourth dimension is represented by NOVELTY. On this dimension, appraisals of novelty and unpredictability are compared with expectedness or familiarity. Fontaine et al. (2007) found that surprise was associated

more with the NOVELTY dimension than the other emotions they analysed. This dimension includes features such as “raised eyebrows”, “jaw dropping”, and “confirmed expectations”. It is the profiling of these underlying dimensions that forms the basis of the main aim of GRID, which is to construct a world atlas of emotion terms.

7.5. *The Componential Emotion Approach vs. the Dimensional Emotion Approach*

To what extent do the componential emotion approach and the dimensional emotion approach offer compatible structures of the emotion domain? Fontaine and Scherer (2013) highlight the congruence between these two approaches that they find in their results, which demonstrates that the four-dimensional solution forms a stable structure that also provides a good representation of the componential data. The four-dimensional structure was mostly achieved even after the removal of each of the five emotion components one at a time. More specifically, the elimination of a single emotion component had no influence on the VALENCE and the POWER dimensions. The AROUSAL dimension is to some extent affected by the removal of the *feelings* component, showing that this dimension is characterised largely by feelings. Similarly, the elimination of the *appraisal* component has a moderate effect on the NOVELTY dimension, demonstrating the relative salience of this dimension in appraisal.

The scale and flexibility of the GRID paradigm, based on its wide range of 24 emotions and 144 features, has been taken advantage of by a number of studies focusing on both the componential and dimensional representations of emotion structure in cross-cultural studies of emotions. Studies that feature the componential approach include an investigation into the meaning of happiness in Japan and the United States (Ishii 2013). Consistent with the hypothesis, it was shown that happiness is relatively more characterised by pride in the US, because pride is a socially disengaging emotion that should be more associated with individualism, and that happiness is related more to love in Japan due to love being more of a socially engaging emotion denoting collectivism; however, crucially, this was found for the *bodily reactions* component, but not for the *appraisals*, *expressions*, *action*, and *feelings* components, suggesting that differences in happiness across cultures may be specific to certain types of features. Mortillaro et al. (2013) investigated the meaning of pride between Southern and Northern Italians on the *appraisals*, *feelings* and *expressions* components. Consistent with predictions, the results for the *appraisal* component showed that Southern Italians were more likely than Northern Italians to include other in-group members in their appraisals of pride. Predictions were also confirmed for the *feelings* component, with Southern Italians associating pride with negative feelings and Northern Italians evaluating pride in terms of positive feelings. However, contrary to predictions, the negative feelings that the Southern Italians associated with pride were not accompanied by more regulation strategies by Southern Italians. This pattern of results is consistent with the Southern Italians being relatively more collectivistic than the Northern Italians. There have also been a number of studies employing the GRID

methodology that feature the dimensional approach. In our research we have observed, for example, that British English *fear* is characterised by higher POWER than its Polish counterpart, *strach* (Lewandowska-Tomaszczyk and Wilson 2013), and that British English *pride* has a more negative VALENCE than the Polish equivalent, *duma* (Wilson and Lewandowska-Tomaszczyk 2017).

7.6. Analyses on GRID Features, Components and Dimensions

It is important to note that the selection of GRID features, components and dimensions in analyses comparing emotions in different cultures is determined by the specific elements of the comparative analyses that are deemed important. As specific emotions are influenced by certain cultural features, the necessity of such selection rather than broad spectrum analyses involving all features, components and dimensions, becomes clear, as illustrated in our own studies. For example, the POWER dimension is central to the study comparing British English *fear* vs. Polish *strach* outlined above as it is argued that this dimension is a function of the FIGHT (active opposition) that is relatively more salient in *fear*, as opposed to *strach*, which is characterised by FRIGHT (inactivity, freezing) (see Lewandowska-Tomaszczyk and Wilson 2013). In contrast, we have shown that the NOVELTY dimension is more fundamental to the comparison of British English *love* vs. the Polish equivalent, *miłość* (Lewandowska-Tomaszczyk and Wilson 2015). Specifically, the finding that *love* is more novel than *miłość* is consistent with the unpredictable, sudden nature of passionate love being more salient in *love* in comparison with more stable, long-lasting, companionate and patriotic types of love and love for God, which are relatively more characteristic of *miłość*. Components are more relevant than dimensions when it comes to the application of cultural variations in emotions in practical spheres such as social robotics. This is illustrated in our study highlighting differences in the expressive profile of GRID emotion features between British English and Polish that, we argue, need to be tuned accordingly if such robots are to successfully interact in social settings in these two cultures (Wilson and Lewandowska-Tomaszczyk 2014; Lewandowska-Tomaszczyk and Wilson 2016).

It is far beyond the scope of the present paper to provide a detailed account of which GRID dimensions, components or features are relevant to which emotions. However, in an attempt to provide deeper understanding, the following section outlines the relevance of the VALENCE dimension and a component pertaining to *desire to act* vs. *desire not to act* in analyses comparing British English *compassion* and its Polish counterpart, *współczucie*.

7.7. British English versus Polish Compassion

Componential and dimensional elements can be used effectively in tandem in GRID analyses to determine cross-cultural differences in emotion concepts. We have demonstrated the effective employment of such a combination in the analyses of

British English vs. Polish compassion - *compassion* and *współczucie*, respectively (Lewandowska-Tomaszczyk and Wilson 2016, in press).

Compassion is similar to sympathy as it is a response to the plight of others; however, it is characterised relatively more by a desire to help, which may or may not be manifested behaviourally (Gladkova 2010). The key feature of compassion is that it can be associated with both positive and negative VALENCE, which appears to be determined to some extent by individualism vs. collectivism. The independent, autonomous nature of individualism means that individualists such as the British are likely to have a self-directed focus in the presence of a suffering individual, which is likely to highlight the possible help that they can provide. In contrast, the outward orientation of focus, due to the salience of interpersonal relationships in more collectivistic cultures, such as Poland, is likely to highlight the pain and suffering of an individual who is in need of help and therefore sadness, sorrow and distress are possibly more foregrounded in *współczucie*⁴. An obvious prediction resulting from this is that British English *compassion* is characterised as relatively more positive than Polish *współczucie*. Additionally, the inward focus on the possible personal help given to a person who is suffering is likely to influence the extent to which that person would like to initiate an active response. The analyses that we have performed to compare *compassion* and *współczucie* in terms of VALENCE and the desire to initiate action are a good example of how the analysis of GRID components and dimensions can complement each other.

To compare *compassion* and *współczucie* on the VALENCE dimension a 2 x 2 Anova was performed on the positive VALENCE GRID features vs. the negative VALENCE GRID features (see Lewandowska-Tomaszczyk and Wilson (2016, in press) for further details on this analysis). Consistent with what one would expect on the basis of the relatively more individualistic British culture compared with the more collectivistic Polish culture, the results showed that *compassion* was characterized by more positive VALENCE than *współczucie*.

As predicted, the results of the componential analyses were consistent with the dimensional analyses. Apart from representing the six major components of emotion (i.e. *appraisals, bodily reactions, facial, vocal or gestural expressions, action, feelings, and emotion regulation*) that are outlined above, other, minor, more subordinate components can be represented by GRID features. The subordinate component identified as representing the desire to initiate action is *desire to act (wanted to go on with what doing, required immediate response, wanted to undo what was happening, felt urge to be attentive, felt urge to be active, wanted to act, wanted to tackle situation)* and opposes its opposite, *desire not to act (wanted to submit to situation, felt urge to stop what doing, wanted to do nothing, lacked motivation to do anything, lacked motivation to pay attention to what was going on)*. All of these features are members of the

⁴ Although Poland cannot be deemed to be an individualistic culture, it scores 60 on the individualistic vs. collectivistic dimension, which means that it is more collectivistic in relative terms than Britain, which is considered to be an extremely individualistic society and has a score of 89 on this dimension (Hofstede, Country Comparison).

GRID *action* component, apart from *required immediate response*, which is included in the GRID *appraisals* component. To compare *compassion* and *współczucie* on *desire to act* vs. *desire not to act* a 2 x 2 Anova was performed on the means of these profiles of features (see Lewandowska-Tomaszczyk and Wilson 2016, in press). The results showed that there was a significantly greater *desire to act* for *compassion* than *współczucie*. The greater *desire to act* for *compassion* compared with *współczucie* is consistent with its more positive VALENCE. If an individualist conceptualises compassion more in terms of an inward focus on the help that they can offer, it is likely that it is characterised by both more positive VALENCE and a greater desire to help the suffering individual.

8. Online Emotions Sorting Methodology

The emotions sorting methodology involves the categorisation of emotion terms into groups. In the online version these terms are categorised on the computer desktop. This methodology can be deemed to be complementary to the componential and dimensional emotion approach of the GRID instrument with respect to the meaning of emotion terms. The main advantage of the online emotions sorting methodology is the number of emotion terms it can accommodate (135 vs. 24 GRID emotion terms). It can also enable a more enriched, elaborate understanding of relationships both within and between emotion clusters.

8.1. Selection of Terms

The selection of emotion terms was carried out on the basis of egalitarian principles⁵. To ensure the independence of the choice of emotion terms in each language, the selection of the British English and Polish emotion terms was carried out independently. This meant that neither language was dominant over the other and that neither was used as a point of reference to the other, ensuring a selection of British English emotion terms that are independent of the Polish emotion terms. Various references and Internet sources were initially consulted to compile a list of 569 British English emotion terms and 322 Polish emotion terms. Then each of these emotion terms was placed in one of the following categories: happiness/joy, love, anger, hate, fear, sadness, surprise, compassion, contentment, guilt, shame and pride. The authors then had consultations to decide on the emotion terms that were the most prototypical of the original lists of emotion terms in the respective languages. Care was taken to ensure that each of the above emotion categories was represented as equally as possible. There were 200 emotion terms in the list of the most prototypical British emotions and 199 emotion terms in the list of the most prototypical Polish emotions.

⁵ Due to space restrictions, what is presented in this section is a summary of the procedure followed to select the emotion terms.

In the next phase, the final list of 135 emotions was determined on the basis of participant ratings of prototypicality. There were 29 Polish participants (mean age = 33.8 years, 15 females) and 22 British English participants (mean age = 51.4 years, 11 females). The emotion prototypicality ratings task was the same for the British English and Polish participants, who were required to rate each of the emotion terms (200 for the British English participants and 199 for the Polish participants) on sheets of paper on the basis of the extent to which they deemed them to be emotions on a 9-point scale. The 135 British English emotion terms and the 135 Polish emotion terms that had the highest mean prototypicality scores were selected as the emotion terms to be used in the main experiment.

8.2. Methodology

Participants were informed that the study was concerned with finding out about how people think some emotions “go together” and other emotions belong in different categories. More detailed instructions regarding the specific sorting task were as follows:

You will be presented with 135 emotions on the computer screen. We'd like you to sort these emotions into categories representing your best judgement about which emotions are similar to each other and which are different from each other. There is no one correct way to sort the emotions – make as few or as many categories as you wish and put as few or as many emotions in each group as you see fit. This study requires careful thought and you therefore need to think carefully about which category each emotion belongs to rather than just quickly put emotions in categories without much thought.

There were 58 British English participants (27 females, mean age = 42.7 years) and 58 Polish participants (27 females, mean age = 35.8 years).

8.3. NodeXL

We employ NodeXL (Smith et al. 2010), a network graph tool that is integrated in Microsoft Excel to perform analyses on social networks, to create graphical representations of the Polish and British English co-occurrence emotion matrices. The connection strength in our NodeXL graphs represent the co-occurring frequency of the emotion terms in the online emotions sorting data, and are hence referred to as *co-occurrences* or *interconnections*. For example, Figure 2, which is explained in more detail in the next section, shows that there is a co-occurrence value of 48 between *sympathy* and *empathy*, which is greater than the value of 28 between *empatia* ‘empathy’ and *sympatia* ‘sympathy, fondness’ (see Figure 3), meaning that there is a closer proximity between the English emotions than the corresponding Polish emotions. To illustrate more fully the nature of the co-occurrence matrices that are the basis of such diagrams, an example of the British English co-occurrence matrix pertaining to British English EMPATHY/SYMPATHY/COMPASSION cluster is shown in Figure 1. For example,

the relatively high co-occurrence value of 48 between *sympathy* and *empathy* in the co-occurrence matrix (Figure 1) is represented as a short line between these two emotions in the diagram in Figure two. This means that 48 British English participants placed these two emotions in the same group in the online emotions sorting task.

	Caring	Compassion	Concern	Empathy	Fondness	Love	Pity	Sympathy
Caring		43	26	37	42	31	19	37
Compassion	43		30	40	32	25	23	39
Concern	26	30		32	17	12	22	30
Empathy	37	40	32		29	22	28	48
Fondness	42	32	17	29		39	15	27
Love	31	25	12	22	39		9	21
Pity	19	23	22	28	15	9		30
Sympathy	37	39	30	48	27	21	30	

Figure 1: British English EMPATHY/SYMPATHY/COMPASSION cluster co-occurrence matrix

8.4. British English versus Polish Compassion

The value of the online emotions sorting methodology can be seen when one assesses what the analyses on compassion add to the GRID results on compassion. Specifically, the online emotions sorting results for compassion offer additional possible reasons for the more positive VALENCE of *compassion* in comparison with *współczucie* in the GRID results. Apart from the possible individualism-collectivism basis to this as the GRID results suggest, the online emotions sorting results point to two further possibilities, namely the structure of emotion clusters and the interconnection between *współczucie* and *politowanie* ‘compassion with a sense of superiority’.

The first point to note in the possible emotion cluster structure basis to the more positive VALENCE of *compassion* vis-à-vis *współczucie* is the meaning of *sympatia*, which possesses uniquely positive VALENCE in Polish and is more of a member of the Polish LOVE cluster rather than the Polish EMPATHY/COMPASSION cluster, which contrasts with the central position of *sympathy* in the British English EMPATHY/SYMPATHY/COMPASSION cluster. This is clearly demonstrated in the comparison in Figure 3 between the relatively higher co-occurrences between *sympatia* ‘sympathy, fondness’ and *miłość* ‘love, affection’ (32) and between *sympatia* ‘sympathy, fondness’ and *lubienie* ‘liking, fondness’ (43) vis-à-vis the lower interconnections between *sympatia* ‘sympathy, fondness’ and *empatia* ‘empathy’ (28) and between *sympatia* ‘sympathy, fondness’ and *współczucie* ‘compassion, sympathy’ (16), and the opposite pattern in Figure 2: lower interconnections between *sympathy* and *love* (21), and between *sympathy* and *fondness* (27), compared with higher interconnections between *sympathy* and *empathy* (48), and between *sympathy* and *compassion* (39). Although English

sympathy does include the positive elements of liking, when used in the negative sense it typically refers to Emotion Events expressing an irreversible loss (*grief*), similar to Pol. *współczucie*. The relatively less presence of *sympatia* in the Polish EMPATHY/COMPASSION cluster means that *współczucie* encompasses a greater semantic space than that covered by *compassion* in its respective cluster and *współczucie* is more negative as a consequence. This is most directly visible in the higher co-occurrences between *współczucie* and the Polish sadness cluster emotions in comparison with the interconnections between *compassion* and the British English sadness cluster emotions. It can also be seen in the higher interconnections between the LOVE cluster emotions and *compassion* in comparison with the MIŁOŚĆ cluster emotions and *współczucie*.

The second possible reason for more negative VALENCE effects for Polish *współczucie* centers on the linguistic link between *współczucie* and the unambiguously negative *politowanie* (pity combined with a sense of contempt and superiority). This is illustrated in the interconnection value of 19 between *współczucie* ‘compassion, sympathy’ and *politowanie* ‘compassion with a sense of superiority’ in Figure 3.

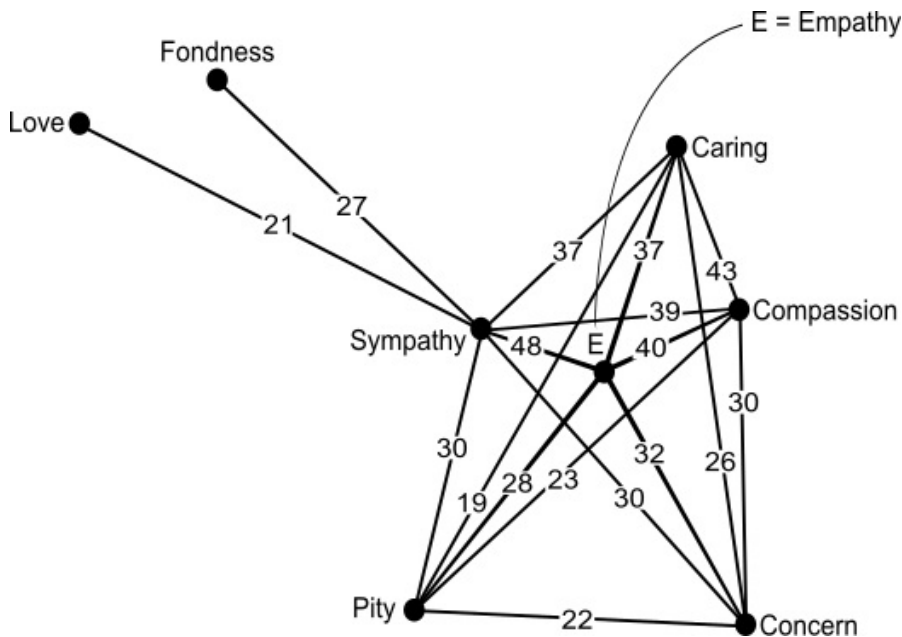


Figure 2: British English EMPATHY/SYMPATHY/COMPASSION Cluster

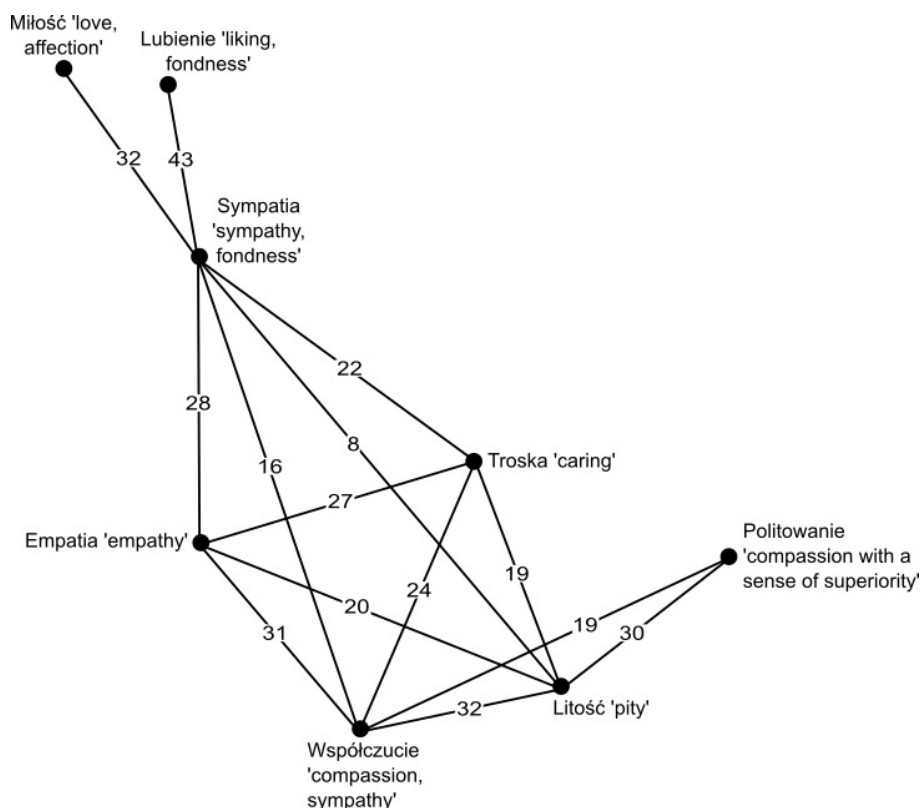


Figure 3: Polish EMPATHY/COMPASSION Cluster

9. Corpus Linguistics methodology

As can be inferred from the above discussion, emotions can be investigated by a number of distinct, but frequently overlapping or complementary methods. The linguistic methods applied in the present study employed the use of large corpus data of British English and Polish to identify quantitative similarities and differences in the statistics of language expressions containing the emotion terms studied, as well as Polish-to-English and English-to-Polish translated texts from parallel corpora, in which, apart from the nominal emotion concepts used, adjectival and verbal collocations were studied, as well as the Adjectival and Verbal emotion forms. This list is not exhaustive, again for practical reasons. The concepts of Cognitive Linguistics are applied to identify similarities and differences in particular emotion conceptualizations in English and Polish, particularly the construal of schemata pertaining to the expression of emotions and *metaphoric uses* with reference to the emotion concepts discussed.

Parallel Polish-to-English corpus data are presented in the paper to support the claim that a translated text involves firstly equivalent structures from the core, as well as successively more and more peripheral areas of the SL-TL similarity space. The value of the relevant *meaning granularity* and *schematicity* is also investigated in the TL text, and juxtaposed to the original. Examples are drawn from the concordance materials referring to a number of *event clusters* in English and Polish, inter alia *Emotion Events* (Lewandowska-Tomaszczyk and Wilson 2013), collocational profiles of relevant items, and confronted with the results of the construal qualitative analysis and questionnaire-based methodology, which reveals high interconnectivity links between terms and structures within the same cluster.

