

Zbitki wielowyrazowe w polskim eurolektie – analiza gatunkowa formuliczności dyskursu administracyjnego. Badanie pilotażowe¹

Łucja Biel, Dariusz Koźbiał, Katarzyna Wasilewska

Uniwersytet Warszawski

1. Wstęp

Eurolekt jest nową „europejską” hybrydową odmianą języka polskiego, stosowaną w kontekście unijnym, będącą wynikiem komunikacji odbywającej się za pośrednictwem tłumaczy. Wykształcenie się nowej odmiany jest konsekwencją pełnienia przez język polski funkcji jednego z 24 języków urzędowych Unii Europejskiej (UE). Hybrydowość eurolektu wiąże się z wielojęzycznym i wieloetapowym procesem tworzenia tekstów przez instytucje unijne, skutkującym przemieszaniem języków, kultur, dyskursów i konwencji komunikacyjnych.

Celem niniejszego tekstu jest zbadanie dystrybucji i funkcji najczęstszych zbitek wielowyrazowych w polskim eurolektie na tle krajowego języka administracyjnego. Metoda zbitek wielowyrazowych będzie zastosowana do analizy wewnętrznego i zewnętrznego zróżnicowania gatunkowego eurolektu w celu oceny, w jakim stopniu proces tłumaczenia oraz wielojęzyczność wpływają na formuliczność unijnego dyskursu

1. Badanie sfinansowane ze środków Narodowego Centrum Nauki przyznanych w ramach projektu badawczego SONATA BIS nr 2014/14/E/HS2/00782 pt. „Eurolekt – unijna odmiana języka polskiego i jej wpływ na polszczyznę urzędową”.

administracyjnego. Dodatkowym celem badania jest ocena przydatności użytej metody w badaniach nad dyskursem specjalistycznym, w tym szczególnie z uwzględnieniem języka polskiego. Z racji bardzo ograniczonych badań z zastosowaniem tej metody na gruncie polskim badanie ma charakter pilotażowy. Korpusowe nachylenie projektu w stronę wyspecjalizowanej komunikacji administracyjnej, która podlega procesom zmiany językowej na styku kultur, stanowi odniesienie do badań Profesor Anny Duszak nad komunikacją międzykulturową (Duszak 1998; Duszak, Gajek i Okulska red. 2006) z perspektywy kontaktu językowego (Duszak i Okulska red. 2004; Duszak, Jabłoński i Leńko-Szymańska red. 2017) w obszarze instytucjonalnym (np. Duszak, House i Kumięga red. 2010; Duszak i Kowalski red. 2015 itp.).

2. Zróżnicowanie gatunkowe eurolektu

Gatunek (ang. *genre*) jest obecnie jednym z podstawowych pojęć językoznawstwa funkcjonalnego, zwłaszcza w nurtach związanych z analizą dyskursu (zob. Miller 1984) i lingwistyką tekstu (zob. Bachtin 1986). Znaczenie tego terminu ewoluowało w ciągu ostatnich kilku dekad i różni się w zależności od ram teoretycznych. Gatunek uważa się m.in. za typ zdarzenia komunikacyjnego (Duszak 1998: 218), „elastyczne makrostruktury” (ang. *flexible macrostructures*) (Cap i Okulska 2013: 4) i sposób (współ)działania językowego (Fairclough 2003: 17) w określonym celu komunikacyjnym łączącym wspólnotę dyskursu (Swales 1990: 58)². W rezultacie gatunek charakteryzuje się skonwencjonalizowanym użyciem języka polegającym na powtarzającym się użyciu stosunkowo stabilnych i rozpoznawalnych schematów, co rodzi określone oczekiwania wspólnoty dyskursu co do formy, jak i treści tekstów danego gatunku. Można założyć, iż stopień konwencjonalizacji będzie

2. Przegląd definicji dotyczących gatunku zob. szerzej Cap i Okulska (red. 2013).

wysoki w przypadku dyskursu administracyjnego będącego przedmiotem niniejszego opracowania.

