

Beata Jarosz
Uniwersytet Marii Curie-Skłodowskiej w Lublinie
ORCID: <https://orcid.org/0000-0002-3251-9734>
e-mail: beata.jarosz@mail.umcs.pl

Internetowy słownik języka zawodowego polskich dziennikarzy prasowych. Koncepcja teaurusu dziedzinowego typu wordnet – preliminaria*

**An online dictionary of the professional language
of Polish press journalists.
The concept of domain-specific wordnet-like thesaurus –
preliminary comments**

Abstrakt

W artykule przedstawiono wstępną koncepcję dziedzinowego teaurusu typu wordnet zawierającego nomenklaturę branżową używaną przez polskich dziennikarzy prasowych w wewnątrzgrupowej komunikacji w trakcie wykonywania zadań zawodowych. Spośród ok. 2600 specjalistycznych nazw stosowanych od XVIII w. do współczesności jedynie niewielki wycinek uwzględniono w leksykonach, co uzasadnia potrzebę stworzenia e-słownika. W oparciu o testowe prace wykonane przez autorkę w ośrodku CLARIN-PL oraz w odniesieniu do metodyki tworzenia wordnetów ogólnych, w tym do koncepcji polskiej Słowsieci, wstępnie sprecyzowano zawartość dziennikarskiego teaurusu, zakres informacji uwzględnianych w artykule hasłowym i określono roboczy katalog domen oraz relacji semantycznych, dzięki którym będą porządkowane branżowe jednostki leksykalne. Zaproponowano też sposób łączenia zasobów z polskim wordnetem ogólnym. Artykuł inicjuje dyskusję nad istotą i metodyką tworzenia wordnetów dziedzinowych, których budowanie ma – w związku z brakiem polskich projektów tego typu i jedynie kilkoma projektami na świecie (np. ArchiWordNet, Jur-WordNet) – charakter eksperymentalny.

Słowa kluczowe: wordnet, wordnet dziedzinowy, język zawodowy dziennikarzy, Słowsieć, PlWordNet

* Artykuł powstał w związku z realizacją projektu „Internetowy słownik języka zawodowego polskich dziennikarzy prasowych – możliwości i ograniczenia” finansowanego przez Narodowe Centrum Nauki w ramach programu Miniatura (nr 2022/06/X/HS2/00242).

Abstract

The article presents an initial concept of the domain-specific wordnet-like thesaurus containing special nomenclature used by Polish press journalists in intra-group communication in the course of performing professional tasks. Of approximately 2,600 specialized lexical units used from the 18th century to the present day, only a fraction have so far been included in dictionaries, which justifies the need to build an e-dictionary. Based on test work done by the author at the CLARIN-PL center (Poland), and with regard to the methodology of creating general wordnets, including the concept of the Polish general wordnet (*Slowosieć*), the content of the journalistic thesaurus and the scope of information included in headword entries were preliminarily specified. Additionally, the author defined a preliminary catalogue of semantic domains and semantic relations by which specific lexical units will be ordered. Moreover, the method of connecting resources with the PIWordNet was proposed. The article initiates discussion on the essence and methodology of building domain-specific wordnets, the construction of which is experimental due to the lack of this kind of Polish projects and only a few global ones (e.g. *ArchiWordNet*, *Jur-WordNet*).

Keywords: wordnet, domain-specific wordnet, professional language of journalists, *Slowosieć*, PIWordNet

1. O potrzebie stworzenia słownika

Najnowsze badania nad językiem zawodowym polskich dziennikarzy prasowych rozumianym jako specyficzny subkod używany w wewnątrzgrupowej komunikacji w trakcie wykonywania zadań zawodowych (tj. zbierania informacji, opracowywania materiałów, udostępniania ich odbiorcom) dowodzą, że wiele swoistych jednostek leksykalnych stosowanych w redakcjach na przestrzeni ponad dwóch stuleci (tzn. od XVIII w. do współczesności) jeszcze nigdy nie zostało zarejestrowanych w słownikach i opracowaniach (Jarosz 2020). Spośród ok. 2600 branżowych nazw¹ wyekscerpowanych z rozmaitych źródeł zastanych (materiały pisane i audiowizualne, drukowane i elektroniczne)² i wywołanych (zarejestrowane rozmowy redakcyjne i kwestionariusze ankiet) (Jarosz 2023: 45–56) jedynie niewielki wycinek

¹ Używając sformułowania *nazwa {branżowa}*, mam na myśli terminologiczne i nie-terminologiczne jednostki leksykalne (wyrazy oraz połączenia wyrazowe) stosowane przez dziennikarzy prasowych w kontaktach zawodowych i desygnujące zjawiska swoiste dla tej profesji (np. *art* ‘artykuł’, *depeszowiec* ‘redaktor opracowujący depesze agencyjne’, *główka* ‘zdjęcie portretowe’, *kolumna niekieszowa* ‘strona gazety z materiałami tekstowymi’).

² Tekstami nasyonymi w wysokim lub wystarczającym stopniu leksyką branżową używaną przez dziennikarzy prasowych w różnych okresach (począwszy od XVIII w.) są: wspomnienia (reporterów, fotoreporterów, redaktorów, korektorów, a nawet drukarzy), fachowe periodyki (w tym „Prasa” 1930–1939, „Prasa Polska” 1947–1990, „Zeszyty Prasoznawcze” od 1960 r. [zwłaszcza numery wydane przed 1980, kiedy pismo było specjalistyczne, ale nie *stricte* naukowe], a nawet „Refleks” 1966–1972 [biuletyn Centralnej Agencji Fotograficznej]), poradniki i podręczniki dziennikarstwa, publikacje okolicznościowe (np. księga jubileuszowa „Kuriera Warszawskiego” z 1896 r.) i materiały dostępne w internecie (w tym dokumenty

funkcjonuje w dostępnych leksykonach³. Dość wspomnieć, że w tematycznych tezaurusach opracowanych przez Mieczysława Kafla (1953) oraz zespoły ekspertów kierowane przez Juliana Maślankę (1976) i Walerego Pisarka (2006) odnotowano jedynie – odpowiednio – 180 (7%), 450 (17%) i 150 (6%) nazw należących do wspomnianej nomenklatury⁴. Autorzy uwzględnili jednak głównie jednostki terminologiczne, upowszechnione w wysokim stopniu (np. *advertorial*, *depeszowiec*, *mutacja*, *wywiad*), a przecież w kręgu specjalistów stosowane są też profesjolektyzmy o nieustabilizowanym znaczeniu, wąskim zasięgu użycia i emocjonalnym nacechowaniu (np. Grabias 2019). Te wszystkie argumenty w powiązaniu z częściową nieaktualnością informacji zawartych w wymienionych publikacjach uzasadniają potrzebę skonstruowania nowego słownika, który pozwoli z jednej strony dostarczyć adeptom zawodu, dziennikarzom, ekspertom i naukowcom odpowiedni zasób danych na temat współczesnej leksyki branżowej używanej w redakcjach prasowych, z drugiej zaś – udostępnić rozproszone i w dużej mierze nieznanne nazewnictwo stosowane w takich ośrodkach zwłaszcza w XIX i XX w.