10. Lexical signalling of emotions

Not all emotions have evolved distinctive verbal correlates. In a number of cases what is expressed by verbal correlates are Emotion Clusters. As to the numbers of emotion terms present and emerging in language, our materials include a few hundred of forms in English and a comparable number in Polish. It may be interesting to note that there is a majority of negative emotion terms on the lists. People have developed to signal *some negative* emotions more frequently than other negative emotions or than positive ones mainly in the context of a lowered degree of control or manifestation of power relationships (pain, disgust, anger). Positive emotions are more frequently signalled in terms of *Emotion Clusters* and can cover consequential or associated properties such as, for example, the position of *love* within the LOVE, JOY and HAPPINESS cluster.

11. Cognitive Linguistics

11.1. Emotion prototypes

We assume that concepts in general have a *prototype-periphery* structure. Some of them are vague, meaning that their sense is not crisp and determinate. Some others are ambiguous in different contexts, that is their cognitive content undergoes modulation in context. Sense modulations can be a consequence of a selection of a particular *perspective* taken on a scene by the viewer/speaker, and, connected with it, the selection of a different *construal* of the observed event (cf. Langacker 1987) in the conceptual structure, namely the way a given event is portrayed by the language user. Lexical shifts in meaning are accounted for in cognitive linguistics by assuming a prototype conception of lexical concepts, with no set of necessary and sufficient features, but with fuzzy boundaries. On the other hand, semantic differences between different construal types expressed in terms of varying syntactic patterns, as say between an active and passive sentence or nominal and verbal constructions, are captured precisely by differences in selected syntactic construal elements of a scene.

11.2. Construal of emotions in figurative language

Both intra-domain (metonymic) and cross-domain mapping are cognitive linguistic instruments of focal importance that can be perceived as “entrenched conceptual patterns” in figure of thought, particularly metaphor (Grady 2007: 196). We also come across the phenomenon of *metaphonymy* such as *skakać z radości* - *jumping (with joy)* for *joy* (Goossens 2004), combining metaphor and metonymy.

We accept and use a meaning description in terms of semantic components and their physical physiological-psychological correlates, but make an attempt to enrich them with the analysis of *mental imagery* expressed in language, predominantly in terms of figurative uses. Metaphor is a cross-domain mapping, in which a conceptual domain is understood as a mental structure of related concepts expressing a body of knowledge. Metaphor refers to the understanding of one idea from a domain, or the whole domain (target domain), in terms of another (a source domain). A source domain is usually more basic and physically grounded, which functions as a mapping site for a given target domain; for example, the metaphor *Tom is a lion* is based on a source domain which includes the wide concept of lions, their looks, behavior, habitat, associations and evaluation, and can be taken as a source for a number of creative metaphorical mappings, although it is most frequently the mapping site for the conventional metaphor for Tom's courageous behavior. Emotions too can be mapped onto by a number of source domains, some of which are used, as will be discussed further, with *emotion* concepts.

Master metaphors (Lakoff et al. 1991) for emotions comprise a number of Source Domains. The basic one assumes that emotions are entities within a person, more particularly:

EMOTIONS ARE ENTITIES WITHIN A PERSON (Pol. *Napełniło mnie to radością*, Eng. *I was filled with joy*), EMOTIONS ARE PHYSICAL FORCES (Pol. *Wyladował na niej swoją złość*, Eng. *He discharged all his anger on her*.) or EMOTIONS ARE THINGS.

Whereas some emotion metaphors overlap in languages, some others possess a culture and language-bound character: for example, in comparison with Polish *dotknięty*, which corresponds to the feeling of *hurt*, English *touched* may signify *sorrow* and/or *compassion*. In a fuller account of emotion research methodology, cognitive theories of metaphor and metonymy developed by numerous scholars, also in Poland (e.g. Bierwiałczonek 2013), would need to be accounted for (compare comments to noun collocations in Table 1 and Table 2 in section 12.2 below). The focus of the present paper is primarily on the less frequently used emotion research methods and sources of data.

12. Language corpora and the quantitative data

12.1. Materials

In order to extend the context of the use of emotion terms in English and Polish, we resort to large corpus data, particularly collocations and their frequencies. By an-

alyzing authentic language, we can detect shifts in meaning for the same linguistic form and we can also describe the contexts which support such shifts. Based on the frequency of occurrence, corpus-based methods let us statistically determine which linguistic meanings are most salient. The materials we use come from several sources: the British National Corpus (100 million words) and the National Corpus of Polish (NKJP), which contains 300 million units of balanced data and a growing reference corpus exceeding one and a half billion units at present. We conduct automatic analyses of word frequencies and lexical (adjectival, verbal, and, when possible, nominal) collocations of the emotion words in spoken and written texts (Pęzik 2014). We also manually extract contexts of relevant words, the axiological charge of emotions (positive-negative) and relevant metaphors (cf. Lakoff and Johnson 1980, Lakoff and Kövecses 1987, Kövecses 2000). The manual analyses we conduct are annotations for metaphoricity and particular metaphorical scenario membership.

The BNC and NKJP used are either of a comparable size or normalized to identical values. We calculate how many times a word or collocation occurs per one hundred million words in a process called 'normalization'. This enables comparison across these differently sized large datasets. They cover well-balanced language materials of different genres and styles, including both written and spoken (ca. 10 %) conversational data. The search tools Wordsmith Tools, SlopeQ (<http://tnij.org/slopeq1>), and HASK were applied to generate frequencies of occurrences and (parts of speech-sensitive) collocations (Pęzik 2015). Corpus studies are carried out most commonly by looking at words and their contexts (so-called KWIC searches: Key Word In Context) in large collections of authentic natural language, compiled from written and oral sources. Quantitative data sets are compared to see if an observed phenomenon (e.g. a co-occurrence) is significantly more frequent than another; frequencies of lexical co-occurrences are also used for descriptive purposes and exemplification.

The overall frequency of all emotion terms is higher in Polish than in English, which can be interpreted as typological differences between the two languages with respect to the part-of-speech preference patterns rather than evidence of the linguistic preferences of the relevant speakers and writers⁶. Spoken language, although clearly marking emotional layers of meaning, will often perform this indirectly with a more constrained use of explicit emotion terms.

A word of caution referring to the reaction type frequencies should also be added in connection with the corpus methodology. Corpora have their limitations. There are

⁶ See Wierzbicka (1992, 1994) and Dziwirek and Lewandowska-Tomaszczyk (2010) for a discussion of the part-of-speech based differences of expressing emotions between Polish and English. As is also found in Lewandowska-Tomaszczyk and Wilson (2013), Polish emotion terms are more frequently expressed in some types of discourse as adjectives, while in their English (translational) equivalents it is the corresponding nominal structures that are preferred in the examined data as, for example, in: Lennie dropped his head *in shame* at having forgotten./Lennie spuścił głowę *zawstydzony* tym, że się zapomniał. He lowered his head *in shame*/ Opuścił głowę *zażenowany*. And yet, in larger samples, the proportions are different: BNC (100 mln segments): *ashamed* 1023, *with shame* 49, *in shame* 36, *of shame* 135; NKJP (ca. 240 mln segments): Adj *zawstydzony** 787, Prep N *ze wstydem* 223, *ze wstydu* 473).

tools available to generate frequencies of individual items and phrases, concordances with expanded contexts, collocations and keywords. There are also encouraging results of automatic metaphor identification (e.g. Stefanowitch and Gries 2007). However, when it comes to semantic and pragmatic annotations of meanings in use, particularly in large corpora, adequate corpus tools are not yet fully developed. Furthermore, one should be advised to exercise some caution with regard to far-reaching generalizations about direct correlation patterns between frequencies of use of particular items and basic cultural dimensions. Batteries of diverse multiple methods are recommended to capture both crucial and more nuanced inter-cultural preferences.

12.2. Collocations

The extraction of collocational combinations is an important methodological tool in corpus analysis. Collocations show the closest lexical units in terms of contextual distance, which are used with particular emotion words. The interesting property of this aspect of corpus methodology is the fact that, as consonant with the *Iconicity Principle* (Haiman 1985), closeness of forms signifies also closeness in meaning. This means that the closer a form is to another one in an utterance and the more frequent such combinations are in discourses, the closer are the semantic links between them.

We resort to the data generated by the collocator HASK and Colosaurus, comparing the collocations of synonymous forms, which are developed by Piotr Pęzik (2014) for the Polish and English corpora. HASK generates lists of all collocates based on the respective corpora and grouped according to main parts of speech: Nouns, Adjectives, and Verbs, and with respect to the term position referring to the collocate (preceding or following). The collocation window ranges from three positions to the left and to the right and the minimal frequency is set to 5 occurrences of the phrase. Tables 1 and 2 below present top noun collocates of the Adjectives *współczujący* ‘compassionate’ and *compassionate*, respectively, in English and Polish.

A number of *association scores* are computed for each collocational combination: *t*-score, mutual information, chi-square, log-log, Dunning’s log-likelihood, and Dice score⁷

#	Collocate	POS	A	TTEST	MI3	English equivalent
1	spojrzenie	noun	29.0	5.3	17.2	look
2	Serce	noun	17.0	4.0	13.4	heart
3	Twarz	noun	10.0	2.9	10.8	face
4	Głos	noun	10.0	2.8	9.7	voice
5	Mina	noun	8.0	2.7	10.8	facial gesture
6	Miłość	noun	6.0	2.2	8.9	love
7	Ton	noun	5.0	2.1	10.5	tone
8	Gest	noun	5.0	2.1	10.1	gesture
9	westchnienie	noun	4.0	1.9	11.9	sigh
10	rozumienie	noun	4.0	1.9	9.1	understanding

Table 1. Top noun collocates of Polish *współczujący* ‘compassionate’

⁷ All association scores can be generated with the publicly available tools at <http://pelcra.clarin-pl.eu/>.

The noun collocates of Pol. *współczujący* ‘compassionate’ in Table 1, particularly those referring to body parts, are characteristic for the construction of the metonymic portrayal of compassion, e.g. *miał współczujące serce dla każdej niedoli* ‘he had a compassionate heart for all the misery’, *szukając pokrzepienia i ucieczki w cierpliwej, współczującej twarzy swojej kobiety* ‘searching reinvigoration and escape in his woman’s patient, compassionate face’. A similar regularity can be identified in English, e.g. *Andrew’s compassionate brown eyes registered shock, and then regret* (Table 2).

#	Collocate	POS	A	TTEST	MI3	Polish equivalent
1	Leave	N%	25.0	4.9	16.5	urlop okolicznościowy
2	Grounds	N%	7.0	2.5	10.1	negatywne okoliczności dla przyznania wizy imigracyjnej
3	circumstance	N%	4.0	1.7	6.8	negatywne okoliczności dla przyznania urlopu (uniwersytet)
4	Attitude	N%	3.0	1.4	5.6	postawa współczująca
5	look/eye	N%	3.0	1.3	5.5	spojrzenia/wzrok współczujące/y
6	Home	N%	4.0	1.0	5.1	dom pełen współczucia
7	Man	N%	7.0	0.8	6.1	współczujący człowiek
8	Woman	N%	3.0	0.0	3.2	współczująca kobieta

Table 2. Top noun collocates of English *compassionate*

As observed in the corpus data, Polish *współczując** (-y, -a, -e) has a higher frequency of use⁸ and a wider range of, more general, applications, while English, apart from describing people’s and communities’ (*home*) attitude, is clearly more specialized – used as part of terminological phraseology, applicable in a number of cases in which official policy is to help society (employees, immigrants, students). This active attitude in English is also marked in the results of the GRID method.

13. Corpus data interpretation

There are certain difficulties in coding emotion scenario identification. Unlike the GRID methodology in which the questionnaire is administered to native speakers of a language to get their native judgments, in the corpus data, language involving an emotion is used with reference to a variety of different experiencers – not only English in English texts or Polish in the Polish corpus data. Corpus examples do not necessarily describe Polish or English people’s experiences of this emotion. The experiencers vary – they can be from different cultures and backgrounds. Nevertheless, they are described in Polish and English, respectively, in terms of an English or a Polish perception of this emotion. When we use corpus data, we assume that (native) speakers use their own language-specific conceptualizations even when they talk about and describe other people or events they have not personally experienced.

⁸ One of the reasons, although not significant, of the higher raw frequency is the fact that the participial form *współczując*, Eng. ‘sympathizing/expressing compassion (participle)’, is also included in the count.

14. Spoken corpus and linguistic emotion expression

To investigate the processes of linguistic emotion expression an analysis of spoken corpus materials is also undertaken. Particular phraseology and lexical uses (e.g. *You dirty old bugger; leave her alone!; You dirty little faggot! Call the manager! I'm tired of listening to this*) correspond to clusters of emotional meanings (*anger* in this case) and uncover possible states of arousal in terms of emotional language use. Another spoken language corpus tool identifies particular utterance *prosodic profiles* (Pęzik 2015). The prosodic profiles generated for utterances make it possible to identify the intensity of emotion expression in particular contexts in terms of pitch dynamics and amplitude. The observation we make with reference to prosodic differences (in a number of cases more intensive emotion expression for some Polish emotions) can be interpreted with reference to the theory of *politeness* profiles (Brown and Levinson 1987; Scollon 1996; Scollon and Wong Scollon 1995), which differ in Polish and English. So-called positive politeness in Polish culture involves preferences for full personal emotion expression, while English *negative politeness*, more individualistically inclined, generates an interactional attitude that is relatively less imposing.

15. Contrastive analysis and parallel corpus

Translation involves a number of cycles of the *re-conceptualization* of an original source language (SL) message, expressed eventually in a target language (TL) (Lewandowska-Tomaszczyk 2010). The translator's *construal* of an original scene contributes to a modified (TL directed) mental model of the original SL scene. Translation thus has to do with various *re-conceptualization* processes, which are expressed in terms of *qualitative* and *quantitative parameters*. Qualitative changes primarily cover differences between Source Culture and Target Culture world-views, related to properties specific to a given culture. Apart from the *qualitative changes*, *quantitative linguistic* parameters such as the frequency of occurrence of a language form, its combinatorics with other items in discourse, as well as patterns of semantic similarity, oppositeness and inclusion, all contribute to the language specific character of SL and TL forms. Collocational patterns studied cross-linguistically are revealing when analyzed in parallel (translated) texts. They contribute to the identification of typological differences between text types of the languages compared, as well as those between the original and the translated text.

15.1. PARALELA tools

We resort to fuller contexts and, apart from the analysis of words and sentences, we investigate equivalence patterns as discussed above, looking primarily at the Polish-to-English and English-to-Polish parallel corpora as collected at the PELCRA Corpus and Computational Linguistics Lab of the Institute of English Studies at Lodz Univer-

sity (pelcra-clarin.pl.eu). The corpora represent a variety of genres and styles; however, it is primarily literary texts that provide the most extensive and varied language materials for emotion discussion. Recorded spoken materials, mainly spontaneous conversational data, provide additional enrichment of the direct properties of emotional language. However, as they are non-surreptitiously recorded, the range of emotional expressions and topics might not be equivalent to natural, unrecorded language.

PARALELA (<http://paralela.clarin-pl.eu/>) is a corpus search engine developed in the CLARIN-PL project for a large collection of annotated Polish-English parallel texts (Pęzik 2014, 2015). The publicly available Paralela index currently contains over 200 m words in close to 6 m parallel segments from Polish-English and English-Polish text translations. More than 8 million words in the corpus were manually aligned and annotated for equivalence types. PARALELA supports both mono- and bilingual corpus queries formulated in the SlopeQ query syntax. Search results can be filtered with facets and query filters. Since both the Polish and English segments in the corpus were part-of-speech tagged, it is also possible to formulate queries for terms and complex phrases matching specific morphosyntactic criteria. For example, a concept corresponding to Polish *współczucie* obtains more granular shades of meaning in translational context, either in terms of its cluster polysemic chain including *litość* ‘pity’ or *troska* ‘concern’ or can be replaced by a hyperonym (e.g. the word *uczucie* ‘emotion’) as in:

(1) Pol. Halleck wahał się, wyczuwając mądrość i współczucie w słowach Tueka.

(1a) Eng. Halleck hesitated, sensing wisdom and sympathy in Tuek’s words.

(2) Pol. Ogarnęło go niezmierne współczucie dla niej.

(2a) Eng. A profound compassion for her swept through him.

(3) Pol. nigdy nie skąpił swojego pobłażliwego współczucia

(3a) Eng. he had always been treated with indulgent commiseration

(4) Pol. w jego głosie niewiele było współczucia

(4a) Eng. without sounding sorry

(5) Pol. Aragorn popatrzył na niego łagodnie i ze współczuciem.

(5a) Eng. Aragorn looked at him with kindly with pity.

(6) Pol. W jego oczach jakoś nie było współczucia.

(6a) Eng. There was no empathy in his eyes.

(7) Pol. nie znajdowała dla niego współczucia

(7a) Eng. she had been rendered incapable of any emotion for her son

The results of our analysis confirm the thesis of *cluster equivalence* (Lewandowska-Tomaszczyk 2017a, 2017b) manifested in translational data. Cluster equivalence is different from that found in lexicographic reference sources. Instead of word-based equivalence patterns as identified in dictionaries, it assumes the emer-

gence of patterns based on the resemblance of clusters of concepts, manifesting different forms of internal composition. These facts further contribute to the thesis concerning translational reconceptualizations of meaning and can account for partial incommensurability (Lakoff 1987) and untranslatability of texts in distinct languages.

16. Conclusions: Semantic approximation and emotion clustering

Emotion terms are, according to the principle of Semantic Approximation, discussed in Section 2, inter-substitutable in discourse and in translation within the same broad emotional cluster of concepts. At the same time, this inter-substitution brings about definite effects. The replacement of *pity* for *compassion*, *joy* for *love* or *anxiety* for *fear* in one language, or rendering the English form *pride* into Polish *pycha* or choosing the Polish term *duma*, introduces, as predicted by the Re-conceptualization Principle (Section 15), a substantial conceptual change or modification, in other words, a reconceptualization of the original sense.

Conceptual EMOTION clusters themselves enter larger constructions with other clusters to form various forms of EMOTIONAL, COGNITIVE and/or VOLITIONAL clusters (Dziwirek and Lewandowska-Tomaszczyk 2010). In the case of *compassion*:

Prototypical *compassion* Emotion concept

Eng. *sympathise with sb* > partly corresponds to < Pol. *współczuć komuś* in terms of the COMPASSION/WSPÓŁCZUCIE clusters

Eng. *compassion* and *feel compassion with/for* > corresponds to Pol. *współczucie* and *współczuć*, respectively

Megacluster of emotive and cognitive meanings

Pol. *Współczuję, że*; Eng. */I am sorry that/I sympathise about*

Megacluster of emotive, cognitive and volitional meanings

Eng. *sympathy* > *sympathize* – identified in the polysemic Eng. COMPASSION/LOVE megacluster – corresponds, respectively, to Pol. *sympatia* and *Rozumiem, popieram ten pomysł/zachęcam do działań w tym kierunku* ‘I sympathize (i.e. understand, support, align, encourage) with the idea’

17. Conclusions: Multiple methodologies and sources of data

The complexity inherent in the mapping of emotion concepts between cultures demands the employment of complementary methodologies to identify the nuanced differences that exist between, for example, British English *compassion* and Polish *współczucie*. The emotion terms and features of the GRID instrument allow emotion prototype structure to be analyzed on the basis of dimensions and components. With respect to the former, we have shown that *compassion* has a more positive VALENCE than *współczucie*. In terms of the latter, further analyses on the emotion features in the

sub-component *desire to act* vs. *desire not to act* produced consistent results demonstrating that *compassion* is characterized relatively more than *współczucie* by the wish to engage in action. However, other paradigms, such as the online emotion sortings methodology and the cognitive corpus linguistics methodology, are necessary to determine possible reasons underlying these effects. The online emotions sorting methodology provides enriched information about the meaning of emotions by analyzing the relationships between emotions both within and across emotion clusters. For example, apart from showing, consistent with the GRID results, that *compassion* has a more positive VALENCE than *współczucie* on the basis of the higher interconnections between the British English LOVE cluster emotions and *compassion* and the lower co-occurrences between the British English SADNESS cluster emotions and *compassion* relative to the corresponding relationships for *współczucie*, the online emotion sortings results offer possible explanations for this pattern. The first of these centers on the wider semantic space, including negative elements, covered by *współczucie* in its respective cluster as a consequence of the relative greater inclusion of *sympatia* in the the MIŁOŚĆ (love) cluster. The conceptual proximity between *współczucie* and the negative *politowanie* offers a second possible reason.

The corpus data also uncover interesting correlations. Polish *współczucie* ‘compassion’ is used with respect to a wider range of referents and contexts, while English *compassion*, as observed in the selection of collocates of the relevant adjectives (Tables 1 and 2), is shown to be used in terminology addressing people’s needs. These corpus results are corroborated in the GRID method and the online sorting task.

Due to the reasons of space the present paper does not aspire to present all existing methods of emotion research. We did not deal in detail with cognitive linguistics findings in extensive metaphor and metonymy studies (Lakoff, Kövecses 1987; Bierwiałczonek 2013; Mikołajczuk 2009). Nor did we specify, for example, prosodic and paralinguistic properties accompanying spoken language in emotional interaction. Rather, the focus of the paper was a presentation and discussion of the less frequent methodologies (GRID, online emotions sorting, and corpus analysis) we have been employing to shed a new light on emotion concepts, in particular the meanings of Polish and English compassion clusters in contrast.

To suggest a wider perspective, it can be concluded that the research experience in the study of cross-linguistic emotion concepts strongly supports a multimethod approach to investigate such complex phenomena, and present ways of taking advantage of qualitative and quantitative sources of data from linguistic, psychological and culture studies paradigms. Attempts to uncover nuances of a relationship between external reality world, mind, language, and culture with respect to the concepts of the ‘third kind’ such as emotions, in which concrete, physical and biological sensations closely intertwine with more abstract, cognitive phenomena and their linguistic and culture-bound expression, will only be successful when such an approach is adopted. The present paper highlights some new aspects of research in emotion systems in contrast. An answer to the question as to whether it is possible to develop a fully comprehensive model of emotions that would account for all cognitive, social, cultural and lin-

guistic conditioning as well as interrelationships between all emotion clusters in one language or across languages, needs further detailed analyses and certainly exceeds the scope of the present discussion.

