Przegląd Sejmowy
nr 2(175)/2023, s. 97–118; https://doi.org/10.31268/PS.2023.178

Zygmunt Tobor*
Mateusz Zeifert**

Korpusy językowe jako narzędzie w procesie tworzenia prawa

Language Corpora as a Tool in the Law-Making Process

The article discussed linguistic corpora as a tool in the law-making process. First, the authors explained the concept of a linguistic corpus, i.e. an electronic set of texts used for linguistic research. Then they discussed the previous applications of linguistic corpora in legal contexts. The authors then presented possible applications of linguistic corpora in the law-making process. They indicated that not only practical but also normative reasons speak for their use. These tools may contribute to a more effective implementation of the principles of legislative technique, regulated in the binding *Zasady techniki prawodawczej* [*Principles of legislative technique*], that is most of all the principle of communicativeness (in this context the basic role is played by general corpora) and the principle of terminological consistency (in this context, it is necessary to build a specialised corpus of texts of binding legal acts).

Keywords: legislation, legislative technique, linguistic corpora, statutory interpretation, legal language

---

W artykule omówiono korpusy językowe jako narzędzia w procesie tworzenia prawa. W pierwszej kolejności autorzy wyjaśnili pojęcie korpusu językowego, tj. elektronicznego zbioru tekstów wykorzystywanego do badań językoznawczych. W drugiej kolejności omówili dotychczasowe użycia korpusów językowych w kontekstach prawnych i kolejno zaprezentowali ich możliwe zastosowania w procesie tworzenia prawa. Wskazali przy tym, że za stosowaniem korpusów językowych przemawiają względy nie tylko praktyczne, lecz także normatywne. Narzędzia te mogą się przyczynić do skuteczniejszej realizacji zasad techniki legislacyjnej, uregulowanych

* **Prof. dr hab. Zygmunt Tobor**
Uniwersytet Śląski w Katowicach, Wydział Prawa i Administracji, Polska
University of Silesia in Katowice, Faculty of Law and Administration, Poland
zygmunt.tobor@us.edu.pl, https://orcid.org/0000-0001-6541-8502

** **Dr Mateusz Zeifert**
Uniwersytet Śląski w Katowicach, Wydział Prawa i Administracji, Polska
University of Silesia in Katowice, Faculty of Law and Administration, Poland
mateusz.zeifert@us.edu.pl, https://orcid.org/0000-0003-3266-3397

Artykuł został udostępniony na licencji Creative Commons – Uznanie Autorstwa 3.0 Polska (CC BY 3.0 PL)

w obowiązujących *Zasadach techniki prawodawczej*, czyli przede wszystkim zasady komunikatywności (podstawową rolę odgrywają korpusy ogólne) i zasady konsekwencji terminologicznej (nieodzowne jest zbudowanie specjalistycznego korpusu składającego się z tekstów obowiązujących aktów prawnych).

Słowa kluczowe: legislacja, technika prawodawcza, korpusy językowe, interpretacja prawa, język prawny

I. Wprowadzenie

Korpusy językowe są narzędziem metodologicznym, które w ostatnich 20 latach zrewolucjonizowało niemal wszystkie obszary badań językoznawczych[1]. Metodologia korpusowa przyczyniła się do gwałtownego rozwoju przetwarzania języka naturalnego (ang. *Natural Language Processing*, NLP) skutkującego m.in. powstaniem automatycznych translatorów, urządzeń rozpoznających mowę czy programów *chatbot*. Nic dziwnego, że wykorzystanie korpusów jest obecnie żywo dyskutowane również w kontekście stosowania prawa. Przykładowo, od 2016 r. stanowi ono jeden z najbardziej aktualnych tematów w amerykańskiej teorii prawa.

W niniejszym artykule poruszono zagadnienie użycia korpusów językowych w procesie tworzenia prawa. Jak dotąd pozostawało ono raczej w cieniu, a poruszane było wyłącznie przez językoznawców (w szczególności tłumaczy) w kontekście tworzenia wielojęzycznego prawa Unii Europejskiej [dalej: UE]. Autorzy żywią przeświadczenie, że temat ten zasługuje na znacznie większe zainteresowanie. Korpusy stanowią nieocenione narzędzie dla wszystkich uczestników procesu tworzenia tekstów prawnych, a zwłaszcza dla osób bezpośrednio zajmujących się ich redagowaniem. Są w stanie istotnie się przyczynić do realizacji kluczowych zasad techniki prawodawczej, takich jak zasada komunikatywności czy zasada konsekwencji terminologicznej. W przyszłości zaś mogą posłużyć do stworzenia wysoce wyspecjalizowanych narzędzi informatycznych, które zrewolucjonizują pracę legislatorów.

Ponieważ omawiane zagadnienie jest stosunkowo nowe, a w Polsce do tej pory w ogóle nie było dyskutowane, niniejszy tekst ma w dużej mierze charakter popularyzatorski. W trakcie jego tworzenia czerpano z dorobku językoznawstwa, a także zagranicznej teorii prawa (głównie amerykańskiej i niemieckiej). W pierwszej kolejności wyjaśniono pojęcie korpusu językowego. W drugiej kolejności przedstawiono zwięzłe podsumowanie dotychczasowych zastosowań korpusów w nauce prawa i praktyce prawniczej. Wreszcie uzasadniono użycie korpusów w procesie redagowania tekstów prawnych, a także dokonano przeglądu potencjalnych sposobów ich wykorzystania.

[1] I. Kida, *Introduction to Corpus Linguistics*, „Linguistica Silesiana" 2013, nr 34, s. 140.

II. Pojęcie korpusu językowego

Korpus językowy to elektroniczny zbiór autentycznych tekstów danego języka (lub języków) o znacznej wielkości, zazwyczaj liczonej w milionach czy nawet miliardach słów[2]. Zbiór taki może mieć postać „surową" (ang. *raw*) bądź przetworzoną, zwaną „znakowaną", „indeksowaną" lub „anotowaną" (ang. *tagged*). W tym drugim przypadku korpus jest opracowany pod względem lingwistycznym, np. każde słowo ma przypisane oznaczenie części mowy i części zdania (tzw. indeksowanie morfologiczno-syntaktyczne), różne formy słowa zostają zredukowane do formy podstawowej (tzw. lematyzacja), poszczególne fragmenty tekstów zawierają metrykę metadanych itd.[3] Korpusy znakowane oferują bez porównania więcej możliwości badawczych, ale ich tworzenie jest dużo bardziej pracochłonne (mimo częściowej automatyzacji wynikającej z wykorzystywania w tym celu specjalistycznego oprogramowania)[4]. Najlepsze rezultaty osiąga się za pomocą ręcznego znakowania, co jednak jest możliwe jedynie w przypadku stosunkowo niewielkich korpusów[5]. Istotną postulowaną cechą korpusu jest reprezentatywność[6]. Powinien on bowiem jak najlepiej odzwierciedlać dany język (języki). Cel ten osiąga się poprzez zachowanie odpowiednich proporcji tekstów pochodzących z różnych rejestrów. Nie wyklucza to możliwości tworzenia korpusów specjalistycznych, zawierających jedynie teksty należące do określonego rejestru, dialektu, specjalistycznej odmiany języka czy pochodzące z określonego okresu itd.[7]

Pierwszy komputerowy korpus powstał w latach sześćdziesiątych XX w., jednak popularność temu narzędziu przyniósł dopiero gwałtowny rozwój technologii komputerowej w latach dziewięćdziesiątych XX w.[8] Dzisiaj językoznawstwo korpusowe stanowi istotną gałąź lingwistyki, a korpusy stały się nieodłączną częścią pracy badawczej m.in. przy tworzeniu słowników, tłumaczeniach, nauce języków obcych, maszynowym

---

[2] Przykładowo popularny anglojęzyczny Corpus of Contemporary American English (<https://www.english-corpora.org/>, dostęp 17 VIII 2022) zawiera 560 mln słów, a Narodowy Korpus Języka Polskiego (<http://www.nkjp.pl/>, dostęp 17 VIII 2022) ponad 1,5 mld słów. Największym istniejącym korpusem jest Google Books Ngram (<http://books.google.com/ngrams/>, dostęp 17 VIII 2022), który zawiera ok. 155 mld słów.

[3] B. Lewandowska-Tomaszczyk, *Podstawy językoznawstwa korpusowego*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2005, s. 61–94; I. Kida, *op. cit.*, s. 135.

[4] I. Kida, *op. cit.*, s. 138–139.

[5] Przykładowo, Narodowy Korpus Języka Polskiego jest znakowany automatycznie, jednak w jego ramach wydzielono niewielki podzbiór korpusu (ok. 1 mln słów) znakowany ręcznie, służący m.in. do „wytrenowania" oprogramowania dokonującego znakowania automatycznego. *Confer* Ł. Degórski, A. Przepiórkowski, *Ręcznie znakowany milionowy podkorpus NKJP*, [w:] *Narodowy Korpus Języka Polskiego*, red. A. Przepiórkowski, M. Bańko, R. Górski, B. Lewandowska-Tomaszczyk, Wydawnictwo Naukowe PWN, Warszawa 2012, s. 51 i n.

[6] I. Kida, *op. cit.*, s. 139–140.

[7] B. Lewandowska-Tomaszczyk, *op. cit.*, s. 32–33, 61–94; I. Kida, *op. cit.*, s. 135–136.

[8] Ł. Biel, *Corpus-Based Studies of Legal Language for Translation Purposes: Methodological and Practical Potential*, [w:] *Reconceptualizing LSP. Online Proceedings of the XVII European LSP Symposium 2009*, red. C. Heine, J. Engberg, Aarhus School of Business, Aarhus University, Aarhus 2010, s. 1; I. Kida, *op. cit.*, s. 133–134.

przetwarzaniu języka naturalnego, a także w trakcie studiów historyczno-językowych, literackich czy socjolingwistycznych[9]. Wykorzystanie korpusów językowych do celów badawczych jest możliwe dzięki rozmaitym narzędziom, procedurom i algorytmom informatycznym, takim jak: wyszukiwanie interaktywne, listy frekwencyjne, konkordancje i kolokacje[10]. Wyszukiwanie stanowi podstawowe narzędzie każdego użytkownika współczesnych technologii informatycznych, przy czym w przypadku korpusu może ono przyjmować bardzo zaawansowane postacie. Listy frekwencyjne określają częstotliwość występowania wyrazów w korpusie, a zatem – przy założeniu reprezentatywności korpusu – w danym języku. Konkordancje to zbiory przykładów użycia wybranego wyrazu w kontekście, zgodnie z kryteriami wprowadzonymi przez badacza. Zazwyczaj prezentuje się je w formacie KWIC (ang. *key words in context*), tzn. poszukiwany wyraz jest wyświetlany na środku generowanej listy, otoczony z obydwu stron przez najbliższy kontekst zdaniowy. Umożliwia to wygodne porównanie wielu przykładów użycia i pozwala wnioskować na temat jego znaczenia. Kolokacje to stale powtarzające się kombinacje słów występujące w swoim sąsiedztwie. Ich uwzględnienie pozwala m.in. określić, w jakim otoczeniu dany wyraz występuje najczęściej, lub wyszukać przykłady użycia wyrazu w określonym kontekście. Wyszukiwanie, listy frekwencyjne, konkordancje i kolokacje to standardowe narzędzia uwzględniane w interfejsach większości powszechnie dostępnych korpusów[11]. Lingwiści wykorzystują do swoich analiz również bardziej zaawansowane środki, w tym specjalistyczne aplikacje, języki programowania, testy statystyczne, bazy danych itd.[12]

III. Dotychczasowe zastosowania korpusów językowych w kontekstach prawnych

Zastosowania korpusów językowych w szeroko rozumianym kontekście prawnym można podzielić na dwie kategorie.