2. Forma i charakter słownika

Zgodnie z założeniami humanistyki cyfrowej i z aktualnymi tendencjami leksykograficznymi nowoczesny słownik powinien być genetycznie elektroniczny (*digitally-born dictionary*), funkcjonować online i spełniać potrzeby zróżnicowanych grup odbiorców (por. Müller-Spitzner, Koplenig 2014). O potencjale nowoczesnej leksykografii i lingwistyki informatycznej świadczy katalog polskich e-leksykonów obejmujący m.in.: *Wielki słownik języka polskiego* (<https://wsjp.pl>), *Słowosieć* (<http://plwordnet.pwr.wroc.pl/wordnet/>), *Słownik pojęciowy języka staropolskiego* (<https://spjs.ijp.pan.pl/ukladSemant/index>), *Elektroniczny słownik języka polskiego XVII i XVIII wieku* (<https://sxvii.pl/>), *Elektroniczny słownik hydronimów Polski* (<https://eshp.ijp.pan.pl>) czy *Verbel. Słownik paradygmatów polskich frazeologizmów czasownikowych* (<http://uwm.edu.pl/verbel/>). Przywołane leksykony są rozmaicie zbudowane, zawierają heterogeniczne typy danych prezentowane

redakcyjne, rozmowy prowadzone na forach dyskusyjnych i w mediach społecznościowych, filmy tematyczne z portalu YouTube).

³ Niektóre nazwy bywają notowane w słownikach ogólnych, a w poradnikach i podręcznikach dziennikarstwa czasem zamieszczane są minileksykony (np. Chyliński, Russ-Mohl [red.] 2008).

⁴ Szczegółowe informacje na temat nazw branżowych zawartych w wymienionych publikacjach znajdują się w monografii poświęconej profesjolektowi polskich dziennikarzy prasowych (Jarosz 2023 [podrozdział 2.2.]).

w różnorodnie projektowanych artykułach hasłowych i mają zróżnicowane funkcjonalności. Metodyka tworzenia takich słowników wciąż ewoluuje, dostępnych jest wiele narzędzi i nieustannie pojawiają się nowe rozwiązania⁵. Zdobyte techniki są jednak wykorzystywane głównie do konstruowania leksykonów ogólnych i normatywnych, rzadziej – onomastycznych, gwarowych i idiolektalnych⁶. Z kolei tezaursy dostarczające informacji o słownictwie branżowym są albo klasyczne (drukowane, bazujące na tradycyjnych modelach leksykograficznych), albo amatorskie (ograniczające się do prostej listy nazw uzupełnionych podstawowymi informacjami o znaczeniu⁷), albo wielojęzyczne (przystosowane do potrzeb tłumaczy, obejmujące zwykle tylko hasła i definicje w różnych wersjach językowych⁸). Z tego względu na etapie przygotowania do stworzenia e-słownika zawierającego swoistą nomenklaturę dziennikarską należy podjąć szereg istotnych decyzji dotyczących przede wszystkim wyboru rodzaju leksykonu, jego struktury, zakresu uwzględnianych informacji, budowy artykułu hasłowego, sposobu prezentacji danych, zestawu funkcjonalności umożliwiających projektowanej grupie odbiorców wykonywanie konkretnych działań⁹ (por. Töpel 2014).

Nowoczesny słownik nazw branżowych powinien moim zdaniem nie tylko zawierać informacje o znaczeniu poszczególnych jednostek leksykalnych, ale też uświadamiać użytkownikom semantyczne powiązania między nimi. Wiedza fachowa jest bowiem uporządkowana i nazwy reprezentujące jej składniki nie funkcjonują w izolacji, a są ze sobą połączone siecią rozmaitych zależności. Z tego względu idealnym rozwiązaniem jest słownik relacyjny, tzn. tezaursu typu wordnet, w którym jednostki leksykalne są grupowane w zbiory semantyczne i opisywane w odniesieniu do innych jednostek leksykalnych (Fjeld, Nygaard 2009). Z punktu widzenia współczesnych przeobrażeń i tendencji ważny jest też fakt, że wordnety są bazami danych, które wykorzystuje się w przetwarzaniu języka naturalnego (np. Smith, Fellbaum 2004), co stwarza dodatkowe – poza tradycyjnymi – możliwości wykorzystania słownika.

⁵ Jednym z nowszych zbiorów leksykograficznych jest e-słownik polskich zapożyczeń w jidysz (<http://polonjid-dictionary.clarin-pl.eu>).

⁶ Pierwszym i chyba jak dotąd jedynym e-leksykonem idiolektalnym jest *Internetowy słownik języka Cypriana Norwida* (<http://www.slownikjezykanorwida.uw.edu.pl>).

⁷ Przykładem jest Wiki-leksykon profesjolektu wojskowego (https://pl.wiktionary.org/wiki/Indeks:Polski_-_Żargon_wojskowy).

⁸ Mowa np. o *Interdyscyplinarnym słowniku wielojęzycznym* (<http://www.imd.pk.edu.pl>).

⁹ Skonstruowanie takiego słownika wymaga ścisłej współpracy z programistami, gdyż dopiero integracja ich kompetencji z wiedzą lingwistyczną umożliwia zbudowanie funkcjonalnej struktury pozwalającej udostępnić informacje językowe w sposób adekwatny i przejrzysty (por. Żmigrodzki 2015).

3. Dziennikarski wordnet dziedzinowy – wstępna koncepcja

Przeszukanie dostępnych zasobów dowodzi, że ani w Polsce, ani na świecie nie powstał dotychczas wordnet zawierający swoistą nomenklaturę używaną przez dziennikarzy prasowych. Najprawdopodobniej nie zostały również podjęte prace nad samą koncepcją takiego tezaursusa, dlatego jej opracowanie wymaga nie tylko namysłu i rozstrzygnięcia najważniejszych kwestii, ale też rozpoznania najnowszych narzędzi oraz możliwości ich implementacji i adaptacji. Wydawałoby się bowiem, że skoro od kilkunastu lat w wielu krajach powstają jedno-, dwu- i multijęzyczne wordnety¹⁰, a największy wśród nich, tj. oparta na zmodyfikowanej teorii pionierskiego Princeton WordNetu polska Słowosieć (PIWordNet), liczy niemal 200 tys. słów, to podjęcie kluczowych decyzji nie stanowi problemu. Okazuje się jednak, że o ile metodyka tworzenia wordnetów ogólnych zawierających współczesne, upowszechnione słownictwo jest przejrzysta i dobrze opracowana (Piasecki [et al.] 2009), o tyle konstruowanie wordnetów dziedzinowych jest zagadnieniem mającym skromną literaturę.