References

- BIERWIACZONEK Bogusław (2013): *Metonymy in language, thought and brain*. – Sheffield: Equinox.
- BOND Michael H. (2002): Reclaiming the individual from Hofstede's ecological analysis – A 20-year odyssey: Comment on Oyserman et al. – *Psychological Bulletin* 128: 73–77.
- BREWER Marillyn B., CHEN Ya-Ru. (2007): Where (Who) are collectives in collectivism? Toward conceptual clarification of individualism and collectivism. – *Psychological Review* 114: 133–151.
- BROWN Penelope, LEVINSON Stephen C. (1987): *Politeness: Some universals in language usage*. – Cambridge: Cambridge University Press.
- COSMIDES Leda, TOOBY John (2000): Evolutionary psychology and the emotions. – [In:] Michael LEWIS, Jeannete M. HAVILAND-JONES (eds.): *Handbook of emotions* (2nd ed). – New York: Guilford Press, 91–115.
- DAMEN Louise (1987): *Culture learning: The fifth dimension on the language classroom*. – Reading, MA: Addison-Wesley.
- DIENER Ed, DIENER Marissa (1995): Cross-cultural correlates of life satisfaction and self-esteem. – *Journal of Personality and Social Psychology* 68: 653–663.
- DZIWIREK Katarzyna, LEWANDOWSKA-TOMASZCZYK Barbara (2010): *Complex emotions and grammatical mismatches*. – Berlin: Mouton de Gruyter.
- EARLEY P. Christopher, GIBSON Cristina B. (1998): Taking stock in our progress on individualism – Collectivism: 100 years of solidarity and community. – *Journal of Management* 24: 265–304.
- EKMAN Paul (2004): *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life*. – New York: Henry Holt.
- ELLSWORTH Phoebe C., SCHERER Klaus R. (2003): Appraisal processes in emotion. – [In:] Richard J. DAVIDSON, Klaus R. SCHERER, H. Hill GOLDSMITH (eds.): *Handbook of affective sciences*. – New York: Oxford University Press, 572–595.
- FAUCONNIER Gilles, TURNER Mark (1998): Conceptual integration networks. – *Cognitive Science* 22: 2, 133–187.
- FISKE Alan P. (2002): Using individualism and collectivism to compare cultures – A critique of the validity and measurement of the constructs: Comment on Oyserman et al. – *Psychological Bulletin* 128, 78–88.
- FONTAINE Johnny J.R. (2013): Dimensional, basic emotion, and componential approaches to meaning in psychological emotion research. – [In:] Johnny J.R. FONTAINE, Klaus R. SCHERER, Cristina SORIANO (eds.): *Components of emotional meaning: A sourcebook*. – Oxford: Oxford University Press, 31–45.
- FONTAINE Johnny J.R., POORTINGA Ype H., SETIADI Bernadette, MARKAM Suprapti S. (2002): Cognitive structure of emotion terms in Indonesia and The Netherlands. – *Cognition and Emotion* 16 (1), 61–86.
- FONTAINE Johnny J.R., SCHERER Klaus R., ROESCH E.B., ELLSWORTH Phoebe C. (2007): World of emotions is not two-dimensional. – *Psychological Science* 18 (12), 1050–1057.
- FONTAINE Johnny J.R., SCHERER Klaus R. (2013): The global meaning structure of the emotion domain: Investigating the complementarity of multiple perspectives on meaning. – [In:] Johnny J.R. FONTAINE, Klaus R. SCHERER, Cristina SORIANO (eds.): *Components of emotional meaning: A sourcebook*. – Oxford: Oxford University Press, 106–125.

- FONTAINE Johnny R. J., SCHERER Klaus R., SORIANO Cristina (2013): *Components of emotional meaning: A sourcebook*. – Oxford: Oxford University Press.
- GEERAERTS Dirk (2002): The scope of diachronic onomasiology. – [In:] Vilmos AGEL, Andreas GARDT, Ulrike HASS-ZUMKEHR, Thorsten ROELCKE (eds.): *Das Wort. Seine strukturelle und kulturelle Dimension. Festschrift für Oskar Reichmann zum 65. Geburtstag*. – Tübingen: Niemeyer, 29–44.
- GENDRON Maria, ROBERTSON Debi, VAN DER VYVER Jacoba Marietta, BARRETT Lisa Feldman (2014): Perceptions of emotion from facial expressions are not culturally universal: Evidence from remote culture. – *Emotion* 14 (2), 251–262.
- GLADKOVA Anna (2010): Sympathy, compassion, and empathy in English and Russian: A linguistic and cultural analysis. – *Culture and Psychology* 16 (2), 267–285.
- GOOSSENS Luis (2004): Metaphonymy: The interaction of metaphor and metonymy in expressions for linguistic action. – [In:] Rene DIRVEN, Ralf PORINGS (eds.): *Metaphor and metonymy in comparison and contrast*. – Berlin: Mouton de Gruyter, 349–378.
- GRADY Joseph E. (2007): Metaphor. – [In:] Dirk GEERAERTS, Hubert CUYCKENS (eds.): *The Oxford handbook of Cognitive Linguistics*. – Oxford: Oxford University Press, 187–213.
- HAIMAN John (ed.) (1985): *Iconicity in syntax*. – Amsterdam–Philadelphia: John Benjamins.
- HOFSTEDE Geert (1980): *Culture's consequences: International differences in work-related values*. – Beverly Hills: Sage.
- HOFSTEDE Geert (2016): *Country Comparison*. – <https://www.geert-hofstede.com/countries.html>, accessed 12.02.2016.
- ISHII Keiko (2013): The meaning of happiness and in Japan and the United States. – [In:] Johnny J.R. FONTAINE, Klaus R. SCHERER, Cristina SORIANO (eds.): *Components of emotional meaning: A sourcebook*, Oxford: Oxford University Press, 473–476.
- KATZ Jerrold J. (1978): Effability and Translation. – [In:] Franz GEUENTHNER, M. GEUENTHNER (eds.): *Meaning and translation. Philosophical and linguistic approaches*. – New York: University Press, 191–234.
- KLUCKHOHN Clyde, KELLY William H. (1945): The concept of culture. – [In:] Ralph Linton (ed.): *The science of man in the world culture*, New York: Columbia University Press, 78–105.
- KÖVECSES Zoltán (2000): *Metaphor and emotion: Language, culture, and body in human feeling*. – Cambridge: Cambridge University Press.
- KRAWCZAK Karolina (2014): Shame, embarrassment and guilt: Corpus evidence for the cross cultural structure of social emotions. – *Poznań Studies in Contemporary Linguistics* 50, 441–475.
- LAKOFF George (1987): *Women, fire, and dangerous things: What categories reveal about the mind*. – Chicago: University of Chicago Press.
- LAKOFF George (1996): *Moral politics – How liberals and conservatives think*. – Chicago: Chicago University Press.
- LAKOFF George, JOHNSON Mark (1980): *Metaphors we live by*. – Chicago: Chicago University Press.
- LAKOFF George, KÖVECSES Zoltán (1987): The cognitive model of anger inherent in American English. – [In:] Dorothy HOLLAND, Naomi QUINN (eds.): *Cultural models in language and thought*. – Cambridge: Cambridge University Press, 195–221.
- LAKOFF George, ESPENSON Jane, SCHWART Alan (1991): *Mater metaphor list* (Second Edition). – Berkeley, CA: Cognitive Linguistics Group University of California at Berkeley.
- LAKOFF George, JOHNSON Mark (1999): *Philosophy in the flesh – The embodied mind & its challenge to western thought*. – New York: Basic Books.
- LANGACKER Ronald W. (1987, 1991): *Foundations of cognitive grammars*, vol. I and vol. II. – Stanford: Stanford University Press.
- LANGACKER Ronald W. (2014): Culture and cognition, lexicon and grammar. – [In:] Masataka YAMAGUCHI, Dennis TAY, Benjamin BLOUNT (eds.): *Approaches to language, culture and cognition: The intersection of cognitive linguistics and linguistic anthropology*. – New York: Palgrave MacMillan, 27–49.

- LEWANDOWSKA-TOMASZCZYK Barbara (2010): Re-conceptualization and the emergence of discourse meaning as a theory of translation. – [In:] Barbara LEWANDOWSKA-TOMASZCZYK, Marcel THELEN (eds.): *Meaning in translation*. – Frankfurt a. Main: Peter Lang, 105–148.
- LEWANDOWSKA-TOMASZCZYK Barbara (2011): Transcultural patterns in communication. – [In:] Adam BEDNAREK, Iwona WITCZAK PLISIECKA (eds.): *Interdisciplinary approaches to communication studies*. – Łódź: Łódź International Studies Academy Press, 7–22.
- LEWANDOWSKA-TOMASZCZYK Barbara (2012): Explicit and tacit – An interplay of the qualitative and quantitative approaches to translation. – [In:] Michael OAKES, Meng JI (eds.): *Quantitative methods in corpus-based translation studies*. – Amsterdam: Benjamins, 3–33.
- LEWANDOWSKA-TOMASZCZYK Barbara (2017a): Cluster equivalence, general language, and language for specific purposes. – [In:] Marcin GRYGIEL (ed.): *Cognitive approaches to specialist language*. – Newcastle upon Tyne: Cambridge Scholars Publishing, 384–418.
- LEWANDOWSKA-TOMASZCZYK Barbara (2017b) Partial perception and approximate understanding. – *Research in Language* 15 (20), 129–152.
- LEWANDOWSKA-TOMASZCZYK Barbara, WILSON Paul A. (2013): English ‘fear’ and Polish ‘strach’ in contrast: GRID approach and Cognitive corpus linguistic methodology. – [In:] Johnny FONTAINE, Klaus SCHERER, Cristina SORIANO (eds.): *Components of emotional meaning: A sourcebook*. – Oxford: Oxford University Press, 425–436.
- LEWANDOWSKA-TOMASZCZYK Barbara, WILSON Paul A. (2016): Compassion, empathy and sympathy expression features in affective robotics. 7th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), 65–70, doi: 10.1109/CogInfoCom.2016.7804526.
- LEWANDOWSKA-TOMASZCZYK Barbara, WILSON Paul A. (in press): Compassion cluster expression features in affective robotics from a cross-cultural perspective. – [In:] Peter BARANYI, Richard KLEMPOUS, Jan NIKODEM (eds.): *Cognitive Infocommunications, theory and applications*, Singapore: Springer, 202–227.
- MARKUS Hazel Rose, KITAYAMA Shinobu (1991): Culture and the self: Implications for cognition, emotion, and motivation. – *Psychological Review* 20, 568–579
- MIKOŁAJCZUK Agnieszka (2009): *Obraz radości we współczesnej polszczyźnie*. – Warszawa: Wydawnictwo Semper.
- MIKOŁAJCZUK Agnieszka (2012): A cross-linguistic perspective (Polish vs. English) on the conceptualisation of zadowolenie (satisfaction/being glad, contentment, pleasure). – [In:] Paul A. WILSON (ed.): *Dynamicity in emotion concepts*, Łódź: Studies in Language, vol. 25, Frankfurt am Main: Peter Lang, 325–347.
- MIKOŁAJCZUK Agnieszka (in preparation): *WSTYD i DUMA. Z badań nad konceptualizacją przeżyć samowiedomościowych w języku polskim*, Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.
- MORTILLARO Marcello, RICCI-BITTI Pio E., BELLETTI Guglielmo, GALATI Dario (2013): Pride is not created equal: Variations between Northern and Southern Italy. – [In:] Johnny J.R. FONTAINE, Klaus R. SCHERER, Cristina SORIANO (eds.): *Components of emotional meaning: A sourcebook*. – Oxford: Oxford University Press, 366–376.
- NIEDENTHAL Paula, KRAUTH-GRUBER Silvia, RIC François (2006): *Psychology of emotion*. – New York: Psychology Press.
- NOWAKOWSKA-KEMPNA Iwona (1995): *Konceptualizacja uczuć w języku polskim*. – Warszawa: WSP; Wiedza Powszechna.
- OSGOOD Charles E., MAY William, MIRON Murray S. (1975): *Cross-cultural universals of affective meaning*. – Urbana, IL.: University of Illinois Press.
- OYSERMAN Daphna, COON Heather M., KEMMELMEIER Markus (2002): Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analysis. – *Psychological Bulletin* 128, 3–72.

- PEZIK Piotr (2014): Graph-based analysis of collocational profiles. – [In:] Vida JESENŠEK, Peter GRZYBEK (eds.): *Phraseologie im Wörterbuch und Korpus/Phraseology in dictionaries and corpora*, ZORA 97. – Maribor–Bielsko–Biała–Budapest–Kansas–Praha: Filozofska fakuteta, 227–43.
- PEZIK Piotr (2015): Spokes – a Search and Exploration Service for Conversational Corpus Data. – [In:] *Selected papers from CLARIN 2014*, 99–109. Linköping Electronic Conference Proceedings. Linköping University Electronic Press, Linköpings universitet. http://www.ep.liu.se/ecp_article/index.en.aspx?issue=116;article=009.
- REIPS Ulf Dietrich (2002): Standards for internet-based experimenting. – *Experimental Psychology* 49, 243–256.
- SCHERER Klaus R. (1984): On the nature and function of emotion: a component process approach. – [In:] Klaus SCHERER, Paul EKMAN: *Approaches to emotion*. – Hillsdale, NJ: Erlbaum, 293–317.
- SCHERER Klaus R. (2005): What are emotions? And how can they be measured? – *Social Science Information* 44, 693–727.
- SCHIMMACK Ulrich, OISHI Shigehiro, DIENER Ed (2005): Individualism: A valid and important dimension of cultural differences between nations. – *Personality and Social Psychology Review* 9: 17–31.
- SCOLLON Ron (1996): Discourse identity, social identity, and confusion in intercultural communication. – *Intercultural Communication Studies* VI: 1, 1–16.
- SCOLLON Ron, WONG SCOLLON Suzanne (1995): *Intercultural communication: A discourse approach*. – Oxford: Basil Blackwell.
- SHAVER Phillip, SCHWARTZ Julia, KIRSTON Donald, O’CONNOR Cary (1987): Emotion knowledge: Further exploration of a prototype approach. – *Journal of Personality and Social Psychology* 52, 1061–1086.
- SHAVER Phillip R., WU Selley, SCHWARTZ Julia C. (1992): Cross-cultural similarities and differences in emotion and its representation. – [In:] Margaret S. CLARK (ed.): *Emotion: Review of personality and social psychology*. – Newbury Park, CA: Sage, 175–213.
- SMITH Mark, CENI Arber, MILIC-FRAYLING Natasa, SHNEIDERMAN Ben, MENDES Eduarda Rodrigues, LESKOVEC Jure, DUNNE Cody (2010): NodeXL: A free and open network overview, discovery and exploration add-in for Excel 2007/2010/2013/2016, <http://nodexl.codeplex.com/> from the Social Media Research Foundation, <http://www.smrfoundation.org>.
- STEFANOWITCH Anatol, GRIES Stefan Th. (2007): *Corpus-based approaches to metaphor and metonymy*. – Berlin: Mouton de Gruyter.
- TRIANDIS Harry C. (1995): *Individualism and collectivism*. – Boulder, CO: Westview Press.
- TRIANDIS Harry C. (2001): Individualism-collectivism and personality. – *Journal of Personality* 69, 907–924.
- WIERZBICKA Anna (1985): Different cultures, different languages different, speech acts: English vs. Polish. – *Journal of Pragmatics* 9, 145–178.
- WIERZBICKA Anna (1992): *Semantics, culture and cognition: Universal human concepts in culture-specific configurations*. – Oxford: Oxford University Press.
- WIERZBICKA Anna (1994). Emotion, language and cultural scripts. – [In:] Shinobu KITAYAMA, Hazel Rose MARKUS (eds.): *Emotion and culture: Empirical studies of mutual influence*. – Washington, D.C.: American Psychological Association, 133–197.
- WIERZBICKA Anna (1999): *Emotions across languages and cultures: Diversity and universals*. – Cambridge: Cambridge University Press.
- WIERZBICKA Anna (2004): Preface: ‘Bilingual lives, bilingual experience’. – Special issue of the *Journal of Multilingual and Multicultural Development* 25 (2/3), 94–104.
- WIERZCHOWSKI Józef (1980): *Semantyka językoznawcza*. – Warszawa: PWN.
- WILSON Paul (ed.) (2012): *Dynamicity of emotion concepts* [Lodz Studies in Language vol. 25]. – Frankfurt a. Main: Peter Lang.

- WILSON Paul, LEWANDOWSKA-TOMASZCZYK Barbara (2012): The Nature of emotions. – [In:] Paul W. WILSON: *Dynamicity in emotion concepts* [Lodz Studies in Language vol. 25]. – Frankfurt a. Main: Peter Lang, 13–36.
- WILSON Paul A., LEWANDOWSKA-TOMASZCZYK Barbara, NIJYA Yu (2013): Happiness and contentment in English and Polish. – [In:] Johnny FONTAINE, Klaus SCHERER, Cristina Soriano (eds.): *Components of emotional meaning: A sourcebook*. – Oxford: Oxford University Press, 477–481.
- WILSON Paul, LEWANDOWSKA-TOMASZCZYK Barbara (2014): Affective robotics: Modelling and testing cultural prototypes. – *Cognitive Computation* 6 (4), 814–840.
- WILSON Paul A., LEWANDOWSKA-TOMASZCZYK Barbara (2017): Pride in British English and Polish: A cultural-linguistic perspective. – [In:] Farzad SHARIFIAN (ed.): *Advances in Cultural Linguistics*, Singapore: Springer, 247–288.
- YIK Michelle S.M., RUSSELL James A., BARRETT Lisa Feldman (1999): Structure of self-reported current affect: Integration and beyond. – *Journal of Personality and Social Psychology* 77, 600–619.

doi: <http://dx.doi.org/10.17651/bptj.74.13>

ANNA KOZŁOWSKA

Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie

Źródła w badaniach idiolektów

Streszczenie

Artykuł wskazuje źródła wykorzystywane w różnego rodzaju badaniach idiolektów oraz omawia rolę poszczególnych typów źródeł. Choć głównym przedmiotem zainteresowania są autonomiczne studia nad językiem jednostek, praca uwzględnia również badania idiolektów podporządkowane opisowi języka ogólnego oraz jego odmian, a także studia z zakresu historii języka i psycholingwistyki. Pokazana jest zależność wykorzystywanych w badaniach źródeł od rodzaju i celów opisu, a także od obecnego w nim sposobu rozumienia idiolektu.

Słowa kluczowe: idiolekt, język jednostki, język pisarza, językoznawstwo, koncepcja języka

Sources in idiolect research

Abstract

The article addresses language data sources used in various types of idiolect research. It discusses the role of particular types of sources in these studies. Despite being focused on autonomous studies on the language of individuals, the article also takes into account research on idiolects that is subordinate to the description of standard language and its varieties, as well as studies in the history of language and psycholinguistics. The article presents the dependence of the sources on the types and purposes of the description and on the way idiolect is understood.

Keywords: idiolect, language of an individual, language of a writer, linguistics, concept of language

1. Idiolekt i jego opis

1.1. Pojęcie idiolektu i idiostylu

Celem niniejszego artykułu jest przegląd różnych rodzajów źródeł wykorzystywanych w opisach idiolektów oraz ocena ich roli i przydatności w tych badaniach. W tytule nie bez powodu pojawiło się określenie *idiolekt*, a nie równie popularny,

jeśli nie popularniejszy we współczesnych pracach *idiostyl* – wbrew częstej praktyce utożsamiania ze sobą dwóch kryjących się za tymi leksemami pojęć, łączenia ich w jeden koncept (*idiolektostyl*, zob. Kudra 2006), a nawet podważania zasadności tego typu dystynkcji terminologicznych, proponuję ich rozróżnianie. I tak przez *idiolekt* (*język osobniczy, język indywidualny, mowę jednostkową*) będę rozumiała repertuar środków językowych, jakimi dysponuje mówiący, podczas gdy *idiostyl* (zwany inaczej *stylem osobniczym, stylem indywidualnym, stylem autora czy jednostki*) – to sposób organizacji jego różnego typu wypowiedzi. Różnice między idiolektem a idiostylem dotyczą więc zarówno statusu zjawiska, jak i jego zakresu. Idiolekt to potencja, a zatem byt statyczny, niejednorodny i abstrakcyjny, natomiast idiostyl – jego funkcjonalna odmiana, ukonkretniona i zależna od kontekstu (zob. też Witosz 2009a). W takim ujęciu idiolekt współtworzą rozmaicie sprofilowane funkcjonalnie idiostyle, czyli indywidualne sposoby kształtowania określonego rodzaju tekstów: artystycznych, naukowych, użytkowych, publicystycznych itp. Idiolekt nie jest jednak ich sumą, ponieważ mieści się w innej sferze: odwołując się do tradycyjnej terminologii lingwistycznej, należałoby wiązać idiolekt z de Saussure’owskim pojęciem *langue individuelle* (zob. Wołowska 2013), czyli indywidualnym „skarbcem” każdego użytkownika, a idiostyl lokować w planie normy ujmowanej tak, jak postulował Eugenio Coseriu – jako zespół reguł określających możliwości realizacji systemu¹.