Po pierwsze, teksty prawne i prawnicze służą jako przedmiot badań językoznawczych. W tym celu są tworzone korpusy specjalistyczne składające się – w zależności od celu – z tekstów aktów prawnych, orzeczeń, czasopism prawniczych, umów itp. Wykorzystuje się je m.in. w dziedzinie tłumaczeń prawniczych, do tworzenia leksykonów i tezaurusów, do badań socjolingwistycznych, komparatystycznych i diachronicznych czy w lingwistyce kryminalistycznej[13]. Tego typu podejście jest szczególnie

[9] B. Lewandowska-Tomaszczyk, M. Bańko, R. Górski, M. Łaziński, P. Pęzik, A. Przepiórkowski, *Narodowy Korpus Języka Polskiego: geneza i dzień dzisiejszy*, [w:] *Narodowy Korpus Języka Polskiego*, red. A. Przepiórkowski, M. Bańko, R. Górski, B. Lewandowska-Tomaszczyk..., s. 3–5; I. Kida, *op. cit.*, s. 140–141.

[10] B. Lewandowska-Tomaszczyk, *op. cit.*, s. 33–41.

[11] *Confer* np. Narodowy Korpus Języka Polskiego (wyszukiwarka PELCRA).

[12] B. Lewandowska-Tomaszczyk, *op. cit.*, s. 95–131.

[13] Ł. Biel, *Corpus-Based Studies...*, s. 4–6.

rozpowszechnione w Europie, gdzie za główną motywację przyjęto poprawę jakości wielojęzycznej legislacji UE[14]. Tytułem przykładu warto tu wspomnieć chociażby o międzynarodowym projekcie badawczym *Eurolect Observatory Project*, w którym wzięli udział naukowcy z kilkunastu krajów członkowskich UE[15]. Jego głównym celem było stworzenie wielojęzycznego korpusu porównawczego złożonego z tekstów unijnych dyrektyw i krajowych aktów implementujących. Umożliwia on m.in. szczegółowe badanie wpływu prawodawstwa unijnego na język krajowych aktów prawnych w perspektywie porównawczej. W tym kontekście trzeba również przywołać korpus *JRC Collection of the Acquis Communautaire* stworzony przez Wspólne Centrum Badawcze (ang. *Joint Research Centre*) działające przy Komisji Europejskiej. Składa się on z tekstów aktów prawa unijnego oraz orzeczeń Trybunału Sprawiedliwości UE w 22 językach. Umożliwia porównanie ze sobą dowolnych dwóch wersji językowych (co daje 231 kombinacji) i pod tym względem jest unikatem na skalę światową. Stworzono go do celów badawczych oraz jako wsparcie dla unijnych służb tłumaczeniowych[16]. Obecnie korpusy są w zasadzie nieodzownym narzędziem translatoryki, w tym również translatoryki prawniczej. Wykorzystuje się je w badaniach teoretycznych, znanych jako *LTS* (ang. *Legal Translation Studies*), a także w praktyce tłumaczeniowej i w trakcie edukacji tłumaczy[17]. W tej dziedzinie doniosłą rolę odgrywają specjalistyczne korpusy wielojęzyczne: porównawcze (zbiory korpusów jednojęzycznych zbudowanych według tych samych kryteriów) i równoległe (zbiory korpusów złożonych z tekstów oryginalnych i ich tłumaczeń)[18]. Warto wspomnieć, że jedną z pionierek prawniczej translatoryki korpusowej jest polska badaczka z Uniwersytetu Warszawskiego – Łucja Biel[19].

Po drugie, korpusy zaczynają być wykorzystywane przez prawników w procesie interpretacji prawa. Praktyka ta narodziła się w Stanach Zjednoczonych Ameryki, gdzie podstawową dyrektywą wykładni prawa jest *ordinary meaning rule* stanowiąca w przybliżeniu połączenie zasady pierwszeństwa wykładni językowej z zasadą

---

[14] F. Vogel, H. Hamann, I. Gauer, *Computer-Assisted Legal Linguistics: Corpus Analysis as a New Tool for Legal Studies*, „Law & Social Inquiry" 2018, nr 43/4, s. 1350–1357; Ł. Biel, *Corpus-Based Studies...*, s. 4 i n.

[15] Eurolect, <https://www.unint.eu/en/research/research-projects/33-page/490-eurolect-observatory-project.html>, dostęp 22 II 2022.

[16] Ł. Biel, *Corpus-Based Studies...*, s. 4; R. Steinberger, B. Pouliquen, A. Widiger, C. Ignat, T. Erjavec, D. Tufis, D. Varga, *The JRC-Acquis: a Multilingual Aligned Parallel Corpus with 20+ Languages*, [w:] *Conference Proceedings: the 5th International Conference on Language Resources and Evaluation – Proceedings, Paris (France)*, European Language Resources Association (ELRA), Paris 2006, <https://publications.jrc.ec.europa.eu/repository/handle/JRC32786>, s. 2147, dostęp 22 II 2022; R. Steinberger, M. Ebrahim, A. Poulis, M. Carrasco-Benitez, P. Schlüter, M. Przybyszewski, S. Gilbro, *An Overview of the European Union's Highly Multilingual Parallel Corpora*, „Language Resources & Evaluation" 2014, nr 48, s. 679–707.

[17] F. Prieto Ramos, *The Use of Corpora in Legal and Institutional Translation Studies: Directions and Applications*, „Translation Spaces" 2019, nr 8, s. 2.

[18] Ł. Biel, *Corpus-Based Studies...*, s. 3; F. Prieto Ramos, *op. cit.*, s. 2.

[19] *Confer* m.in. Ł. Biel, *Lost in the Eurofog: the Textual Fit of Translated Law*, Peter Lang, Frankfurt 2014.

domniemania języka potocznego[20]. W amerykańskim orzecznictwie, w tym w orzecznictwie amerykańskiego Sądu Najwyższego (ang. *Supreme Court of the United States*), są znane przypadki wykorzystania korpusów językowych do ustalenia „zwykłego" znaczenia wyrażenia ustawowego[21]. Przypadki te – choć jak dotąd dość odosobnione – wywołały spore poruszenie w amerykańskiej nauce prawa. W latach 2016–2021 opublikowano kilkadziesiąt artykułów naukowych poświęconych zagadnieniu stosowania korpusów językowych w interpretacji prawa. Znajdziemy wśród nich m.in. reinterpretacje znanych kazusów z wykorzystaniem korpusów, analizy semantyczne, praktyczne wskazówki dla prawników sięgających po korpusy oraz – przede wszystkim – ożywioną debatę na temat dopuszczalności używania korpusów jako narzędzi interpretacyjnych. Wybrane wątki tej debaty zostaną poruszone w dalszej części niniejszego artykułu, natomiast szczegółowe omówienie amerykańskich doświadczeń z korpusami znajdzie Czytelnik w tekście *Korpusy językowe jako narzędzie interpretacji prawa. Amerykańska teoria i praktyka*[22]. Do tej pory w polskich realiach korpusy nie były wykorzystywane przez sądy do interpretacji prawa[23]. W nauce prawa zaledwie je dostrzeżono[24].

---

[20] B. Slocum, *Ordinary Meaning: a Theory of the Most Fundamental Principle of Legal Interpretation*, University of Chicago Press, Chicago–London 2016, s. 1–26; L. Solan, *The Language of Statutes*, University of Chicago Press, Chicago–London 2010, s. 50–81; A. Grabowski, *Siła argumentu interpretacyjnego z prawniczego domniemania języka (znaczenia) potocznego*, „Zagadnienia Naukoznawstwa" 2016, nr 3(209), s. 357–370.

[21] *Confer* m.in. [Wyrok Sądu Najwyższego stanu Utah z dnia 14 sierpnia 2015 r. w sprawie State *vs*. Rasabout 356 P.3d 1258, 1260 [Utah 2015]]; [Wyrok Sądu Najwyższego Stanów Zjednoczonych z dnia 1 marca 2011 r. w sprawie FCC *vs*. AT&T, Inc. 562 U.S. [2011]]; [Wyrok Sądu Najwyższego Stanów Zjednoczonych z dnia 21 czerwca 2018 r. w sprawie Lucia *vs*. SEC 138 U.S. [2018]]; [Wyrok Sądu Najwyższego Stanów Zjednoczonych z dnia 26 czerwca 2018 r. w sprawie Carpenter *vs*. United States 585 U.S. [2018]].

[22] Z. Tobor, M. Zeifert, *Korpusy językowe jako narzędzie interpretacji prawa. Amerykańska teoria i praktyka*, „Archiwum Filozofii Prawa i Filozofii Społecznej" 2020, nr 4(25), s. 80–92.