Komunikacja w Unii Europejskiej charakteryzuje się dużym zróżnicowaniem gatunkowym i wielojęzycznością. Znaczący wpływ na kształtowanie gatunków unijnych ma proces tłumaczenia i inne formy wielojęzycznej mediacji. W ramach badania analizie poddane zostały trzy podstawowe gatunki w specjalistycznej komunikacji unijnej: akty normatywne, orzeczenia oraz sprawozdania.

2.1. Akty normatywne

Akty normatywne (akty prawne) zajmują nadrzędne miejsce w hierarchii gatunków prawniczych (por. Biel 2014: 21). Uważa się je za gatunek primarny (Bhatia 2006) i konstytutywny (Kjær 2000), gdyż determinuje on warstwę pojęciową gatunków pochodnych i stanowi podstawę dla praktyki prawniczej. Akty normatywne pełnią funkcję preskryptywną i normatywną. Są instrumentem tworzenia prawa, służą kreowaniu norm prawnych i regulowaniu zachowań, a także określają prawa i obowiązki adresatów norm. Celem aktów prawnych jest uporządkowanie otaczającej nas skomplikowanej rzeczywistości za pomocą przepisów prawnych (Bhatia 2006: 6). Akty normatywne kreują zatem świat modelowy, który jest konfrontowany z rzeczywistością w gatunkach pochodnych, takich jak orzecznictwo (Bhatia 2006: 6). Autorem aktów normatywnych jest prawodawca – podmiot instytucjonalny i zbiorowy. W procesie stanowienia prawa uczestniczą różne podmioty, w tym podmioty, które przygotowują projekty aktów, podejmują decyzje prawotwórcze i promulgują teksty prawne (zob. Gizbert-Studnicki 1986: 52). Odbiorcami aktów normatywnych są w pierwszej kolejności podmioty będące adresatami norm (Gizbert-Studnicki 1986: 52), lecz grono faktycznych odbiorców jest znacznie szersze. Unijne akty prawne są tłumaczone przez wewnętrzne służby tłumaczeniowe instytucji i w przeciwieństwie do

innych gatunków po przetłumaczeniu mają status równorzędnych, autorytatywnych i wiążących wersji językowych (por. Šarčević 1997: 64).

Spośród gatunków prawniczych akty normatywne stanowią najczęściej badany gatunek, zarówno w ujęciu jakościowym, jak i ilościowym, głównie ze względu na swoją rangę oraz powszechną dostępność. Są one przedmiotem badań³ głównie pod kątem techniki prawodawczej na gruncie nauk prawnych (np. Wronkowska i Zieliński 1993, 2012; Malinowski 2010, 2012), aspektów językowych i terminologicznych (np. Gizbert-Studnicki 1986; Jadacka 1996; Hałas 2005; Malinowski 2006; Choduń 2007, Łapa 2008, Przetak 2015), jak i zagadnień translatorskich (Jopek-Bosiacka 2006; Biel 2014, 2017). Akty prawne charakteryzują się wysokim stopniem standaryzacji i są dość szczegółowo regulowane przez zasady techniki prawodawczej na gruncie polskim i analogiczne zalecenia na gruncie unijnym (np. Wronkowska i Zieliński 1993, 2012; Malinowski 2010, 2012). Do podstawowych cech gatunkowych aktów prawnych zalicza się m.in. struktury warunkowe (hipotetyczne), depersonalizację, stronę bierną i nominalizację, odesłania intra- i intertekstualne, wysoką częstość występowania terminologii prawnej (zob. szerzej Biel 2014).

2.2. Orzeczenia

Orzeczenia są to czynności procesowe sądu o charakterze decyzyjnym (Cioch i Studzińska 2012: 281). Orzeczenia stanowią w głównej mierze instrument sądowego stosowania prawa służący interpretacji (wykładni) norm prawnych, regulowaniu zachowania stron postępowania, a także zapewnianiu jednolitego stosowania prawa (zob. Rzucidło-Grochowska 2015: 62). W zależności od grupy odbiorców orzeczenia sądowe pełnią funkcję dyrektywną, informacyjną lub wyjaśniającą (argumentacyjną).

3. Z racji mnogości opracowań dla gatunku aktów normatywnych przegląd literatury ograniczony został do wybranych najważniejszych pozycji dotyczących polskiego języka prawnego.