Trzeba tu dodać, że w polskiej lingwistyce dominuje inna, „idiograficzna” koncepcja idiostylu, rozumianego jako zespół właściwości i tendencji nacechowanych, charakteryzujących sposób wypowiadania się danej jednostki i odróżniających go od tekstów innych mówiących, właściwości będących rezultatem pewnego wyboru i kombinacji, a w przypadku osób kreatywnych – także i tworzenia elementów językowych. Znajduje ona jednak zastosowanie tylko w analizach tekstów kreatywnych i idiolektów (a właściwie idiostylów) pisarzy, nie nadaje się natomiast do opisywania rezultatów działalności mownej innych, mniej twórczych autorów².

1.2. Zakres badań nad idiolektami

Zagadnienia języka i stylu jednostek nie sposób sprowadzić do tradycyjnego problemu tzw. języka pisarzy, nawet jeśli kategorię „pisarzy” potraktowalibyśmy szeroko – jak to się czasem robi – i zaliczylibyśmy do tego grona nie tylko twórców literatury pięknej, ale także autorów prac naukowych i popularnonaukowych, osoby pa-

¹ O różnicach między idiolektem a idiostylem pisałam szerzej w pracy: Kozłowska 2013b: 33–46.

² Inną sprawą pozostaje to, czy rzeczywiście da się wskazać zespół takich własności, które byłyby charakterystyczne dla jednostki mówiącej niezależnie od wybranego przez nią stylistycznego i gatunkowego wzorca wypowiedzi. Zapewne w większości przypadków możliwe byłoby rozpoznanie tendencji czy cech ponadjednoodmianowych, przejawiających się w więcej niż jednym idiostylu (w skrajnych, mało prawdopodobnych sytuacjach, zwłaszcza u jednostek słabo „przełączających” style – nawet cech wspólnoodmianowych), ale trzeba zauważyć, że w rozmaitych typach tekstów realizują się one na różne sposoby i zwykle pełnią w nich odmienne funkcje. W działalności mownej jednostki pojawiają się również elementy niesfunkcjonalizowane, takie jak np. jąkanie, wady wymowy itp., ale są one powiązane nie tyle z idiolektem czy idiostylem, ile z ich realizacją.

rające się tzw. literaturą faktu, jak również tłumaczy, publicystów, kaznodziejów itp. kreatywne językowo podmioty. Idiolekty stanowią przecież jeszcze przedmiot zainteresowania historyków języka, dialektologów, badaczy języka grup społecznych (socjolektów), języka dzieci oraz innych biolektów; nie sposób pominąć również uczonych opisujących bilingwizm i wielojęzyczność. Idiolekt badają także przedstawiciele tych nurtów lingwistycznych, które deklarują wprawdzie dążenie do uchwycenia właściwości języka ogólnego, ale jego analizę rozpoczynają od obserwacji własnego poczucia językowego. Pierwszy z takich nurtów to generatywizm, traktujący intuicję rodzimego użytkownika języka jako świadectwo jego kompetencji, wyraz tzw. języka wewnętrznego (ang. *internal language*, *I-language*); drugi – semantyka introspekcyjna, w obrębie której podstawową i deklarowaną wprost metodą analizy znaczeń jest konstruowanie sztucznych zdań diagnostycznych i rozstrzyganie o ich (nie)akceptowalności (por. Bogusławski 2009: 124–150).

1.3. Cele badań idiolektów

Wśród lingwistycznych zastosowań wiedzy o języku jednostki zarysowują się dwie odmienne postawy: w pierwszej chodzi o „prześledzenie indywidualnych cech mowy autorów na tle języka ogólnego danego okresu [...] (ostatecznym przedmiotem zainteresowania jest więc twórca oglądany przez język)”; w drugiej

język autora (lub utworu) traktowany jest jako dokument (świadectwo) zjawisk językowych epoki: ostatecznym przedmiotem zainteresowania jest w tym wypadku język ogólny, a teksty literackie³ pełnią funkcję dokumentu (Grzegorzczkova 2007: 159).

Każdy z tych kierunków myślenia o idiolektcie uwypukla inne jego elementy: w analizach języka jednostek wybitnych, głównie pisarzy, zasadniczym celem jest uchwycenie komponentów wyróżniających danego mówiącego, co ma pomóc w rozumieniu jego wypowiedzi⁴; w pracach historyków języka, dialektologów, badaczy języka dzieci oraz innych biolektów czy socjolektów opisywane są cechy ponadindywidualne, pozwalające wyabstrahować z tekstów osobniczych tzw. normę uzualną – czyli stan typowy, charakteryzujący całą grupę użytkowników języka, którą reprezentuje badana jednostka. Ze względu na odmiennosć celów opracowania odwołujące się do idiolektów różnią się doбором źródeł oraz sposobem ich wykorzystania.

2. Typy źródeł wykorzystywane w badaniach idiolektów

Z przeglądu prac dotyczących idiolektów wynika, że wykorzystywane są w nich następujące typy źródeł⁵:

³ Renata Grzegorzczkova pisała o tekstach literackich, ale zasadę tę można poszerzyć także na inne typy wypowiedzi.

⁴ Opisy takie będę nazywała *autonomicznymi*.

⁵ Kryterium podziału zastosowanym na najwyższym poziomie proponowanej typologii jest sposób pozyskania źródeł.

I. Naturalne (tj. nie stworzone specjalnie do potrzeb opisu) wypowiedzi jednostek:

1. teksty pisane:

- 1.1. teksty artystyczne,
- 1.2. prywatne teksty użytkowe: listy i notatki, a w odniesieniu do współczesności – teksty opublikowane w Internecie,
- 1.3. teksty naukowe,
- 1.4. teksty publicystyczne (opublikowane w różnych mediach, także elektronicznych),
- 1.5. teksty retoryczne;

2. różnorodne wypowiedzi ustne, najczęściej utrwalone w postaci nagrań audio/wideo bądź stenografu.

II. Dane uzyskane w wyniku specjalnych procedur i eksperymentów lingwistycznych:

1. sztuczne teksty diagnostyczne,
2. wyniki ankiet (najczęściej pisemnych) i wywiadów,
3. rezultaty eksperymentów lingwistycznych.

III. Dane pośrednie bądź już opracowane:

1. słowniki idiolektów i indeksy leksyki autorów;
2. słowniki i korpusy języka ogólnego;
3. studia językoznawcze;
4. świadectwa osób trzecich.

2.1. Typy źródeł w poszczególnych rodzajach badań nad idiolektami

Wymienione rodzaje danych mają rozmaite zastosowania w różnych typach badań nad idiolektami; niejednakowy jest też ich status i rola w tych badaniach.

2.1.1. Autonomiczne studia nad idiolektami odwołują się w znakomitej większości do tekstów pisanych, rzadziej – o czym będzie jeszcze mowa – do danych opracowanych już w postaci słowników czy studiów lingwistycznych. W tradycyjnym ujęciu przedmiotem analizy były przede wszystkim utwory reprezentujące styl artystyczny bądź retoryczny. Ponieważ w ostatnim czasie krąg bohaterów prac analitycznych bardzo się poszerzył i coraz częściej obiektem odrębnych analiz staje się np. idiolekt uczonego, polityka, dziennikarza, kaznodziei czy nawet tzw. celebryty, badacze nie koncentrują się już tylko na dziełach literackich, lecz uwzględniają – w zależności od specyfiki działalności opisywanego autora – m.in. teksty publicystyczne i/lub naukowe, co prowadzi do charakterystyki idiostylów innych niż tylko artystyczny. Dość powszechną praktykę stanowi też uzupełnianie obserwacji dotyczących języka danej jednostki o wnioski płynące z oglądu jej wypowiedzi użytkowych: listów i notatek. Te dane najczęściej mają charakter dodatkowy, stanowią tło dla analiz, choć zdarzają się również osobne opracowania na temat języka listów czy prywatnych zapisków⁶.

⁶ Ciekawe wydaje się tu ograniczenie gatunkowe – nie powstały jeszcze chyba prace o języku podań, które wyszły spod pióra nawet bardzo znanych autorów, albo sporządzonych przez nich list zakupów; pytanie, czy kiedyś opisu doczekają się np. ich posty na portalach społecznościowych.

Wskazana przed chwilą tendencja stawia w nowym świetle kilka ważnych problemów praktycznych, związanych z pozyskiwaniem danych, jak również teoretycznych, dotyczących samego paradygmatu opisu.

Po pierwsze, uświadomienie sobie stylistycznej niejednorodności idiolektu wymaga zrewidowania dotychczasowej praktyki opisu języka osobniczego, zwłaszcza w odniesieniu do tych jednostek, których działalność mowna jest bardzo zróżnicowana⁷. Zwykle albo wybiera się jeden z typów wypowiedzi autorskich (np. artystyczne czy naukowe) i na nim koncentruje uwagę, traktując pozostałe jako tło, albo łączy się co najmniej dwa idiostyle i opisuje je jako nieodróżnicowaną całość – w przekonaniu, że teksty powstałe pod piórem jednego autora, choć różne funkcjonalnie, pozostają w relacji spójności, a przynajmniej niesprzeczności, i oświetlają się nawzajem. Obie te strategie, choć sankcjonowane tradycją badawczą i przede wszystkim praktyczne, wydają się jednak niewystarczające – charakterystyka idiolektu powinna z jednej strony obejmować względną całość językowej twórczości danego mówiącego (także w porządku diachronicznym), z drugiej zaś – uwzględniać jej wewnętrzną dywersyfikację. Rzecz jasna, nie sposób poddać obserwacji naprawdę każdej wypowiedzi, jaką wybrany, nawet stosunkowo małomówny użytkownik języka wygłasza lub zapisuje w trakcie swojego życia, a nawet gdyby kiedyś stało się to technicznie możliwe (badaczom przychodzą przecież w sukurs coraz lepsze narzędzia umożliwiające gromadzenie i przetwarzanie dużych korpusów danych), trzeba pogodzić się z ograniczeniami natury zasadniczej: teksty zawsze poświadczają tylko pewien wycinek kompetencji. Idzie jednak o to, aby dobór źródeł, na podstawie których tę kompetencję się opisuje, był reprezentatywny dla wszystkich odmian języka, jakimi posługuje się czy posługiwał bohater danego opracowania.

Po drugie, tendencja do charakteryzowania języka różnych postaci życia publicznego w niektórych przypadkach wiąże się z koniecznością odejścia od tekstu pisanego i zajęcia się wypowiedziami ustnymi, a także porzucenia tradycyjnych narzędzi analitycznych, wypracowanych dla potrzeb opisu tekstu, na rzecz „stylistyki dyskursu” i proponowanej przez nią „orientacji pozatekstowej” (zob. Witosz 2009b), innymi słowy i w pewnym uproszczeniu – wymagają przeniesienia środka ciężkości z produktu działalności mownej opisywanego użytkownika języka na zdarzenia komunikacyjne, w których on uczestniczy czy które inicjuje. Te pomysły w obrębie autonomicznych badań nad idiolektami stanowią wciąż pewną nowość, zwłaszcza w sferze studiów materiałowych, ale zmiany prowadzą chyba właśnie w tym kierunku.

⁷ Zróżnicowanie stylistyczne najlepiej ujawnia się w idiolektach użytkowników szczególnie sprawnych, biegle „przełączających kody”, najsłabiej lub nawet wcale – w języku osobniczym tych użytkowników, którzy opanowali jedynie pierwszą przyswajaną przez mówiących odmianę języka ojczystego – potoczną. Zdarza się też, że jakiś nosiciel idiolektu, bardzo twórczy w obrębie jednego idiostylu, prezentuje znikomą kreatywność w innym, czyli realizuje w nim wszystkie główne cechy stylowe charakterystyczne dla praktyki swego czasu i/lub środowiska i nie wprowadza żadnych modyfikacji. Opis takiego nieinnowacyjnego idiostylu ma wówczas znaczenie jedynie ze względu na potrzebę uchwycenia całego repertuaru możliwości stylistycznych mówiącego.

Po trzecie, w związku ze wskazanymi zjawiskami należałoby się na nowo przyjrzeć źródłom analiz pod względem edytorskim. W badaniach języka pisarzy najwyższe cenione, bo pozwalające uchwycić najwięcej cech osobniczych, były zawsze studia oparte na autorskich rękopisach, ewentualnie – autoryzowanych pierwodrukach. Zasadę wykorzystywania tych źródeł, które najwierniej zdają sprawę z językowej swoistości jednostki, niewątpliwie trzeba zachować. Przekształcenia, jakim podlega współczesna komunikacja, i rosnąca rola mediów elektronicznych każą jednak postawić na nowo zagadnienie wyboru tekstu źródłowego. Rodzi się pytanie, co należałoby uznać za źródło np. w przypadku wypowiedzi wielokrotnie reprodukowanych w różnych, wydawałoby się – równorzędnych wersjach czy hipertekstów, których kształt nigdy nie jest ostateczny i każdorazowo zależy od decyzji odbiorcy. Z kolei nowe spojrzenie na problematykę tożsamości, ujmowanie jej jako zjawiska płynnego, *in statu nascendi*, podważa tradycyjną koncepcję autora jako osoby o tożsamości ustalonej i względnie stałej (zob. Witosz 2004). Problemy te pojawiały się oczywiście także wcześniej, ale wydaje się, że uwzględnienie tekstów innych niż tylko literackie, przede wszystkim tekstów w wersji ustnej oraz źródeł internetowych, radykalnie je wyostrza, prowadząc do konieczności przemyślenia fundamentalnych pojęć.

Jak już wspominałam, źródłem wiedzy o idiolektach, a także materiałem źródłowym autonomicznych studiów nad nimi bywają także opracowane już dane lingwistyczne. Dane takie mają z różnych względów ograniczoną przydatność, a – co za tym idzie – dość wąski zakres zastosowania. Najbardziej pożyteczne i, jak się wydaje, stosunkowo najczęściej wykorzystywane są słowniki i indeksy leksyki autorów; ponieważ jednak takie opracowania przygotowuje się zazwyczaj jedynie dla najwybitniejszych i najbardziej wpływowych twórców, krąg ich profesjonalnych użytkowników w praktyce zawęża się do badaczy języka Adama Mickiewicza (zob. Górski, Hrabec 1962–1983), Jana Chryzostoma Paska (zob. Koneczna, Doroszewski 1965–1973), Jana Kochanowskiego (zob. Kucała 1994–2012) oraz Cypriana Norwida (zob. Puzyńska, Korpysz 2009–)⁸, a także kilku pisarzy, których leksyka została opracowana w postaci słowników cząstkowych: tematycznych (Cyprian Norwid, Stefan Żeromski, zob. Puzyńska 1993; Chojak 1994; Teleżyńska 1994; Korpysz, Puzyńska 1998; Kadyjewska, Korpysz, Puzyńska 2000; Handke 2002–1014)⁹ i dyferencyjnych (Stanisław Wyspiański, zob. Śliwiński 2016) oraz indeksów i list frekwencyjnych (zob. omówienie w: Kozłowska 2013a). Dodatkową trudność może stanowić również zaproponowany w pracach leksykograficznych sposób uporządkowania i opisu materiału – osobie stawiającej własne pytania i dążącej do swoich celów niejednokrotnie wygodniej mieć do czynienia z „surowymi” danymi tekstowymi bądź co najwyżej z konkordancją niż z rezultatami metodologicznych i interpretacyjnych wyborów innych uczonych.

Pozostałe typy źródeł już opracowanych pełnią tylko funkcję pomocniczą: ze słowników, korpusów i studiów dotyczących języka ogólnego czy jakiegś jego odmiany można wprawdzie czasem wydobyć np. przykłady autorskich użyc, trudno tu

⁸ *Internetowy słownik języka Cypriana Norwida* jest nadal w przygotowaniu.

⁹ Słowniki języka pisarzy (całościowe i tematyczne) szczegółowo omówił Tomasz Korpysz (2010).

natomiast szukać kompletności czy nawet reprezentatywności danych; ze względu na zasadniczo odmienne cele badawcze wykorzystuje się je zwykle zatem tylko do opisu relacji między danym idiolektem (przede wszystkim jego warstwą leksykalną) a językiem ogólnym. Ponieważ najwyżej ceni się dane bezpośrednie, pisemne bądź ustne informacje osób trzecich dotyczące zarówno idiolektu, jak i idiosylów czy tekstów interesującego nas podmiotu mówiącego, mogą jedynie uzupełniać wyniki analiz materiałowych, ewentualnie informować o recepcji jego działalności mownej.

2.1.2. Także badacze procesów rozwojowych języka uwzględniają głównie różnego typu dokumenty pisane, w tym literackie, oraz wyniki przeprowadzonych już prac lingwistycznych: słowniki, korpusy oraz źródła, które Zenon Klemensiewicz nazywał interpretacyjnymi: dawne studia gramatyczne, poradniki, istniejące opracowania naukowe: monografie i syntezy (zob. Klemensiewicz 1982: 663–667). Ze względu na cel prac historycznojęzykowych rola tych danych, a także sposób ich wykorzystania, różnią się jednak od wartości i funkcji, jakie przypisuje się im w obrębie autonomicznych badań nad idiolektami.

Po pierwsze, choć dla niektórych okresów rozwoju języka dysponujemy przede wszystkim czy wręcz tylko tekstami artystycznymi czy innymi świadectwami ponadprzeciętnej sprawności językowej (np. kazaniami), nie stanowią one tu głównej, a nawet w żaden sposób uprzywilejowanej podstawy opisu. Irena Bajerowa proponowała, aby z opracowywanego materiału, mającego reprezentować pewną średnią językową, wyłączać teksty autorów wybitnych, a także utwory poetyckie (zob. Bajerowa 1988: 10–12). Uczona uważała również, że podstawą wniosków historycznojęzykowych powinny być nie rękopisy, a teksty drukowane, które pozwalają lepiej „ścierać i spłaszczać indywidualności języków osobniczych na rzecz ogólności i przeciętności języka ogólnego” (Bajerowa 1988: 12; por. też Borawski 1982: 346–347).

Po drugie, słowniki i opracowania lingwistyczne, które dostarczają informacji o stanie i/lub rozwoju języka ogólnego czy jego odmian, w historii języka ceni się oczywiście zdecydowanie wyżej niż w studiach ukierunkowanych na charakterystykę poszczególnych idiolektów; dodatkowym ich atutem jest to, że dokumentują one nie tylko właściwości samego języka, ale także świadomość użytkowników i badaczy z danego okresu.

W badaniach diachronicznych dokonuje się właśnie zasadnicza zmiana dotycząca wykorzystania wypowiedzi mówionych. Jeszcze Zenon Klemensiewicz pisał, że stanowią one ważne źródło wiedzy o ewolucji języka, ale w odniesieniu do przeszłości można je wyławiać tylko wtórnie i pośrednio z dokumentów pisanych (zob. Klemensiewicz 1982: 663–664). W miarę upływu czasu i rozwoju techniki liczba materiałów do najnowszej historii języka utrwalonych już nie w postaci tekstów, ale nagrań audio i wideo znacząco wzrasta. Można więc przypuszczać, że w przyszłości badania historycznojęzykowe będą w coraz większym zakresie sięgać również źródeł o charakterze mówionym.

2.1.3. O ile dwa wskazane powyżej typy opisu opierają się głównie na naturalnych wypowiedziach użytkowników języka, o tyle we wszystkich pozostałych nurtach odwołujących się do pojęcia idiolektu – czyli w badaniach dotyczących współczesnego

języka narodowego i jego odmian, takich jak dialekty, socjolekty i biolekty, oraz fenomenów bilingwizmu i wielojęzyczności – uwzględnia również dane pozyskane w specjalnych, sztucznie skonstruowanych sytuacjach komunikacyjnych. Większość z nich zakłada istnienie pewnego, choćby minimalnego dystansu między eksploratorem a respondentem (zob. Kaś 1986: 120). Wyjątkowe pod tym względem są procedury stosowane przez przedstawicieli nurtu introspekcyjnego, którzy – jak już wspominałam – wyrokuje o (nie)akceptowalności skonstruowanych przez siebie wypowiedzi, odwołując się do własnej intuicji językowej na drodze samoobserwacji. Takim postępowaniu przyświeca przekonanie, że badacz jest także użytkownikiem języka swojego czasu, a zatem sam może pełnić funkcję informatora.

W badaniach zmierzających do pozyskania informacji od osób trzecich stosuje się przede wszystkim ukierunkowane wywiady na określone tematy oraz rozmaite ankiety i kwestionariusze. Pierwsza z wymienionych metod, tradycyjnie stosowana w dialektologii, znajduje zastosowanie także w przygotowywaniu charakterystyk socjolektów i biolektów (zwłaszcza w opisach języka dzieci). Moderowana przez eksploratora rozmowa, kiedyś zapisywana bądź stenografowana, dziś – zwykle utrwalana w postaci nagrań, z jednej strony pozwala zebrać naprawdę wiele różnorodnych danych językowych, które można później ze sobą konfrontować, weryfikować i uściślać, z drugiej – umożliwia obserwację informatorów oraz ich charakterystykę, nie tylko ściśle lingwistyczną, ale także socjalną i kulturową.

Jedną z najczęściej stosowanych obecnie metod pozyskiwania informacji jest ankietowanie, które zajmuje poczesne miejsce nie tylko w dialektologii (w postaci kwestionariuszy dialektologicznych), ale przede wszystkim w socjo-, etno- i psycholingwistyce¹⁰. Na popularność tego źródła we współczesnej nauce wpłynęła przede wszystkim oferowana przez nie szansa zgromadzenia bardzo wielu danych w stosunkowo krótkim czasie, spotęgowana dziś jeszcze przez możliwości Internetu, który bardzo ułatwił zbieranie informacji od strony technicznej. Efektywność i przydatność źródeł ankietowych zależy przede wszystkim od sposobu przygotowania pytań oraz doboru informatorów. Kwestie te są przedmiotem fachowych opracowań i nie będę ich tutaj szczegółowo omawiała. Chciałabym tylko zwrócić uwagę, że pozyskiwanie danych za pomocą ankiet ma dwa zasadnicze ograniczenia:

1. Ankieta dostarcza informacji w istocie wcześniej zaprojektowanych czy przynajmniej przewidzianych przez badającego. Rzecz jasna, ankietę można uzupełniać o pytania otwarte, ale one również muszą być przemyślane; ich przydatność osłabia reszta czynnik, o którym będzie mowa w punkcie 2.