[23] Pojawiają się natomiast w opiniach biegłych językoznawców lub w argumentacji stron. *Confer* np. Wyrok Sądu Apelacyjnego w Białymstoku z dnia 26 czerwca 2015 r., I ACa 230/15, LEX nr 1765926: „Biegła wyjaśniła, iż podstawą sporządzania opinii z zakresu językoznawstwa jest tzw. «korpus języka polskiego». Jest to zrównoważony tematycznie i gatunkowo zbiór tekstów polskich służących do zilustrowania użycia słów i ich znaczeń. Składa się z tekstów książek, czasopism, druków ulotnych i akcydensowych (np. reklam, instrukcji obsługi, regulaminów, ulotek wyborczych), stron internetowych oraz tekstów mówionych, stanowiąc autentyczny materiał językowy, na którego podstawie opisuje się znaczenia słów i konstrukcji. Na podstawie tego rodzaju bazy, biegła oceniła rozumienie słowa w przestrzeni komunikowania publicznego. Z tych przyczyn Sąd Apelacyjny, nie podziela zarzutu nieprzydatności opinii biegłej w przedmiotowej sprawie i naruszenia art. 233 § 1 k.p.c. przy jej ocenie". *Confer* także Wyrok Sądu Apelacyjnego w Łodzi z dnia 13 listopada 2014 r., I ACa 709/14, LEX nr 1554764; Wyrok Wojewódzkiego Sądu Administracyjnego w Warszawie z dnia 19 marca 2019 r., VI SA/Wa 1895/18, LEX nr 2696759; Wyrok Sądu Okręgowego w Warszawie z dnia 22 sierpnia 2018 r., X Ka 721/18, LEX nr 2566284.

[24] A. Choduń, *Aspekty językowe derywacyjnej koncepcji wykładni prawa*, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin 2018, s. 257–261; Z. Tobor, M. Zeifert, *op. cit.*, *passim*; A. Malinowski, *Pomiar czytelności polskiego tekstu prawnego*, „Państwo i Prawo" 2017, nr 5, s. 3–18.

Warto wszakże odnotować pierwsze próby zastosowania analiz korpusów językowych do rozwiązania praktycznych problemów interpretacyjnych[25].

IV. Korpusy językowe w procesie tworzenia prawa

Zagadnienie wykorzystania korpusów w procesie tworzenia prawa jak dotąd było dyskutowane głównie na łamach publikacji językoznawczych w kontekście wielojęzycznego prawa unijnego. W przekonaniu autorów temat ten zasługuje na zainteresowanie wszystkich osób zaangażowanych w proces prawodawczy, w szczególności legislatorów, a także teoretyków prawa i teoretyków legislacji. W dalszej części artykułu wskazano zatem na potencjalne sposoby wykorzystania korpusów jako narzędzi tworzenia prawa. Dla czytelności wywodu osobno omówiono korpusy ogólne i korpusy specjalistyczne. Zaznaczyć wypada, że omówienie to ma charakter wstępny i należy je traktować jako przyczynek do dalszej dyskusji.

1. Korpusy ogólne

W pierwszej kolejności zostanie omówione wykorzystanie korpusów ogólnych, takich jak przywoływany już, dostępny publicznie, Narodowy Korpus Języka Polskiego [dalej: NKJP]. Zdaniem autorów użycie tego rodzaju korpusów jako pomocy przy redagowaniu tekstu prawnego znajduje głębokie uzasadnienie normatywne i aksjologiczne. W *Zasadach techniki prawodawczej* [dalej: ZTP] prawodawca postanowił, że „w ustawie należy posługiwać się poprawnymi wyrażeniami językowymi (określeniami) w ich podstawowym i powszechnie przyjętym znaczeniu”[26]. Bez względu na znane kontrowersje wokół mocy wiążącej tego aktu prawnego i formalnych konsekwencji jego naruszenia[27] należy uznać, że przepis ten nakłada na tworzącego ustawę (i odpowiednio inne akty prawne[28]) obowiązek w zakresie doboru odpowiedniego słownictwa. W teorii legislacji wskazuje się, że przepis ten wyraża zasadę komunikatywności, która jest

[25] E. Kubicka, K. Grec, W. Leszczak, M. Mitrew, *Wykorzystanie słowników w interpretacji prawniczej – implikacje praktyczne*, „Kwartalnik Sądowy Apelacji Gdańskiej” 2015, nr 1, s. 39–40, 42, 44–45; M. Zeifert, Z. Tobor, *Linguam novit curia? Analiza korpusowa w prawie karnym – przyczynek do dyskusji*, „Forum Polityki Kryminalnej” 2021, nr 1, s. 12–25.

[26] § 8 ust. 1 Załącznika do rozporządzenia Prezesa Rady Ministrów z dnia 20 czerwca 2002 r. w sprawie „Zasad techniki prawodawczej”, tekst jedn. Dz.U. z 2016 r. poz. 283.

[27] M. Błachut, J. Kaczor, *Typy uzasadnień obowiązywania zasad techniki prawodawczej w polskim porządku prawnym*, [w:] *Folia Iuridica Wratislaviensis*, red. L. Górnicki, Prawnicza i Ekonomiczna Biblioteka Cyfrowa, Wrocław 2012; S. Wojtczak, *Zasady techniki prawodawczej – luka w prawie?*, „Państwo i Prawo” 2005, nr 1; S. Wronkowska, *Zasady techniki prawodawczej, czy rzeczywiście luka w prawie?*, „Państwo i Prawo” 2005, nr 2.

[28] *Confer* § 132 (rozporządzenia), § 141 (uchwały i zarządzenia), § 143 (akty prawa miejscowego) ZTP.

„odbiciem zasad wykładni o identycznej treści"[29]. Powstaje pytanie o metody zapewnienia realizacji tej normy, a więc źródła ustalenia poprawności wyrażeń językowych oraz ich podstawowego i powszechnie przyjętego znaczenia. W tym zakresie są aktualne ustalenia zawarte w literaturze w odniesieniu do interpretacji tekstów prawnych. Poszukując znaczenia powszechnego (w dyskursie interpretacyjnym często określanego również jako „potoczne"[30]), legislator – tak samo jak interpretator – ma do dyspozycji własną intuicję językową, słowniki języka polskiego, pomoc specjalisty (językoznawcy) oraz korpus ogólny[31].

Bez większego ryzyka można założyć, że najczęstszym źródłem wiedzy jest intuicja językowa legislatora. Na ogół sprawdza się ona znakomicie, gdy chodzi o ocenę poprawności gramatycznej zdania czy dopuszczalności semantycznej użycia danego wyrazu[32]. Pojęcie podstawowego i powszechnie przyjętego znaczenia zawiera w sobie jednak element frekwencji, a intuicja językowa nie sprawdza się za dobrze w ocenie częstotliwości określonego użycia. Ludzie, w tym również prawnicy, są narażeni na błąd poznawczy znany jako fałszywy konsensus (ang. *false consensus bias*), tzn. mają tendencję do zawyżonej oceny częstości własnych poglądów i preferencji. Dotyczy to również uzusu językowego[33]. Jeżeli zatem legislator przeważnie używa jakiegoś wyrazu w określonym znaczeniu, będzie miał skłonność do traktowania tego znaczenia jak podstawowe i powszechnie przyjęte.

Kolejnym źródłem wiedzy o znaczeniu słów są słowniki. W teorii legislacji zazwyczaj właśnie do nich się odsyła przy omawianiu § 8 ZTP[34]. W odniesieniu do wymaganego przez ten przepis przymiotu poprawności używanych wyrażeń słowniki są bez wątpienia najlepszym źródłem wiedzy. W polskiej leksykografii od dawna dominuje przekonanie, że zadaniem słowników jest m.in. kształtowanie normy językowej[35]. Natomiast jeśli chodzi o ustalenie, które znaczenie wyrazu jest podstawowe i powszechnie przyjęte,

---

[29] G. Wierczyński, *Redagowanie i ogłaszanie aktów normatywnych. Komentarz*, Wolters Kluwer, Warszawa 2016, s. 98. Zasada ta została wyrażona na bardziej ogólnym poziomie w § 6 ZTP: „Przepisy ustawy redaguje się tak, aby dokładnie i w sposób zrozumiały dla adresatów zawartych w nich norm wyrażały intencje prawodawcy".

[30] A. Grabowski, *op. cit.*, s. 363.

[31] L. Solan, T. Gales, *Corpus Linguistics as a Tool in Legal Interpretation*, „ Brigham Young University Law Review" 2017, nr 6, s. 1331–1336; S. Mouritsen, *Hard Cases and Hard Data: Assessing Corpus Linguistics as an Empirical Path to Plain Meaning*, „Columbia Science and Technology Law Review" 2011, nr 13, s. 172–201; A. Bielska-Brodziak, *Interpretacja tekstu prawnego na podstawie orzecznictwa podatkowego*, Wolters Kluwer, Warszawa 2009, s. 25–39.

[32] L. Solan, T. Gales, *op. cit.*, s. 1333–1334.

[33] *Ibidem*; L. Solan, T. Rosenblatt, D. Osherson, *False Consensus Bias in Contract Interpretation*, „Columbia Law Review" 2008, nr 5(108), s. 1282–1285.

[34] G. Wierczyński, *op. cit.*, s. 97; M. Dębska, *Zasady techniki prawodawczej. Komentarz*, Lexis Nexis, Warszawa 2013, komentarz do § 8, <https://sip-1lex-1pl-15d274sv506ca.han.bg.us.edu.pl/#/commentary/587689724/487164?tocHit=1&cm=RELATIONS>, dostęp 11 III 2022; A. Malinowski, *Rozdział X. Jednoznaczność i precyzja słów w tekście prawnym*, [w:] *Zarys metodyki pracy legislatora. Ustawy – akty wykonawcze – prawo miejscowe*, red. A. Malinowski, Lexis Nexis, Warszawa 2009, s. 290.

[35] P. Żmigrodzki, *Wprowadzenie do leksykografii polskiej*, Wydawnictwo Uniwersytetu Śląskiego, Katowice 2009, s. 111–113.

sprawa jest bardziej skomplikowana. W teorii legislacji wskazywano, że znaczeniem podstawowym jest to, które w słowniku występuje na pierwszym miejscu[36] lub jako jedno z pierwszych[37]. Zauważmy zatem, że pomiędzy tymi dwiema wskazówkami zachodzi istotna różnica. Nie powinno to jednak dziwić, zważywszy, że również w leksykografii brak stanowczych ustaleń w tej materii. W polskiej tradycji przyjęło się znaczenie podstawowe umieszczać jako pierwsze[38]. Tym niemniej w niektórych nowszych słownikach zdecydowano się podporządkować kolejność opisywanych znaczeń frekwencji ich występowania w tekstach źródłowych[39]. Odrębnym źródłem wątpliwości jest samo pojęcie znaczenia podstawowego. W leksykografii definiuje się je jako „to, które «pierwsze przychodzi do głowy» rodowitemu użytkownikowi języka polskiego, kiedy słyszy dany wyraz"[40]. Jest to zatem kwestia subiektywnej, intuicyjnej wręcz oceny autora określonego hasła, której nie da się wyeliminować poprzez zastosowanie jednolitej metodologii. Warto wspomnieć, że autorzy haseł słownikowych – tak jak cała reszta populacji – są narażeni na wspomniany powyżej błąd fałszywego konsensusu. W rezultacie nie tylko opisy znaczeń, lecz także ich kolejność będą się różnić w zależności od słownika[41]. Wydaje się zatem, że słowniki jako narzędzie ustalania podstawowego i powszechnie przyjętego znaczenia słów nie są pozbawione mankamentów.

