

Zygmunt Tobor<sup>1</sup>, Mateusz Zeifert<sup>2</sup>

Uniwersytet Śląski w Katowicach

# Korpusy językowe jako narzędzie interpretacji prawa. Amerykańska teoria i praktyka

## 1. Wprowadzenie

Interpretacja prawa to fascynujące i niezwykle złożone zjawisko. Stulecia praktyki prawniczej i dyskursu naukowego doprowadziły do wytworzenia pewnych zasad (zwanymi też *kanonami*, *maksymami*, *regułami*, *dyrektywami* itp.), które mają kierować działalnością interpretatora. Przychylamy się do poglądu, że zasady te mają wartość raczej heurystyczną niż normatywną i najlepiej postrzegać je jako „narzędzia”<sup>3</sup>, które „istnieją po to, by wyjaśniać i pomagać, a nie ograniczać czy hamować”<sup>4</sup>. Jednym z takich narzędzi, które stało się dostępne dopiero niedawno dzięki postępowi technologicznemu, są korpusy językowe. W ostatnich kilku latach korpusy zaczęły być wykorzystywane przez amerykańskie sądy, w tym Sąd Najwyższy Stanów Zjednoczonych. Wywołało to gwałtowną reakcję przedstawicieli amerykańskiej teorii prawa i doktryny prawniczej. Od 2016 r. w amerykańskiej literaturze obserwujemy lawinowy wzrost liczby publikacji poświęconych temu zagadnieniu. Znajdziemy wśród nich m.in. reinterpretacje klasycznych problemów interpretacyjnych z wykorzystaniem korpusów, teoretyczne analizy stosowania korpusów w wykładni prawa, praktyczne wskazówki dla prawników sięgających po korpusy oraz liczne argumenty za i przeciw. Pojawiają się nawet głosy o powstaniu nowej dyscypliny naukowej: *law and corpus linguistics* (LCL)<sup>5</sup>. W literaturze polskiej użycie korpusów jako narzędzia wykładni prawa nie było jak dotąd szerzej omawiane<sup>6</sup>. Niniejszy artykuł ma zatem przede wszystkim charakter sprawozdawczy. Omawiamy w nim użyteczność korpusów językowych dla celów wykładni prawa

<sup>1</sup> Numer ORCID: 0000-0001-6541-8502. Adres e-mail: [zygmunt.tobor@us.edu.pl](mailto:zygmunt.tobor@us.edu.pl)

<sup>2</sup> Numer ORCID: 0000-0003-3266-3397. Adres e-mail: [mateusz.zeifert@us.edu.pl](mailto:mateusz.zeifert@us.edu.pl)

<sup>3</sup> M. Mullins, *Tools, Not Rules: The Heuristic Nature of Statutory Interpretation*, „Journal of Legislation” 2004/1, s. 3 i n.

<sup>4</sup> D. Lowe, Ch. Potter, *Understanding Legislation: A Practical Guide to Statutory Interpretation*, Oksford–Londyn 2018, s. 30.

<sup>5</sup> J. Phillips, J. Egbert, *Advancing Law and Corpus Linguistics: Importing Principles and Practices from Survey and Content-Analysis Methodologies to Improve Corpus Design and Analysis*, „Brigham Young University Law Review” 2017/6, s. 1589.

<sup>6</sup> Jednym z nielicznych wyjątków jest praca: E. Kubicka, K. Grec, W. Leszczak, M. Mitew, *Wykorzystanie słowników w interpretacji prawniczej – implikacje praktyczne*, „Kwartalnik Sądowy Apelacji Gdańskiej” 2015/1, s. 25–50.

w oparciu na doświadczeniach amerykańskich. Pomijamy przy tym inne potencjalne zastosowania korpusów w kontekście prawniczym<sup>7</sup>.

## 2. Charakterystyka językoznawstwa korpusowego

Językoznawstwo korpusowe jest częścią językoznawstwa komputerowego (*computational linguistics*). Zajmuje się analizą danych językowych zgromadzonych w korpusach, tj. komputerowych zbiorach tekstów. Jest wyrazem funkcjonalnego (a więc nie formalnego) podejścia do badania języka, a teoretycznej podbudowy dostarcza mu obecnie m.in. socjolingwistyka i językoznawstwo kognitywne. Początki językoznawstwa korpusowego sięgają lat 50. XX w. i wiążą się z amerykańską odmianą strukturalizmu zwaną dystrybucjonizmem. Zasadnicza krytyka ze strony największego ówczesnego autorytetu językoznawczego, Noama Chomsky'ego spowodowała jednak, że metodę tę szybko zarzucono. Ponownie zaczęto do niej sięgać dopiero w latach 70. i 80. XX w., kiedy epoka hegemonii N. Chomsky'ego dobiegła końca. Od lat 90. XX w. językoznawstwo korpusowe przeżywa gwałtowny rozwój, związany z coraz łatwiejszym dostępem do komputerów.

Korpus tworzy się z autentycznych tekstów języka (jednego lub kilku). Jego istotną cechą jest wielkość – korpus powinien zawierać znaczną liczbę słów, liczoną w milionach. Bardziej precyzyjne określenie tej liczby jest niemożliwe ze względu na stały postęp technologiczny<sup>8</sup>. Kolejną postulowaną cechą jest reprezentatywność. Korpus powinien jak najlepiej odzwierciedlać dany język. Oznacza to m.in., że powinien zawierać – w odpowiednich proporcjach – teksty pochodzące z różnych rejestrów językowych. Nie wyklucza to możliwości tworzenia korpusów specjalistycznych, zawierających jedynie teksty należące do określonego rejestru, dialektu, specjalistycznej odmiany języka itd. W praktyce reprezentatywność jest trudna do uzyskania. Przykładowo, względy techniczne powodują, że w większości korpusów nadreprezentowane są teksty pisane, a niedoreprezentowane – teksty mówione. Wreszcie, korpus musi posiadać formę elektroniczną. Warto tu wspomnieć o istotnym rozróżnieniu na korpusy „surowe” (*raw*), inaczej: nieindeksowane lub nieanotowane oraz korpusy indeksowane lub anotowane (*tagged*). Te pierwsze zawierają wyłącznie elektroniczny zapis tekstów. Te drugie zawierają także dodatkowe informacje, w tym dane lingwistyczne, np. przypisane do każdego słowa oznaczenie części mowy lub części zdania (tzw. indeksowanie morfologiczno-syntaktyczne), transkrypcję fonetyczną czy tzw. lematyzację<sup>9</sup>. Korpusy indeksowane oferują bez porównania więcej możliwości badawczych, ale ich tworzenie jest dużo bardziej pracochłonne (mimo częściowej automatyzacji, wynikającej z wykorzystywania w tym celu specjalistycznego oprogramowania).

<sup>7</sup> Korpusy bywają wykorzystywane m.in. do ustalenia znaczenia inkryminowanej wypowiedzi (sprawy o zniesławienie itp.), do ustalenia waloru oryginalności określonej nazwy (sprawy z zakresu prawa własności intelektualnej itd.) czy do identyfikacji autora tekstu (ekspertyzy kryminalistyczne). Por. B. Lewandowska-Tomaszczyk (red.), *Podstawy językoznawstwa korpusowego*, Łódź 2005, s. 270–279; S. Mouritsen, *Corpus Linguistics in Legal Interpretation: An Evolving Interpretative Framework*, „International Journal of Language & Law” 2017/6, s. 69–70. Językoznawstwo korpusowe jest obecnie wiodącą metodą wykorzystywaną w legilingwistyce i translatoryce prawniczej. Por. przykładowo, międzynarodowy projekt badawczy *Eurolect Observatory Project*, <https://www.unint.eu/en/research/research-projects/33-page/490-eurolect-observatory-project.html>, dostęp: 3.01.2020 r. Te niezwykle interesujące kwestie zostaną w niniejszym opracowaniu pominięte.