2. Omawiana metoda pozwala na uzyskanie tylko tych danych, które respondent chce i w danym momencie może świadomie przekazać. Niezależnie od formy ankiety osoba badana doświadcza przecież stresu, czasem wstydu bądź innego rodzaju oporu przed ujawnieniem informacji; może wreszcie zwyczajnie zapomnieć o danym zjawisku albo celowo wprowadzać eksploratora w błąd.

¹⁰ W poradnikach dotyczących ankietowania odróżnia się kwestionariusze od ankiet; na potrzeby niniejszego artykułu będę jednak stosowała te określenia wymiennie.

Drugie wspomniane powyżej ograniczenie nie pojawia się w ostatnim eksperymentalnym źródle danych wykorzystywanym w badaniach idiolektów, a mianowicie w różnego rodzaju testach i próbach psycholingwistycznych i neurolingwistycznych, które – w odróżnieniu np. od badań dialektologicznych, socjo- i etnolingwistycznych – mogą prowadzić nie tylko do wniosków dotyczących mowy zbiorowości, ale także do charakterystyk poszczególnych jej reprezentantów (tzw. studia przypadków). Wprawdzie omawiana metoda ma swoje oczywiste ograniczenie dotyczące badanych jednostek, ponieważ testy można przeprowadzać tylko na żyjących użytkownikach języka, oraz techniczne, jako że jej stosowanie wymaga zwykle zaawansowanej aparatury, pozwala jednak za to uzyskać także te dane, których informator nie chce albo nie może podać świadomie (np. nie pamięta o nich). Eksperymenty przeprowadzane przez lingwistów mają różne formy i cele. Ze względu na omawiany problem warto zwrócić uwagę na następujące ich cechy:

1) wagę faktów o charakterze negatywnym (błędy, dysfunkcje, anomalie), które stanowią istotny przedmiot obserwacji;

2) rolę parametrów mierzalnych, takich jak np. czasowa charakterystyka procesów związanych z mówieniem i rozumieniem, natężenie impulsów, intensywność reakcji;

3) wykorzystywanie do opisu języka i komunikacji człowieka wniosków płynących z obserwacji zachowań komunikacyjnych innych zwierząt.

Dopełnieniem wszystkich metod polegających na stosowaniu specjalnych procedur/eksperymentów lingwistycznych jest zwykle obserwacja uczestnicząca, która polega na rejestrowaniu i analizowaniu swobodnych zachowań językowych osób mówiących w sytuacji dłuższego kontaktu.

2.2. Zależność źródeł od paradygmatów badawczych

Rola poszczególnych typów źródeł wiąże się nie tylko z możliwościami, jakimi dysponuje nauka danego czasu, i z celami opisu idiolektu; ich waga jest zależna również – choć w mniejszym stopniu – od przyjętej koncepcji języka osobniczego. W praktyce rolę tego czynnika osłabia brak pogłębionej świadomości metodologicznej autorów wielu prac, którzy często przyjmują intuicyjne rozumienie podstawowych kategorii, oraz fakt, że większość paradygmatów badawczych, jakie można wyodrębnić w opisach języka jednostki, rysuje się wyraźnie tylko w obrębie autonomicznych badań nad idiolektami, zwłaszcza kreatywnymi.

2.2.1. Paradygmaty badań idiolektów

W badaniach nad idiolektami można mówić o czterech zasadniczych paradygmatach czy modelach opisu, nie zawsze wyraźnie rozłącznych i w konkretnych pracach czasem ze sobą związanych. Są to:

1) model historycznojęzykowy, obecny np. w dawniejszych monografiach języka autorów, które, choć pozornie autonomiczne, w istocie mają przede wszystkim ilustrować procesy rozwojowe polszczyzny ogólnej lub jej regionalnych odmian, a idio-

lekt traktują jako „indywidualny system”, będący jednostkowym wariantem (adaptacją) języka ogólnego;

2) model strukturalno-stylistyczny, nastawiony przede wszystkim na poszukiwanie indywidualnego, często świadomie kreowanego systemu funkcjonalnego, który obejmuje zespół elementów uporządkowanych, wewnętrznie powiązanych i służących określonej celowi, oraz wiedzę o sposobach ich używania;

3) model generatywny, proponujący dynamiczną koncepcję idiolektu jako kompetencji językowo-komunikacyjnej jednostki, jej mentalnej dyspozycji do wytwarzania i rozumienia tekstów, wiedzy o sposobach używania języka w różnych sytuacjach społecznych;

4) model kognitywny, który widzi w idiolekcie nośnik autorskiego obrazu świata, indywidualny „sposób widzenia pewnych fragmentów rzeczywistości pozajęzykowej poprzez język” (Wojciechowska 2000: 8).

2.2.2. Źródła w poszczególnych paradygmatach

Zarówno w badaniach utrzymanych w paradygmacie historycznojęzykowym, jak i w strukturalistycznych oraz kognitywnych podstawowe źródło informacji stanowią teksty (także – choć w mniejszym zakresie – już opracowane), które – nawet jeśli są zróżnicowane stylistycznie i gatunkowo, co w ostatnim czasie stało się właściwie regułą we wszystkich trzech wymienionych podejściach metodologicznych – poświadczają przecież tylko wybrane językowe umiejętności i cechy autora. Punkt zwrotny w studiach nad językiem osobniczym stanowiła natomiast wprowadzona przez Noama Chomsky’ego i rozszerzona przez Della Hymesa idea modelowania kompetencji językowej/komunikacyjnej idealnego użytkownika języka¹¹. W tym ujęciu niewystarczająca okazała się bowiem obserwacja cech zarejestrowanych wypowiedzi, a zatem rezultatów działań mownych jednostki – chodziło przecież o opis jej umiejętności czy możliwości, które nie zawsze i nie w pełni muszą się ujawniać w znanych badaczom tekstach. Kompetencja językowo-komunikacyjna to swego rodzaju ukryty, bezpośrednio niedostępny poznaniu mechanizm albo – jak pisał swego czasu Henryk Borek – „czarna skrzynka”:

Znamy jej „wejścia” w postaci języka zbiorowiskowego, jakim się dana jednostka posługuje, i znamy „wyjścia” tej „czarnej skrzynki” w postaci gotowych wypowiedzi jednostki, jednak to, co się wewnątrz niej dzieje, jakie są potencjalne możliwości językowe jednostki, może być tylko pośrednio odzwzorowywane [...] (Borek 1988: 19–20).

Aby wymodelować sposób funkcjonowania „czarnej skrzynki”, potrzebne są specjalnie zaprojektowane eksperymenty, ewentualnie ankiety, które pozwolą nie tylko poświadczyć zjawiska z różnych powodów nieobecne w tekstach autorstwa

¹¹ Koncepcja Chomsky’ego nie dotyczy wprawdzie języka jednostki sensu stricto, ponieważ obiektem zainteresowania staje się tu nie konkretny mówiący, a użytkownik modelowy, nieistniejący w rzeczywistości idealny reprezentant klasy; wywarła ona jednak spory wpływ na badania nad idiolektami (zob.: Kozłowska 2013b: 24–30).

danego mówiącego, ale także w jakimś stopniu odpowiedzieć na pytanie o zakres jego kompetencji. Ze względu na wskazywane już ograniczenia w stosowaniu tego rodzaju metod – a przynajmniej między innymi dlatego – koncepcja idiolektu jako kompetencji komunikacyjnej pojedynczego użytkownika zakorzeniła się przede wszystkim na gruncie psycho- i socjolingwistyki; w obrębie autonomicznych studiów nad idiolektami wciąż dominują pozostałe trzy paradygmaty metodologiczne, odwołujące się do analizy wypowiedzi. Niestety, zdarza się również, że trudność w opisywaniu „kompetencyjnie” zdefiniowanego języka autorów skutkuje wcale nierzadko spotykaną w pracach rozbieżnością między deklarowaną teorią lingwistyczną a rzeczywistą podstawą i metodą analizy – idiolekt rozumiany podobno jako „kompetencja komunikacyjna jednostki” omawiany bywa w istocie w tradycyjny, statyczny sposób, poprzez charakterystykę realizowanych w tekstach środków językowych.

3. Jedno źródło czy wiele źródeł? Wnioski i postulaty badawcze

Jest rzeczą oczywistą, że dobór źródeł powinien korespondować z typem opisu, przyjętymi celami oraz z paradygmatem badawczym, w który wpisują się prowadzone prace. Czy jednak w charakterystyce języka jednostki wystarczy ograniczyć się do jednego typu źródła, czy też trzeba uwzględnić większą ich liczbę? W dziejach badań nad idiolektami, zwłaszcza kreatywnymi, udzielano na to pytanie różnych odpowiedzi, zależnych od sposobu rozumienia „języka pisarzy”. Tradycyjne opracowania zwykle opierały się jedynie na utworach artystycznych, chociaż już Zenon Klemensiewicz zaznaczał, że nie można utożsamiać języka autora z językiem jego utworów, i proponował w związku z tym odróżnianie „złoża własnego”, które nie stanowi przedmiotu refleksji osoby mówiącej i służy spontanicznej „stylizacji samorzutnej”, oraz „złoża postronnego” – „zasobu zdobywanego przez osobnika świadomie” (Klemensiewicz 1961: 208), wykorzystywanego w „stylizacji umyślnej” (zob. Klemensiewicz 1961: 207–208). W przypadku języka pisarza podstawowy cel badawczy miało stanowić wykrycie zasobów „złoża własnego”, innymi słowy – rozpoznanie cech „języka poety”, a nie „języka utworów poety” (zob. Klemensiewicz 1961). Jeśli tak, nie można jednak ograniczać się tylko do utworów literackich, które są przecież z natury „stylizowane umyślnie”. Ze względu na coraz lepiej uświadamiane dziś wewnętrzne zróżnicowanie każdego idiolektu (przede wszystkim funkcjonalno-stylowe, gatunkowe, społeczne, ale także terytorialne i czasowe) komplikuje się też Klemensiewiczowski podział zasobów językowych jednostki na „własne” i „postronne” – te ostatnie winny obejmować nie tylko elementy przyswajane dla wykorzystania w twórczości artystycznej, ale także obecne w tekstach naukowych, urzędowych, w komunikacji z członkami grup zawodowych, hobbystycznych itp. Innymi słowy, z uwagi na przemiany, jakim podlegają w ostatnim czasie badania nad idiolektem oraz jego sama koncepcja, należy postulować możliwie jak najszerszą dywersyfikację źródeł wiedzy o języku jednostki. Ich wielość stwarza szansę na zbudowanie opisu wielostronnego,

zdającego sprawę z wewnętrznego zróżnicowania idiolektu (ciągle przecież brak prac pokazujących różne idiosstyle tego samego autora!).

Postulat zróżnicowania źródeł dotyczy przede wszystkim tego nurtu badań nad idiolektami, który nazwałam autonomicznym. Uwzględnienie wszystkich dostępnych danych, także ustnych, może bowiem służyć głównie pokazywaniu wewnętrznej niejednorodności języka jednostki. W opisach podporządkowanych lingwistycznym charakterystykom grup, które reprezentują badani użytkownicy języka, sprawa przedstawia się nieco inaczej, ponieważ opisy te są z natury ukierunkowane na uchwycenie jednej odmiany języka – nawet jeśli dany informator włada kilkoma jego wariantami społecznymi czy terytorialnymi, w centrum uwagi znajduje się przede wszystkim (co nie znaczy – tylko) ten, który stanowi przedmiot analizy. I w tego rodzaju pracach warto jednak odwoływać się do różnych źródeł informacji: obserwacji wypowiedzi, ankiet, eksperymentów. Daje to możliwość pozyskania względnie kompletnych danych oraz ich weryfikacji.

Bibliografia

- BAJEROWA Irena (1988): Badanie język osobniczego jako metodologiczny problem historii języka ogólnego. – [w:] Jerzy BRZEZIŃSKI (red.): *Język osobniczy jako przedmiot badań lingwistycznych*. – Zielona Góra: Wydawnictwo WSP w Zielonej Górze, 7–14.
- BOGUSŁAWSKI Andrzej (2009): *Myśli i gwiazdce i regule*. – Warszawa: BEL Studio.
- BORAWSKI Stanisław, 1982, Typowość i wzorcowość w studiach historycznojęzykowych. – *Język Polski* 4–5, 344–348.
- BOREK Henryk (1988): Co możemy wiedzieć o języku osobniczym?. – [w:] Jerzy BRZEZIŃSKI (red.): *Język osobniczy jako przedmiot badań lingwistycznych*. – Zielona Góra: Wydawnictwo WSP w Zielonej Górze, 15–21.
- CHOJAK Jolanta, red. (1994): *Słownictwo estetyczne Cypriana Norwida*. – Warszawa: Wydawnictwa UW.
- GÓRSKI Konrad, HRABEC Stefan, red. (1962–1983): *Słownik języka Adama Mickiewicza*. – Wrocław: Zakład Narodowy im. Ossolińskich, t. 1–11.
- GRZEGORCZYKOWA Renata (2007): *Wstęp do językoznawstwa*. – Warszawa: Wydawnictwo Naukowe PWN.
- HANDKE Kwiryna, red. (2002–2014): *Słownictwo pism Stefana Żeromskiego*. – Kraków–Poznań: UNIVERSITAS – Wydawnictwo Naukowe UAM, t. 1–16.
- KADYJEWSKA Anna, KORPYSZ Tomasz, PUZYŃSKA Jadwiga (2000): *Chrześcijaństwo w pismach Cypriana Norwida*. – Warszawa: Wydawnictwa UW.
- KĄŚ Józef (1986): Metody statystyczne w badaniach dialektologicznych. – *Acta Universitatis Lodziensis, Folia Linguistica* 12, 117–123.
- KLEMENSIEWICZ Zenon (1961): Jak charakteryzować język osobniczy? – [w:] Zenon KLEMENSIEWICZ, *W kręgu języka literackiego i artystycznego*. – Warszawa: Wydawnictwo Naukowe PWN, 204–214 [artykuł z 1946 r.].
- KLEMENSIEWICZ Zenon (1982): Zagadnienia i założenia historii języka polskiego. – [w:] Zenon KLEMENSIEWICZ, *Składnia, stylistyka, pedagogika językowa*. – Warszawa: Wydawnictwo Naukowe PWN, 639–697 [artykuł 1956 r.].
- KONECZNA Halina, DOROSZEWSKI Witold, red. (1965–1973): *Słownik języka Jana Chryzostoma Paska*. – Wrocław: Zakład Narodowy im. Ossolińskich, t. 1–2.

- KORPYSZ Tomasz (2010): Słowniki języka autorów jako typ opracowań leksykograficznych. – *Poradnik Językowy* 4, 51–71.
- KORPYSZ Tomasz, PUZYNNIA Jadwiga (1998): – *Wolność i niewola w pismach Cypriana Norwida*. – Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- KOZŁOWSKA Anna (2013a): Indeks leksyki pisarza jako typ opracowania leksykograficznego. – *Prace Filologiczne* LXIV, cz. 1, 147–158.
- KOZŁOWSKA Anna (2013b): *Od psalmów słowiańskich do rzymskich medytacji. O stylu artystycznym Karola Wojtyły*. – Warszawa: Wydawnictwo UKSW.
- KUCAŁA Marian, red. (1994–2012): *Słownik polszczyzny Jana Kochanowskiego*. – Kraków: Wydawnictwo IJP PAN, t. 1–5.
- KUDRA Andrzej (2006): Idiolekt, idiostyl czy idiolektostyl?. – [w:] Magdalena PIETRZAK (red.): *Osoba i osobowość – czynniki je kształtujące. Materiały z konferencji 9–11 maja 2005 r.* – Łódź: Archidiecezjalne Wydawnictwo Łódzkie, 209–221.
- PUZYNNIA Jadwiga, red. (1993): *Słownictwo etyczne Cypriana Norwida*, część 1: *prawda, fałsz, kłamstwo*. – Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- PUZYNNIA Jadwiga, KORPYSZ Tomasz (2009–): *Internetowy słownik języka Cypriana Norwida*. – <http://www.sownikjezykanorwida.uw.edu.pl>.
- ŚLIWIŃSKI Władysław, red. (2016): *Słownik osobliwości leksykalnych Stanisława Wyspiańskiego na materiale utworów dramatycznych*. – Kraków: Libron.
- TELEŻYŃSKA Ewa (1994): *Nazwy barw w twórczości Cypriana Norwida*. – Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- WITOSZ Bożena (2004): Lingwistyczna koncepcja idiolektów a problem tożsamości. – [w:] Włodzimierz BOLECKI, Ryszard NYCZ (red.): *Narracja i tożsamość (I). Narracje w kulturze*. – Warszawa: Wydawnictwo IBL PAN, 128–143.
- WITOSZ Bożena (2009a): Język osobniczy/idiolekt i/a (?) styl indywidualny/idiostyl. – [w:] Bożena WITOSZ, *Dyskurs a stylistyka*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego, 249–255.
- WITOSZ Bożena (2009b): Stylistyka dyskursu – nowe horyzonty badań nad stylem autora. – [w:] Krzysztof MAĆKOWIAK, Cezary PIĄTKOWSKI (red.): *Język i styl twórcy w kręgu badań współczesnej humanistyki*. – Zielona Góra: Oficyna Wydawnicza Uniwersytetu Zielonogórskiego, 312–320.
- WOJCIECHOWSKA Anna (2000): *Magdaleny z Kossaków Samozwaniec widzenie świata*. – Warszawa: Wydawnictwo Naukowe PWN.
- WOŁOWSKA Katarzyna (2013): Parole potentielle w teorii Ferdynanda de Saussure’a. – *Linguistica Copernicana* 2, 77–90.

doi: <http://dx.doi.org/10.17651/bptj.74.14>

EWA CISZEWSKA-JANKOWSKA
Uniwersytet Śląski, Katowice

Le futur gnomique

Czas przyszły gnomiczny w języku francuskim

Streszczenie

Celem artykułu jest opis gnomicznego użycia czasu przyszłego *futur simple* w języku francuskim. Może on wyrażać prawdy ogólne ponadczasowe w przysłowia, sentencjach, maksymach oraz innych formach wypowiedzi o znaczeniu ogólnym. Autorka porównuje użycia gnomiczne czasu przyszłego i teraźniejszego, a także wskazuje na podobieństwa i różnice między gnomicznym a czasowym (podstawowym) użyciem czasu przyszłego. Przeprowadzona analiza pozwala na stwierdzenie, że użycie gnomiczne różni się od innych szczególnych użyć czasu *futur simple*.

Słowa kluczowe: *futur simple*, czas przyszły, gnomiczny, przysłowia

The French *Futur gnomique*

Abstract

The present paper focuses on the description of the gnomic uses of the French Future Simple tense (*futur simple*). It can express common truths in proverbs, sayings, and in other forms of expression of general meaning. The Author compares the gnomic uses of the Future and Present tenses, and points out the similarities and differences that exist between the gnomic and primary uses of the French Future tense. The results of the analysis show that the gnomic use of the French Future Simple tense differs from its other special uses.

Key words: *futur simple*, future tense, gnomic, proverbs

Quand on analyse l'emploi des formes temporelles, on distingue les valeurs principales, de base, et les valeurs particulières, modales ou non, qui résultent de la combinaison des traits propres à la forme temporelle en question avec les éléments contextuels et les traits caractéristiques de la situation d'énonciation.

Certains linguistes considèrent qu'on ne peut pas assigner une valeur de vérité à un énoncé qui se rapporte à l'avenir (cf. Lyons 1977). De ce fait, le temps futur appartiendrait au domaine modal du non certain, contrairement au temps présent et aux temps passés qui feraient partie du domaine du certain, du réalisé (Desclés 1994). Nous adhérons¹ aux arguments de Gosselin (2005 et 2010) qui critique ces approches modales et constate qu'on peut douter de nombreux faits présents ou passés et éprouver des certitudes quant à l'avenir. Cette certitude peut être facilement exprimée et même soulignée dans la langue : *Je suis certain qu'il viendra* (2010 : 15–16).

Parmi les différents emplois du futur simple (FS), l'emploi gnomique paraît bien marginal. En général, il est mentionné dans les travaux qui évoquent les valeurs du FS, mais parfois, on refuse de le considérer comme un emploi particulier en constatant qu'il s'agit d'un emploi typique et temporel (Barceló & Bres 2006 : 107). Très souvent, on n'entre pas trop dans les détails et on s'interroge surtout sur son caractère modal. Cependant, l'emploi gnomique paraît beaucoup plus complexe qu'on ne le pense, et son analyse approfondie pourrait permettre de mieux comprendre les relations entre la valeur principale et les valeurs particulières des formes temporelles. Avec le présent article, nous avons l'intention d'apporter quelques précisions concernant cet emploi.

Dans notre analyse, nous nous appuyerons sur un corpus de proverbes, dictons, maximes et sentences recueillis dans différents dictionnaires (voir : Sources d'exemples). Nous n'allons pas entrer ici dans la polémique largement débattue (cf. entre autres Kleiber 1988, Anscombe 1994) concernant la définition du proverbe et d'autres formes sentencieuses car la distinction entre elles n'est pas pertinente pour notre étude.