Celem komunikacyjnym orzeczeń jest przekonanie stron postępowania do zasadności decyzji podjętej przez sąd (funkcja informacyjna i wyjaśniająca), a także nałożenie obowiązku zastosowania się do niej (funkcja dyrektywna orzeczeń) (zob. Salmi-Tolonen 2005: 61–64, Szczyrbak 2014: 118–127). Oprócz stron postępowania, będących prymarnymi odbiorcami orzeczeń, wymienić można odbiorców wtórnych, którymi są m.in. podmioty władzy wykonawczej i ustawodawczej, a także odbiorców dalszych, czyli prawników, badaczy prawa, opinię publiczną i polityków (funkcja informacyjna, wyjaśniająca orzeczeń) (zob. Rzucidło-Grochowska 2015: 62). Autorami orzeczeń są organy sądowe – podmioty instytucjonalne i zbiorowe działające na podstawie autorytetu i uprawnień nadanych im w ustawie zasadniczej (w przypadku sądów krajowych) bądź źródła prawa pierwotnego (w przypadku sądów unijnych). W procesie stosowania prawa główną rolę odgrywają sędziowie wchodzący w skład sędziowski rozpatrujący poszczególne sprawy.

Orzeczenia jako gatunek prawniczy nie zostały jeszcze dogłębnie zbadane pod kątem aspektów językowych i terminologicznych. Badacze zajmowali się jednak wybranymi zagadnieniami związanymi z językiem orzeczeń, takimi jak: strategię argumentacyjne (np. Mazzi 2007, 2010, 2015; Szczyrbak 2009, 2014), elementy wartościujące (np. Goźdz-Roszkowski 2013), wykładniki modalności (np. Cheng i Cheng 2014; Szczyrbak 2017) i zagadnienia frazeologiczne (np. Jablonkai 2010; Goźdz-Roszkowski 2011; Breeze 2013; Pontrandolfo 2016; Tománková 2016). Nieliczni badacze zajmowali się również szerszym opisem orzeczeń jako gatunku (np. Bhatia 1993; Vazquez-Orta 2010; Szczyrbak 2014), a także przekładem orzeczeń (np. Engberg 1999; Jopek-Bosiacka 2006; Matulewska 2007). Charakterystyczną cechą orzeczeń jest wysoki stopień standaryzacji i formalizacji. Aby orzeczenie zostało uznane za spełniające wymogi formalne, powinno zawierać wiele elementów określonych we właściwych przepisach prawa (Bładowski 2012: 24). Podstawowymi cechami tego gatunku są m.in. depersonalizacja, pasywizacja, częste występowanie spójników wynikowych, częste odesłania

do aktów prawnych i istniejącego orzecznictwa (intertekstualność) wraz z odesłaniami intratekstualnymi (Vazquez-Orta 2010), a także bogactwo wyrażeń modalizujących, które odgrywają istotną rolę w argumentacji sądów, gdyż wzmacniają siłę perswazji występujących w orzeczeniach uzasadnień, decydując tym samym o wartości rozstrzygnięcia sądu (Gapska 2009).

2.3. Sprawozdania

Sprawozdania stanowią grupę tekstów o funkcji informacyjnej, które zawierają opisy działalności podmiotów, opisy procesów lub zjawisk w określonej dziedzinie. Niektóre z nich zawierają same dane liczbowe, z kolei inne to wielostronicowe opracowania zawierające elementy multimodalne. W projekcie EUROLEKT badane są wyłącznie sprawozdania opracowane przez organy administracji publicznej, co pozwala zaklasyfikować je do podgatunku sprawozdań administracyjnych. Ich autorami są urzędnicy instytucji i agencji unijnych oraz polskich ministerstw. Celem sprawozdań administracyjnych co do zasady jest komunikowanie wiedzy specjalistycznej za pośrednictwem danych faktograficznych. Często obowiązek sporządzania sprawozdań wynika z przepisów prawa (rozporządzeń, dyrektyw, ustaw). Część z nich ma jednak również na celu kreowanie pozytywnego obrazu instytucji i jest narzędziem komunikowania się instytucji z szerokim gronem odbiorców. W tej grupie sprawozdań stosuje się więcej elementów multimodalnych.