<sup>8</sup> Przykładowo popularny anglojęzyczny *Corpus of Contemporary American English* (<https://www.english-corpora.org/>) zawiera 560 mln słów, a Narodowy Korpus Języka Polskiego (<http://www.nkjp.pl/>) ponad 1,5 mld słów. Największym istniejącym korpusem jest *Google Books Ngram* (<http://books.google.com/ngrams>), który zawiera ok. 155 mld słów.

<sup>9</sup> Lematyzacja, inaczej hasłowanie, oznacza proces redukcji różnych form danego wyrazu do jego formy podstawowej (tzw. lemma), np. mianownika liczby pojedynczej dla rzeczowników, bezokolicznika dla czasowników, por. B. Lewandowska-Tomaszczyk (red.), *Podstawy...*, s. 35–36.

Wielu badaczy podnosi też problem pewnej arbitralności i nieuniknionych błędów w procesie indeksowania tekstów zawartych w korpusie<sup>10</sup>.

Wykorzystanie korpusów dla celów badawczych możliwe jest dzięki różnym narzędziom, procedurom i algorytmom informatycznym takim jak: wyszukiwanie interaktywne, listy frekwencyjne, konkordancje i kolokacje<sup>11</sup>. Wyszukiwanie jest podstawowym narzędziem każdego użytkownika współczesnych technologii informatycznych, przy czym w przypadku korpusu może przyjmować bardzo zaawansowane postacie. Listy frekwencyjne określają częstotliwość występowania wyrazów w danym korpusie, a zatem – przy założeniu reprezentatywności korpusu – w danym języku. Wykorzystuje się je m.in. do tworzenia słowników frekwencyjnych i do nauki języków obcych. Konkordancje to zbiory przykładów użycia danego wyrazu w kontekście, zgodnie z wprowadzonymi przez badacza kryteriami. Zazwyczaj prezentuje się je w formie KWIC (*key words in context*), tzn. poszukiwany wyraz wyświetlany jest 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. Konkordancje są wykorzystywane m.in. przy tworzeniu słowników, tezaurusów i glosariuszy, gdyż doskonale zastępują tradycyjne fiszki. Kolokacje to stale powtarzające się kombinacje słów występujące w swoim sąsiedztwie. Ich uwzględnienie pozwala m.in. określić, w jakim kontekście dany wyraz występuje najczęściej lub wyszukać przykłady użycia wyrazu w określonym kontekście. Wyszukiwanie, listy frekwencyjne, konkordancje i kolokacje to standardowe narzędzia uwzględnione w interfejsie większości powszechnie dostępnych korpusów<sup>12</sup>. Lingwiści do swoich analiz wykorzystują również bardziej zaawansowane środki, w tym specjalistyczne aplikacje, języki programowania, testy statystyczne, bazy danych itd.<sup>13</sup>

### 3. Korpusy w amerykańskim orzecznictwie

Podstawową zasadą amerykańskiej praktyki interpretacji prawa jest *ordinary meaning rule*, stanowiąca w przybliżeniu połączenie zasady pierwszeństwa wykładni językowej z zasadą domniemania języka potocznego<sup>14</sup>. Pojęcie *ordinary meaning* („zwykłe znaczenie”<sup>15</sup>) jest różnie rozumiane. Jak ironicznie zauważają autorzy: „nie znamy zwykłego znaczenia «zwykłego znaczenia»”<sup>16</sup>. W ujęciu intencjonalistycznym stanowi ono zakres dopuszczalnych znaczeń danego wyrażenia językowego. Po ustaleniu takiego zakresu należy posłużyć się metodami pozajęzykowymi w celu sprecyzowania, które

<sup>10</sup> B. Lewandowska-Tomaszczyk (red.), *Podstawy...*, s. 32–33, 61–94.

<sup>11</sup> B. Lewandowska-Tomaszczyk (red.), *Podstawy...*, s. 33–41.

<sup>12</sup> Por. np. Narodowy Korpus Języka Polskiego (wyszukiwarka PELCRA).

<sup>13</sup> B. Lewandowska-Tomaszczyk (red.), *Podstawy...*, s. 95–131.

<sup>14</sup> „Absent some reason for deviation, such as words with technical or special legal meanings, the language used in legal texts should be viewed as corresponding with the language used in nonlegal communications”. S. Gries, B. Slocum, *Ordinary Meaning and Corpus Linguistics*, „Brigham Young University Law Review” 2017/6, s. 1419. Literatura na ten temat jest ogromna. Por. przykładowo: B. Slocum, *Ordinary Meaning: A Theory of the Most Fundamental Principle of Legal Interpretation*, Chicago 2015; T. Lee, S. Mouritsen, *Judging Ordinary Meaning*, „Yale Law Journal” 2018/4, s. 813 i n.; L. Solan, *The Language of Statutes: Laws and Their Interpretation*, Chicago 2010; A. Scalia, B. Garner, *Reading Law: The Interpretation of Legal Texts*, St. Paul (MN) 2012. Por. także w języku polskim A. Grabowski, *Siła argumentu interpretacyjnego z prawniczego domniemania języka (znaczenia) potocznego*, „Zagadnienia Naukoznawstwa” 2016/3, s. 357 i n.

<sup>15</sup> W niniejszym artykule *ordinary meaning* tłumaczymy jako „zwykłe znaczenie” i celowo rezygnujemy z terminów stosowanych w polskim piśmiennictwie („znaczenie potoczne”, „znaczenie powszechne” itd.), aby uniknąć ewentualnych nieporozumień.

<sup>16</sup> T. Lee, S. Mouritsen, *Judging Ordinary...*, s. 798.

z dopuszczalnych znaczeń odpowiada intencji prawodawcy. W ujęciu tekstualistycznym zwykle znaczenie to sposób, w jaki dane wyrażenie najczęściej jest rozumiane. Ustalenie go jest celem interpretatora i – pomijając szczególne przypadki (np. absurdalne rezultaty) – ma charakter wiążący. A zatem w obu dominujących teoriach interpretacyjnych zwykle znaczenie pełni znaczącą – choć różną – funkcję<sup>17</sup>. Poważne kontrowersje wiążą się z pytaniem, w jaki sposób należy ustalać zwykle znaczenie tekstu prawnego. Do niedawna jedyną alternatywą dla sędziowskiej intuicji były słowniki. Użycie słowników w interpretacji prawa doczekało się ogromnej literatury, zawierającej liczne głosy krytyczne, dotyczące zarówno samej natury słowników, jak również praktyki ich wykorzystywania<sup>18</sup>. Warto odnotować, że od kilkunastu lat dyskusja na temat słowników toczy się również w polskiej literaturze<sup>19</sup>.