W ostatnich dwóch dekadach podejmowano co prawda próby zbudowania takich leksykonów, ale jest ich niewiele i dotyczą jedynie kilku wybranych dziedzin, w tym medycyny (angielski Medical WordNet; Smith, Fellbaum 2004), prawa (włoski Jur-WordNet; Sagri [et al.] 2004), architektury (angielsko-włoski ArchiWordNet; Bentivogli [et al.] 2004) i turystyki (turecki Tourism WordNet; Arıcan [et al.] 2021). Część tego typu projektów nie została jednak sfinalizowana, co uniemożliwia ocenę koncepcji i zastosowanych rozwiązań¹¹, a tym samym korzystanie z konkretnych, gotowych modeli. Nieliczne doświadczenia i rozstrzygnięcia opisane w publikacjach twórców wymienionych tezaursusów stanowią więc jedynie punkt wyjścia, zwłaszcza że na polskim gruncie nie powstał dotąd żaden wordnet dziedzinowy, który mógłby być wzorem do naśladowania lub przynajmniej prototypem do udoskonalenia.

W zarysowanym kontekście staje się jasne, że wordnet zawierający branżową leksykę stosowaną w środowisku prasowym będzie innowacją nie tylko

¹⁰ Obecnie – jak wynika z wykazu opracowywanego przez Global WordNet Association – funkcjonuje ich blisko 80 (<<http://globalwordnet.org/resources/wordnets-in-the-world/>>, dostęp: 21.05.2023).

¹¹ Przykładem może być wspomniany Medical WordNet, który miał zawierać nomenklaturę medyczną i stanowić dziedzinowe uzupełnienie angielskiego wordnetu ogólnego. Koncepcja została wstępnie opracowana (Smith, Fellbaum 2004) i podjęto kilka prób realizacji projektu, ale konkretne efekty nie zostały dotąd udostępnione (Zotova [et al.] 2023).

ze względu na zgromadzony i udostępniony zbiór danych. Myśląc o stworzeniu takiego słownika, chciałabym w niniejszym tekście sformułować kilka podstawowych założeń. Opisanych poniżej rozstrzygnięć dokonałam w oparciu o testowe prace zrealizowane w marcu 2023 r. w ośrodku CLARIN-PL we Wrocławiu¹², z uwzględnieniem dotychczasowej wiedzy na temat konstruowania leksykonów branżowych i wordnetów ogólnych oraz w nawiązaniu do koncepcji polskiej Słowosieci. Artykuł jest wstępem do dalszych prac i inicjuje równocześnie dyskusję nad istotą wordnetu dziedzinowego, którego budowanie ma charakter eksperymentalny.

3.1. Źródła i zawartość

Zasób danych, który będzie stanowić prymarną bazę wordnetu dziennikarskiego, obejmuje wspomniany zbiór ok. 2600 ciągłych i nieciągłych jednostek leksykalnych używanych w środowisku prasowym w różnych fazach kształtowania się języka zawodowego, tj. od XVIII w. do współczesności (Jarosz 2022; też: Jarosz 2023 [rozdział III]). Nazwy branżowe i ilustracje tekstowe dokumentujące ich funkcjonowanie zostały zgromadzone tradycyjnymi metodami, tzn. w wyniku niezautomatyzowanej kwerendy źródeł, klasycznej ekscerpcji słów, wyimków oraz informacji leksykalno-semantycznych¹³.

Takie założenia odbiegają od koncepcji i metodyki tworzenia wordnetów ogólnych, w których uwzględnia się – jak nadmieniałam – słownictwo współczesne, archaizmy funkcjonują w bardzo ograniczonym zakresie¹⁴, a hasła, przykłady użycia i informacje językowe są pozyskiwane z istniejących korpusów¹⁵ przy użyciu narzędzi do automatycznego wydobywania danych z tekstu (Piasecki [et al.] 2009; Maziarz [et al.] 2014). W przypadku profesjolektu dziennikarzy prasowych nie istnieje jednak adekwatny korpus tematyczny

¹² W ośrodku CLARIN-PL funkcjonującym przy Politechnice Wrocławskiej powstaje zaawansowana infrastruktura do przetwarzania języka naturalnego i prowadzone są prace nad Słowosiecią. Podczas trzytygodniowego pobytu badawczo-konsultacyjnego w tym konsorcjum mogłam więc przedyskutować z ekspertami różne rozwiązania i przetestować najnowsze narzędzia. Dziękuję za tę możliwość i merytoryczną pomoc dr. hab. inż. Maciejowi Piaseckiemu, polskiemu koordynatorowi europejskiej grupy CLARIN, i kierowanemu przez Niego zespołowi.

¹³ Działania te zostały zrealizowane w trakcie prac nad monografią (Jarosz 2023).

¹⁴ Archaiczne jednostki leksykalne uwzględniane są tylko w uzasadnionych przypadkach, np. gdy często występują we współczesnych tekstach jako kontekst czy metafora (por. Dobrowolska-Pigoń [et al.] 2020: 11).

¹⁵ Korpusy internetowe są też wykorzystywane w tworzeniu e-słowników niebędących wordnetami (np. Żmigrodzki [et al.] 2018: 20).

zawierający materiały nasycone w odpowiednim stopniu leksyką branżową¹⁶ (por. Bańko 2001; Gruzca 2008). Wordnet dziedzinowy musi więc – przynajmniej w początkowej fazie – opierać się na danych wyekscerpowanych w sposób tradycyjny, ponieważ jeśli bazowałyby na niesprofilowanych korpusach, znalazłyby się w nim jedynie nazwy funkcjonujące w powszechnym obiegu. W rezultacie zgromadzony zasób leksyki fachowej niewiele różniłby się pod względem jakościowym i ilościowym od zbioru znajdującego się choćby w *Słowniku terminologii medialnej* (Pisarek [red.] 2006) i w takiej perspektywie tworzenie nowego tezaursusa byłoby nieracjonalne. Fakt ten przesądził o zmodyfikowaniu standardowej strategii budowania wordnetów ogólnych.

Z kolei włączenie do katalogu haseł archaizmów branżowych jest uzasadnione z kilku powodów. Przede wszystkim najwcześniejszy odnaleziony przeze mnie leksykon zawierający fachowe słownictwo stosowane w ośrodkach prasowych pochodzi z lat 50. XX w. (Kafel 1953). Biorąc więc pod uwagę, że pojedyncze, swoiste nazwy funkcjonowały już w XVIII w., profesjolekt dziennikarzy ukonstytuował się w XIX stuleciu (Jarosz 2022) i leksyka z tych okresów nie została dotąd zarejestrowana w żadnym słowniku, konieczne jest moim zdaniem uwzględnienie jej w projektowanym wordnetcie. Pozwoli to udostępnić różnym grupom odbiorców nieznane dotąd dane językowe i umożliwi wielokierunkowe badanie branżowej nomenklatury używanej w ośrodkach prasowych.