Trzecią możliwością jest odwołanie się do eksperta, czyli językoznawcy. W kontekście interpretacyjnym rozwiązanie to jest bardzo kontrowersyjne ze względów procesowych i z uwagi na zasadę *iura novit curia* (pol. sędzia zna prawo). W efekcie sądy niezwykle rzadko decydują się na zasięgnięcie opinii biegłego językoznawcy[42]. W kontekście legislacyjnym te wątpliwości nie występują. Przepisy ustawy o języku polskim wprost przewidują możliwość zasięgnięcia opinii Rady Języka Polskiego przez wiele organów władzy publicznej „w wypadku wystąpienia w toku czynności urzędowych istotnych wątpliwości dotyczących użycia języka polskiego"[43]. Tego typu pomoc zewnętrzna musi być zastrzeżona dla szczególnych sytuacji. W niektórych ustawodawstwach europejskich współpraca legislatorów z lingwistami przybrała formy instytucjonalne. Przykładowo, w Szwajcarii proces tworzenia przepisów wspiera wyspecjalizowana jednostka Kancelarii Federalnej (niem. *Bundeskanzlei*), tzw. Centralna Służba Językowa (niem. *Zentrale Sprachdienste*), a także interdyscyplinarna Wewnątrzadministracyjna Komisja

---

[36] D. Szafrański, *Komentarz do § 8*, [w:] *Zasady techniki prawodawczej w zakresie aktów prawa miejscowego. Komentarz praktyczny z wzorami oraz przykładami*, red. D. Szafrański, W. Białończyk, A. Bielecki, Ł. Kasiak, J. Piecha, C.H.Beck, Warszawa 2016, <https://sip-1legalis-1pl-1lp6svawk0089.han.bg.us.edu.pl/document-view.seam?documentId=mjxw62zogiztkmjtgi2dqmjoobqxalrtgqytgmzvge4q#tabs-metrical-info>, dostęp 11 III 2022.

[37] G. Wierczyński, *op. cit.*, s. 97.

[38] P. Żmigrodzki, *op. cit.*, s. 76.

[39] *Ibidem*, s. 169.

[40] *Ibidem*, s. 76.

[41] *Ibidem*, s. 78.

[42] A. Bielska-Brodziak, *op. cit.*, s. 26.

[43] Artykuł 14 ust. 1 w zw. z art. 4 Ustawy z dnia 7 października 1999 r. o języku polskim, tekst jedn. Dz.U. z 2021 r. poz. 672.

Redakcyjna (niem. *Verwaltungsinterne Redaktionskommission*, *VIRK*)[44]. Powstaje jednak pytanie, skąd ekspert, czyli językoznawca, będzie czerpał wiedzę na temat podstawowego i powszechnie przyjętego znaczenia słów? Wnioskując na podstawie ogólnych trendów we współczesnym językoznawstwie, w tym aktualnej metodologii tworzenia słowników, można stwierdzić, że będą to przede wszystkim korpusy językowe. Ustalenie to prowadzi do kolejnego punktu.

Ostatnią opcją jest wykorzystanie korpusów językowych, które zawierają „surowy" materiał językowy, „nieskażony" ocenami i preferencjami leksykografa czy samego legislatora. Pozwalają one na stosowanie metod ilościowych, w tym określenie rzeczywistej frekwencji użycia słowa w danym znaczeniu[45]. W skrócie, umożliwiają zastąpienie intuicji obserwacją oraz odzwierciedlają to, co w języku jest centralne, typowe, normalne i przewidywalne[46]. Wydaje się, że w zakresie, w którym omawiany przepis formułuje nakaz używania wyrazów „w ich podstawowym i powszechnie przyjętym znaczeniu"[47], korpusy dają lepszą gwarancję jego realizacji niż metody alternatywne, tj. własna intuicja językowa legislatora i słowniki. Warto w tym miejscu przywołać stosowny przykład.

Zgodnie z ustawą nie uważa się za koszty uzyskania przychodów „*kosztów reprezentacji*, w szczególności poniesionych na usługi gastronomiczne, zakup żywności oraz napojów, w tym alkoholowych"[48]. Użyte w tym przepisie (a także w jego poprzedniej wersji) słowo „reprezentacja" spowodowało, że pojawiły się poważne wątpliwości w praktyce organów podatkowych i w orzecznictwie sądów administracyjnych. Przez wiele lat sądy definiowały znaczenie tego wyrazu jako 'okazałość, wystawność', powołując się na dyrektywę języka potocznego i różne słowniki języka polskiego. Zgodnie z *Uniwersalnym słownikiem języka polskiego* pod redakcją Stanisława Dubisza „reprezentacja" to: „1. grupa osób występująca w czyimś imieniu, reprezentująca czyjeś interesy, przedstawicielstwo; 2. okazałość, wystawność w czyimś sposobie życia, związana ze stanowiskiem, pozycją społeczną; 3. grupa zawodników reprezentująca w zawodach barwy swojego kraju, klubu lub miasta"[49]. Takie rozumienie zostało ostatecznie zakwestionowane przez Naczelny Sąd Administracyjny z uwagi na kwestie systemowe

[44] F. Vogel, H. Hamann, I. Gauer, *op. cit.*, s. 1350; Ł. Iluk, J. Iluk, *Szwajcarski sposób na jasność i komunikatywność aktów normatywnych na przykładzie procedury redagowania konstytucji kantonu Zurych z 2005 r.*, „Państwo i Prawo" 2020, nr 3, s. 81.

[45] L. Solan, *Can Corpus Linguistics Help Make Originalism Scientific?*, „The Yale Law Journal Forum" 2016, nr 126, s. 62–63.

[46] Ł. Biel, *Corpus-Based Studies...*, s. 2.

[47] § 8 ust. 1 Załącznika do rozporządzenia Prezesa Rady Ministrów z dnia 20 czerwca 2002 r. w sprawie „Zasad techniki prawodawczej", tekst jedn. Dz.U. z 2016 r. poz. 283.

[48] Artykuł 23 ust. 1 pkt 23 Ustawy z dnia 16 lipca 1991 r. o podatku dochodowym od osób fizycznych, tekst jedn. Dz.U. z 2020 r. poz. 1426, ze zm. oraz art. 16 ust. 1 pkt 28 Ustawy z dnia 15 lutego 1992 r. o podatku dochodowym od osób prawnych, tekst jedn. Dz.U. z 2020 r. poz. 1406, ze zm.

[49] *Uniwersalny słownik języka polskiego*, t. 3, red. S. Dubisz, Wydawnictwo Naukowe PWN, Warszawa 2008, s. 932. Podobne definicje znajdziemy w słowniku pod redakcją Witolda Doroszewskiego. *Vide* strona internetowa Słownika Języka Polskiego PWN, <https://sjp.pwn.pl/doroszewski/reprezentacja;5489765.html>, dostęp 23 II 2022.

i funkcjonalne[50]. W dominującym obecnie ujęciu „reprezentacja" oznacza „działanie w celu wykreowania oraz utrwalenia pozytywnego wizerunku podatnika wobec innych podmiotów"[51]. W przekonaniu autorów wspomniane kontrowersje pojawiły się na skutek niezastosowania dyrektywy z § 8 ZTP przy redakcji spornego przepisu. Niezależnie od tego, co wynika z lektury słowników, żadne z konkurujących znaczeń słowa „reprezentacja" nie ma charakteru podstawowego i powszechnie przyjętego. Przekonuje o tym rudymentarna analiza korpusowa. Po pierwsze, w losowej próbce 100 wystąpień słowa „reprezentacja" w zbalansowanym podkorpusie NKJP próżno szukać użycia odpowiadającego któremuś ze wskazanych wyżej znaczeń[52]. Po drugie, wśród najczęstszych kolokacji (współwystąpień) tego słowa zdecydowanie wiodą prym wyrażenia mające kontekst sportowy i polityczny[53]. Bez względu zatem na to, czy intencją prawodawcy było nadanie słowu „reprezentacja" w ustawach podatkowych znaczenia zawierającego element wystawności czy też element przedstawicielstwa, żadne z tych znaczeń nie stanowi „podstawowego i powszechnie przyjętego" znaczenia tego słowa, a zatem dyrektywa komunikatywności z § 8 ZTP została naruszona.

Oczywiście prawodawca może tak postąpić. Dyrektywa komunikatywności wyrażona w § 8 ZTP nie ma charakteru bezwzględnego. Jak wiele postanowień tego aktu prawnego paragraf ten ma raczej status zasady niż reguły prawnej. Należy się zgodzić z poglądami, zgodnie z którymi w konfrontacji z tą dyrektywą pierwszeństwo mają zasady precyzji i adekwatności, również wyrażone w ZTP[54]. Odejście od podstawowego i powszechnie przyjętego znaczenia nie może być jednak dowolne. Powinno wynikać z potrzeby realizacji innych określonych wartości. Ponadto prawodawca ma obowiązek zadbać, by wybrane przez niego niekonwencjonalne znaczenie danego wyrażenia było zrozumiałe dla adresatów tekstu. Może to uczynić przykładowo poprzez umieszczenie wyrażenia w odpowiednim kontekście czy też poprzez wprowadzenie jego definicji legalnej. W przypadku słowa „reprezentacja" wymogi te nie zostały spełnione, o czym świadczy długa i burzliwa historia sporów interpretacyjnych toczonych przed sądami administracyjnymi.

Zaprezentowany przykład pokazuje dobitnie, jak różne mogą być w praktyce rezultaty analizy słownikowej i analizy korpusowej. Przypomnijmy, że dominująca przez

---

[50] H. Filipczyk, *Koszty reprezentacji w rozumieniu art. 23 ust. 1 pkt 23 ustawy z 26.07.1991 r. o podatku dochodowym od osób fizycznych. Glosa do wyroku NSA z dnia 25 stycznia 2012 r., II FSK 1445/10*, „Glosa" 2012, nr 4, s. 113–119. *Confer* też Wyrok Naczelnego Sądu Administracyjnego z dnia 25 stycznia 2012 r., II FSK 1445/10, LEX nr 1102780.

[51] Wyrok Wojewódzkiego Sądu Administracyjnego w Warszawie z dnia 29 czerwca 2020 r., III SA/Wa 2042/19, LEX nr 3064883.

[52] Strona internetowa Narodowego Korpusu Języka Polskiego, <http://nkjp.pl/poliqarp/nkjp300/query/>, dostęp 23 II 2022.