Słowniki jako narzędzie do ustalania zwykłego znaczenia tekstu prawnego zyskały jednak ostatnio poważnego konkurenta w postaci korpusów językowych. Według niektórych autorów symboliczną datą jest rok 1998, kiedy to sędzia Stephen Breyer w sprawie *Muscarello v. United States*<sup>20</sup> przeprowadził analizę dwóch elektronicznych baz danych, zawierających teksty publicystyczne w celu ustalenia zwykłego znaczenia wyrazu *to carry* („nosić”, „nieść”). Analiza doprowadziła większość orzekającą do kontrowersyjnego wniosku, że takim znaczeniem jest „przewozić w samochodzie”<sup>21</sup>. Innym pionierskim orzeczeniem był wyrok w sprawie *United States v. Costello*<sup>22</sup> z 2012 r., w którym sędzia Richard Posner badał znaczenie wyrazu *to harbor* („udzielać schronienia”, „ukrywać”) za pomocą wyszukiwarki Google, wpisując różne frazy zawierające ten wyraz. Na podstawie tej analizy uznał, że zwykle znaczenie *to harbor* odnosi się do ukrywania pewnych osób ze względu na ich przynależność do określonej kategorii zasługującej na ochronę, a nie po prostu do ukrywania się<sup>23</sup>. We wskazanych orzeczeniach sądy nie korzystały z korpusów językowych w znaczeniu opisanym powyżej, lecz z innych elektronicznych baz danych. Przyjęte przez ich autorów metodologie budzą poważne zastrzeżenia, z których część zostanie omówiona poniżej<sup>24</sup>.

Początkowo wykorzystanie korpusów językowych *sensu stricto* było domeną sądów stanowych. W sprawie *In re Adoption of Baby E.Z.*<sup>25</sup> sędzia Thomas Lee przeprowadził rozległą analizę popularnego korpusu *Corpus of Contemporary American English* (COCA) w celu ustalenia, czy wyraz *custody* („opieka”, „piecza”, „kuratela”) odnosi się

<sup>17</sup> S. Gries, B. Slocum, *Ordinary Meaning...*, s. 1422–1441; S. Mouritsen, *Hard Cases and Hard Data Assessing Corpus Linguistics as an Empirical Path to Plain Meaning*, „Columbia Science and Technology Law Review” 2011/1, s. 163–179. Więcej na temat intencjonalizmu i tekstualizmu w literaturze polskiej zob. B. Brzeziński, *Wykładnia prawa podatkowego i orzekanie w sprawach podatkowych w państwach anglosaskich*, „Kwartalnik Prawa Podatkowego” 2004/4, s. 21–50; B. Brzeziński, *Współczesne amerykańskie teorie wykładni prawa*, „Państwo i Prawo” 2006/7, s. 22–39; B. Brzeziński, *Zasady wykładni prawa podatkowego w krajach anglosaskich*, Warszawa 2007; Z. Tobor, *W poszukiwaniu intencji prawodawcy*, Warszawa 2013.

<sup>18</sup> Por. przykładowo: J. Kirchmeier, S. Thumma, *The Lexicon Has Become a Fortress: The United States Supreme Court's Use of Dictionaries*, „Buffalo Law Review” 1999/47, s. 227–302; Ph. Rubin, *War of the Words: How Courts Can Use Dictionaries in Accordance with Textualist Principles*, „Duke Law Journal” 2013/60, s. 167–206; E. Aprill, *The Law of the Word: Dictionary Shopping in the Supreme Court*, „Arizona State Law Journal” 1998/30, s. 277–336.

<sup>19</sup> Por. np. M. Zieliński, *Wykładnia prawa. Zasady, reguły, wskazówki*, Warszawa 2012, s. 274–289; Z. Tobor, P. Żmigrodzki, A. Bielska-Brodziak, *Co każdy prawnik o słownikach wiedzieć powinien*, „Przegląd Sądowy” 2008/7–8, s. 79–95; A. Bielska-Brodziak, Z. Tobor, *Słowniki a interpretacja tekstów prawnych*, „Państwo i Prawo” 2007/5, s. 20–33.

<sup>20</sup> 524 U.S. [1998].

<sup>21</sup> S. Mouritsen, *The Dictionary Is Not a Fortress: Definitional Fallacies and a Corpus-Based Approach to Plain Meaning*, „Brigham Young University Law Review” 2010/5, s. 1946–1948.

<sup>22</sup> 666 F.3d 1040 [7th Circuit 2012].

<sup>23</sup> S. Gries, B. Slocum, *Ordinary Meaning...*, s. 1442–1453.

<sup>24</sup> Por. pkt 4. niniejszego artykułu.

<sup>25</sup> 266 P.3d 702, 724–25 [Utah 2011].

w swoim zwykłym znaczeniu do adopcji dziecka. Analiza wykazała, że dziesięciokrotnie częściej używany bywa w kontekście rozwodu, a zatem „opieka nad adoptowanym dzieckiem” nie stanowi jego zwykłego znaczenia. W późniejszej sprawie *State v. Rasabout*<sup>26</sup> ten sam sędzia odwołał się do korpusu, rozstrzygając że *to discharge* („wyrzelić”) oznacza zazwyczaj wyrzelenie pojedynczego pocisku, a nie opróżnienie całego magazynku. W sprawie *People v. Harris*<sup>27</sup> problem interpretacyjny dotyczył tego, czy wyraz *information* („informacja”) obejmuje także fałszywe lub niekompletne oświadczenia. Analizy korpusowe przeprowadzili zarówno większość orzekająca, jak i autor zdania odrębnego, jednak całkowicie odmiennie zinterpretowali ich wyniki. Większość wskazała na występujące w korpusie przykłady wyrażen typu *false information* („fałszywa informacja”) czy *inaccurate information* („nie dokładna informacja”). Autor zdania odrębnego wskazał z kolei – na podstawie tego samego korpusu – że wyraz *information* najczęściej występuje bez jakichkolwiek przymiotników i oznacza wówczas – co do zasady – informację prawdziwą<sup>28</sup>.

W federalnym Sądzie Najwyższym językoznawstwo korpusowe początkowo było wykorzystywane jedynie pośrednio, poprzez odwołania do prac naukowych zawierających analizy korpusowe. W sprawie *FCC v. AT&T, Inc.*<sup>29</sup> sędzia Ruth Ginsburg w ustnych motywach powołała się na opracowanie Neala Goldfarba, dotyczące znaczenia przymiotnika *personal* („osobisty”) w kontekście osób prawnych<sup>30</sup>. Natomiast w sprawie *Lucia v. SEC*<sup>31</sup> sędzia Clarence Thomas powołał się na analizę Jennifer Mascott, dotyczącą znaczenia problematycznego wyrażenia „*Officers of the United States*”<sup>32</sup>. Ponieważ problem dotyczył tekstu amerykańskiej konstytucji, analiza przeprowadzona została w oparciu na korpusie złożonym z tekstów „ojców-założycieli” (*the Framers*) z lat 1783–1789, co pozwoliło ustalić historyczne znaczenie tego zwrotu. W zdaniu odrębnym do wyroku w sprawie *Carpenter v. United States*<sup>33</sup> sędzia C. Thomas poszedł krok dalej i przeprowadził własną analizę aż sześciu korpusów zawierających historyczne teksty. Jej celem było wykazanie, że w czasach formułowania konstytucji nie znano pojęcia *reasonable expectation of privacy* („uzasadnione roszczenie do prywatności”), które posłużyło większości orzekającej za podstawę rozstrzygnięcia<sup>34</sup>.

#### 4. Korpusy w amerykańskiej teorii prawa

Pojawienie się analiz korpusowych w uzasadnieniach orzeczeń sądowych wywołało ożywioną reakcję przedstawicieli amerykańskiej teorii prawa. Pierwsze propozycje wykorzystania korpusów do ustalania zwykłego znaczenia tekstów ustawowych wysuwano już dekadę temu<sup>35</sup>, jednak zainteresowanie tym tematem wzrosło gwałtownie dopiero

<sup>26</sup> 356 P.3d 1258, 1260 [Utah 2015].

<sup>27</sup> 885 N.W.2d 832, 838–839 [Michigan 2016].

<sup>28</sup> Por. L. Solan, T. Gales, *Corpus Linguistics as a Tool in Legal Interpretation*, „Brigham Young University Law Review” 2017/6, s. 1344–1349.