3.2. Zasób informacji w artykule hasłowym

W funkcji hasła będzie występować zlematyzowana ciągła lub nieciągła jednostka leksykalna, która zostanie odpowiednio scharakteryzowana, a poszczególne rodzaje informacji będą uporządkowane w obrębie artykułu hasłowego, przy czym jego wizualny kształt zostanie zaprojektowany w przyszłości. Formę słownikową uzupełniać będzie definicja opracowana na podstawie zrealizowanej już klasycznej analizy tekstów źródłowych i ujawniająca specjalistyczne znaczenie w sposób zrozumiały dla różnych grup

¹⁶ Od kilku lat powstaje korpus ChronoPress (<https://chronopress.clarin-pl.eu>), który zawiera obecnie [stan na lipiec 2023 r.] ok. 100 tys. tekstów prasowych z lat 1945–1964 (łącznie 30 mln słów), ale finalnie będzie obejmować materiały prasowe z okresu 1918–2018. Użytkownicy mogą m.in. wygenerować (1) zestaw kolokacji i wyimków źródłowych zawierających dany leksem oraz (2) wizualizację chronologiczną prezentującą wskaźniki frekwencji użycia słowa w poszczególnych latach (Pawłowski 2017). Korpus może być wykorzystywany w rozmaitych badaniach, zwłaszcza lingwistycznych i medioznawczych, ale nie jest adekwatnym zasobem w przypadku języka zawodowego dziennikarzy. Profesjolekt stanowi bowiem tworzywo komunikatów formułowanych w kontaktach wewnątrzgrupowych, redakcyjnych, w dużej mierze nieoficjalnych, nie zaś materiałów publikowanych w mediach.

odbiorców. W artykule hasłowym z pewnością znajdą się też podstawowe informacje etymologiczne i gramatyczne, dane o wariantach fonetycznych i pisownianych oraz ilustracje tekstowe wraz z odniesieniem bibliograficznym.

Jednostki leksykalne zostaną dodatkowo opisane za pomocą co najmniej pięciu rodzajów znaczników (tzw. tagów):

- 1) chronologicznych, informujących o poświadczonych źródłowo okresach, w których nazwa była stosowana (tzn. <#XVIII>, <#XIX>, <#XX>, <#XXI>);
- 2) specjalistycznych, uwypuklających funkcjonowanie nazwy w ramach różnych nomenklatur branżowych (np. <#drukarski>, <#fotograficzny>);
- 3) ekspresywnych, dookreślających charakter nacechowania (np. <#zartobliwy>, <#pogardliwy>);
- 4) frekwencyjnych, uściślających zakres użycia (np. <#rzadki>, <#indywidualny>);
- 5) genetycznych, informujących o pochodzeniu nazwy (np. <#angielski>, <#francuski>).

Wymienione kategorie znaczników odpowiadają właściwie tradycyjnym rodzajom kwalifikatorów stosowanych w słownikach, przy czym kwalifikatory chronologiczne zwykle mają inną postać (np. *dawny*, *przestarzały*) (Kurkiewicz 2007: 30–31). Dzięki wprowadzeniu wypisanych powyżej oznaczeń użytkownicy będą mogli wyfiltrować zestaw haseł spełniający określone przez nich kryteria (np. nazwy o niemieckiej proweniencji używane w XX w. w zakładach poligraficznych zostaną wyodrębnione dzięki kombinacji tagów <#drukarski>, <#XX>, <#niemiecki>).

Zaproponowany zestaw danych stanowiących podstawę artykułu hasłowego częściowo odpowiada modelowi opisywania jednostek leksykalnych w Słowosieci. Pewną odmiennością będzie dokładne definiowanie nazw, ponieważ w wordnetach ogólnych objaśnienia mają zwykle postać krótkich glos (Maziarz [et al.] 2014: 80), ale też ich brak nie stanowi problemu, gdyż – zgodnie z koncepcją słownika relacyjnego – znaczenie wynika z powiązań z innymi jednostkami leksykalnymi (Miller [et al.] 1990: 241). Pozostawienie haseł w wordnecie dziedzicznym bez odpowiedniej definicji informującej o branżowym sposobie rozumienia jest jednak w moim przekonaniu ryzykowne. Niuanse semantyczne są bowiem na tyle istotne, że brak objaśnień może mieć poważne konsekwencje, np. prowadzić do porażki komunikacyjnej (por. Sagri [et al.] 2004: 307).

Innowacyjna jest też propozycja scharakteryzowania haseł pod względem temporalnym, ponieważ w wordnetach ogólnych nie wprowadza się informacji diachronicznych. W przypadku wordnetu dziennikarskiego uwzględnienie takich danych jest jednak merytorycznie uzasadnione – w opracowaniach naukowych można bowiem znaleźć nieadekwatne zapisy na temat momentu

pojawienia się konkretnych nazw na polskim gruncie. Przykładowo: rzeczownik *reporter* zdaniem niektórych badaczy pojawił się w polszczyźnie dopiero na przełomie lat 60. i 70. XIX w. (tak np. Żyrek-Horodyska 2019: 33). Klasyczna kwerenda prasy z tego stulecia dowodzi natomiast, że słowem tym operowano już w latach 40. (np. „Rozmaitości” 1944, nr 39). Czasami rozbieżności oscylują wokół półwiecza, co można zilustrować przykładem anglicyzmu *lid*, który występuje w źródłach z lat 50. XX w. (np. „Prasa Polska” 1959, nr 2), ale został niedawno uznany za zapożyczenie z przełomu XX i XXI w. (tak: Olkowska 2011: 241). Podanie w artykułach hasłowych informacji na temat czasu funkcjonowania poszczególnych nazw pozwoli więc skorygować dotychczasowe, błędne lub nieprecyzyjne ustalenia. W wielu przypadkach takie dane zostaną zresztą udostępnione po raz pierwszy, gdyż w odniesieniu do jednostek leksykalnych używanych współcześnie na ogół nie podejmuje się prób określenia ich genezy. Analiza źródeł uświadamia natomiast, że część z nich funkcjonuje na polskim gruncie od ponad stulecia, czego potwierdzeniem może być romanizm *faits divers* stosowany w rodzimej prasie już w latach 80. XIX w. (np. „Ateneum” 1886, nr 2).