[53] Strona internetowa Narodowego Korpusu Języka Polskiego, <http://nkjp.uni.lodz.pl/collocations.jsp?qfield=reprezentacja**&dummystring=%C4%85%C4%84%C4%87%C4%86%C4%99%C4%98%C5%82%C5%81%C5%84%C5%83%C3%B3%C3%93%C5%9B%C5%9A%C5%BA%C5%B9%C5%BC%C5%BB>, dostęp 23 II 2022.

[54] G. Wierczyński, *op. cit.*, s. 96.

lata linia orzecznicza opierała się w dużej mierze na słownikowej definicji znaczenia wyrazu „reprezentacja", opisywanego jako „okazałość, wystawność w czyimś sposobie życia, związana ze stanowiskiem, pozycją społeczną"[55]. W cytowanym powyżej popularnym słowniku pod redakcją S. Dubisza definicja ta występuje na drugim miejscu (z trzech). Na tej podstawie można by twierdzić, iż wyraża ona jedno z podstawowych i powszechnie przyjętych znaczeń rzeczonego wyrazu[56]. Wniosek taki w żaden sposób nie koresponduje jednak z przeprowadzoną analizą korpusową. Przypomnijmy, że w losowej próbce 100 użyć słowa „reprezentacja" nie znalazło się ani jedno, które prezentowałoby wskazane wyżej znaczenie słownikowe. Również analiza kolokacji spornego wyrazu ujawniła niezbicie, że najczęściej pojawiającymi się jego znaczeniami są znaczenia odwołujące się do kontekstów sportowego i politycznego. Przeciwnicy stosowania korpusów w interpretacji prawa podnoszą, że częstotliwość występowania danego znaczenia nie przesądza o jego poprawności czy statusie znaczenia podstawowego[57]. Wobec tego należy wskazać na dwie kwestie. Po pierwsze, frekwencja jest obecnie uważana przez większość autorów za zasadnicze – choć nie jedyne – kryterium poprawności językowej[58]. Po drugie, w kontekście omawianego przykładu analiza korpusowa miała charakter negatywny. Jej celem nie było ustalenie, co stanowi podstawowe i powszechnie przyjęte znaczenie słowa „reprezentacja", a jedynie wykazanie, że znaczenia funkcjonujące w dyskursie prawniczym nie przejawiają takich cech.

W § 8 ZTP zawarto także inne, bardziej szczegółowe dyrektywy doboru słownictwa. Wskazano mianowicie, że w ustawie należy unikać posługiwania się:

„1) określeniami specjalistycznymi, o ile ich użycie nie jest powodowane zapewnieniem należytej precyzji tekstu;

2) określeniami lub zapożyczeniami obcojęzycznymi, chyba że nie mają dokładnego odpowiednika w języku polskim;

3) nowo tworzonymi pojęciami lub strukturami językowymi, chyba że w dotychczasowym słownictwie polskim brak jest odpowiedniego określenia"[59].

W ocenie autorów korpusy mogą posłużyć jako przydatne narzędzia realizacji również i tych dyrektyw. W przypadku wyrażeń specjalistycznych (profesjonalizmów) korpus może przede wszystkim pomóc w ich identyfikacji. Zapoznanie się z próbką zdań zawierających dane słowo pozwoli ustalić, czy ma ono charakter specjalistyczny; pomocna może się tu okazać dostępna w darmowym interfejsie NKJP opcja zawężenia wyszukiwania do określonych źródeł (np. tylko literatura naukowa albo odwrotnie – tylko język mówiony). Analiza korpusowa będzie przydatna także w razie poszukiwania

[55] *Uniwersalny słownik języka polskiego*, t. 3..., s. 932.

[56] *Ibidem*, s. 97.

[57] J. Ehrett, *Against Corpus Linguistics*, „The Georgetown Law Journal Online" 2017, nr 108, s. 62–64. Na tę kwestię zwraca również uwagę A. Choduń, *op. cit.*, s. 257–262.

[58] *Confer* U. Zdunek, *Rola kryterium frekwencyjnego w ocenie poprawności leksykalnej w słownikach poprawnej polszczyzny*, niepublikowana rozprawa doktorska obroniona na UW, Warszawa 2013, <https://depotuw.ceon.pl/bitstream/handle/item/626/Doktorat%20ca%C5%82o%C5%9B%C4%875.pdf?sequence=1>, dostęp 10 II 2022.

[59] § 8 ust. 2 ZTP.

synonimów wyrażeń specjalistycznych. W przypadku określeń i zapożyczeń obcojęzycznych, a także nowo utworzonych pojęć i struktur językowych (neologizmów, neosemantyzmów) korpus pomoże ustalić, na ile dane wyrażenie jest obecne w języku polskim. Szczególną rolę odgrywają tu cechy reprezentatywności i aktualności korpusów. Z uwagi na elektroniczną formę i możliwość stałego dodawania kolejnych źródeł korpusy mają większą szansę nadążać za rozwojem języka niż tradycyjne słowniki, które są opracowywane i publikowane w konkretnym momencie[60].

Wreszcie, korpusy mogą pomóc uniknąć nieświadomego stosowania słów wieloznacznych. W literaturze przedmiotu przestrzega się przed lekceważeniem tego ryzyka i przed poprzestawaniem na intuicji językowej legislatora. Doradza się sięganie do słowników języka polskiego „w celu poznania potencjalnych znaczeń słów występujących w tym tekście"[61]. Podobną rolę mogą odgrywać korpusy. Wydają się przy tym mieć nad słownikami tę przewagę, że dają o wiele bardziej zniuansowany, rzeczywisty obraz języka. W rezultacie pozwalają wychwycić nawet drobne różnice semantyczne, które niekoniecznie znajdują odzwierciedlenie w definicjach słownikowych.

2. Korpusy specjalistyczne

Opisane powyżej sposoby użycia korpusów ogólnych co do zasady odpowiadają ich wykorzystaniu w procesie wykładni prawa, które od kilku lat jest przedmiotem ożywionej debaty, zwłaszcza w Stanach Zjednoczonych Ameryki. Warto jednak zwrócić uwagę, że w językoznawstwie wykorzystuje się nie tylko korpusy ogólne, takie jak NKJP, lecz także – a może nawet przede wszystkim – korpusy specjalistyczne, tj. korpusy stworzone w określonym celu badawczym, zawierające teksty dobrane według odpowiedniego kryterium (np. kryterium rejestru językowego, kryterium pochodzenia tekstów, kryterium autora itd.). Liczne korpusy tego typu są wykorzystywane w translatoryce prawniczej. Przykładem jest wspomniany powyżej korpus *JRC Collection of the Acquis Communautaire* składający się z aktów unijnego prawa wtórnego w 22 językach[62]. Takie korpusy są zwykle budowane w celach językoznawczych i składają się z tekstów o bardzo różnym – z prawnego punktu widzenia – charakterze, przykładowo: aktów prawnych, orzeczeń, artykułów naukowych, umów, testamentów, mów sądowych, zeznań świadków itd.[63]

W kontekście tworzenia prawa szczególnie istotne znaczenie mają korpusy zbudowane wyłącznie z tekstów obowiązujących aktów prawnych (nowego znaczenia

[60] W teorii językoznawczej wyróżnia się w tym kontekście korpusy referencyjne, które w momencie utworzenia zostają „zamknięte", i korpusy monitorujące, które są stale wzbogacane o nowy materiał językowy. *Confer* I. Kida, *op. cit.*, s. 134–135. Narodowy Korpus Języka Polskiego stanowi korpus referencyjny, a jego tworzenie zostało zakończone w 2012 r.

[61] A. Malinowski, *Rozdział X...*, s. 284.

[62] Listę istniejących prawniczych korpusów specjalistycznych podano w pracy F. Vogel, H. Hamann, I. Gauer, *op. cit.*, s. 1350–1355.

[63] *Ibidem*; Ł. Biel, *Corpus-Based Studies...*, s. 3–6; F. Prieto Ramos, *op. cit.*, s. 4–9.

nabiera tu termin *corpus iuris*). Dostępność tego typu korpusów jest obecnie bardzo słaba[64]. Wyjątek stanowi korpus prawa szwajcarskiego, zawierający kompletny zbiór tekstów szwajcarskich aktów prawnych w trzech oficjalnych językach (niemieckim, francuskim i włoskim)[65]. Korpus ten został stworzony przez duet uczonych z Uniwersytetu Zuryskiego (niem. *Universität Zürich*), tj. Stefana Höflera i Kyoko Sugisaki, i docelowo ma służyć nie tylko celom badawczym, lecz także praktycznym, tj. wsparciu służb legislacyjnych. Korpusy złożone z tekstów prawnych otwierają przed legislatorem wiele niezwykle przydatnych możliwości. Zespół niemieckich badaczy wskazuje, że oferują one wiarygodne odpowiedzi na istotne pytania praktyki legislacyjnej, takie jak:

- „W jaki sposób język (słownictwo) był używany w przeszłości i jak się go używa obecnie w różnych dziedzinach prawa?
- Jak często, przez kogo i gdzie są używane poszczególne wyrażenia i jakie znaczenie jest im przypisywane?
- Jakie prawidłowości językowe (ang. *speech patterns*) mają specyficzne znaczenie w ramach wspólnoty prawniczej?
- Jakie słowa zazwyczaj współwystępują w konkretnych zestawieniach (kolokacje)?
- Jakie wyrażenia są używane w podobnych, ale nie identycznych kontekstach (quasi-synonimy) i powinny być szczegółowo zdefiniowane w ustawie?"[66].

W omawianym kontekście jedną z kluczowych funkcji korpusów specjalistycznych jest tworzenie konkordancji, tj. zbiorów przykładów użycia danego wyrazu, zgodnie z kryteriami wprowadzonymi przez badacza. Konkordancje są nieocenioną pomocą w tłumaczeniu aktów prawnych i w wielojęzycznej legislacji[67]. Autorzy są jednak zdania, że mogą się one okazać bardzo pomocne również przy redagowaniu jednojęzycznych tekstów prawnych. Legislator zamierzający użyć danego wyrazu może w łatwy sposób prześledzić przykłady jego dotychczasowego użycia w obowiązujących aktach prawnych. Mianowicie, wprowadzając do projektu ustawy słowo „reprezentacja", może uzyskać informację o wszystkich wystąpieniach tego słowa w aktualnie obowiązujących aktach prawnych. Ich prezentacja w formacie KWIC umożliwia wygodne porównanie nawet bardzo wielu przykładów. Tego typu analiza może dodatnio wpłynąć na spójność terminologiczną i semantyczną systemu prawnego. Przypomnijmy, że zasada konsekwencji terminologicznej jest uznawana za „fundamentalną zasadę dobrej roboty prawodawczej"[68]. Została ona zresztą wprost wyrażona w ZTP:

„§ 9. W ustawie należy posługiwać się określeniami, które zostały użyte w ustawie podstawowej dla danej dziedziny spraw, w szczególności w ustawie określanej jako „kodeks" lub „prawo".