<sup>29</sup> 562 U.S. [2011].

<sup>30</sup> D. Tankersley, *Beyond the Dictionary: Why Sua Sponte Judicial Use of Corpus Linguistics Is Not Appropriate for Statutory Interpretation*, „Mississippi Law Journal” 2018/4, s. 642, przypis 6.

<sup>31</sup> 138 U.S. [2018].

<sup>32</sup> J. Mascott, *Who Are “Officers of the United States”?*, „Stanford Law Review” 2018/70, s. 443–564.

<sup>33</sup> 585 U.S. [2018].

<sup>34</sup> Zdanie odrębne sędziego C. Thomasa do orzeczenia *Carpenter v. United States*, s. 7, [https://www.supremecourt.gov/opinions/17pdf/16-402\\_h315.pdf](https://www.supremecourt.gov/opinions/17pdf/16-402_h315.pdf), dostęp: 17.12.2019 r. Można dodać, że analiza ta miała postać całkiem rudymenarną, tzn. sprowadzała się do przeszukania korpusów pod kątem występowania określonej frazy z wynikiem negatywnym.

<sup>35</sup> Por. L. Solan, *The New Textualists’ New Text*, „Loyola of Los Angeles Law Review” 2005/38, s. 2059 i n.; S. Mouritsen, *The Dictionary...*

w ostatnich kilku latach. W latach 2016–2019 opublikowano kilkadziesiąt artykułów naukowych poświęconych tej kwestii. Wielu autorów wyraża przekonanie, że językoznawstwo korpusowe wnosi do dyskursu interpretacyjnego nieoceniony pierwiastek empiryczny: „metody korpusowe przenoszą proces ustalania zwykłego znaczenia z «czarnej skrzynki» wewnętrznych wrażeń sędziego na światło rzeczywistości, czyniąc dyskusję na temat zwykłego znaczenia namacalną i wymierną”<sup>36</sup>. Sprawia to, że interpretacja prawa staje się przedsięwzięciem bardziej naukowym, przewidywalnym, weryfikowalnym i obiektywnym<sup>37</sup>.

Zalety korpusów omawia się najczęściej na tle innych źródeł wiedzy o tzw. zwykłym znaczeniu. Po pierwsze, korpusy przeciwstawia się intuicji językowej sędziego. Intuicja językowa (kompetencja językowa) sprawdza się na ogół znakomicie, gdy chodzi o ocenę poprawności gramatycznej zdania, czy dopuszczalności semantycznej użycia danego wyrazu, jednak zawodzi, gdy w grę wchodzi ocena częstotliwości określonego użycia<sup>38</sup>. Sędzia, jak każdy użytkownik języka naturalnego, posługuje się właściwym tylko dla siebie językiem osobniczym (idiolektem), który podświadomie czyni punktem odniesienia takiej oceny. Ponadto, skłonny jest dostrzegać przypadki szczególne, a nie dostrzegać przypadków typowych. Najpoważniejszym zastrzeżeniem wobec intuicji sędziego, jako źródła zwykłego znaczenia, jest fakt, że ma ona charakter z założenia introspektywny – nie poddaje się jakiegokolwiek zewnętrznej ocenie<sup>39</sup>. W kontekście stosowania prawa jest to zarzut bardzo poważny. Przewaga korpusów nad intuicją jest oczywista i ma dwojaki charakter. Korpusy są zbiorami tekstów pochodzących od ogromnej rzeszy użytkowników języka, co usuwa ryzyko idiosynkrazji. Ponadto, dostęp do korpusów jest powszechny, a każdy etap analizy korpusowej jest jawny – można go prześledzić, zweryfikować, a w razie potrzeby nawet powtórzyć<sup>40</sup>.

Po drugie, korpusy przeciwstawia się słownikom<sup>41</sup>. Jak wspominaliśmy, problem wykorzystania słowników w interpretacji prawa został w amerykańskiej literaturze omówiony bardzo gruntownie. Słowniki zdobyły popularność wśród sędziów w okresie sukcesów tekstualizmu jako metody interpretacyjnej<sup>42</sup>. Stanowią niewątpliwie postępowanie w stosunku do intuicji sędziego – oferują dostępne dla wszystkich definicje słów, sformułowane przez niezależnych ekspertów (leksykografów). Definicjom zawartym w słownikach zarzuca się jednak akontekstowość i pomijanie subtelności semantycznych właściwych dla naturalnego użycia języka<sup>43</sup>. Większość słów posiada kilka różnych definicji w danym słowniku i nie istnieje kryterium wyboru jednej z nich jako odpowiedniej w danej sprawie. Co więcej, różne słowniki zawierają różne definicje i różną liczbę

<sup>36</sup> S. Mouritsen, *The Dictionary...*, s. 1970.

<sup>37</sup> S. Mouritsen, *Corpus Linguistic...*, s. 68; J. Phillips, J. Egbert, *Advancing Law...*, s. 1589.

<sup>38</sup> L. Solan, T. Gales, *Finding Ordinary Meaning in Law: The Judge, the Dictionary or the Corpus?*, „Brooklyn Law School Legal Studies” 2016/474, s. 1–3.

<sup>39</sup> S. Mouritsen, *Hard Cases...*, s. 174–180.

<sup>40</sup> S. Mouritsen, *Hard Cases...*, s. 201 i n.

<sup>41</sup> S. Mouritsen, *The Dictionary...*; L. Solan, T. Gales, *Finding Ordinary Meaning...*, s. 3 i n.; S. Gries, B. Slocum, *Ordinary Meaning...*, s. 1417 i n.

<sup>42</sup> Por. m.in. E. Aprill, *The Law...*, s. 278–281; J. Brudney, L. Baum, *Oasis or Mirage: The Supreme Court's Thirst for Dictionaries in the Rehnquist and Roberts Eras*, „William & Mary Law Review” 2013/55, s. 494–502; L. Solan, *Can Corpus Linguistics Help Make Originalism Scientific*, „Yale Law Journal” 2016/26, s. 1464.

<sup>43</sup> „Dictionary definitions are acontextual, whereas the meaning of sentences depends critically on context, including all sorts of background understandings” (sędzia Richard Posner w uzasadnieniu wyroku w sprawie *United States v. Costello*. W literaturze funkcjonują plastyczne metafory słownika jako „muzeum słów” (por. F.H. Easterbrook, *Text, History, and Structure in Statutory Interpretation*, „Harvard Journal of Law & Public Policy” 1994/17, s. 67), czy „zoo słów” (por. A.R. Randolph, *Dictionaries, Plain Meaning, and Context in Statutory Interpretation*, „Harvard Journal of Law & Public Policy” 1994/17, s. 74).

definicji. W rezultacie, wybór słownika, nierzadko całkowicie arbitralny, może determinować rozstrzygnięcie sądowe. W skrajnych przypadkach problem interpretacyjny zostaje sprowadzony do sporu o właściwy słownik (tzw. *dictionary shopping*)<sup>44</sup>.

Korpusy górują nad słownikami w kilku aspektach. Po pierwsze, ukazują słowa w szerszym kontekście zdaniowym<sup>45</sup>. Po drugie, umożliwiają określenie frekwencji użycia wyrazu w danym znaczeniu<sup>46</sup>. Tym samym pozwalają ustalić, które znaczenie wieloznacznego wyrazu jest najbardziej powszechne. Słowniki takich informacji nie zawierają, natomiast częste wyciąganie wniosków w tym zakresie z kolejności umieszczenia definicji danego wyrazu jest zwykłym nieporozumieniem<sup>47</sup>. Po trzecie, znaczenie płynące z korpusów nie jest zapośredniczone przez osoby trzecie (leksykografów), których subiektywne oceny determinują liczbę i treść definicji słownikowych<sup>48</sup>. Analizując korpus, sędzia operuje na rzeczywistych przykładach użycia słów i sam wyciąga wnioski co do ich znaczenia. Innymi słowy, korpus pozwala sędziemu stać się „leksykografem chwili”<sup>49</sup>. Po czwarte, korpusy – dzięki możliwości aktualizacji – mają większą szansę nadążać za dynamicznym rozwojem języka, wiążącym się z popularnością nowoczesnych form komunikacji międzyludzkiej<sup>50</sup>.