3.3. Domeny semantyczne i relacje

Jednostki leksykalne uwzględnione w wordncie dziedzinowym będą – podobnie jak w Słowsieci – uporządkowane pod względem tematycznym i zostaną powiązane gęstą siecią relacji. Analiza katalogu domen semantycznych zdefiniowanych dla polskiego wordnetu ogólnego wykazała jednak, że konieczne są modyfikacje. O ile bowiem branżowe nazwy używane przez dziennikarzy prasowych mogłyby zostać przyporządkowane do domen typu WYTWORY (*fotoesej*), LUDZIE (*depeszowiec*), MIEJSCA (*korektornia*), JEDNOSTKI MIAR (*szpalta*), ZDARZENIA (*operatywka* ‘zebranie redakcyjne’) i CZYNNOSCI (*snapować* ‘wysłać depeszę’), o tyle wyodrębnione w Słowsieci dla rzeczowników ogólnych domeny takie, jak ZWIERZĘTA, JEDZENIE i CZĘŚCI CIAŁA, pozostałyby w przypadku porządkowania nomenklatury dziennikarskiej zbiorami pustymi. W miejsce zbędnych kategorii trzeba będzie natomiast wprowadzić nowe domeny, w tym: NIEPRAWIDŁOWOŚCI (*pustak* ‘niezapełnione miejsce na stronie’) i INSTRUMENTARIUM (*makieta* ‘szablon z układem gazety’). Takie modyfikacje są ściśle związane z faktem, że specjalistyczna, fachowa konceptualizacja zjawisk jest w pewnych zakresach odmienna od potocznego postrzegania rzeczywistości. O ile więc wyraz *rotacja* oznaczający drukarską maszynę rotacyjną zostałby w wordncie ogólnym przypisany do kategorii WYTWORY (jako rezultat pracy konstruktora), o tyle na poziomie

branżowym byłaby to kategoryzacja nieadekwatna. Wytworem pracy dziennikarza prasowego jest bowiem – uogólniając – materiał tekstowy lub graficzny, a nie maszyna – ona jest narzędziem, czyli mieści się w domenie INSTRUMENTARIUM.

Próby uporządkowania niewielkiego zbioru swoistych rzeczowników dowiodły ponadto, że podczas konstruowania wordnetu dziedzinowego niezbędne będzie stworzenie katalogu subdomen. W przypadku wspomnianej kategorii LUDZIE chodziłoby o wprowadzenie podkategorii: DZIENNIKARZE, CZŁONKOWIE REDAKCJI, WSPÓLPRACOWNICY, ODBIORCY, które – co oczywiste – są nieadekwatne do opisu języka ogólnego. Z kolei dla domeny INSTRUMENTARIUM zasadne będzie wydzielenie subdomen AKCESORIA i URZĄDZENIA. Oczywiście nowe domeny i ich nazwy są wstępnymi propozycjami. Dopiero podczas tworzenia słownika okaże się, czy są one wystarczające, a przypisane im etykiety – w pełni informacyjne.

W porządkowaniu dziennikarskiej nomenklatury przydatne będą typowe dla wordnetów ogólnych, w tym dla Słowsieci, podstawowe relacje takie, jak hiperonimia/hiponimia, meronimia/holonimia i antonimia (Dobrowolska-Pigoń [et al.] 2016). Wśród istotnych zmian, jakie trzeba wdrożyć, jest ustanowienie relacji, którą określiłam wstępnie jako SYNONIMIA TEMPORALNA. Pozwala ona łączyć współczesne swoiste jednostki leksykalne z archaicznymi odpowiednikami, które – jak już zaznaczyłam – zostaną uwzględnione w wordnecie w pełnym wymiarze. Specyfika leksyki branżowej przemawia ponadto za wprowadzeniem relacji KOMPRESJA, dla której należałoby zdefiniować kilka podtypów, w tym: KOMPRESJA FONETYCZNA (*feuilleton* → *felieton*), KOMPRESJA DERYWACYJNA (*konferencja prasowa* → *konfpres*), KOMPRESJA SKŁADNIOWA (*people magazine* → *people*) (Jarosz 2023: 797–798; por. Lejčik 1981). Co ważne, rozstrzygnięcia te zostały poczynione podczas eksperymentalnego łączenia kilkunastu nazw branżowych z domeny LUDZIE wykonanego w aplikacji Wordnetloom, dzięki której tworzona jest polska Słowsieć (Piasecki [et al.] 2013). Weryfikacja pełnego katalogu relacji używanych w wordnecie ogólnym będzie jednak możliwa dopiero w trakcie prac nad dziedzinowym tezauresem.

3.4. Przykład artykułu hasłowego

Na podstawie powyższych ustaleń można skonstruować przykładowy artykuł hasłowy, który należy traktować jako mikrozaczątek wordnetu dziennikarskiego¹⁷.

Lemat: redaktor naczelny
Część mowy: rzeczownik
Definicja: osoba zarządzająca redakcją, podejmująca kluczowe decyzje wydawnicze, kadrowe, finansowe itd.
Domena: LUDZIE
Subdomena: CZŁONKOWIE REDAKCJI
Tagi: <#XIX>, <#XX>, <#XXI>, <#powszechny>

Ilustracje źródłowe¹⁸:

<#XIX> „Pod ten czas kilku z dawniejszych współpracowników ustąpiło z »Kurjera«, zachowując jednak z **redaktorem naczelnym** zawsze najlepsze stosunki” [KKW/149].

<#XX> „Omówienie programowej wypowiedzi **naczelnego redaktora** »Czasu« w pierwszym, tak dla nas cennym numerze znajdzie czytelnik w innym miejscu” [ST/12].

<#XXI> „Ja na miejscu **redaktorów naczelnych** bardziej bym inwestował w dziennikarzy czy nawet redaktorów wizualnych, którzy mają się zajmować infografikami, niż grafików” [Pr/2006/4].

Relacje¹⁹:

[MERONIMIA] **redakcja**

[SYNONIMIA TEMPORALNA] **redaktor en chef** <#XIX>, <#XX>, <#francuski>

[SYNONIMIA MIĘDZYREJESTROWA] **nadnaczelnny** <#XXI> <#pogardliwy>

[SYNONIMIA MIĘDZYREJESTROWA] **naczelnik** <#XXI> <#zartobliwy>

[KOMPRESJA]