Le FS est habituellement défini comme exprimant la postériorité par rapport au moment de l'énonciation. Il sert à situer un procès dans le futur de celui qui parle. Si quelqu'un énonce :

(1) Je vois que les hirondelles volent bas, donc il pleuvra bientôt.

il exprime un jugement personnel concernant le temps qu'il fera dans un avenir très proche. La prévision vaut uniquement pour la situation spacio-temporelle dans laquelle se trouve l'énonciateur.

De même, avec

(2) Les hommes ont été, sont et seront menés par les événements. (Voltaire)

Voltaire fait une remarque générale concernant les hommes, il parle de leur comportement dans le passé et il prévoit leur comportement futur : *seront menés* marque un fait situé dans la postériorité par rapport au moment où Voltaire énonce cette phrase.

La situation est différente lorsque quelqu'un dit :

(3) Selon le proverbe, hirondelle volant haut, le temps sera beau, hirondelle volant bas, bientôt il pleuvra.

ou bien :

¹ Cf. Ciszewska-Jankowska (2014 : 31–34).

- (4) Comme l'a dit Voltaire, les hommes ont été, sont et seront menés par les événements.

Celui qui parle est bien le locuteur du proverbe (3), mais il n'en est pas l'auteur ; « en termes de polyphonie, il n'est pas l'énonciateur du principe qui y est attaché » (Anscombe 1994 : 100). Le locuteur n'est pas responsable du contenu de ce proverbe, de sa forme, des mots employés, de leur combinaison (cf. Kleiber 1999 : 54) et, évidemment, il n'est pas responsable non plus des formes temporelles employées. Il en est de même en ce qui concerne les maximes, les pensées, les sentences, les aphorismes, qui ont un auteur spécifique, contrairement aux proverbes et dictons, dont l'auteur « est quelque chose comme une conscience linguistique collective » (Anscombe 1994 : 100). C'est l'auteur (au sens large de ce terme) qui a choisi la forme du FS, alors que le locuteur la répète seulement telle quelle. Il ne peut pas y avoir de relation temporelle directe. Dans (3), la forme *sera* marque la postériorité par rapport à la situation *hirondelle volant haut* et la forme *pleuvra*, par rapport à la situation *hirondelle volant bas*. Dans (4), la forme du FS indique la permanence du comportement humain dans l'avenir. Les énoncés à valeur générale tels que : proverbes, dictons, maximes, aphorismes, sentences, etc. constituent des unités indépendantes, autonomes, et la référence temporelle reste figée à l'intérieur de ces énoncés.

Essayons de déterminer les traits caractéristiques de l'emploi gnomique. Il est étroitement lié aux formes sentencieuses, ce qui résulte de la signification même de l'adjectif *gnomique*. Selon le dictionnaire *Larousse*, cet adjectif exprime des vérités morales sous forme de proverbes ou de maximes, et, lorsqu'il s'applique à une forme verbale telle qu'un temps ou un mode, il sert à marquer un fait général d'expérience. L'emploi gnomique n'est pas propre seulement au FS ; on distingue aussi le présent, le passé simple et le passé composé gnomiques. Pour Imbs (1968 : 47), l'emploi gnomique est l'occasion de faire un rapprochement entre le FS et le passé simple. Il constate que les deux formes sont toujours accompagnées de *ne...jamais* :

- (5) À l'égard des voleurs on ne sera jamais assez prudent.
 (6) Reprenez vos esprits ; et souvenez-vous bien qu'un dîner réchauffé ne valut jamais rien.

ou d'une forme grammaticale équivalente comme par exemple l'interrogation oratoire dans :

- (7) Un homme qui meurt de soif se rafraîchira-t-il, au souvenir d'une eau qu'il a bue jadis ?

Imbs (1968 : 48) remarque qu'il ne s'agit pas de vérités générales conçues de façon intemporelle. En ce qui concerne le passé simple, il s'agit de vérités d'expérience (ainsi le passé nous enseigne telle ou telle chose), alors que dans le cas du FS, on parle de conseils tirés d'expérience, mais formulées pour l'avenir. De cette façon, cet emploi se rapproche des valeurs modales du futur. Il faut tout de suite observer que lorsque Imbs parle du présent, du passé simple ou du passé composé gnomiques, il

n'évoque pas du tout la possibilité d'interprétation modale. Le présent gnomique est rangé directement du côté des valeurs temporelles. Ainsi on peut penser que la valeur modale qu'il entrevoit n'est pas liée à la valeur gnomique en général, mais seulement à la valeur gnomique du FS.

Remarquons aussi que l'emploi du futur gnomique ne se limite pas aux constructions évoquées par Imbs. L'analyse du corpus démontre qu'il peut apparaître dans des phrases à des structures très variées, par exemple :

- (8) Le loup mourra dans sa peau.
- (9) Qui vivra verra.
- (10) Graissez les bottes d'un vilain, il dira qu'on les lui brûle.
- (11) À la fin le renard sera moine.
- (12) Quand il tonne le premier jour de mai, les vaches auront du bon lait.

Le futur gnomique est le plus fréquent dans des phrases avec une subordonnée conditionnelle de type *Si p, q*, ou dans des phrases qui se laissent facilement réduire à la structure *Si p, q*. La condition peut être alors exprimée par un impératif, une proposition relative, un gérondif ou un autre élément lexical.

- (13) Graissez les bottes d'un vilain, il dira qu'on les lui brûle. = Si vous graissez les bottes d'un vilain,...
- (14) Tel qui rit vendredi... = Si quelqu'un rit vendredi...
- (15) Hirondelle volant haut, le temps sera beau... = Si l'hirondelle vole haut,...
- (16) Quand il tonne le premier jour de mai,... = S'il tonne le premier jour de mai,...

Rocci (2000 : 268) remarque que certains usages du FS, notamment le futur gnomique ou le futur d'habitude dans les énoncés ayant des sujets génériques, présentent une structure conditionnelle cachée.

Dans ses travaux sur le sens des proverbes, Kleiber (2000 : 52) défend l'idée que tous les proverbes ont un sens implicatif, même ceux dans lesquels, à première vue, il est impossible de retrouver une implication. Cela est dû au fait que le sens d'un proverbe n'est jamais entièrement « celui de la combinaison de ses constituants » (Kleiber 2000 : 51–52) et son interprétation standard n'est jamais donnée compositionnellement. Ainsi, si on tient compte uniquement du sens littéral de (8), il est difficile de le transformer en une structure de type *Si p, q*. Par contre, si l'on prend en considération le sens de ce proverbe, l'interprétation implicative devient tout à fait naturelle : si quelqu'un est méchant par nature, il ne se corrigera pas. (8) a un sens implicatif en tant que proverbe, mais non en tant qu'une simple phrase générique (Kleiber 2000 : 52). D'autres genres (maximes, dictons, sentences, etc.) ou énoncés à caractère général, dans lesquels est employé le futur gnomique, n'ont pas forcément un sens figuré. Parfois, leur structure donne facilement lieu à une implication, comme dans (12), mais dans d'autres cas, l'interprétation implicative est impossible, comme dans (4). Par conséquent, on ne peut pas affirmer que le futur gnomique soit lié uniquement à des énoncés présentant une structure conditionnelle cachée.

Dans les énoncés à caractère général, on emploie habituellement le temps présent. Il convient parfaitement à ce type d'énoncés car il peut englober toutes les époques : le présent, le futur et le passé. (cf. Riegel, Pellat, Rioul 1994). En même temps, le présent « fait figure de représentation temporelle “ neutre ”, c'est-à-dire qu'il représente le procès à un mode personnel, mais sans l'inscrire dans une temporalité quelconque » (Gouvard 1999 : 74). Ce trait du présent s'accorde avec le sens général des proverbes, dictons, maximes et sentences qui renvoient à des situations valables à tout moment (Gouvard 1999 : 74). Le futur gnomique, quant à lui, renvoie seulement à l'époque postérieure, par conséquent, son emploi est beaucoup plus restreint que celui du présent. On pourrait même avoir l'impression qu'il empiète sur le territoire du présent gnomique et prend sa place. Nous estimons que ce n'est pas le cas et que le FS joue son propre rôle, celui d'un temps qui se réfère à la postériorité. Pour le vérifier, nous allons examiner les énoncés à valeur générale où le FS peut alterner avec le présent et nous allons indiquer les différences entre les deux emplois. Là où la substitution du FS par le présent est impossible, nous allons essayer d'en expliquer la raison.

En effet, dans un certain nombre d'énoncés, le futur gnomique peut être substitué par le présent sans qu'il y ait une différence sensible du sens général :

- (17) Rira / rit bien qui rira / rit le dernier.
- (18) Calomniez, calomniez, il en restera / reste toujours quelque chose.
- (19) Jamais d'août la sécheresse, n'amènera / n'amène la richesse.
- (20) L'orgueilleux aimera / aime mieux se perdre que de demander son chemin.

D'ailleurs, les versions au présent fonctionnaient ou fonctionnent encore comme des variantes de (17), (18) et (19). La substitution du futur gnomique par le présent sans changement notable du sens est possible dans les énoncés qui se réfèrent à des situations ou des comportements valables aussi bien au futur qu'au passé ou au présent. Béchade (1986 : 54) appelle ce futur de permanence et remarque qu'il est toujours accompagné d'un complément de temps. Les compléments de temps employés le plus souvent sont *toujours* ou *jamais* ; même si dans certains cas, le complément est absent, on peut facilement l'ajouter pour rendre compte du caractère permanent de ce qui est exprimé dans l'énoncé :

- (21) Rira toujours bien qui rira le dernier.
- (22) L'orgueilleux aimera toujours mieux se perdre que de demander son chemin.

Avec le FS, on souligne que le comportement ou la situation décrite se poursuivra ou se répétera dans l'avenir. La conviction qu'il en sera ainsi résulte de l'expérience; la répétition de certains phénomènes dans le passé et dans le présent permet de tirer la conclusion quant à leur réapparition dans l'avenir.

Dans un nombre considérable d'énoncés, l'emploi du présent à la place du futur gnomique est possible, il entraîne cependant un certain changement de sens. Certains

proverbes et dictons sont utilisés aussi bien avec le présent qu'avec le FS, parfois avec des différences plus ou moins importantes dans la forme pour assurer le rythme et la rime.

- (23) Qui sème l'injustice, récoltera / récolte la haine.
- (24) Chassez-le par la porte, il rentrera / rentre par la fenêtre.
- (25) Faites du bien à un vilain, il vous fera / fait dans la main.
- (26) Quand le 1^{er} mai la pluie oint, il n'y aura pas le moindre coing.
- (26a) S'il pleut le 1^{er} mai, peu de coings ; s'il pleut le 2, ils sont véreux ; s'il pleut le 3, y en a pas.
- (27) Viendra le temps où la vache aura besoin de sa queue.
- (27a) Il vient un temps que les vaches ont besoin de leur queue.

Le FS est employé lorsqu'on désigne la conséquence postérieure d'un comportement, d'une situation ou d'une action particulière. Il est très fréquent dans les dictons météorologiques, ce qui est tout à fait naturel, vu leur fonction. En observant un phénomène naturel (pluie, neige, soleil, gel, sécheresse, orage, etc.) ou un comportement d'un animal (*hirondelle volant haut...*), on prévoit le temps qu'il fera dans l'avenir, ou bien on prévoit l'influence d'un tel phénomène sur les récoltes futures. Le FS est aussi habituel dans les proverbes et dictons qui ont une fonction didactique. Ils jouent le rôle d'une recommandation ou d'un avertissement : on prévoit des conséquences défavorables ou favorables d'un comportement actuel. On pourrait paraphraser ce type d'énoncés par : *si maintenant (à tel moment précis) p, alors dans l'avenir q*. L'emploi du FS fait que la conséquence de la situation actuelle ou du comportement décrit devient inévitable. Damourette et Pichon (1970 : 394) remarquent à ce propos que le FS est employé « quand il s'agit d'exprimer que la condition une fois remplie, le phénomène se produira inéluctablement ». Avec le présent, on souligne davantage le caractère habituel des phénomènes décrits :

- (23a) Il arrive habituellement que, si quelqu'un sème l'injustice, il récolte la haine.

Il peut également indiquer la rapidité de la conséquence :

- (24a) Il arrive habituellement que, si vous le chassez par la porte, il rentre immédiatement par la fenêtre.

On tire un enseignement du fait de la répétition de la relation où telle cause produit habituellement tel effet. Le raisonnement suit le cheminement suivant : *il arrive habituellement que si p, alors q, par conséquent, si maintenant (à tel moment précis) p, alors dans l'avenir q*. Ainsi, lorsque l'énoncé porte sur la conséquence située dans l'avenir, le FS rend cet énoncé plus direct et en même temps plus catégorique.

On peut également constater une différence de sens lorsqu'on remplace le FS par le présent dans les énoncés suivants :

- (28) Le temps n'est pas un loup, il ne fuira ($\neq$ fuit) pas dans les bois.
 (29) On conquerra ($\neq$ conquiert) le monde entier par la parole, mais non par un sabre tiré.
 (30) On reconnaît vite l'arbre qui portera ($\neq$ porte) des fruits.
 (31) L'ennemi ne se changera ($\neq$ change) pas en ami, ni le son en farine.

Dans ces exemples, le FS n'indique pas des faits qui doivent inéluctablement se produire dans l'avenir, comme c'est le cas dans (23) ; il signale qu'un fait ou une situation peut éventuellement avoir lieu. Les verbes au FS de (28)–(29) peuvent être paraphrasés par « *peut + infinitif* » :

- (28a) Le temps n'est pas un loup, il ne peut pas fuir dans les bois.
 (29a) On peut conquérir le monde entier par la parole, mais non par un sabre tiré.
 (30a) On reconnaît vite l'arbre qui peut porter des fruits.
 (31a) L'ennemi ne peut pas se changer en ami, ni le son en farine.

Clédat (1907 : 312) oppose ce type du FS (qu'il appelle futur d'habitude) aussi bien au présent qu'au futur « ordinaire ». Par rapport à ce dernier, le futur d'habitude ne sert pas à localiser un fait dans la période postérieure au moment d'énonciation, mais il a pour fonction d'indiquer la capacité du sujet, sa nature en général, sans référence à un moment bien précis. Clédat remarque que « l'idée de la réalisation future de l'action disparaît presque devant celle de la capacité présente de la réaliser » (Clédat 1907 : 312). Avec le présent, on insiste toujours sur le caractère habituel d'un fait, alors qu'avec le FS, on met en évidence son caractère potentiel. Ce type du FS apparaît non seulement dans des énoncés à valeur générale, mais il peut également caractériser des personnes ou objets spécifiques :

- (32) C'est un homme excellent : il rendra service même à ses ennemis. (Clédat 1907 : 312)

Il rendra service équivaut ici à *il peut rendre service, il est capable de rendre service* (Clédat 1907 : 312). Quel que soit le sujet (générique ou spécifique), le FS permet d'exprimer son trait caractéristique.

Certains énoncés à valeur générale se rapportent au futur et la forme du FS ne peut en aucun cas être remplacée par le présent. Il s'agit en premier lieu de situations où l'on met en parallèle des faits présents ou passés et des faits futurs. Le FS s'oppose alors soit au présent, soit à un temps passé : le passé composé (le plus souvent), le passé simple ou l'imparfait.

- (33) Qui a bu boira.
 (34) Le linge l'a apporté, le linceul l'emportera.
 (35) Tout fut autrui, tout sera autrui.
 (36) Avant notre venue, rien ne manquait au monde ; après notre départ, rien ne lui manquera.

Le FS a ici la même fonction que dans des énoncés à sujet spécifique :

- (37) Je ne dirai rien de M. Gadagne, homme content de lui, qui a fait du chemin et en fera encore. (Billy André, *Introïbo*, p. 43)

Dans tous les cas, il indique des faits postérieurs par rapport au moment à partir duquel s'effectue la comparaison.

Le FS est la seule forme qui puisse être employée dans les énoncés qui portent uniquement sur l'avenir. Tel est le cas dans les exemples suivants :

- (38) Qui vivra verra.
 (39) Quand les chats siffleront, à beaucoup de choses nous croirons.
 (40) Quand les vaches sont couchées, toutes de même côté, il fera mauvais.
 (41) Aide-toi, Dieu t'aidera.

L'emploi du présent à la place du FS changerait complètement le sens de ces énoncés. Dans (38), il ne serait plus question de l'inconnu que réserve l'avenir, et dans (39), une supposition absurde se transformerait en réalité. Dans (40) et (41), les deux faits pourraient être interprétés comme simultanés, ce qui ferait que (40) cesserait d'être une prévision pour le futur et (41) perdrait sa valeur didactique et moralisante.

Dans une situation spécifique, le présent peut être employé pour évoquer le futur lorsqu'il est accompagné d'un complément circonstanciel de temps à référence future. C'est le complément qui permet de situer le procès dans le futur :

- (42) Nous déménageons l'année prochaine.

On peut cependant observer que le présent se combine assez mal avec ce type de compléments dans des énoncés à valeur générale.

- (43) C'est toujours l'année prochaine que le paysan deviendra (?devient) riche.
 (44) À la fin saura-t-on (?sait-on) qui a mangé le lard.
 (45) Au jour de la résurrection, on te demandera (?demande) quels sont tes actes, et non quel est ton père.

Cela est dû au fait que pour pouvoir indiquer le futur, le présent doit se trouver dans un énoncé qui soit relié au moment de la parole (cf. Riegel, Pellat, Rioul 1994 : 300). Or, comme le locuteur n'est pas l'auteur d'un énoncé à valeur générale, celui-ci est forcément détaché du moment de la parole. L'interprétation habituelle est également exclue ; ces énoncés indiquent des faits uniques, qui ne peuvent pas se répéter, et dont on espère la réalisation dans le futur. Par conséquent, le FS ne peut pas être remplacé ici par le présent.

Le FS possède de nombreuses valeurs modales ; il peut notamment exprimer un ordre, un conseil ou une suggestion. Avec cette fonction, le FS peut apparaître dans des

énoncés à valeur générale, cependant il est relativement rare (une dizaine d'occurrences dans tout le corpus dont certaines venant de la Bible).

(46) Quand reviendra la Saint-Henri, tu planteras ton céleri.

(47) Tu n'emploieras ton labeur qu'en terre de bonne senteur.

(48) Tu aimeras ton prochain comme toi-même.

Le pronom *tu* a ici la valeur générique et ne renvoie pas à un individu défini. Selon Maingueneau (1999 : 24), le *tu* générique remplace ici le sujet universel (le *on*) et sert à « personnaliser » des énoncés à valeur générique. Pour obtenir le sens injonctif, il faut que le locuteur jouisse d'une certaine autorité et qu'il puisse imposer sa volonté. Comme on l'a déjà remarqué plusieurs fois, dans des énoncés à valeur générale, le statut du locuteur est bien particulier : il n'est pas l'auteur de l'énoncé, mais il est évident qu'il transmet ce que dit la « sagesse populaire » ou une autorité divine. L'emploi du pronom *tu* avec le FS reproduit une situation d'énonciation naturelle dans laquelle celui qui parle a le pouvoir d'ordonner quelque chose à l'interlocuteur, et celui-ci reconnaît son autorité et est prêt à exécuter cet ordre. Quand on remplace le FS des énoncés (46)–(48) par le présent, ils perdent leur caractère injonctif au profit de la valeur habituelle.

L'analyse que nous avons effectuée démontre que l'emploi du futur gnomique n'est pas homogène. D'un côté, on observe des occurrences où le FS, accompagné des adverbes *toujours* ou *jamais*, peut se référer non seulement à l'avenir, mais aussi au présent et au passé, exactement comme le présent. D'autre part, on relève de nombreux emplois où le FS joue son propre rôle, celui d'un temps qui indique la postériorité. Le FS dans des énoncés à valeur générale fonctionne alors de la même façon que le FS dans son emploi principal (temporel). Dans d'autres cas encore, il peut prendre la valeur injonctive ou indiquer le caractère potentiel du sujet. Ce sont des valeurs particulières que l'on rencontre aussi, et même avant tout, dans des énoncés non gnomiques. Le FS aurait donc sa valeur temporelle de base et de nombreuses valeurs particulières parmi lesquelles la valeur injonctive, la valeur habituelle (potentielle) et la valeur gnomique qui, à son tour, pourrait avoir la valeur temporelle, injonctive ou habituelle. Une telle approche des valeurs du FS, et de la valeur gnomique en particulier, semble cependant peu convaincante. Si les emplois du FS dans des énoncés à valeur générale se laissent tous ramener aux valeurs temporelle, injonctive ou habituelle, on peut se demander s'il est justifié de distinguer l'emploi gnomique comme une valeur particulière du FS. On peut se poser également la question de savoir si la nuance de sens que l'on perçoit entre les exemples (1) et (3) ou entre (2) et (4) est due à l'emploi du FS. Le caractère général de ces énoncés ne découle en aucun cas des traits aspectuo-temporels du FS. Les linguistes remarquent souvent à ce propos que ce caractère résulte des éléments génériques tels que par exemple le pronom relatif *qui* ou l'article défini (cf. entre autres Barcelo & Bres 1996). Mais ces éléments, à leur tour, peuvent être ambigus et leur interprétation dépend plus d'une fois du contexte. Ce qui détermine l'interprétation gnomique, c'est l'absence de relation spatio-temporelle

entre le locuteur et ce qui est prédiqué dans l'énoncé. Une telle situation a lieu lorsque le locuteur n'est pas l'auteur de cet énoncé. Cela fait que pour chaque locuteur (et chaque auditeur) les faits décrits par le FS appartiennent à l'avenir et le FS lui-même acquiert le caractère générique.

L'emploi gnomique diffère de tous les autres emplois particuliers par le fait qu'il ne résulte pas de la combinaison des traits inhérents de la forme temporelle avec les éléments contextuels et les traits caractéristiques de la situation de communication. Seuls les éléments contextuels et les caractéristiques de la situation de communication déterminent l'interprétation générique. Les propriétés de la forme temporelle ne contribuent pas à créer l'effet de sens gnomique.