Po trzecie, korpusy przeciwstawia się badaniom ankietowym (*surveys*). W latach 70. XX w. psychologowie i lingwiści (m.in. Eleanor Rosch, William Labov) przeprowadzili szereg pionierskich badań nad mechanizmem kategoryzacji. Zasadniczo polegały one na zadawaniu ankietowanym pytań o językową klasyfikację określonych zjawisk (obrazów, kształtów, kolorów, pojęć). Wyniki wykazały, że ludzka kategoryzacja opiera się – przynajmniej częściowo – na podobieństwie do najbardziej typowego okazu danego pojęcia (tzw. prototypu), a nie na weryfikacji kryterialnych cech danego pojęcia<sup>51</sup>. W teorii prawa wyrażono pogląd, że zwykle znaczenie to nic innego, jak właśnie prototyp pojęcia ustawowego<sup>52</sup>. Niektórzy autorzy zaproponowali zatem aplikację podobnych metod do rozstrzygania wątpliwości interpretacyjnych<sup>53</sup>. Ich oczywistą przewagą nad słownikami jest podbudowa empiryczna. Natomiast ich przeprowadzenie w praktyce napotyka liczne trudności: wymaga czasu, pieniędzy, zaangażowania osób trzecich (najlepiej ekspertów), a sama ankieta musi być sformułowana niezwykle starannie, by uniknąć sugerowania odpowiedzi i zminimalizować różnego rodzaju błędy poznawcze<sup>54</sup>. Trudności tych można uniknąć zastępując ankietę korpusami. Są one powszechnie dostępne, bezpłatne, gwarantują wyższy stopień obiektywizmu, a przyswojenie podstaw metodologii językoznawstwa korpusowego nie przekracza możliwości sędziego<sup>55</sup>.

Choć wykorzystanie korpusów jako narzędzia interpretacyjnego cieszy się od kilku lat dużym zainteresowaniem teorii prawa i praktyki prawniczej, pojawiają się także głosy

<sup>44</sup> Por. np. E. Aprill, *The Law...*, s. 277; J. Brudney, L. Baum, *Oasis...*, zwłaszcza s. 566 i n.; J. Kirchmeier, S. Thumma, *The Lexicon...*, zwłaszcza s. 264–276; Ph. Rubin, *War...*, zwłaszcza s. 174–177; L. Solan, T. Gales, *Corpus Linguistics...*, s. 1334–1336.

<sup>45</sup> S. Gries, B. Slocum, *Ordinary Meaning...*, s. 1441.

<sup>46</sup> L. Solan, *Can Corpus Linguistics Help...*, s. 62–63.

<sup>47</sup> S. Mouritsen, *The Dictionary...*, s. 1926–1938.

<sup>48</sup> L. Solan, *Can Corpus Linguistics Help...*, s. 62–63.

<sup>49</sup> L. Solan, *Can Corpus Linguistics Help...*, s. 60.

<sup>50</sup> A. Wang, *Googling for Meaning: Statutory Interpretation in the Digital Age*, „Yale Law Journal” 2016/125, s. 267–280.

<sup>51</sup> Por. na ten temat M. Zeifert, *Wyrok Trybunału Sprawiedliwości Unii Europejskiej z perspektywy kognitywnej teorii prototypu. Studium przypadku*, „Archiwum Filozofii Prawa i Filozofii Społecznej” 2020/2, s. 109–120.

<sup>52</sup> Por. np. L. Solan, *Law, Language, and Lenity*, „William & Mary Law Review” 1998/57, s. 67–68.

<sup>53</sup> Por. C. Cunningham, J. Levi, G. Green, J. Kaplan, *Plain Meaning and Hard Cases*, „Yale Law Journal” 1993/103, s. 1588 i n.

<sup>54</sup> S. Mouritsen, *Hard Cases...*, s. 181–190.

<sup>55</sup> S. Mouritsen, *Hard Cases...*, s. 180 i n.; T. Lee, S. Mouritsen, *Judging Ordinary...*, s. 865–868.

krytyczne. Zdaniem jednego z autorów „badania korpusowe w najlepszym razie mogą okazać się czasochłonnym urozmaiceniem, w najgorszym – doprowadzić do nietrafnych rozstrzygnięć, które nie poddają się weryfikacji”<sup>56</sup>. Formułowane uwagi podzielić można na dwie grupy: dotyczące samych korpusów oraz dotyczące sposobów ich używania przez sędziów.

Do pierwszej grupy należy zarzut, że korpusy, jako zbiory rzeczywistych przypadków użycia języka, zawierają również użycia niekonwencjonalne, metaforyczne, potoczne czy zwyczajnie niepoprawne. Metodologia tworzenia korpusów wyklucza jakiegokolwiek wartościowanie, czy „cenzurowanie” uwzględnianych przykładów, dlatego korpusy – w przeciwieństwie do słowników – nie przyczyniają się do kształtowania normy językowej<sup>57</sup>. Wskazuje się także, że częstotliwość użycia wyrazu w określonym znaczeniu niekoniecznie oznacza, że jest to znaczenie najbardziej podstawowe – może wynikać z kilku innych czynników<sup>58</sup>. Sam proces tworzenia korpusu nasuwa pewne wątpliwości. Jest nie tylko całkowicie niezależny od sędziego, ale także niedostępny dla jego świadomości. Tymczasem dobór tekstów tworzących korpus w oczywisty sposób determinuje wyniki przyszłych analiz. Mamy zatem do czynienia ze swoistym „outsourcingiem parametrycznym”<sup>59</sup>. Zwraca się również uwagę, że korpusy różnią się między sobą: rozmiarami, procentowym udziałem poszczególnych rejestrów języka, wiekiem tekstów źródłowych itd. Ostatecznie więc analizy korpusowe cierpią na taką samą arbitralność jak analizy słownikowe<sup>60</sup>.

Lista zarzutów należących do drugiej grupy jest znacznie dłuższa. Najbardziej radykalni przeciwnicy korpusów twierdzą, że korzystanie z nich wypacza naturę interpretacji prawa, wprowadza bowiem kryterium ilościowe w miejsce kryterium jakościowego, i tym samym minimalizuje rolę sędziów, która zostaje sprowadzona do liczenia wyników w wyszukiwarce<sup>61</sup>. Taki „zwodniczy empiryzm” zaciemnia rzeczywistość, dając mylne wrażenie, że interpretacja prawa jest przedsięwzięciem deskryptywnym i obiektywnym<sup>62</sup>. Jednocześnie rodzi obawy co do niezawisłości sędziowskiej i odpowiedzialności sędziów za podejmowane rozstrzygnięcia.