---

[64] F. Vogel, H. Hamann, I. Gauer, *op. cit.*, s. 1350.

[65] *Ibidem*, s. 1354; S. Höfler, K. Sugisaki, *Constructing and Exploiting an Automatically Annotated Resource of Legislative Texts*, „Zurich Open Repository and Archive" 2014, s. 175.

[66] F. Vogel, H. Hamann, I. Gauer, *op. cit.*, s. 1350. [Tłumaczenie z języka angielskiego – tłumaczenie własne].

[67] S. Höfler, K. Sugisaki, *op. cit.*, s. 178–179.

[68] Uchwała Sądu Najwyższego z dnia 30 kwietnia 2003 r., I KZP 8/03, OSNKW 2003, nr 5–6, poz. 41. *Confer* też G. Wierczyński, *op. cit.*, s. 103–109.

§ 10. Do oznaczenia jednakowych pojęć używa się jednakowych określeń, a różnych pojęć nie oznacza się tymi samymi określeniami”[69].

Ze względu na objętość systemu prawnego realizacja wyżej wymienionych dyrektyw, w szczególności zaś dyrektywy wyrażonej w § 10 ZTP, napotyka istotne trudności praktyczne. Ogromna liczba obowiązujących tekstów prawnych wyklucza możliwość bezpośredniego zapoznania się z nimi wszystkimi przez jedną osobę. Również dostępne powszechnie narzędzia informatyczne, tj. prawnicze elektroniczne bazy danych (zarówno publiczne, jak i komercyjne), nie oferują dogodnego sposobu zestawienia różnych wystąpień określonego wyrazu. Legislator traktujący poważnie zalecenie z § 10 ZTP jest skazany na żmudne przeglądanie poszczególnych tekstów w formie elektronicznej, względnie sprawdzanie wyników uzyskanych w wyszukiwarce elektronicznej bazy danych. Przykładowo, po wpisaniu w wyszukiwarkę Systemu Informacji Prawnej „Lex” hasła „reprezentacja” wyświetla się 15 369 wyników w kategorii „akty prawne”, w tym 1259 wyników w podkategorii „Dz.U. i M.P.”[70] Przy takiej liczbie analiza uzyskanych rezultatów byłaby zadaniem niezwykle czasochłonnym. Właściwości interfejsu sprawiają, że nie ma możliwości wyświetlenia wyłącznie zdań zawierających żądane hasło. Każdorazowo jest konieczne otwarcie danego aktu prawnego w nowym oknie i prześledzenie podświetlonych wystąpień hasła przy użyciu przycisków „poprzedni” – „następny”, względnie kolejne wyszukiwanie. Dodatkowe utrudnienie stanowi fakt, że wyszukiwarka nie różnicuje tytułów aktów normatywnych, preambuł, treści tekstu prawnego czy dopisków redakcyjnych (o charakterze nieoficjalnym). W rezultacie znaczna część uzyskanych wyników w rzeczywistości nie stanowi przykładów użycia słowa „reprezentacja” w tekstach prawnych, a jedynie przykłady jego użycia w nagłówkach poszczególnych artykułów dopisanych przez redaktorów wydawnictwa, które to nie mają żadnego znaczenia z punktu widzenia realizacji zasady konsekwencji terminologicznej. Jeszcze gorzej przedstawia się sytuacja w odniesieniu do niekomercyjnych, prawniczych elektronicznych baz danych. Internetowy System Aktów Prawnych [dalej: ISAP], prowadzony przez Kancelarię Sejmu[71], w ogóle nie ma opcji swobodnego przeszukiwania treści aktów prawnych, a w prowadzonym w formie elektronicznej Dzienniku Ustaw[72] po wpisaniu słowa „reprezentacja” wyświetla się jedynie lista aktów prawnych (w liczbie 353)[73], w których ono występuje. Dalsza analiza wymagałaby otwarcia każdego z tych aktów w osobnym oknie i zapoznania się z jego treścią (ewentualnie za pomocą wyszukiwarki).

Konkordancja, którą można by wygenerować przy użyciu omawianego korpusu specjalistycznego, stanowi sposób bez porównania bardziej efektywny. Po wpisaniu poszukiwanej frazy wyświetlałaby się lista wszystkich jej wystąpień w systemie prawnym

[69] § 9–10 ZTP.

[70] Stan na 17 VIII 2022.

[71] Internetowy System Aktów Prawnych, <https://isap.sejm.gov.pl/>, dostęp 17 VIII 2022.

[72] Dziennik Ustaw, <https://dziennikustaw.gov.pl/DU>, dostęp 17 VIII 2022.

[73] Stan na 17 VIII 2022.

w dogodnym do analizy formacie KWIC. Interfejs korpusu powinien dawać możliwość zawężenia wyników do aktów określonego rodzaju, wykluczenia tytułów czy tekstów preambuł itd. Legislator mógłby się zapoznać z uzyskanymi wynikami bez potrzeby otwierania kolejnych okien czy nawigowania po poszczególnych aktach prawnych. Oczywiście ostatecznie zachowanie konsekwencji terminologicznej zależy od decyzji redakcyjnej legislatora i nawet najlepsze narzędzie do analizy nie gwarantuje jej podjęcia. Wydaje się jednak, że w obecnym stanie rzeczy realizacja tej zasady jest znacznie utrudniona, a możliwości techniczne wynikające z rozwoju techniki komputerowej pozostają w tym zakresie niewykorzystane.

Na koniec warto wspomnieć o ogromnym potencjale drzemiącym w różnego rodzaju zautomatyzowanych narzędziach do analizy i redakcji tekstów prawnych. Dziedzina ta jak dotąd pozostaje głównie w sferze *legislative science-fiction*, choć w kilku krajach europejskich (Szwajcaria i Niemcy) są prowadzone w tym zakresie pionierskie projekty badawcze.

W Niemczech grupa uczonych z Uniwersytetu w Heidelbergu (niem. *Ruprecht-Karls-Universität Heidelberg*) od kilku lat pracuje nad korpusem złożonym z niemieckich tekstów prawnych, wyposażonym w przyjazny dla legislatorów interfejs graficzny i umożliwiającym m.in. wyszukiwanie danych o częstotliwości użycia wprowadzanej frazy, a także śledzenie odnośników w tekstach prawnych czy proponowanie alternatywnych sformułowań[74].

Prawdziwie rewolucyjne narzędzia opracowują szwajcarscy uczeni z Uniwersytetu w Zurychu. Punkt wyjścia stanowi dla nich wspomniany powyżej trzyjęzyczny korpus prawa szwajcarskiego. Jest to korpus indeksowany, tzn. opatrzony różnego rodzaju informacjami lingwistycznymi, w tym na temat struktury tekstu i cech morfosyntaktycznych jego poszczególnych elementów. Ponadto korpus zawiera informacje odnoszące się do treści zgromadzonych w nim tekstów, przykładowo wyróżnia przepisy przejściowe, przepisy nowelizujące, preambuły itd.[75] Połączenie tych informacji pomaga tworzyć nowatorskie narzędzia informatyczne wspierające pracę legislatora.

Jednym z takich narzędzi jest program rozpoznający definicje legalne. Jak pokazuje praktyka, zadanie to może przysporzyć problemów nawet wykształconemu prawnikowi i nierzadko budzi kontrowersje w procesie stosowania prawa[76]. W prawie szwajcarskim, analogicznie do polskiego, występuje kilka rodzajów definicji legalnych, w tym zamieszczone w specjalnym „słowniczku" ustawowym, zamieszczone w tekście merytorycznym oraz definicje nawiasowe[77]. Każdy rodzaj definicji wymagał stworzenia odmiennego algorytmu, który kieruje się relewantnymi cechami morfosyntaktycznymi, umiejscowieniem w strukturze tekstu, słowami kluczowymi itd.[78] Skuteczność omawianego narzędzia

[74] F. Vogel, H. Hamann, I. Gauer, *op. cit.*, s. 1350.

[75] S. Höfler, K. Sugisaki, *op. cit.*, s. 176.

[76] A. Bielska-Brodziak, *op. cit.*, s. 44–46.

[77] S. Höfler, A. Bünzli, K. Sugisaki, *Detecting Legal Definitions for Automated Style Checking in Draft Laws*, „Technical Reports in Computational Linguistics" 2011, s. 4–7.

[78] *Ibidem*, s. 4–8.

wynosi – w zależności od rodzaju definicji – od 81 do 99%[79]. Automatyczne rozpoznawanie definicji legalnych może być wykorzystane na kilka sposobów. Sygnalizuje ono legislatorowi, że dane słowo (np. wprowadzone do projektu lub wpisane w wyszukiwarkę) ma definicję legalną i informuje o jej treści. Ponadto umożliwia zautomatyzowane wychwytywanie błędów w budowie definicji. Przykładowo w szwajcarskiej teorii legislacji przyjmuje się, że wyraz powinien zostać zdefiniowany, jeżeli występuje w danym tekście co najmniej trzykrotnie. Omawiany detektor automatycznie sprawdza zachodzenie tego warunku i informuje o nim legislatora[80].

Wątek ten prowadzi do kolejnego narzędzia informatycznego, którym jest zautomatyzowany detektor błędów legislacyjnych. Program ten rozpoznaje przypadki naruszenia zasad techniki prawodawczej, zarówno pochodzących z oficjalnego rządowego podręcznika, jak i z innych źródeł, np. prac naukowych czy dobrych praktyk legislacyjnych[81]. W fazie testowej posłużył on do badania niemieckojęzycznej wersji korpusu prawa szwajcarskiego, a docelowo ma wspomagać prace legislacyjne[82]. Sam proces modelowania błędów legislacyjnych stanowił największe wyzwanie badawcze. Szwajcarskie zasady techniki prawodawczej, podobnie jak ich polski odpowiednik, zawierają zarówno konkretne reguły w rodzaju „każdy punkt wyliczenia kończy się średnikiem", jak i bardziej abstrakcyjne zasady typu „jedno zdanie w sensie gramatycznym powinno wyrażać jedną myśl". Formalizacja tych pierwszych wymagała zazwyczaj odwołania się do cech strukturalnych lub morfosyntaktycznych tekstu i w tym sensie nie różniła się zasadniczo od analogicznych narzędzi wykorzystywanych przy tworzeniu dokumentacji technicznej. W przypadku tych drugich była konieczna bardziej zaawansowana analiza treściowa odwołująca się do cech charakterystycznych dla tekstów prawnych[83]. Po zidentyfikowaniu błędu program zaznacza wadliwy fragment tekstu i wyświetla treść naruszonej zasady techniki prawodawczej[84]. Warto nadmienić, że procedury identyfikacji błędów są również wykorzystywane do oceny jakości obowiązujących aktów prawnych. Uzyskane dzięki nim wyniki mogą przykładowo służyć do przewidywania sporów interpretacyjnych lub formułowania wniosków *de lege ferenda*[85].