L'emploi gnomique est bien un emploi particulier aussi bien du FS que d'autres formes temporelles, mais la façon dont cet effet de sens est produit nous oblige à le situer à un autre niveau, en dehors de la valeur principale et des autres valeurs particulières.

Références

- ANSCOMBRE Jean-Claude (1994) : Proverbes et formes proverbiales : valeur évidentielle et argumentative. – *Langue française* 102, 95–107.
- BARCELÓ Gérard Joan, BRES, Jacques (2006) : *Les Temps de l'indicatif en français*. – Paris : Ophrys.
- BÉCHADE Hervé-D. (1986) : *Syntaxe du français moderne et contemporain*. – Paris : PUF.
- CISZEWSKA-JANKOWSKA Ewa (2014) : *Le futur antérieur et ses emplois. Analyse contextuelle*. – Katowice : Wydawnictwo Uniwersytetu Śląskiego.
- CLÉDAT Léon [1907] (1973) : Le futur à la place du présent. – [w:] *Mélanges Chabaneau*. – Genève : Slatkine Reprints, 311–314.
- DAMOURETTE Jacques, PICHON Édouard [1936] (1970) : *Des mots à la pensée. Essai de grammaire de la langue française*, t. 5 – Paris : Éditions d'Artrey.
- DESCLÉS Jean-Pierre (1994) : Quelques concepts relatifs au temps et à l'aspect pour l'analyse des textes. – [w:] *Etudes cognitives*, t. 1. *Sémantique des catégories d'aspect et de temps*. – Warszawa : Slawistyczny Ośrodek Wydawniczy, 57–88.
- GOSSELIN Laurent (2005) : *Temporalité et modalité*. – Bruxelles : De Boeck-Duculot.
- GOSSELIN Laurent (2010) : *Les modalités en français. La validation des représentations*. – Amsterdam–New York : Rodopi.
- GOUVARD Jean-Michel (1999) : Les adages du droit français. – *Langue française* 123, 70–84.
- IMBS Paul (1968) : *L'emploi des temps verbaux en français moderne*. – Paris : Klincksieck.
- KLEIBER Georges (1988) : Sur la définition du proverbe. – *Recherches germaniques* 2, 232–252.
- KLEIBER Georges (1999) : Les proverbes : des dénominations d'un type « très très spécial ». – *Langue française* 123, 52–69.
- KLEIBER Georges (2000) : Sur le sens des proverbes. – *Langages* 139, 39–58.
- Larousse : <http://www.larousse.fr/dictionnaires/francais> (accessible : 30.11.2017).
- LYONS John (1977) : *Semantics*. – Cambridge University Press.
- MAINGUENEAU Dominique (1999) : *L'énonciation en linguistique française*. – Paris : Hachette.
- RIEGEL Martin, PELLAT Jean-Christophe, RIOUL René (1994) : *Grammaire méthodique du français*. – Paris : PUF.

ROCCI Andrea (2000) : L'interprétation épistémique du futur en italien et en français : une analyse procédurale. – *Cahiers de linguistique française* 22, 241–274.

Sources d'exemples

DOURNON Jean-Yves (1992) : *Mini-encyclopédie des proverbes et dictons de France*. – Paris : France Loisirs.

Frantext : <http://www.frantext.fr> (accessible : 30.11.2017).

MALOUX Maurice (1960) : *Dictionnaire des proverbes, sentences et maximes*. – Paris : Librairie Larousse.

MONTREYNAUD Florence, PIERRON Agnès, SUZZONI François (1989) : *Dictionnaire de proverbes et dictons*. – Paris : Dictionnaires Le Robert.

Sprawy administracyjne

LXXV ZJAZD POLSKIEGO TOWARZYSTWA JĘZYKOZNAWCZEGO W ROKU 2017

W dniach 18–19 września 2017 r. w Łodzi odbył się LXXV Zjazd Polskiego Towarzystwa Językoznawczego. Zarząd PTJ funkcję organizatora powierzył prof. UŁ Rafałowi Zarębskiemu. Obrady odbywały się w gmachu Wydziału Filologicznego Uniwersytetu Łódzkiego przy ulicy Pomorskiej 171/173. Honorowy patronat nad LXXV Zjazdem PTJ objął Rektor Uniwersytetu Łódzkiego, prof. dr hab. Antoni Różalski, dziekan Wydziału Filologicznego prof. dr hab. Joanna Jabłkowska oraz Komitet Językoznawstwa Polskiej Akademii Nauk. Dodatkowo LXXV Zjazd PTJ patronatem objął Prezydent Miasta Łodzi Hanna Zdanowska. Patronami medialnymi było Radio Łódź oraz TVP3 Łódź. Tematem LXXV Zjazdu były *Źródła w badaniach językoznawczych*. Wygłoszono osiem referatów podczas dwóch sesji plenarnych i przedstawiono cztery plakaty podczas sesji plakatowej. Ponadto wygłoszono trzydzieści sześć referatów w sekcjach. Zgodnie z tradycją przed rozpoczęciem obrad drugiego dnia Zjazdu odprawiona została msza święta za zmarłych członków PTJ.

Pierwszego dnia odbyło się Walne Zgromadzenie członków PTJ, którego prowadzenie powierzono prof. UŁ Rafałowi Zarębskiemu. Przewodnicząca Towarzystwa prof. dr hab. Krystyna Kleszczowa poinformowała o działalności Zarządu w czasie od poprzedniego zjazdu w Zielonej Górze. Ponadto przedstawiła sprawozdanie z działalności Zarządu z jego kończącej się trzyletniej kadencji. Stosowne sprawozdania złożyli prof. IJP PAN Rafał Górski (Komisja Rewizyjna) oraz prof. dr hab. Elżbieta Chrzanowska-Kluczevska (skarbnik). Zaproponowała podniesienie składki członkowskiej do 100 zł. Sprawę przedyskutowano oraz poddano głosowaniu, w wyniku którego składkę podwyższono przy jednym głosie sprzeciwu oraz dwóch wstrzymujących się. Sprawy związane z *Biuletynem Polskiego Towarzystwa Językoznawczego* przedstawiła prof. dr hab. Renata Przybylska, redaktor naczelny tego czasopisma. Dr Leszek Szymański zaprezentował zaś funkcjonowanie strony PTJ na Facebooku.

Zgodnie z przyjętym obyczajem w trakcie Walnego Zgromadzenia przewodnicząca Zarządu, prof. dr hab. Krystyna Kleszczowa, wymieniła członków PTJ, którzy zmarli w ostatnim roku, a byli to: prof. dr hab. Władysław Cyran, dr Krystyna

Długosz-Kurczabowa, dr hab. Antonina Grybosiowa, prof. UŚ, prof. dr hab. Halina Grzmil-Tylutki, prof. dr hab. Roman Kalisz, prof. dr hab. Violetta Koseska-Toszeza, dr hab. Maria Przybysz-Piwko, prof. UW, dr Halina Olszewska-Michalczykowa, dr Jarosław Szuta, mgr Anna Karasiowa oraz mgr Joanna Zamościńska-Kucała. Wspomnienia o zmarłych wygłosili: dr hab. Barbara Kudra, prof. UŁ o prof. dr hab. Władysławie Cyranie, prof. dr hab. Krystyna Waszakowa o dr Krystynie Długosz-Kurczabowej, dr hab. Aleksandra Niewiara, prof. UŚ o dr hab. Antoninie Grybosiowej, prof. UŚ, dr Krystyna Data o prof. dr hab. Halinie Grzmil-Tylutki, prof. dr hab. Piotr Stalmaszczyk o prof. dr hab. Romanie Kaliszu, dr hab. Roman Roszko, prof. IS PAN o prof. dr hab. Violetcie Koseskiej-Toszewej, prof. dr hab. Krystyna Waszakowa o dr hab. Marii Przybysz-Piwko, prof. UW, prof. dr hab. Krystyna Kleszczowa o dr Halinie Olszewskiej-Michalczykowej, dr Marek Kaszewski o dr. Jarosławie Szucie, a dr Joanna Duska o mgr Annie Karasiowej i mgr Joannie Zamościńskiej-Kucałowej. Pamięć zmarłych uczczono minutą ciszy.

Sekretarz Zarządu PTJ dr Ewelina Grześkiewicz odczytała listę trzynastu nowych członków przyjętych podczas posiedzeń Zarządu. Jednogłośnie przyjęty został także wniosek o nadanie statusu honorowego członka prof. dr hab. Maciejowi Grochowskiemu, prof. dr hab. Januszowi Siatkowskiemu oraz dr Krystynie Wilczewskiej.

W związku z upływającą kadencją obecnego Zarządu PTJ odbyły się wybory nowego Zarządu. Pięcioosobowa komisja skrutacyjna pilnowała, by wybory przebiegały zgodnie ze Statutem PTJ. Poddano głosowaniu jedenaście nazwisk, oddano pięćdziesiąt cztery ważne głosy. Po naradzie ukonstytuował się nowy Zarządu PTJ: prof. dr hab. Piotr Stalmaszczyk (przewodniczący), dr hab. Rafał Górski, prof. IJP PAN (zastępca przewodniczącego), dr Krystyna Data (sekretarz), prof. dr hab. Elżbieta Chrzanowska-Kluczevska (skarbnik), członkowie: dr hab. Tomasz Mika, prof. UAM, dr hab. Małgorzata Nowakowska, prof. UP, prof. Krystyna Waszakowa, dr hab. Rafał Zarębski, prof. UŁ. Odbyły się także wybory uzupełniające do Komisji Rewizyjnej w związku z wyborem dr. hab. Rafała Górskiego, prof. IJP PAN do Zarządu PTJ. W wyniku głosowania w skład Komisji Rewizyjnej weszła dr hab. Anna Kozłowska.

Na zakończenie zakomunikowano, że LXXVI Zjazd PTJ odbędzie się w Gdańsku, a głównym organizatorem zostanie dr hab. Katarzyna Wojan, prof. UG. Tematem będą *Kontakty językowe: historia i współczesność*.

Ewelina Grześkiewicz
sekretarz Zarządu PTJ

Lista nowych członków PTJ

Dominik Chomik (UG)
Ewa Ciszewska-Jankowska (UŚ)
Magdalena Czachorowska (UKW)
Joanna Darda-Gramatyka (UP)
Ewa Kaptur (UAM)
Małgorzata Karczevska (UZ)
Irmina Kotlarska (UZ)
Anna Krzywicka-Ustrzycka (UO)
Marta Noińska (UG)
Dorota Rojszczak-Robińska (UAM)
Gabriela Szewczyk (UP)
Sławomir Wacewicz (UMK)
Przemysław Żywicznyński (UAM)

PROGRAM LXXV ZJAZDU POLSKIEGO TOWARZYSTWA
JĘZYKOZNAWCZEGO, 18–19 WRZEŚNIA 2017, ŁÓDŹ

Źródła w badaniach językoznawczych

PONIEDZIAŁEK, 18 września 2017 r.

Obrady plenarne

Renata Przybylska, *Próba typologii źródeł do badań nad współczesnym językiem polskim*
Barbara Lewandowska-Tomaszczyk, Paul A. Wilson, *Źródła danych i podstawy metodologiczne semantycznych badań kontrastywnych pojęć wyrażających emocje i uczucia*
Tomasz Mika, *Teoria piśmienności jako kontekst badań historycznojęzykowych. Staropolskie źródła tłumaczone*
Mirosław Bańko, *Korzyści i problemy wynikające z „nadmiaru” danych językowych i z wielości metod ich analizy*

Sesja plakatowa

Dorota Kołodziej, Katarzyna Jasińska, Magdalena Klapper, Maciej Godny, *Słowniki lacińskie – źródło w badaniach nad staropolszczyzną*
Emil Popławski, Ewelina Grześkiewicz, *Źródła „Słownika gwar polskich” PAN*
Adam Dombrowski, *Zasoby cyfrowe przydatne w pracy językoznawcy-polonisty jako źródło pozyskiwania materiału badawczego*

Daniel Śledziński, *Wielowarstwowy model transkrypcji fonematycznej tekstów polskich*

Obrady w sekcjach

Sekcja A

Krystyna Data, *Źródła źródeł w badaniach językoznawczych*

Sławomir Waciewicz, Przemysław Żywicznyński, *Ewolucja języka – pojęcia, praktyka badawcza, źródła danych*

Iwona Dembowska-Wosik, *Dlaczego warto opierać badania wyłącznie na jednym źródle. Zastosowanie metodologii studium przypadku w językoznawstwie*

Olga Lesicka, *Rola źródeł w badaniach nad zapożyczeniami terminologicznymi*

Wanda Małgorzata Stec, *Materiały źródłowe w polskich i rosyjskich badaniach nazewnictwa botanicznego*

Sekcja B

Aleksandra Janowska, *Źródła a typy badań historycznych*

Mariusz Leńczuk, *Rękopiśmienne źródła w badaniach językowych najstarszej doby polszczyzny*

Ewa Woźniak, Rafał Zarębski, *Typ źródeł a wyniki badań – uwagi na marginesie badań nad tendencjami rozwojowymi polskiej leksyki prawnej pod wpływem polskich tłumaczeń Kodeksu Napoleona*

Ewa Deptuchowa, Mariusz Frodyma, *Wpływ rozariuszy na wzbogacenie leksyki staropolskiej*

Dorota Masłej, *Gdy glosy układają się w tekst. „Kazania augustiańskie” jako źródło w badaniach historycznojęzykowych*

Sekcja C

Agata Kwaśnicka-Janowicz, Patrycja Pałka, *„Przewodnik po elektronicznych zasobach językowych dla polonistów” – nowe technologie w badaniach lingwistycznych*

Mirosława Podhajecka, *Google Books jako korpus językowy*

Ewelina Grześkiewicz, *Źródła „Słownika gwar Ostródzkiego, Warmii i Mazur”. Jaki materiał wykorzystuje dialektolog leksykograf*

Łukasz Grabowski, Milena Hebał-Jezierska, *O korpusowych metodach badawczych: próba krytycznej refleksji*

Witold Kieraś, Łukasz Kobyliński, Maciej Ogrodniczuk, *Korpusomat: łatwe tworzenie korpusów językowych z dowolnych źródeł tekstowych*

Sekcja D

Donata Ochmann, *Źródła w badaniach regionalizmów (na przykładzie leksyki Krakowa)*

Maciej Rak, *Testamenty chłopów polskich jako źródło badań dialektologicznych, historycznojęzykowych i etnolingwistycznych (uwagi wstępne)*

- Michał Głuszkowski, *Hierarchizacja źródeł w dialektologicznych badaniach podłużnych (na przykładzie dwujęzycznej społeczności staroobrzędowców w Polsce)*
- Tomasz Kurdyła, *Źródła leksykograficzne do słownika gwarowych gniazd słowotwórczych*
- Angelika Pawlaczyk, *Znaczenie metody biograficznej w badaniach gwary polskich staroobrzędowców*

WTOREK, 19 września 2018 r.

Sekcja A

- Michał Szczyszek, *Źródła w badaniach językowych. Polszczyzna mówiona: gromadzenie, przetwarzanie, analiza danych językowych*
- Anetta Luto-Kamińska, *Słowo w (kon)tekście i w słowniku*
- Katarzyna Wojan, *Bibliografie słowników przekładowych jako narzędzie rekonstrukcji dziejów oraz ocen stanu i dorobku słownictwa narodowego*
- Błażej Osowski, *Hiperyzmy – kiedy źródło nas mami*

Sekcja B

- Urszula Wójcik, *Średniowieczne zasoby onimiczne – niewykorzystane źródło w badaniach historycznojęzycznych? Prolegomena badawcze*
- Dorota Rojszczak-Robińska, *Łacińskie źródła jako kontekst badań historycznojęzycznych*
- Joanna Kulwicka-Kamińska, *Piśmiennictwo religijne Tatarów Wielkiego Księstwa Litewskiego jako źródło badań historycznojęzycznych*
- Agnieszka Piel, *Źródła w studiach nad frazeologią historyczną*

Sekcja C

- Iwona Kosek, Monika Czerepowicka, Sebastian Przybyszewski, *Źródła w tworzeniu elektronicznego słownika paradygmatów frazeologizmów werbalnych*
- Artur Czesak, Łukasz Szałkiewicz, Sebastian Żurowski, *A może jednak „wziąć na tapetę”? O roli źródeł pisanych w tworzeniu internetowego słownika poprawnej polszczyzny*
- Jerzy Tomaszczyk, *Źródła do słownika potocyzmów*
- Mariola Szymczak-Rozlach, *Problem gromadzenia leksyki ekspresywnej w korpusie narodowym języka słowackiego*

Sekcja D

- Anna Kozłowska, *Rola źródeł w badaniu idiolektów*
- Artur Bracki, *Teksty pisarzy dwujęzycznych (dwukulturowych) jako źródło analiz językoznawczych*
- Leszek Szymański, *Wykorzystanie korpusu współczesnej amerykańskiej odmiany języka angielskiego w badaniu interakcji modalności z aspektem*

Rafał Górski, Magdalena Król, *Co o imiesłowach przysłówkowych mówi Narodowy Korpus Języka Polskiego?*

Obrady plenarne

Maciej Grochowski, *Kompetencja językowa, reprezentacja semantyczna a NKJP (na przykładzie stopniowania przymiotników)*

Małgorzata Nowakowska, *Jak analiza dwujęzycznego korpusu prowadzi do odkrycia aspektowych „covert categories”*

Aleksander Kiklewicz, *Korpus internetowy jako źródło informacji lingwistycznej: ograniczenia*

Michaił Kotin, *Możliwości i ograniczenia w zastosowaniu korpusów w badaniach nad gramatyką*

Autorzy

MIROSLAW BAŃKO, profesor zwyczajny, dyrektor Instytutu Języka Polskiego Uniwersytetu Warszawskiego; e-mail: m.banko@uw.edu.pl.. W latach 1991–2000 najpierw kierownik Redakcji Słowników Języka Polskiego w Wydawnictwie Naukowym PWN, potem redaktor naczelny słowników języka polskiego PWN. Zainteresowania naukowe: gramatyka współczesnej polszczyzny, językoznawstwo normatywne, leksykografia, semantyka leksykalna. Autor lub współautor kilkunastu słowników różnego rodzaju oraz trzech monografii: *Z pogranicza leksykografii i językoznawstwa. Studia o słowniku jednojęzycznym* (2001), *Współczesny polski onomatopeikon. Ikoniczność w języku* (2008), *Nie całkiem obce. Zapożyczenia wyrazowe w języku polskim i czeskim* (2016). Członek Polskiego Towarzystwa Naukowego.

MONIKA BIESAGA, doktor, pracownik naukowy Instytutu Języka Polskiego PAN w Krakowie; e-mail: monika.biesaga@interia.pl, https://www.researchgate.net/profile/Monika_Biesaga. Jej zainteresowania badawcze skupiają się wokół komunikacji multimodalnej w leksykografii, semantyki leksykalnej oraz leksykologii, zaś zainteresowania organizacyjne wokół zagadnień promocji nauki (także w Internecie) oraz międzynarodowej współpracy badawczej. Autorka monografii *Historia i retoryka. Wykład akademicki o dziejach w świetle francuskiej teorii dyskursu* (2012) oraz licznych polsko- i angielskojęzycznych artykułów naukowych. Członkini Polskiego Towarzystwa Językoznawczego, Zarządu Głównego Towarzystwa Miłośników Języka Polskiego, European Association of Lexicography (EURALEX); stypendystka DAAD, wykładowczyni, popularyzatorka nauki.

MONIKA CZEREPOWICKA, doktor, adiunkt w Instytucie Polonistyki i Logopedii Uniwersytetu Warmińsko-Mazurskiego w Olsztynie; e-mail: czerepowicka@gmail.com. Autorka artykułów z zakresu polskiej fleksji i składni, monografii *Opis powierzchniowoskładniowy wyrażen niestandardowych typu na lewo, do dziś, po trochu, na zawsze we współczesnym języku polskim* (2006). Współautorka *Słownika elektronicznego jednostek frazeologicznych* SEJF (2012). Redaktor naukowy trzech wydań *Czasownika polskiego* Zygmunta Saloniego (2001, 2002, 2006). Anotatorka korpusu IPI PAN. Jej zainteresowania naukowe koncentrują się wokół rygorystycznego opisu współczesnej polszczyzny, lingwistyki korpusowej oraz leksykografii elektronicznej.

EWA CISZEWSKA-JANKOWSKA, doktor, starszy wykładowca w Instytucie Języków Romańskich i Translatoryki Uniwersytetu Śląskiego w Katowicach; e-mail: ewa.ciszewska@us.edu.pl. Główne zainteresowania naukowe: aspekt, modalność, czasy gramatyczne (w szczególności francuskie czasy przyszłe futur antérieur i futur simple), ich szczególne użycia oraz ich polskie ekwiwalenty. Autorka monografii *Expression de la perfectivité en français contemporain* (2002) i *Le futur antérieur et ses emplois. Analyse contextuelle* (2014).

ALEKSANDRA JANOWSKA, doktor habilitowany, pracownik naukowy w Instytucie Języka Polskiego Uniwersytetu Śląskiego; e-mail: aleksandra.janowska@us.edu.pl. Interesuje się przede wszystkim zagadnieniami związanymi z rozwojem języka, zwłaszcza słowotwórstwem i leksyką. Najważniejsze publikacje: *Funkcje przestrzenne przedrostków czasownikowych w polszczyźnie* (1999), *Polisemia staropolskich czasowników. Źródła, swoistość, konsekwencje* (2007), *Kształtowanie się polskiej klasy przyimków wtórnych* (2015). Jest także współautorką pracy poświęconej derywacji czasownikowej w staropolszczyźnie: *Słowotwórstwo czasowników staropolskich. Stan i tendencje rozwojowe* (2005; współautorstwo Magdalena Pastuch).