> [KOMPRESJA DERYWACYJNA] **rednacz** <#XXI>

> [KOMPRESJA SKŁADNIOWA] **naczelnny** <#XX>, <#XXI>

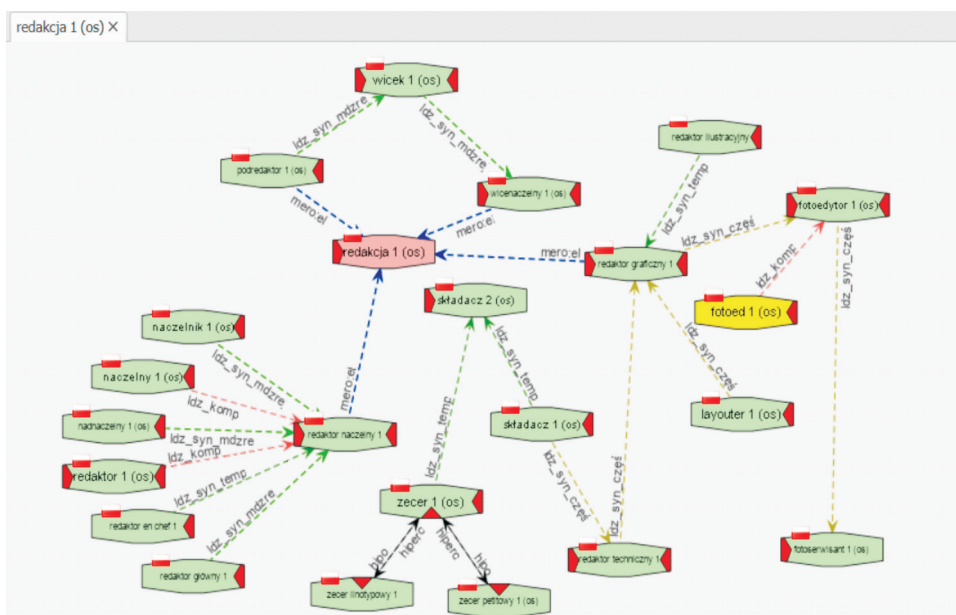
Rozwiązaniem przydatnym w prezentacji uwikłań semantycznych jest wizualizacja prezentująca swoistą strukturę złożoną z kafelków z hasłami oraz łączących je linii wskazujących rodzaj relacji zachodzącej pomiędzy poszczególnymi jednostkami leksykalnymi. W przypadku wyrażenia *redaktor*

¹⁷ Testowe prace wykonane w aplikacji Wordnetloom dowodzą, że opisywanie jednostek leksykalnych w zaprezentowany sposób jest wykonalne.

¹⁸ Oznaczenia w nawiasach kwadratowych odnoszą się do następujących źródeł: KKW/149 – [„Kurier Warszawski”] 1896: 149; ST/12 – Tarnowski [et al.] 1938: 12; Pr/2006/4 – „Press” 2006, nr 4. Podczas prac nad wordnetem konieczne będzie rozstrzygnięcie kwestii ewentualnej modernizacji pisowni w przywoływanych ilustracjach źródłowych.

¹⁹ Wynotowane relacje nie wyczerpują oczywiście wszystkich możliwości (por. przypis kolejny).

naczelny sieć powiązań semantycznych jest oczywiście znacznie bardziej złożona, niż wynikałoby to z zapisów zawartych w powyższym, wstępnie skonstruowanym artykule hasłowym. Do pewnego stopnia uświadamia to grafika przedstawiająca rezultaty testowych prac wykonanych w aplikacji Wordnetloom²⁰ (Ryc. 1). Warto przy tym dodać, że analiza widocznego na wizualizacji gąszczu linii wiążących segmenty powiązane z *nomen collectivum redakcja* pozwala poznać sposób postrzegania redakcyjnej hierarchii przez dziennikarzy prasowych. Oczywiście docelowo zbiór nazw zaprezentowanych na grafie nie funkcjonowałby w izolacji, a stanowiłby część siatki haseł całego leksykonu branżowego, którą użytkownik mógłby sukcesywnie



Legenda: *mero.el* – meronimia, *hipo* – hiponimia, *hipero* – hiperonimia, *ldz_komp* – kompresja, *ldz_syn_mdzre* – synonimia międzyrejestrowa, *ldz_syn_temp* – synonimia temporalna, *ldz_syn_czesc* – synonimia częściowa

Ryc. 1. Robocza wizualizacja uwikłań leksykalno-semantycznych w dziennikarskim wordniecie dziedzinowym

Źródło: opracowanie własne wykonane w aplikacji Wordnetloom.

²⁰ Wizualizacja nie prezentuje wszystkich zależności łączących poszczególne jednostki leksykalne. Prace w aplikacji zostały zrealizowane w celu przetestowania różnych opcji, dlatego nie określałam kompleksowo związków semantycznych. W niektórych przypadkach między dwiema nazwami można zresztą ustanowić więcej niż jedno połączenie. Ze względu na rozpoznawczy charakter działań nie rozpisalam też relacji KOMPRESJA na planowane podtypy. Ich wprowadzenie nie stanowi problemu, ale na tym etapie pracy nie było to konieczne.

rozwiąć, klikając czerwone trójkąty w rogach kafelków²¹. Sygnalizują one bowiem, że dana nazwa jest powiązana jakimiś relacjami z innymi składnikami fachowej nomenklatury.

4. Problemy – wyzwania – perspektywy

Testowe prace nad dziennikarskim wordnetem dziedzinowym ujawniły konieczność sformułowania koncepcji wiązania tego tezaursusa z polską Słowosiecią. Część nazw mających branżową genezę lub specjalistyczny charakter funkcjonuje wszak w powszechnym obiegu i zostało uwzględnionych w wordnecie ogólnym, czego przykładem wymienione powyżej wyrażenie *redaktor naczelny*. Automatyczne porównanie katalogu swoistych określeń stosowanych w środowisku prasowym z siatką haseł uwzględnionych w Słowosieci dowiodło, że 778 jednostek leksykalnych występuje w obu zbiorach (ok. 30% ogółu branżowej nomenklatury), przy czym spora ich część posiada w wordnecie ogólnym definicje nieodpowiadające fachowemu sposobowi rozumienia²². Jest to sytuacja typowa (Fjeld, Nygaard 2009), zwłaszcza jeśli weźmiemy pod uwagę funkcjonowanie w leksykonie branżowym neosemantyzmów, którym przypisuje się sens niemający żadnych punktów wspólnych z powszechnym znaczeniem (np. *balkonik* ‘urządzenie do poruszania się dla osób starszych i z dysfunkcjami ruchu’; *dz. pras.* ‘nadytuł’).

Wstępnie rozważono więc kilka rozwiązań, wykorzystując wiedzę na temat metod mapowania wordnetów ogólnych (tzn. polskiego i angielskiego; Rudnicka [et al.] 2012) i mając na uwadze opisane już propozycje łączenia wordnetu ogólnego z dziedzinowym (np. Sagri [et al.] 2004). Podstawowym warunkiem, jaki będzie musiał zostać spełniony, jest zachowanie swoistości obu tezaursusów ze względu na scharakteryzowane już odrębne katalogi domen i relacji semantycznych. Model wiązania zasobów musi być więc nieinwazyjny, nienaruszający integralności obu wordnetów, ale dostatecznie informacyjny. Jedną z możliwych strategii jest połączenie haseł funkcjonujących w poszczególnych tezaursusach nowymi, odpowiednio zdefiniowanymi relacjami, np. SYNONIMIA MIĘDZYODMIANOWA (gdy x ma podobne znaczenie, ale w wordnecie dziedzinowym jest ono wyspecjalizowane) czy SYNONIMIA MIĘDZYODMIANOWA CZĘŚCIOWA (gdy potoczne znaczenie x pokrywa się

²¹ Naturalnie, podczas prac nad wordnetem mogą zostać opracowane nowe formy prezentacji i ujawniania/ukrywania danych.