Przede wszystkim jednak zwraca się uwagę, że używanie korpusów jest trudne: „nie polega po prostu na wpisaniu poszukiwanej frazy w okienko i ocenie wyników”<sup>63</sup>. Po pierwsze, należy opanować oprogramowanie umożliwiające przeszukiwanie korpusu. Wymaga to od sędziego umiejętności technicznych, które normalnie nie są w tym zawodzie niezbędne. Po drugie, należy zastosować odpowiednią metodologię. Nie jest to zadanie łatwe, ponieważ samo językoznawstwo korpusowe – w przeciwieństwie np. do leksykografii – jest dyscypliną młodą i niedojrzałą metodologicznie<sup>64</sup>. Analiza korpusowa składa się z wielu kroków, a każdy z nich cechuje wysoki stopień dyskrecjonalności, która wpływa na uzyskiwane wyniki<sup>65</sup>.

<sup>56</sup> J. Ehrett, *Against Corpus Linguistics*, „Georgetown Law Journal Online” 2019/108, s. 61.

<sup>57</sup> J. Ehrett, *Against Corpus...*, s. 62–65. Autor podaje przykład angielskiego słowa *literally* („dosłownie”), które w ostatnich latach w mowie potocznej zyskało znaczenie dokładnie przeciwstawne („w przenośni”). W polskich realiach pouczające może być prześledzenie przykładów użycia słowa *masakra* w Narodowym Korpusie Języka Polskiego. Wiele spośród pojawiających się tam znaczeń ma charakter przenośny i odległy od znaczeń słownikowych: „masowe zabijanie w okrutny sposób”, „wypadek lub inne tragiczne wydarzenie, w którym jest wielu rannych i zabitych”. Zob. *Internetowy słownik PWN*, <https://sjp.pwn.pl/szukaj/masakra.html>, dostęp: 20.12.2019 r.

<sup>58</sup> D. Tankersley, *Beyond the Dictionary...*, s. 668.

<sup>59</sup> J. Ehrett, *Against Corpus...*, s. 65.

<sup>60</sup> J. Ehrett, *Against Corpus...*, s. 65–68.

<sup>61</sup> C. Hessick, *Corpus Linguistics and the Criminal Law*, „Brigham Young University Law Review” 2017/6, s. 1529 i n.

<sup>62</sup> C. Hessick, *Corpus Linguistics...*, s. 1529 i n.; D. Tankersley, *Beyond the Dictionary...*, s. 667–672.

<sup>63</sup> J. Ehrett, *Against Corpus...*, s. 69.

<sup>64</sup> J. Phillips, J. Egbert, *Advancing Law...*, s. 1598.

<sup>65</sup> D. Tankersley, *Beyond the Dictionary...*, s. 667.



Z uwagi na obszerność korpusów, zapoznanie się z wszystkimi przykładami użycia danego wyrazu jest zazwyczaj niemożliwe. Konieczne jest zatem wprowadzenie filtrów, które zawężą wyniki wyszukiwania. Dobór filtrów jest niczym nieograniczony i bezpośrednio wpływa na wyniki analizy. Przykładowo, we wspomnianej sprawie *Muscarello v. United States*, sędzia S. Breyer usiłował wykazać, że zwykłe znaczenie wyrażenia *to carry a weapon* („nosić broń”) obejmuje „przewozić w samochodzie”. W tym celu przeprowadził quasi-korpusową analizę dwóch elektronicznych baz danych. Analiza polegała na wyszukaniu zdań, które zawierały słowa *to carry* („nosić”), *weapon* („broń”), *vehicle* („pojazd”) i ich synonimy. Taki dobór filtrów spowodował oczywiście potwierdzenie hipotezy, bowiem zdania w ogóle nieodnoszące się do pojazdów były po prostu pomijane w wynikach<sup>66</sup>.

Z uwagi na wskazane trudności metodologiczne, niektórzy autorzy domagają się, by wykorzystanie korpusów w praktyce stosowania prawa zostało uregulowane w formie powszechnie obowiązujących aktów prawnych<sup>67</sup>. Inni postulują, że użycie korpusów powinno być zastrzeżone dla biegłych<sup>68</sup>. Jeszcze inni skłonni są dopuszczać analizy korpusowe jedynie na wniosek strony postępowania<sup>69</sup>. Formulowane bywa pytanie, czy analiza korpusowa powinna być kwalifikowana jako ustalenia co do prawa (*iura novit curia*), czy co do faktów. Ma to doniosłe znaczenie z uwagi na kwestie proceduralne takie jak zasada kontrydiktoryjności (*adversarial system*) czy notoryjność powszechna (*judicial notice*)<sup>70</sup>. Ze względu na odmienność amerykańskiego i polskiego modelu procesu sądowego wątek ten pomijamy. Wydaje się wszakże, że również w polskich realiach status procesowy analizy korpusowej może nasuwać pewne wątpliwości.

## 5. Podsumowanie

Niniejsze opracowanie ma charakter sprawozdawczy. Nie jest naszym celem wszechstronna analiza przywołanych argumentów ani zajmowanie ostatecznego stanowiska w zarysowanym sporze. Ograniczyliśmy się jedynie do kilku uwag płynących z doświadczeń amerykańskich. Jak zostało wspomniane, argumenty krytyczne wobec korpusów dają się podzielić na dwie grupy. Pierwsza obejmuje m.in. obawy co do niezawisłości sędziego, interpretacyjnego „outsourcingu”, unikania odpowiedzialności za wynik interpretacji. Nasuwa się refleksja, że obawy te w jednakowym – jeśli nie większym – stopniu dotyczą innych narzędzi interpretacyjnych, w szczególności słowników, szeroko rozpowszechnionych również w polskiej praktyce orzeczniczej. W tym kontekście korpusy, z powodów wskazanych powyżej, jawią się jako istotny krok naprzód<sup>71</sup>. Druga grupa obejmuje obawy co do sposobu korzystania z korpusów. Niewątpliwie, przeprowadzenie analizy korpusowej wymaga sporo czasu i wysiłku intelektualnego. Jak pokazują doświadczenia amerykańskich sędziów, nietrudno przy tym o popełnienie rażących błędów metodologicznych. Jedną z głównych zalet analiz korpusowych jest jednak właśnie to, że błędy te są bardzo łatwe do wychwycenia<sup>72</sup>. Jeden z czołowych orędowników

<sup>66</sup> Por. L. Solan, T. Gales, *Corpus Linguistics...*, s. 1344–1348; S. Mouritsen, *Corpus Linguistics...*, s. 71–72.

<sup>67</sup> J. Ehrett, *Against Corpus...*, s. 71–72.

<sup>68</sup> S. Gries, B. Slocum, *Ordinary Meaning...*, s. 1470–1472; L. Solan, *Can Corpus Linguistics Help...*, s. 1469 i n.

<sup>69</sup> D. Tankersley, *Beyond the Dictionary...*, s. 672 i n.

<sup>70</sup> D. Tankersley, *Beyond the Dictionary...*, s. 669–672. Por. jednak kontrargumenty wyrażone w: J. Ramer, *Corpus Linguistics: Misfire or More Ammo for the Ordinary-Meaning Canon*, „Michigan Law Review” 2017/2, s. 323, a także T. Lee, S. Mouritsen, *Judging Ordinary...*, s. 868–871.

<sup>71</sup> T. Lee, S. Mouritsen, *Judging Ordinary...*, s. 877.

<sup>72</sup> S. Mouritsen, *Corpus Linguistics...*, s. 74; J. Ramer, *Corpus Linguistics...*, s. 328.

korpusów, sędzia Th. Lee w swojej opinii do sprawy *Utah vs. Rasabout*<sup>73</sup> przedstawił wymowną analogię. Gdy na rynek wchodziły pierwsze elektroniczne prawnicze bazy danych, wielu teoretyków prawa wieszczyło zmierzch profesji prawniczej. Jeden z autorów ostrzegał, że nowe technologie zniweczą umiejętność „myślenia jak prawnik”. Obawy te w żadnej mierze się nie spełniły. Być może zatem – jak postuluje inny autor – trzeba zaakceptować fakt, że na skutek powszechnej dostępności korpusów „myślenie jak prawnik” będzie czasami wymagało umiejętności „myślenia jak językoznawca”<sup>74</sup>.