Jeszcze innym narzędziem jest wielojęzyczna statystyczna konkordancja, która umożliwia porównanie ze sobą różnych tłumaczeń danego wyrażenia. Wykorzystuje się ją w translatoryce prawniczej, co w wielojęzycznej Szwajcarii – podobnie jak w UE – odgrywa zasadniczą rolę[86].

---

[79] *Ibidem*, s. 8.

[80] *Ibidem*, s. 8–9.

[81] S. Höfler, K. Sugisaki, *From Drafting Guideline to Error Detection: Automating Style Checking for Legislative Texts*, [w:] *Proceedings of the Second Workshop on Computational Linguistics and Writing (CL&W 2012): Linguistic and Cognitive Aspects of Document Creation and Document Engineering*, 2012, <https://aclanthology.org/W12-0302>, s. 15, dostęp 23 II 2022.

[82] *Ibidem*, s. 17.

[83] *Ibidem*, s. 9.

[84] S. Höfler, K. Sugisaki, *Constructing and Exploiting...*, s. 178.

[85] *Ibidem*, s. 179.

[86] *Ibidem*.

Omówione powyżej narzędzia informatyczne zostały opracowane na potrzeby obcego systemu prawnego i w odmiennych realiach językowych. Truizmem będzie stwierdzenie, że język niemiecki różni się od polskiego pod względem morfosyntaktycznym, a szwajcarskie zasady techniki prawodawczej jedynie w pewnym zakresie pokrywają się z rodzimymi ZTP. Tym niemniej ustalenia teoretyczne, jak również część rozwiązań technicznych zaproponowanych przez zagranicznych badaczy mają w dużej mierze wymiar uniwersalny[87]. Zdaniem autorów zautomatyzowane narzędzia rozpoznające definicje legalne, identyfikujące błędy redakcyjne czy badające jakość obowiązujących tekstów prawnych znalazłyby zastosowanie również w polskich warunkach. Podkreślenia wymaga zatem to, że – jak pokazują doświadczenia szwajcarskie i niemieckie – warunkiem koniecznym stworzenia tego typu narzędzi jest uprzednie zbudowanie korpusu specjalistycznego składającego się z tekstów obowiązujących aktów prawnych.

Warto na koniec zadać pytanie o to, czy funkcji takiego korpusu nie mogą pełnić istniejące prawnicze elektroniczne bazy danych. Obecnie takimi bazami dysponują zarówno organy publiczne (np. ISAP prowadzony przez Kancelarię Sejmu), jak i wydawnictwa komercyjne (Lex wydawany przez Wolters Kluwer, Legalis – C.H.Beck). Chociaż ich wejściu na rynek towarzyszyły ogromne kontrowersje[88], to dziś stanowią one podstawowe narzędzie pracy nie tylko w kancelariach i sądach, lecz także w urzędach czy przedsiębiorstwach[89]. Podobnie jak omawiane korpusy specjalistyczne bazy te zawierają m.in. teksty obowiązujących aktów prawnych i umożliwiają ich przeszukiwanie z uwzględnieniem przeróżnych kryteriów. Zarazem jednak cechują się istotnymi odmiennościami. Po pierwsze, są ograniczone wbudowanym interfejsem, co oznacza, że użytkownik może przeprowadzać w nich jedynie określony zakres predefiniowanych operacji[90]. Po drugie, zazwyczaj uniemożliwiają pobieranie i prezentowanie większych partii tekstu. Przykładowo, po wpisaniu żądanego wyrazu w wyszukiwarce wyświetli się lista dokumentów (np. aktów prawnych), w których wyraz ten występuje. Nie powstanie jednak konkordancja, która umożliwiłaby wygodne porównanie wszystkich wystąpień w najbliższym kontekście zdaniowym[91]. Po trzecie i najważniejsze, teksty zgromadzone w prawniczych bazach danych nie są indeksowane, a więc składające się na nie elementy nie mają przyporządkowanej informacji strukturalnej, morfosyntaktycznej itd.[92] Innymi słowy, program nie rozpoznaje danego wyrazu jako rzeczownika, czasownika czy przysłówka, podmiotu czy orzeczenia zdania, fragmentu określonego artykułu, rozdziału itd. Bez tego typu metalingwistycznych informacji stworzenie omówionych wyżej narzędzi

[87] S. Höfler, K. Sugisaki, *From Drafting...*, s. 10.

[88] N. Goldfarb, *A Lawyer's Introduction to Meaning in the Framework of Corpus Linguistics*, „Brigham Young University Law Review" 2017, nr 6, s. 1416.

[89] B. Patkowski, *Rozwój Systemów Informacji Prawnej oraz ich rola w kształtowaniu gospodarki opartej na wiedzy (na przykładzie Polski)*, „Prace Komisji Geografii Przemysłu Polskiego Towarzystwa Geograficznego" 2013, nr 21, s. 373–376.

[90] F. Vogel, H. Hamann, I. Gauer, *op. cit.*, s. 1351.

[91] H. Hamann, F. Vogel, *Evidence-Based Jurisprudence Meets Legal Linguistics – Unlikely Blends Made in Germany*, „Brigham Young University Law Review" 2017, nr 6, s. 1498–1499.

[92] F. Vogel, H. Hamann, I. Gauer, *op. cit.*, s. 1351.

automatycznego przetwarzania tekstu jest niemożliwe. Pomimo ogromnej użyteczności praktycznej i niewątpliwych podobieństw do specjalistycznych korpusów językowych istniejące prawnicze bazy danych nie mogą zatem pełnić ich funkcji. Mogą natomiast stanowić wygodne źródło pozyskiwania tekstów do budowy takich korpusów.

V. Podsumowanie

W ostatnich latach korpusy językowe są coraz częściej wykorzystywane do badania języka aktów prawnych (podejście europejskie) oraz do interpretacji prawa (podejście amerykańskie). W naszym przekonaniu mogą one stanowić również niezwykle użyteczne narzędzie w procesie tworzenia prawa, wprost przyczyniając się do realizacji określonych zasad techniki prawodawczej. Korpusy ogólne umożliwiają przede wszystkim realizację zasady komunikatywności tekstu prawnego, a w szczególności ustalenie, które znaczenie danego wyrazu należy uznać za podstawowe i powszechnie przyjęte. Co istotne, nie ma żadnych technicznych przeciwskazań, by już dziś weszły do arsenału legislatora. Barierą jest jedynie *know-how,* którego dostarczenie wymagać będzie współpracy przedstawicieli nauki prawa i językoznawstwa. Korpusy specjalistyczne, zbudowane z tekstów obowiązujących aktów prawnych, mogą się wydatnie przyczynić do realizacji zasady konsekwencji terminologicznej i dać legislatorowi wiarygodny i szczegółowy obraz funkcjonowania dowolnego wyrażenia ustawowego. Nadto mogą posłużyć do stworzenia zautomatyzowanych narzędzi wspierających pracę legislatora. Cel ten jawi się na chwilę obecną jako równie intrygujący, co odległy. Ani w Polsce, ani w większości krajów europejskich taki korpus nie istnieje. Jego budowa stanowi poważne wyzwanie teoretyczne i praktyczne wymagające współpracy prawników, językoznawców i informatyków. Autorzy mają nadzieję, że niniejsza publikacja to pierwszy, skromny, acz niezbędny krok w tym kierunku.

Bibliografia

Źródła

Corpus of Contemporary American English, <https://www.english-corpora.org/>.
Dziennik Ustaw Rzeczypospolitej Polskiej, <https://dziennikustaw.gov.pl/DU>.
Eurolect, <https://www.unint.eu/en/research/research-projects/33-page/490-eurolect-observatory-project.html>.
Google Books Ngram, <http://books.google.com/ngrams/>.
Internetowy System Aktów Prawnych, <https://isap.sejm.gov.pl/>.
Narodowy Korpus Języka Polskiego, <http://www.nkjp.pl/>, <http://nkjp.pl/poliqarp/nkjp300/query/>, <http://nkjp.uni.lodz.pl/collocations.jsp?qfield=reprezentacja**&dummystring=%C4%85%C4%84%C4%87%C4%86%C4%99%C4%98%C5%82%C5%81%C5%84%C5%83%C3%B3%C3%93%C5%9B%C5%9A%C5%BA%C5%B9%C5%BC%C5%BB>.

Strona internetowa Słownika Języka Polskiego PWN, <https://sjp.pwn.pl/doroszewski/reprezentacja;5489765.html>.
Uchwała Sądu Najwyższego z dnia 30 kwietnia 2003 r., I KZP 8/03, OSNKW 2003, nr 5–6, poz. 41.
Ustawa z dnia 16 lipca 1991 r. o podatku dochodowym od osób fizycznych, tekst jedn. Dz.U. z 2020 r. poz. 1426, ze zm.
Ustawa z dnia 15 lutego 1992 r. o podatku dochodowym od osób prawnych, tekst jedn. Dz.U. z 2020 r. poz. 1406, ze zm.
Ustawa z dnia 7 października 1999 r. o języku polskim, tekst jedn. Dz.U. z 2021 r. poz. 672.
Wyrok Naczelnego Sądu Administracyjnego z dnia 25 stycznia 2012 r., II FSK 1445/10, LEX nr 1102780.
Wyrok Sądu Apelacyjnego w Białymstoku z dnia 26 czerwca 2015 r., I ACa 230/15, LEX nr 1765926.
Wyrok Sądu Apelacyjnego w Łodzi z dnia 13 listopada 2014 r., I ACa 709/14, LEX nr 1554764.
[Wyrok Sądu Najwyższego Stanów Zjednoczonych z dnia 1 marca 2011 r. w sprawie FCC *vs.*AT&T, Inc. 562 U.S. [2011]].
[Wyrok Sądu Najwyższego Stanów Zjednoczonych z dnia 21 czerwca 2018 r. w sprawie Lucia *vs.* SEC 138 U.S. [2018]].
[Wyrok Sądu Najwyższego Stanów Zjednoczonych z dnia 26 czerwca 2018 r. w sprawie Carpenter *vs.* United States 585 U.S. [2018]].
[Wyrok Sądu Najwyższego stanu Utah z dnia 14 sierpnia 2015 r. w sprawie State *vs.* Rasabout 356 P.3d 1258, 1260 [Utah 2015]].
Wyrok Sądu Okręgowego w Warszawie z dnia 22 sierpnia 2018 r., X Ka 721/18, LEX nr 2566284.
Wyrok Wojewódzkiego Sądu Administracyjnego w Warszawie z dnia 19 marca 2019 r., VI SA/Wa 1895/18, LEX nr 2696759.
Wyrok Wojewódzkiego Sądu Administracyjnego w Warszawie z dnia 29 czerwca 2020 r., III SA/Wa 2042/19, LEX nr 3064883.
Załącznik do rozporządzenia Prezesa Rady Ministrów z dnia 20 czerwca 2002 r. w sprawie „Zasad techniki prawodawczej", tekst jedn. Dz.U. z 2016 r. poz. 283.