²² Jeśli 1791 swoistych nazw (ok. 70% ogólnego zasobu) nie funkcjonuje w Słowosieci, to jest to dodatkowy argument przemawiający za zasadnością stworzenia wordnetu dziedzinowego.

częściowo z fachowym rozumieniem) (por. Magnini, Speranza 2001). Inną opcją jest wprowadzenie w wordnecie ogólnym odpowiedniego oznaczenia lub etykiety informującej, że hasło x jest składnikiem branżowej nomenklatury. Takie rozwiązanie pozwoliłoby połączyć hasła, które nie mają semantycznych punktów wspólnych (np. przywołane już słowo *balkonik*). Użytkownik miałby jednak możliwość wyświetlenia odnośnych informacji z dziennikarskiego wordnetu albo po kliknięciu odpowiedniego elementu byłby przekierowywany do hasła w dziedzinowym tezaurusie.

Biorąc natomiast pod uwagę szybki rozwój technologiczny i łatwość tworzenia innowacji, ostateczny sposób zorganizowania danych leksykalnych wraz z pozostałymi zasobami zostanie sprecyzowany w trakcie prac nad e-słownikiem. Dokonane w tym momencie wybory już za kilka miesięcy mogą się bowiem okazać nieracjonalne lub niekorzystne z punktu widzenia wdrożonych bądź możliwych do opracowania rozwiązań technicznych. W projektowaniu makrostruktury, mikrostruktury i mediostruktury wordnetu dziedzinowego pożądane jest maksymalne wyzyskanie digitalnego potencjału, dlatego pośpiech decyzyjny jest niewskazany. Bez wątplenia jednak porządkowanie zawartości musi odpowiadać na heterogeniczne potrzeby użytkowników. Kluczowe jest umożliwienie im wszechstronnego wykorzystywania danych i wykonywania różnorodnych operacji, w tym filtrowania hasel pod kątem różnych identyfikatorów, wyodrębniania pożądanych informacji leksykograficznych i samodzielnego wydobywania konkretnych danych (np. generowania zestawu hasel należących do jednej domeny semantycznej albo schematu prezentującego na osi czasu rozwój branżowej leksyki w ramach określonej subdomeny). W celu zapewnienia komfortu korzystania z tezausa warto też umożliwić odbiorcom ukrywanie zasobów, które mogą być dla nich zbędne (np. przynależność nazwy do konkretnej części mowy).

Słownik języka zawodowego polskich dziennikarzy prasowych powinien wykorzystywać najnowsze zdobycze techniki i zawierać komplet informacji leksykograficznych standardowo umieszczanych w artykułach hasłowych. Z uwagi na heterogeniczne potrzeby zróżnicowanych grup odbiorców (dziennikarzy, studentów, badaczy itd.) nazwy reprezentujące określone składniki wiedzy fachowej powinny zostać uporządkowane w sposób ujawniający konceptualną siatkę organizującą doświadczenie zawodowe kumulowane w procesie rozwoju profesji. Niezbędne jest więc ustanowienie powiązań znaczeniowych między swoistymi jednostkami leksykalnymi oraz ich zgrupowanie w obrębie domen i subdomen semantycznych. Wybór wordnetu jest zatem uzasadniony i – co ważne – korzystny z wielu powodów. Przede wszystkim po stworzeniu

pierwszej jego wersji będzie można go aktualizować, weryfikować adekwatność relacji łączących poszczególne nazwy, uzupełniać nowe hasła i dodawać nowe zbiory leksyki (np. swoistą nomenklaturę używaną w środowisku radiowym i telewizyjnym), wprowadzać kolejne znaczniki umożliwiające filtrowanie zasobów według nowych kryteriów, umieszczać kolejne ilustracje źródłowe wyekscerpowane z istniejących korpusów, a nawet powiązać wordnet z korpusem tematycznym, który – w związku z obfitością tekstów nasyconych branżowym słownictwem (zob. przypis 3) – można stworzyć w każdej chwili. Opracowanie wordnetu dziedzinowego daje nieograniczone możliwości również dlatego, że będzie można go wiązać i mapować z właściwie każdym wordnetem, w tym: (1) ogólnym, stworzonym np. na gruncie angielskim, (2) dziedzinowym stycznym, np. dotyczącym poligrafii lub edytorstwa, (3) dziedzinowym symetrycznym, np. dziennikarskim wordnetem angielskim.

Określone w tym tekście założenia stanowią wstęp do opracowania ostatecznej koncepcji dziennikarskiego tezaury i do konkretnych prac, które będą realizowane w najbliższej przyszłości. Zasadność ich podjęcia jest moim zdaniem niepodważalna i może zachęcić innych badaczy do opracowywania kolejnych wordnetów dziedzinowych. Wstępne rezultaty uzyskane podczas testowych działań są bowiem satysfakcjonujące i przekonują, że tworzenie specjalistycznych e-słowników relacyjnych jest wartościową i obiecującą metodą rejestrowania fachowego słownictwa. W dobie rozwoju sztucznej inteligencji tworzenie wordnetów dziedzinowych stanowi pożądaną kierunek gromadzenia informacji o języku, w tym o nomenklaturach branżowych funkcjonujących w różnych środowiskach i obszarach wiedzy oraz techniki.

Literatura

- Arıcan B.N., Özçelik M., Aslan D.B., Sarmış E., Parlar S., Yıldız O.T. (2021): *Creating domain dependent Turkish WordNet and SentiNet*. [W:] *Proceedings of the 11th Global Wordnet Conference 2021, GWC 2021, 18–21 January, 2021*. Red. S. Bosch, C. Fellbaum, M. Griesel, A. Rademaker, P. Vossen. Potchefstroom, s. 243–251.
- Bańko M. (2001): *Z pogranicza leksykografii i językoznawstwa*. *Studia o słowniku jednojęzycznym*. Warszawa.
- Bontivogli L., Bocco A., Pianta E. (2004): *ArchiWordNet: Integrating WordNet with Domain-Specific Knowledge*. [W:] *Proceedings of the 2nd International Global Wordnet Conference*. Red. P. Sojka, K. Pala, P. Smrž, Ch. Fellbaum, P. Vossen. Brno, s. 39–46.
- Chyliński M., Russ-Mohl S. (2008): *Dziennikarstwo*. Warszawa.
- Dobrowolska-Pigoń M., Dziob A., Maziarz M., Piasecki M., Wieczorek J. (2016): *Opis rzeczownikowych relacji leksykalno-semantycznych w Słownosieci 4.0*. Wrocław.
- Dobrowolska-Pigoń M., Dziob A., Nowakowska B., Wieczorek J. (2020): *Procedura korekty Słownosieci*. Wrocław.