W naszym przekonaniu omówione przykłady pokazują dobitnie, że korpusy językowe należy postrzegać jako narzędzie interpretacyjne. Ich rolą jest służyć sędziemu, a nie kierować jego poczynaniami<sup>75</sup>. Efektywne wykorzystanie tego narzędzia nie jest sprawą łatwą. Wymaga „zadawania właściwych pytań, przeprowadzania właściwych wyszukiwań i wyciągania uzasadnionych wniosków zarówno z obecności, jak i nieobecności danych odpowiadających takiemu czy innemu sposobowi użycia słów”<sup>76</sup>. Odpowiedzialność za wykształcenie sędziów w metodach językoznawstwa korpusowego spoczywa na przedstawicielach nauki prawa. W literaturze sformułowano już wiele wartościowych zaleceń w tym zakresie<sup>77</sup>. Należy również pamiętać, że korpusy, jak każde narzędzie, nadają się do rozwiązania tylko niektórych problemów pojawiających się w praktyce interpretacyjnej. Przykładowo, raczej nie pomogą ustalić specjalistycznego (np. prawnego) znaczenia danego wyrazu, nie wyeliminują też problemu nieostrości<sup>78</sup>. Same korpusy nie oferują w tym zakresie żadnych wskazówek, ponownie zatem zaznacza się kluczowa rola nauki prawa<sup>79</sup>.

Na koniec trzeba podkreślić, że pomimo pojawiających się głosów krytycznych, korpusy torują sobie drogę w amerykańskim orzecznictwie, a *law and corpus linguistics* nabiera kształtu jako dyscyplina naukowa. Organizowane są konferencje poświęcone wykorzystaniu korpusów w interpretacji prawa<sup>80</sup>, na uczelniach prawniczych wprowadzane są kursy językoznawstwa korpusowego<sup>81</sup>. Trudno obecnie prognozować docelową skalę tego zjawiska, ale wydaje się, że korpusy językowe mogą przebyć drogę podobną do tej, jaką w latach 90. XX w. przebyły słowniki, stając się podstawowym narzędziem amerykańskiego sędziego<sup>82</sup>. W tej sytuacji zastanawia niemal kompletny brak zainteresowania językoznawstwem korpusowym w polskiej literaturze prawniczej<sup>83</sup>. W polskim orzecznictwie korpusy pojawiają się sporadycznie w sprawach z zakresu prawa własności

<sup>73</sup> J. Ramer, *Corpus Linguistics...*, s. 324.

<sup>74</sup> N. Goldfarb, *A Lawyer's Introduction to Meaning in the Framework of Corpus Linguistics*, „Brigham Young University Law Review” 2017/6, s. 1416.

<sup>75</sup> D. Lowe, Ch. Potter, *Understanding Legislation...*, s. 30.

<sup>76</sup> L. Solan, T. Gales, *Corpus Linguistics...*, s. 1357.

<sup>77</sup> Por. np. S. Mouritsen, *Hard Cases...*, s. 190–201; L. Solan, T. Gales, *Corpus Linguistics...*, s. 1342–1356; N. Goldfarb, *A Lawyer's Introduction...*, s. 2397–1415; J. Phillips, J. Egbert, *Advancing Law...*, s. 1608–1618.

<sup>78</sup> S. Mouritsen, *Hard Cases...*, s. 204–205.

<sup>79</sup> Por. wskazówki sformułowane w pracy: L. Solan, T. Gales, *Corpus Linguistics...*, s. 1342 i n.

<sup>80</sup> Pierwszą cykliczną konferencją tego typu jest *Corpus Linguistics Conference*, organizowana od 2016 r. przez Brigham Young University w Provo, stan Utah (USA).

<sup>81</sup> Pierwszy tego typu kurs prowadzi Stephen Mouritsen. Por. S. Mouritsen, *Corpus Linguistics...*, s. 80–81.

<sup>82</sup> Por. na ten temat np. J. Brudney, L. Baum, *Oasis...*, zwłaszcza s. 494–502.

<sup>83</sup> Por. jednak przypis 5. O językoznawstwie korpusowym wspomina znakomity badacz polskiego języka prawnego, Andrzej Malinowski. Por. A. Malinowski, *Pomiar czytelności polskiego tekstu prawnego*, „Państwo i Prawo” 2017/5, s. 5–6. Korpusy są za to powszechnie wykorzystywane przez językoznawców zajmujących się językiem prawnym i translatoryką prawniczą. Por. np. Ł. Biel, *Wyrażenie tak zwany (tzw.) jako leksykalny wykładnik modalności epistemicznej w języku prawniczym: analiza korpusowa orzecznictwa Sądu Najwyższego*, w: D. Kondratczyk-Przybylska, A. Niewiadomski, E. Walewska (red.), *Język prawny i prawniczy. I Kongres ogólnopolski*, Warszawa 2017, s. 19–45; Ł. Biel, *Wyrażenia przyimkowe w krajowych i unijnych aktach prawnych – analiza korpusowa funkcji przyimków w polskim języku prawnym*, w: D. Kondratczyk-Przybylska, A. Niewiadomski, E. Walewska (red.), *Język polskiego prawa: nowe wyzwania*, Warszawa 2016, s. 9–35.

intelektualnej i dóbr osobistych, jednak wyłącznie w opiniach biegłych językoznawców<sup>84</sup> lub w argumentacji stron<sup>85</sup>. Innymi słowy, korpusy językowe są narzędziem zasadniczo obcym rodzimej teorii prawa i praktyce prawniczej. Zmiana tego stanu rzeczy wymaga dalszych, pogłębianych badań i kolejnych publikacji poświęconych temu zagadnieniu.

### **Linguistic Corpora as a Tool of Statutory Interpretation: American Theory and Practice**

**Abstract:** In American adjudicating practice and theory of statutory interpretation, great attention is paid to the notion of ordinary meaning of legal text. In order to determine the ordinary meaning, judges usually refer to their own linguistic intuition or dictionaries – both these methods give rise to a number of reservations that have been expressed in legal literature for years. In the last few years, courts have also started using linguistic corpora for this purpose. Linguistic corpora are electronic collections of authentic texts in a given language which can be analyzed using IT tools (e.g. searches, frequency lists, concordances, collocations). Corpus research requires considerable linguistic knowledge and technical skills, and in return it offers statistical data that can reveal a lot about the semantic layer of language. The use of corpora by judges – first in state courts, later also in the federal Supreme Court – sparked lively academic discussion. In Polish literature, this issue has not been discussed so far, which is why the article is essentially of a reporting nature. First, an outline of corpus linguistics is presented. Then the history of the use of linguistic corpora by US courts is reconstructed. Finally, the legal discussion about corpora in the US is presented, with particular emphasis on voices approving their use and on critical voices.

**Keywords:** statutory interpretation, corpus linguistics, linguistic corpora, US case law

<sup>84</sup> Por. np. wyrok Sądu Apelacyjnego (dalej: „SA”) w Białymstoku z 26.06.2015 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”. Por. także wyrok SA w Łodzi z 13.11.2014 r. (I ACa 709/14), LEX nr 1554764; wyrok Wojewódzkiego Sądu Administracyjnego w Warszawie z 19.03.2019 r. (VI SA/Wa 1895/18), LEX nr 2696759; wyrok Sądu Okręgowego w Warszawie z 22.08.2018 r. (X Ka 721/18), LEX nr 2566284.