Piśmiennictwo

Biel Ł., *Corpus-Based Studies of Legal Language for Translation Purposes: Methodological and Practical Potential*, [w:] *Reconceptualizing LSP. Online Proceedings of the XVII European LSP Symposium 2009*, red. C. Heine, J. Engberg, Aarhus School of Business, Aarhus University, Aarhus 2010.
Biel Ł., *Lost in the Eurofog: the Textual Fit of Translated Law*, Peter Lang, Frankfurt 2014.
Bielska-Brodziak A., *Interpretacja tekstu prawnego na podstawie orzecznictwa podatkowego*, Wolters Kluwer, Warszawa 2009.
Błachut M., Kaczor J., *Typy uzasadnień obowiązywania zasad techniki prawodawczej w polskim porządku prawnym*, [w:] *Folia Iuridica Wratislaviensis*, red. L. Górnicki, Prawnicza i Ekonomiczna Biblioteka Cyfrowa, Wrocław 2012.
Choduń A., *Aspekty językowe derywacyjnej koncepcji wykładni prawa*, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin 2018.
Degórski Ł., Przepiórkowski A., *Ręcznie znakowany milionowy podkorpus NKJP*, [w:] *Narodowy Korpus Języka Polskiego*, red. A. Przepiórkowski, M. Bańko, R. Górski, B. Lewandowska-Tomaszczyk, Wydawnictwo Naukowe PWN, Warszawa 2012.
Dębska M., *Zasady techniki prawodawczej. Komentarz*, Lexis Nexis, Warszawa 2013.

Ehrett J., *Against Corpus Linguistics*, „The Georgetown Law Journal Online" 2017, nr 108.
Filipczyk H., *Koszty reprezentacji w rozumieniu art. 23 ust. 1 pkt 23 ustawy z 26.07.1991 r. o podatku dochodowym od osób fizycznych. Glosa do wyroku NSA z dnia 25 stycznia 2012 r., II FSK 1445/10*, „Glosa" 2012, nr 4.
Goldfarb N., *A Lawyer's Introduction to Meaning in the Framework of Corpus Linguistics*, „Brigham Young University Law Review" 2017, nr 6, <https://doi.org/10.2139/ssrn.2907485>.
Grabowski A., *Siła argumentu interpretacyjnego z prawniczego domniemania języka (znaczenia) potocznego*, „Zagadnienia Naukoznawstwa" 2016, nr 3(209), <https://doi.org/10.24425/118016>.
Hamann H., Vogel F., *Evidence-Based Jurisprudence Meets Legal Linguistics – Unlikely Blends Made in Germany*, „Brigham Young University Law Review" 2017, nr 6.
Höfler S., Bünzli A., Sugisaki K., *Detecting Legal Definitions for Automated Style Checking in Draft Laws*, „Technical Reports in Computational Linguistics" 2011, <https://doi.org/10.5167/uzh-61648>.
Höfler S., Sugisaki K., *Constructing and Exploiting an Automatically Annotated Resource of Legislative Texts*, „Zurich Open Repository and Archive" 2014, <https://doi.org/10.5167/UZH-96172>.
Höfler S., Sugisaki K., *From Drafting Guideline to Error Detection: Automating Style Checking for Legislative Texts*, [w:] *In Proceedings of the Second Workshop on Computational Linguistics and Writing (CL&W 2012): Linguistic and Cognitive Aspects of Document Creation and Document Engineering*, 2012, <https://aclanthology.org/W12-0302>.
Iluk Ł., Iluk J., *Szwajcarski sposób na jasność i komunikatywność aktów normatywnych na przykładzie procedury redagowania konstytucji kantonu Zurych z 2005 r.*, „Państwo i Prawo" 2020, nr 3.
Kida I., *Introduction to Corpus Linguistics*, „Linguistica Silesiana" 2013, nr 34.
Kubicka E., Grec K., Leszczak W., Mitrew M., *Wykorzystanie słowników w interpretacji prawniczej – implikacje praktyczne*, „Kwartalnik Sądowy Apelacji Gdańskiej" 2015, nr 1.
Lewandowska-Tomaszczyk B., *Podstawy językoznawstwa korpusowego*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2005.
Lewandowska-Tomaszczyk B., Bańko M., Górski R., Łaziński M., Pęzik P., Przepiórkowski A., *Narodowy Korpus Języka Polskiego: geneza i dzień dzisiejszy*, [w:] *Narodowy Korpus Języka Polskiego*, red. A. Przepiórkowski, M. Bańko, R. Górski, B. Lewandowska-Tomaszczyk, Wydawnictwo Naukowe PWN, Warszawa 2012.
Malinowski A., *Pomiar czytelności polskiego tekstu prawnego*, „Państwo i Prawo" 2017, nr 5.
Malinowski A., *Rozdział X. Jednoznaczność i precyzja słów w tekście prawnym*, [w:] *Zarys metodyki pracy legislatora. Ustawy – akty wykonawcze – prawo miejscowe*, red. A. Malinowski, Lexis Nexis, Warszawa 2009.
Mouritsen S., *Hard Cases and Hard Data: Assessing Corpus Linguistics as an Empirical Path to Plain Meaning*, „Columbia Science and Technology Law Review" 2011, nr 13.
Patkowski B., *Rozwój Systemów Informacji Prawnej oraz ich rola w kształtowaniu gospodarki opartej na wiedzy (na przykładzie Polski)*, „Prace Komisji Geografii Przemysłu Polskiego Towarzystwa Geograficznego" 2013, nr 21.
Prieto Ramos F., *The Use of Corpora in Legal and Institutional Translation Studies: Directions and Applications*, „Translation Spaces" 2019, nr 8, <https://doi.org/10.1075/ts.00010.pri>.
Slocum B., *Ordinary Meaning: a Theory of the Most Fundamental Principle of Legal Interpretation*, University of Chicago Press, Chicago–London 2016.
Solan L., *Can Corpus Linguistics Help Make Originalism Scientific?*, „The Yale Law Journal Forum" 2016, nr 126.

Solan L., *The Language of Statutes*, University of Chicago Press, Chicago–London 2010.
Solan L., Gales T., *Corpus Linguistics as a Tool in Legal Interpretation*, „Brigham Young University Law Review" 2017, nr 6.
Solan L., Rosenblatt T., Osherson D., *False Consensus Bias in Contract Interpretation*, „Columbia Law Review" 2008, nr 5(108).
Steinberger R., Ebrahim M., Poulis A., Carrasco-Benitez M., Schlüter P., Przybyszewski M., Gilbro S., *An Overview of the European Union's Highly Multilingual Parallel Corpora*, „Language Resources & Evaluation" 2014, nr 48, <https://doi.org/10.1007/s10579-014-9277-0>.
Steinberger R., Pouliquen B., Widiger A., Ignat C., Erjavec T., Tufis D., Varga D., *The JRC-Acquis: a Multilingual Aligned Parallel Corpus with 20+ Languages*, [w:] *Conference Proceedings: the 5th International Conference on Language Resources and Evaluation – Proceedings, Paris (France)*, European Language Resources Association (ELRA), Paris 2006, <https://publications.jrc.ec.europa.eu/repository/handle/JRC32786>.
Szafrański D., *Komentarz do § 8*, [w:] *Zasady techniki prawodawczej w zakresie aktów prawa miejscowego. Komentarz praktyczny z wzorami oraz przykładami*, red. D. Szafrański, W. Białończyk, A. Bielecki, Ł. Kasiak, J. Piecha, C.H.Beck, Warszawa 2016, <https://sip-1legalis-1pl-1lp6svawk0089.han.bg.us.edu.pl/document-view.seam?documentId=mjxw62zogiztkmjtgi2dqmjoobqxalrtgqytgmzvge4q#tabs-metrical-info>.
Tobor Z., Zeifert M., *Korpusy językowe jako narzędzie interpretacji prawa. Amerykańska teoria i praktyka*, „Archiwum Filozofii Prawa i Filozofii Społecznej" 2020, nr 4(25), <https://doi.org/10.36280/AFPiFS.2020.4.80>.
Uniwersalny słownik języka polskiego, t. 3, red. S. Dubisz, Wydawnictwo Naukowe PWN, Warszawa 2008.
Vogel F., Hamann H., Gauer I., *Computer-Assisted Legal Linguistics: Corpus Analysis as a New Tool for Legal Studies*, „Law & Social Inquiry" 2018, nr 43/4, <https://doi.org/10.1111/lsi.12305>.
Wierczyński G., *Redagowanie i ogłaszanie aktów normatywnych. Komentarz*, Wolters Kluwer, Warszawa 2016.
Wojtczak S., *Zasady techniki prawodawczej – luka w prawie?*, „Państwo i Prawo" 2005, nr 1.
Wronkowska S., *Zasady techniki prawodawczej, czy rzeczywiście luka w prawie?*, „Państwo i Prawo" 2005, nr 2.
Zdunek U., *Rola kryterium frekwencyjnego w ocenie poprawności leksykalnej w słownikach poprawnej polszczyzny*, niepublikowana rozprawa doktorska obroniona na UW, Warszawa 2013, <https://depotuw.ceon.pl/bitstream/handle/item/626/Doktorat%20ca%C5%82o%C5%9B%C4%875.pdf?sequence=1>.
Zeifert M., Tobor Z., *Linguam novit curia? Analiza korpusowa w prawie karnym – przyczynek do dyskusji*, „Forum Polityki Kryminalnej" 2021, nr 1, <https://doi.org/10.31261/FPK.2021.01.04>.
Żmigrodzki P., *Wprowadzenie do leksykografii polskiej*, Wydawnictwo Uniwersytetu Śląskiego, Katowice 2009.