- Fjeld R.V., Nygaard L. (2009): *NorNet – a monolingual wordnet of modern Norwegian*. [W:] *Proceedings of the NODALIDA 2009 workshop WordNets and other Lexical Semantic Resources – between Lexical Semantics, Lexicography, Terminology and Formal Ontologies*. Red. B. Sandford Pedersen, A. Braasch, S. Nimb, R. Vatvedt Fjeld. Tartu, s. 13–16.
- Grabias S. (2019): *Język w zachowaniach społecznych*. Lublin.
- Grucza S. (2008): *Teksty specjalistyczne. Językowe eksponenty wiedzy specjalistycznej*. [W:] *Kulturowy i leksykograficzny obraz języków specjalistycznych*. Red. Ł. Karpiński. Warszawa, s. 181–193.
- Jarosz B. (2020): *Język zawodowy polskich dziennikarzy jako problem badawczy. Uwagi wstępne*. „Język Polski” 1, s. 122–133.
- Jarosz B. (2022): *The professional language of Polish journalists in a diachronic perspective*. „Slovo a Slovesnost” 2, s. 103–121.
- Jarosz B. (2023): *Język zawodowy polskich dziennikarzy prasowych (XIX–XXI)*. Lublin. Online: <https://polona.pl/preview/c4bc10d8-e6ae-465e-879f-efe0bdece667>
- Kafel M. (1953): *Maty ilustrowany słownik techniki wydawniczej*. Warszawa.
- [„Kurier Warszawski”]. (1896): „Kurier Warszawski”. *Książka jubileuszowa ozdobiona 247 rysunkami w tekście 1821–1896*. Warszawa.
- Kurkiewicz J. (2007): *Kwalifikatory w Wielkim słowniku języka polskiego*. [W:] *Nowe studia leksykograficzne*. Red. P. Żmigrodzki, R. Przybylska. Kraków, s. 29–45.
- Lejčik V.M. (1981): *Optimalnaja dlina i optimalnaja struktura termina*. „Voprosy Āzykoznanija” 2, s. 63–73.
- Magnini B., Speranza M. (2001): *Integrating Generic and Specialized Wordnets*. [W:] *Proceedings of the 2nd Conference on Recent Advances in Natural Language Processing*. Tzgov Chark, s. 149–153.
- Maślanka J. (red.) (1976): *Encyklopedia wiedzy o prasie*. Wrocław.
- Maziarz M., Piasecki M., Rudnicka E. (2014): *Słowosieć – polski wordnet. Proces tworzenia tezaurusa*. „Polonica” 34, s. 79–98.
- Miller G.A., Beckwith R., Fellbaum Ch., Gross D., Miller K.J. (1990): *Introduction to WordNet: An On-line Lexical*. „International Journal of Lexicography” 4, s. 235–244.
- Müller-Spitzer C., Kopenlig A. (2014): *Online dictionaries: Expectations and Demands*. [W:] *Using online dictionaries*, Red. C. Müller-Spitzer. Berlin, s. 143–188.
- Olkowska K.M. (2011): *Najnowsze angielskie zapożyczenia cytatowe w polszczyźnie ogólnej początku XXI wieku*. „Studia Językoznawcze” 10, s. 211–256.
- Pawłowski A. (2017): *ChronoPress – Chronological Corpus of Polish Press Texts (1945–1962)*, <https://www.clarin.eu/sites/default/files/Pawlowski-CLARIN2017_paper_35.pdf>, dostęp: 22.07.2023.
- Piasecki M., Marcińczuk M., Ramocki R. (2013). *WordNetLoom: A WordNet development system integrating form-based and graph-based perspectives*. „International Journal of Data Mining Modelling and Management” 5, s. 210–232.
- Piasecki M., Szpakowicz S., Broda B. (2009): *Wordnet from the ground up*. Wrocław.
- Pisarek W. (red.) (2006): *Słownik terminologii medialnej*. Kraków.
- Rudnicka E., Maziarz M., Piasecki M., Szpakowicz S. (2012): *A strategy of Mapping Polish WordNet onto Princeton WordNet*. [W:] *Proceedings of COLING 2012: Posters*. Red. M. Kay, Ch. Boitet. Bombaj, s. 1039–1048.
- Sagri M.T., Tiscornia D., Bertagna F. (2004): *Jur-WordNet*. [W:] *Proceedings of the 2nd International Global Wordnet Conference*. Red. P. Sojka, K. Pala, P. Smrz, Ch. Fellbaum, P. Vossen. Brno, s. 305–310.
- Smith B., Fellbaum Ch. (2004): *Medical WordNet: A New Methodology for the Construction and Validation of Information Resources for Consumer Health*. [W:] *Proceedings of the 20th International Conference on Computational Linguistics, Geneva, 23–27 August 2004*. Red. E. Yuste, S.J. Jekat, A.K. Pantli, G. Massey. Genewa, s. 31–38.

- Tarnowski S. [et al.]. (1938): *Księga pamiątkowa na dziewięćdziesięciolecie dziennika „Czas”*. Kraków 1848 – Warszawa 1938. Warszawa.
- Töpel A. (2014): *Review of research into the use of electronic dictionaries*. [W:] *Using Online Dictionaries*. Red. C. Müller-Spitzer. Boston, s. 13–54.
- Zotova E., Cuadros M., Rigau G. (2023): *Towards the integration of WordNet into ClinIDMap*. [W:] *Proceedings of the 12th Global Wordnet Conference (GWC2023), San Sebastian, Spain, January 23–27, 2023*, <<https://dl.acm.org/doi/pdf/10.3115/1220355.1220409>>, dostęp: 19.04.2023.
- Żmigrodzki P. (2015): *Wielki słownik języka polskiego PAN a obecna sytuacja językoznawstwa polskiego*. „LingVaria” 1, s. 109–119.
- Żmigrodzki P., Bańko M., Batko-Tokarz B., Bobrowski J., Czelakowska A., Grochowski M., Przybylska R., Waniakowa J., Węgrzynek K. (2018): *Wielki słownik języka polskiego PAN. Geneza, koncepcja, zasady opracowania*. Kraków.
- Żyrek-Horodyska E. (2019): *Kartografowie codzienności. O przestrzeni (w) reportażu*. Kraków.