<sup>85</sup> Por. wyrok Naczelnego Sądu Administracyjnego (dalej: „NSA”) z 16.10.2013 r. (II GSK 937/12), LEX nr 1559063: „Uprawniony przecząc twierdzeniom wnioskodawcy podniósł natomiast, że nazwa (...) nie pojawia się w żadnym z ogólnych zbiorów leksykalnych współczesnego języka i nie jest odnotowywane przez korpus języka polskiego, w tym przez Korpus Języka Polskiego PWN. Wnioskodawca nie wykazał, by twierdzenia te były niezgodne ze stanem faktycznym”. Por. także wyrok NSA z 30.05.2019 r. (II GSK 137/18), LEX nr 2690475.

## BIBLIOGRAFIA / REFERENCES:

- Aprill, E.P. (1998). The Law of the Word: Dictionary Shopping in the Supreme Court. *Arizona State Law Journal* 30/2, 275–336.
- Biel, Ł. (2016). Wyrażenia przyimkowe w krajowych i unijnych aktach prawnych - analiza korpusowa funkcji przyimków w polskim języku prawnym. In D. Kondratczyk-Przybylska, A. Niewiadomski, E. Walewska (Eds.), *Język polskiego prawa: nowe wyzwania*. Warszawa: Międzywydziałowe Koło Naukowe Kultury Języka Prawnego i Prawniczego Uniwersytetu Warszawskiego Lingua Iuris.
- Biel, Ł. (2017). Wyrażenie tak zwany (tzw.) jako leksykalny wykładnik modalności epistemicznej w języku prawniczym: analiza korpusowa orzecznictwa Sądu Najwyższego. In D. Kondratczyk-Przybylska, A. Niewiadomski, E. Walewska (Eds.), *Język prawny i prawniczy. I Kongres ogólnopolski*. Warszawa: Międzywydziałowe Koło Naukowe Kultury Języka Prawnego i Prawniczego Uniwersytetu Warszawskiego Lingua Iuris.
- Bielska-Brodziak, A., Tobor, T. (2007). Słowniki a interpretacja tekstów prawnych. *Państwo i Prawo* 5, 20–33.
- Brudney, J., Baum, L. (2013). Oasis or Mirage: The Supreme Court's Thirst for Dictionaries in the Rehnquist and Roberts Eras. *William & Mary Law Review* 55/2, 483–580.
- Cunningham, C.D., Levi, J.N., Green, G.M., Kaplan, J.P. (1993). Plain Meaning and Hard Cases. *Yale Law Journal* 103/6, 1561–1625.
- Ehrett, J.S. (2019). Against Corpus Linguistics. *Georgetown Law Journal Online* 108, 50–73.
- Goldfarb, N. (2017). A Lawyer's Introduction to Meaning in the Framework of Corpus Linguistics. *Brigham Young University Law Review* 2017/6, 1359–1416.
- Grabowski, A. (2016). Siła argumentu interpretacyjnego z prawniczego domniemania języka (znaczenia) potocznego. *Zagadnienia Naukoznawstwa* 52/3/209, 357–370.
- Hessick, C.B. (2017). Corpus Linguistics and the Criminal Law. *Brigham Young University Law Review* 2017/6, 1503–1530.
- Kirchmeier, J., Thumma, S. (1999). The Lexicon Has Become a Fortress: The United States Supreme Court's Use of Dictionaries. *Buffalo Law Review* 47/1, 227–562.
- Kubicka, E., Grec, K., Leszczak, W., Mitew, M. (2015). Wykorzystanie słowników w interpretacji prawniczej – implikacje praktyczne. *Kwartalnik Sądowy Apelacji Gdańskiej* 2015/1, 25–50.
- Lee, T., Mouritsen, S. (2018). Judging Ordinary Meaning. *Yale Law Journal* 127/4, 788–879.
- Lowe, D., Potter, Ch. (2018). *Understanding Legislation: A Practical Guide to Statutory Interpretation*. Oxford: Hart Publishing.
- Malinowski, A. (2017). Pomiar czytelności polskiego tekstu prawnego. *Państwo i Prawo* 5, 3–18.
- Mascott, J. (2018). Who Are “Officers of the United States”. *Stanford Law Review* 70/2, 443–564.
- Mouritsen, S. (2010). The Dictionary is Not a Fortress: Definitional Fallacies and a Corpus-Based Approach to Plain Meaning. *Brigham Young University Law Review* 2010/5, 1915–1980.
- Mouritsen, S. (2011). Hard Cases and Hard Data: Assessing Corpus Linguistics as an Empirical Path to Plain Meaning. *Columbia Science and Technology Law Review* 13/1, 156–205.

- Mouritsen, S. (2017). Corpus Linguistics in Legal Interpretation: An Evolving Interpretative Framework. *International Journal of Language & Law* 6, 67–89.
- Mullins Sr., M.E. (2004). Tools, Not Rules: The Heuristic Nature of Statutory Interpretation. *Journal of Legislation* 30/1, 1–76.
- Phillips, J., Egbert, J. (2017). Advancing Law and Corpus Linguistics: Importing Principles and Practices from Survey and Content-Analysis Methodologies to Improve Corpus Design and Analysis. *Brigham Young University Law Review* 2017/6, 1589–1620.
- Ramer, J.D. (2017). Corpus Linguistics: Misfire or More Ammo for the Ordinary-Meaning Canon. *Michigan Law Review* 116/2, 303–328.
- Rubin, P.A. (2010). War of the Words: How Courts Can Use Dictionaries in Accordance with Textualist Principles. *Duke Law Journal* 60/1, 167–206.
- Scalia, A., Garner, B. (2012). *Reading Law: The Interpretation of Legal Texts*. St. Paul: Thomson/West.
- Slocum, B.G. (2015). *Ordinary Meaning: A Theory of the Most Fundamental Principle of Legal Interpretation*. Chicago: The University of Chicago Press.
- Solan, L.M. (1998). Law, Language, and Lenity. *William and Mary Law Review* 40/1, 57–144.
- Solan, L.M. (2005). The New Textualists' New Text. *Loyola of Los Angeles Law Review* 38/5, 2027–2062.
- Solan, L.M. (2010). *The Language of Statutes: Laws and Their Interpretation*. Chicago: The University of Chicago Press.
- Solan, L.M. (2016). Can Corpus Linguistics Help Make Originalism Scientific. *Yale Law Journal Forum* 126, 57–64.
- Solan, L.M., Gales, T. (2016). Finding Ordinary Meaning in Law: The Judge, the Dictionary or the Corpus? *Brooklyn Law School, Legal Studies Paper* 474, 1–19.
- Solan, L.M., Gales, T. (2017). Corpus Linguistics as a Tool in Legal Interpretation. *Brigham Young University Law Review* 2017/6, 1311–1358.
- Tankersley, D.C. (2018). Beyond the Dictionary: Why Sua Sponte Judicial Use of Corpus Linguistics Is Not Appropriate for Statutory Interpretation. *Mississippi Law Journal* 87/4, 641–678.
- Tobor, T., Żmigrodzki, P., Bielska-Brodziak, A. (2008). Co każdy prawnik o słownikach wiedzieć powinien. *Przegląd Sądowy* 7/8, 79–95.
- Tobor, Z. (2013). *W poszukiwaniu intencji prawodawcy*. Warszawa: Wolters Kluwer Polska.
- Wang, A. (2016). Googling for Meaning: Statutory Interpretation in the Digital Age. *Yale Law Journal Forum* 125, 267–280.
- Zieliński, M. (2012). *Wykładnia prawa. Zasady, reguły, wskazówki*. Warszawa: LexisNexis Polska.