IWONA KOSEK, doktor habilitowany, pracownik naukowo-dydaktyczny w Instytucie Polonistyki i Logopedii Uniwersytetu Warmińsko-Mazurskiego w Olsztynie, kierownik Zakładu Współczesnego Języka Polskiego, współpracownik Sekcji Frazeologicznej Komitetu Językoznawstwa PAN; e-mail: i.kosek@uwm.edu.pl. Zainteresowania naukowe skupiają się wokół gramatyki (fleksji i składni) języka polskiego, frazeologii w ujęciu synchronicznym i diachronicznym, leksykografii. Obecnie kieruje projektem tworzenia elektronicznego słownika paradygmatów frazeologizmów czasownikowych (finansowanym przez NCN). Wybrane publikacje i prace redakcyjne: *Przyczasownikowe frazy przyimkowo-nominalne wo* (1999); *Fleksja i składnia nieciągłych imiennych jednostek leksykalnych* (2008); (współredaktor) *Nowe zjawiska w języku, tekście i komunikacji V – Frazeologizmy* (2015).

MICHAŁ KOTIN, profesor zwyczajny Uniwersytetu Zielonogórskiego, kierownik Zakładu Gramatyki i Historii Języka Niemieckiego UZ; e-mail: michailkotin1@gmail.com. Zainteresowania naukowe: językoznawstwo historyczne, teoria zmian językowych, gramatyka, teoria nominacji, filozofia języka, metodologia językoznawstwa, typologia lingwistyczna. Autor 6 monografii oraz ponad 110 publikacji w czasopismach naukowych, tomach zbiorowych i pokonferencyjnych, licznych recenzji i sprawozdań, redaktor lub współredaktor 6 tomów zbiorowych i pokonferencyjnych. Stypendysta Fundacji im. Aleksandra Humboldta (RFN), członek 6 stowarzyszeń naukowych. Najważniejsze publikacje książkowe to monografie *Die werden-Perspektive und die werden-Periphrasen im Deutschen* (2003); *Die Sprache in statu movendi* (2 tomy) (2005–2007, 2012); książki pod redakcją: *Geschichte und Typologie der Sprachsysteme* (współredaktor Elizaveta Kotorova) (2011); *To Be or not to Be? The Verbum substantivum from Synchronic, Diachronic and Typological Per-*

spectives (współredaktor Richard Whitt) (2015). Członek Polskiego Towarzystwa Językoznawczego od 2016 r.

ANNA KOZŁOWSKA, doktor habilitowany, pracownik Wydziału Nauk Humanistycznych Uniwersytetu Kardynała Stefana Wyszyńskiego w Warszawie, kierownik Zakładu Badań nad Językiem Autorów w Instytucie Filologii Polskiej; email: a.kozłowska@uksw.edu.pl. Główne zainteresowania naukowe: język pisarzy, język religijny, metodologia językoznawstwa, stylistyka, semantyka i składnia. Autorka monografii: *Chrześcijaństwo w pismach Cypriana Norwida* (2000, pod nazwiskiem Anna Kadyjewska; współautorzy: Tomasz Korpysz, Jadwiga Puzynina) oraz *Od psalmów słowiańskich do rzymskich medytacji. O stylu artystycznym Karola Wojtyły* (2013); współredaktorka (wraz z Tomaszem Korpyszem) serii prac poświęconych językowi pisarzy: *Język pisarzy jako problem lingwistyki* (2009), *Język pisarzy: problemy słownictwa* (2011), *Język pisarzy: problemy metajęzyka i metatekstu* (2015).

JOANNA KULWICKA-KAMIŃSKA, doktor habilitowany, adiunkt w Instytucie Języka Polskiego UMK w Toruniu, kierownik Zakładu Historii Języka Polskiego, Centrum Badań Kitabistycznych i Podyplomowego Studium Logopedii UMK; e-mail: Joanna.Kulwicka-Kaminska@umk.pl. Prowadzi prace badawcze z zakresu językoznawstwa diachronicznego i kontrastywnego (historia języka polskiego, w tym staro- i średnio-polski system fleksyjny; słowiańsko-orientalne relacje językowe, m.in. zapożyczenia do polszczyzny z języków arabskiego, tureckiego i perskiego; dialektologia; kitabistyka) i translatoryki (przekład na język polski terminologii religijnej islamu i chrześcijaństwa). Autorka trzech książek, m.in. *Kształtowanie się polskiej terminologii muzułmańskiej* (2004), *Przekład terminologii religijnej islamu w polskich tłumaczeniach Koranu na tle biblijnej tradycji translatorycznej* (2013) oraz ponad stu artykułów naukowych. Redaktor 25 monografii zbiorowych. Należy do: Polskiego Towarzystwa Językoznawczego, Towarzystwa Miłośników Języka Polskiego, Towarzystwa Naukowego w Toruniu, Renaissance Society of America.

BARBARA LEWANDOWSKA-TOMASZCZYK, profesor zwyczajny i kierownik Katedry Badań nad Językiem, Literaturą i Przekładem w Państwowej Wyższej Szkole Zawodowej w Koninie, przez długi czas sprawowała funkcję kierownika Katedry Języka Angielskiego i Językoznawstwa Stosowanego w Uniwersytecie Łódzkim; e-mail: barbara.lewandowska.tomaszczyk@gmail.com. Zainteresowania badawcze: semantyka i pragmatyka kognitywna, językoznawstwo korpusowe oraz jego zastosowania w językoznawstwie kontrastywnym oraz studiach nad przekładem. Jest autorką licznych publikacji, zapraszana na konferencje międzynarodowe oraz do prowadzenia seminariów oraz cyklów wykładów w Polsce i za granicą.

ANGELIKA PAWLACZYK, magister, absolwentka filologii rosyjskiej, doktorantka UMK; e-mail: a_pawlaczyk@wp.pl. Zainteresowania naukowe skupiają się na problemach funkcjonowania rosyjskiej gwary w tekstach pisanych staroobrzędowców oraz doko-

nujących się zmianach w ich dialekcie pod wpływem polskiego otoczenia językowego. Badania terenowe prowadzi wśród staroobrzędowców mieszkających w ośrodku augustowskim oraz Polaków na Ukrainie (Stryj).

MIROSLAWA PODHAJECKA, doktor habilitowany, adiunkt w Instytucie Filologii Angielskiej Uniwersytetu Opolskiego; email: mira.podhajecka@wp.pl. Zainteresowania naukowe obejmują leksykografię polsko-angielską i angielsko-polską, zagadnienia przekładowe oraz problematykę kontaktów językowych. Ważniejsze publikacje: *Russian borrowings in English: A dictionary and corpus study* (2013), *A history of Polish-English / English-Polish bilingual lexicography (1788–1947)* (2016).

SEBASTIAN PRZYBYSZEWSKI, doktor, pracownik Instytutu Polonistyki i Logopedii na Uniwersytecie Warmińsko-Mazurskim w Olsztynie; e-mail: sebastian.przybyszewski@wp.pl. Jego zainteresowania obejmują głównie zagadnienia związane z pragmatyką, semantyką, teorią języka, frazeologią, lingwistyką tekstu i akwizycją języka drugiego. Obecnie zajmuje się zagadnieniami implicytności i literalności w komunikacji oraz zależnościami między pośrednimi aktami mowy a kontekstem. Współredaktor monografii *Nowe zjawiska w języku III. Kontekst a komunikacja* (2011) oraz *Język polski – nie taki obcy* (2015).

MACIEJ RAK, doktor habilitowany, pracownik naukowy Katedry Historii Języka i Dialektologii Wydziału Polonistyki UJ; e-mail: maciej.rak@uj.edu.pl. Główne zainteresowania naukowe: dialektologia polska (zwłaszcza gwary Karpat), etnolingwista. Autor książek: *Słownik frazeologiczny gwary Dębna w Górach Świętokrzyskich* (2005), *Językowo-kulturowy obraz zwierząt utrwalony w animalistycznej frazeologii gwar Gór Świętokrzyskich i Podtatrza (na tle porównawczym)* (2007), *Materiały etnograficzne z Podhala Ignacego Moczydłowskiego* (2011), *Kulturemy podhalańskie* (2015), *Materiały do etnografii Podhala* (2016). Redaktor 6 prac zbiorowych.

DOROTA ROJSZCZAK-ROBIŃSKA, doktor habilitowany, adiunkt w Zakładzie Historii Języka Polskiego Instytutu Filologii Polskiej i Klasycznej UAM w Poznaniu; e-mail: dorroj@amu.edu.pl. Zainteresowania naukowe: historia języka, staropolszczyzna, tekstologia, stylistyka historyczna. Zajmuje się staropolską literaturą religijną, a badania prowadzi na pograniczu językoznawstwa i literaturoznawstwa. Ważniejsze publikacje: *Jak pisano „Rozmyślanie przemyskie”* (2012), *Staropolskie pasje: Rozmyślanie przemyskie, Sprawa chędogą, Karta Rogawskiego. Źródła – język – fabuła* (2016).

PAUL A. WILSON, profesor w Katedrze Języka Angielskiego i Językoznawstwa Stosowanego Uniwersytetu Łódzkiego; e-mail: paul.wilson@uni.lodz.pl. Uzyskał stopień doktora w dziedzinie badań nad współzależnością poznania, uczuć i emocji w Birkbeck (University of London) w 2000 r. Głównym tematem zainteresowań badawczych jest reprezentacja konceptualna pojęć emotywnych w perspektywie kulturowej.

KATARZYNA WOJAN, doktor habilitowany, profesor nadzwyczajny w Instytucie Skandynawistyki Uniwersytetu Gdańskiego; kierownik jednostki Pracownia Języka, Kultury i Gospodarki Finlandii; e-mail: finkw@univ.gda.pl. Zainteresowania naukowe obejmują: fennistykę, językoznawstwo porównawcze słowiańskie i ugrofińskie, semantykę leksykalną, leksykografię, bibliografistykę; twórczyni szkoły metodologiczno-badawczej w zakresie homonemiki międzyjęzykowej. Autorka i współautorka ok. 180 publikacji, w tym 17 książek, m.in. *Wstęp do badań wieloznaczności leksemów w ujęciu kontrastywnym* (2010), *Przypadkowe i nieprzypadkowe wędrówki leksemów* (2010), *Rosyjskie leksemy homonimiczne w teorii językoznawczej i praktyce leksykograficznej* (2011), *Języki sztuczne. Zapotrzebowanie społeczeństw czy fantazja jednostek? Zarys dziejów interlingwistyki* (2015), *Język fiński w teorii i praktyce* (2015). Redaktor naczelna „Studia Rossica Gedanensia”, redaktor „Studia Scandinavica”. Członek m.in. European Acoustics Association, Polskiego Towarzystwa Akustycznego i Polskiego Towarzystwa Językoznawczego.

Lista recenzentów

prof. dr hab. Mirosław Bańko (Uniwersytet Warszawski)
prof. dr hab. Jerzy Bartmiński (Polska Akademia Umiejętności)
prof. dr hab. Zofia Berdychowska (Uniwersytet Jagielloński)
prof. dr hab. Bogusław Bierwiaczonek (Uniwersytet Humanistyczno-Przyrodniczy
im. Jana Długosza w Częstochowie)
prof. dr hab. Anna Bochnakowa (Uniwersytet Jagielloński)
dr hab. Waldemar Czachur, prof. UW (Uniwersytet Warszawski)
prof. dr hab. Janusz Danecki (Uniwersytet Warszawski)
prof. dr hab. Bogusław Dunaj (Państwowa Wyższa Szkoła Zawodowa w Tarnowie)
dr hab. Maciej Eder (Instytut Języka Polskiego PAN)
prof. dr hab. Stanisław Gajda (Uniwersytet Opolski)
prof. dr hab. Iryda Grek-Pabisowa (Instytut Sławistyki PAN)
prof. dr hab. Maciej Grochowski (Uniwersytet Mikołaja Kopernika w Toruniu)
prof. dr hab. Włodzimierz Gruszczyński (Instytut Języka Polskiego PAN)
prof. dr hab. Ewa Jędrzejko (Uniwersytet Śląski)
prof. dr hab. Henryk Kardela (Uniwersytet Marii Curie-Skłodowskiej)
prof. dr hab. Halina Karaś (Uniwersytet Warszawski)
prof. dr hab. Stanisław Koziara (Uniwersytet Pedagogiczny w Krakowie)
prof. dr hab. Barbara Lewandowska-Tomaszczyk (Uniwersytet Łódzki)
dr hab. Tadeusz Lewaszkiewicz, prof. UAM (Uniwersytet im. Adama Mickiewicza
w Poznaniu)
prof. dr hab. Alfred F. Majewicz (Uniwersytet Mikołaja Kopernika w Toruniu)
dr hab. Tomasz Mika, prof. UAM (Uniwersytet im. Adama Mickiewicza w Poznaniu)
prof. dr hab. Rafał Molencki (Uniwersytet Śląski)
prof. dr hab. Anna Pajdzińska (Uniwersytet Marii Curie-Skłodowskiej)
prof. dr hab. Janusz Rieger (Uniwersytet Warszawski)
dr hab. Mirosława Siuciak, prof. UŚ (Uniwersytet Śląski)
prof. dr hab. Marek Stachowski (Uniwersytet Jagielloński)
prof. dr hab. Jolanta Szpyra-Kozłowska (Uniwersytet Marii Curie-Skłodowskiej)
dr hab. Dorota Szumska (Uniwersytet Jagielloński)
prof. dr hab. Jolanta Tambor (Uniwersytet Śląski)

prof. dr hab. Teresa Tomaszek (Uniwersytet im. Adama Mickiewicza w Poznaniu)
prof. dr hab. Jadwiga Waniakowa (Instytut Języka Polskiego PAN)
prof. dr hab. Krystyna Waszakowa (Uniwersytet Warszawski)
dr hab. Ewa Willim (Uniwersytet Jagielloński)
prof. dr hab. Krzysztof Tomasz Witczak (Uniwersytet Łódzki)
prof. dr hab. Bożena Witosz (Uniwersytet Śląski)
prof. dr hab. Bogusław Wyderka (Uniwersytet Opolski)
dr hab. Włodzimierz Wysoczański, prof. UWr (Uniwersytet Wrocławski)
prof. dr hab. Piotr Żmigrodzki (Instytut Języka Polskiego PAN)

Zalecenia edytorskie dotyczące publikowania artykułów w Biuletynie Polskiego Towarzystwa Językoznawczego

Objętość: maksymalnie 15 stron x 1800 znaków = 27000 znaków z bibliografią umieszczoną na dole strony według poniższego wzoru.

Czcionka: Times New Roman, czcionka 12 p.

Marginesy: górny – 2,5, dolny – 2,5, lewy – 2,5, prawy – 2,5.

Interwał: 1,5.

Tabulator (wielkość wcięcia wierszy akapitowych): 1,0 cm.

IMIĘ NAZWISKO Autora: lewy górny róg (KAPITALIKI, pierwsza litera imienia i nazwiska: wersalik, czcionka 12): Jan KOWALSKI

Miejsce zatrudnienia (instytucja, miasto): lewy górny róg pod imieniem i nazwiskiem autora, czcionka 12.

Tytuł: środek, druk półtłusty (bold), czcionka 14.

Tekst: czcionka 12, justowanie do obu marginesów.

Prosimy nie stosować spacji w celu wyrównywania tekstu, przesuwania pojedynczych liter do następnego wersu, a także przy druku rozstrzelonym (druk rozstrzelony wg formatu .doc lub .docx).

Przypisy: u dołu strony, czcionka 10.

Cytaty w tekście:

cytaty krótsze (nie dłuższe niż jeden wiersz) winny być podawane bezpośrednio w tekście artykułu;

cytaty dwuwierszowe lub dłuższe proszę podawać jako odrębny blok, wcięcie 0,75 cm, czcionka 11 p.

Cytowana bibliografia: bezpośrednio w tekście (nie w przypisach), w nawiasie (Autor/Redaktor rok wydania: stronica); np. (Kowalski 1986: 142), (Kowalski 1992: 16–21) (**Nie** Kowalski 1992: 16 i n.!). Jeśli nazwisko autora/redaktora jest wkomponowane w tekst artykułu - notacja typu: „według Kowalskiego (1992: 16–21)...”, „trudno zgodzić się z Kowalskim (1992: 16–21)”.

Tytuły czasopism proszę podawać w pełnym brzmieniu (**Nie** skróty tytułów).

Wykaz bibliografii:

Na końcu artykułu. W wykazie proszę podawać **tylko** bibliografię cytowaną w artykule.

Pozycje bibliograficzne w alfabetych opartych na cyrylicy proszę podawać w oryginalnej pisowni.

Porządek alfabetyczny. Przy kilku publikacjach jednego autora – porządek chronologiczny; jeśli cytowane są dwie (lub więcej) publikacje jednego autora wydane w tym samym roku, należy je rozróżnić literami alfabetu łacińskiego [NOWAK Józef (2003a), NOWAK Józef (2003b)].

Pozycje pisane **cyrylicą (grażdanką)** proszę podawać po wszystkich pozycjach w alfabecie łacińskim.

Format bibliografii:

Autor / Redaktor: NAZWISKO (KAPITALIKI) Imię; po imieniu redaktora: „red.” oddzielone od imienia przecinkiem.

Rok wydania: w nawiasie po imieniu autora / redaktora; przed tytułem: dwukropek.

Tytuł publikacji książkowej / czasopisma: *kursywa*

Wzór:

MAŃCZAK-WOHLFELD Elżbieta (1995): *Tendencje rozwojowe współczesnych założeń angielskich w języku polskim*. – Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

PELC Jerzy, KOJ Leon, red. (1991): *Semiotyka dziś i wczoraj. Wybór tekstów*. – Wrocław: Ossolineum.

ZAREK Józef, red. (2002): *Nadzieje i zagrożenia. Sławistyka i komparatystyka u progu nowego tysiąclecia*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego.

АПРЕСЯН Юрий д., БОГУСЛАВСКАЯ Ольга Ю. [et al.] (1997): *Новый объяснительный словарь синонимов русского языка*. – Школа «Языки русской культуры», Москва.

Tytuł artykułu: Antykwa. Po tytule artykułu: kropka odstęp myślnik odstęp (. –)

Artykuł opublikowany w czasopiśmie: tytuł czasopisma po myślniku.

Artykuł opublikowany w wydawnictwie książkowym: po myślniku, przed tytułem artykułu [w:], a następnie imię i nazwisko redaktora (w tej kolejności)

Po tytule czasopisma / tomu: przecinek stronice (. – *Tytuł*, 72–86.)

Drugi (i następne) wiersz pozycji bibliograficznej: wcięcie 0,75 cm.

Wzór:

ŁUCZYŃSKI Edward (2002): Fleksja języka polskiego z punktu widzenia ontogenezy mowy. – *Biuletyn Polskiego Towarzystwa Językoznawczego LVIII*, 157–165.

MARTIN Richard M. (1991): O metaforze i formie logicznej. – [w:] Jerzy PELC i Leon KOJ (red.): *Semiotyka dziś i wczoraj. Wybór tekstów*, 49–62.

RABIEGA-WIŚNIEWSKA Joanna, RUDOLF Michał (2003a): AMOR – program automatycznej analizy fleksyjnej tekstu polskiego. – *Biuletyn Polskiego Towarzystwa Językoznawczego LXIII*, 175–186.

WYSOCHAŃSKI Włodzimierz (2002): Kategoria symulatywności w językach słowiańskich w ujęciu typologicznym. – [w:] Józef ZAREK (red.): *Nadzieje i zagrożenia. Sławistyka i komparatystyka u progu nowego tysiąclecia*. – Katowice: Wydawnictwo Uniwersytetu Śląskiego, 371–376.

ПАДУЧЕВА Елена В. (1992): О семантическом подходе к синтаксису и генитивном субъекте глагола *быть*. – *Russian Linguistics* 16, 53–63.

Artykuł, streszczenia, wyrazy klucze i bibliografię należy przesłać **w jednym pliku**.

Streszczenia: maksymalnie 900 znaków (pół strony) w języku **angielskim** (z przełożeniem **tytułu**) i **polskim** należy zamieścić **bezpośrednio pod tytułem artykułu** (najpierw angielskie, potem polskie).

Wyrazy klucze: w języku **angielskim i polskim** (nie więcej niż 7 wyrazów) należy zamieścić **pod streszczeniami przed treścią artykułu** (wyrazy klucze w j. angielskim pod streszczeniem angielskim, wyrazy klucze w j. polskim pod streszczeniem polskim).

Wszystkie artykuły są przed złożeniem do druku recenzowane przez dwóch niezależnych recenzentów, którzy pozostają anonimowi.

Artykuł należy wysłać pocztą elektroniczną na adres sekretarza Redakcji **w dwóch formatach – .doc/.docx oraz .pdf:** p.palka@uj.edu.pl

Dodatkowe dokumenty, które należy przesłać na adres sekretarza Redakcji:

- (1) notka o autorze (wzór dostępny na stronie internetowej czasopisma)
- (2) skan wypełnionego oświadczenia o prawach autorskich (formularz dostępny na stronie internetowej czasopisma)

Upraszamy prosimy o załączenie wszystkich dokumentów do jednej wiadomości – w sumie 4 załączniki:

- (1) artykuł, streszczenia, wyrazy klucze, bibliografia w formacie .doc/.docx
- (2) artykuł, streszczenia, wyrazy klucze, bibliografia w formacie .pdf
- (3) notka o autorze
- (4) skan wypełnionego oświadczenia o prawach autorskich