Gatunek sprawozdań nie został dotychczas wyczerpująco opisany w literaturze. Do tej pory badacze zajmowali się głównie analizą sprawozdań z działalności akademickiej (Wharton 2007; Parkinson 2017; Friginal 2013; Hyon 2008), (medialnych) raportów wydarzeń (np. Renkema 1993; van Dijk 1988a, 1988b; zob. także Duszak 1998: 175–187) i raportów biznesowych, np. rocznych sprawozdaniań spółek (Garzone 2006a, 2006b), raportów z badania sprawozdań finansowych i opinii biegłych rewidentów (Flowerdew i Wan 2010; Ditlevsen i Engberg 2007),

raportów rynkowych i doradczych – przygotowywanych zarówno na zlecenie zewnętrznych klientów, jak i na użytek wewnętrzny (Berber i Paulo 1997; Yeung 2007; Nathan 2016). Sprawozdania były badane pod kątem wybranych elementów językowych, np. metafor czy wykładników modalności (m.in. Kruger 1997; Charteris-Black i Ennis 2001; Charteris-Black i Musolff 2003; Guangrong 2009; Spence i Liu 2013; Crosthwaite, Cheung i Jiang 2017).

Sprawozdania administracyjne charakteryzuje styl urzędowo-kancelaryjny. Poziom standaryzacji tekstów jest bardzo niski. W przeciwieństwie do sprawozdań spółek, dla raportów instytucjonalnych z reguły nie istnieją wytyczne ani uzus, które determinowałyby strukturę tekstów lub zakres treści (por. Garzone 2006b: 314). Skutkuje to dużą różnorodnością makrostruktury sprawozdań, które liczą od kilkunastu do kilkudziesięciu stron. W warstwie mikrostruktury w tekstach występuje silna tendencja do stosowania form analitycznych. Przejawia się to wysokim stopniem nominalizacji i wysoką frekwencyjnością przyimków. Często jest także stosowanie form bezosobowych, takich jak strona bierna, konstrukcje czasownikowe z *-no*, *to* oraz z *się*. Przejawem ekonomii komunikacyjnej jest powszechne stosowanie skrótów i wyciężeń.

Stopień konwencjonalizacji i schematyzacji poszczególnych gatunków w obrębie dyskursu administracyjnego zbadamy poniżej w części empirycznej, posługując się zbitkami wielowyrazowymi jako miarą stopnia formuliczności.

3. Podejście frekwencyjne w badaniach nad frazeologią – zbitki wielowyrazowe

Językoznawstwo korpusowe – rozwijające się intensywnie od końca lat 80. – przyczyniło się najpierw do wzrostu zainteresowania frazeologią, a następnie do diametralnej zmiany rozumienia frazeologii. Dzięki nowym narzędziom i metodom językoznawstwo korpusowe wykazało wysoką formuliczność i schematyczność języka, przesuując centrum

uwagi ze słowa na schematy / wzorce (ang. *patterns*) leksykalno-gramatyczne – innymi słowy, językowe prefabrykaty uznawane za podstawową jednostkę znaczeniową (Stubbs 2004: 118). W rezultacie w nurcie **frazeologii korpusowej** nastąpiło przededefiniowanie i rozszerzenie kategorii frazeologii. Dodano nowe typy łączliwości wyrazów, wskazując na ich centralność ze względu na wysoką częstość użycia, jednocześnie marginalizując uprzednio kluczowe typy stałych związków frazeologicznych, takich jak: idiomy, przysłowia, porzekadła czy aforyzmy, z powodu ich rzadkiego użycia w języku (Nesselhauf 2005: 12). W podejściu korpusowym, zwanym podejściem frekwencyjnym (ang. *frequency-based approach*), frazeologizmy identyfikuje się na podstawie ich powtarzalności w tekście, tj. wysokiej częstości, a nie na podstawie kryterium współwystępowania – stopnia zespolenia związku frazeologicznego czy jego kompozycyjności (por. Granger i Paquot 2008: 28–32). Związki frazeologiczne będące w centrum zainteresowania podejścia frekwencyjnego to kolokacje i różnego typu ciągle i nieciągle formy wielowyrzowe, takie jak zbitki wielowyrzowe, ramy frazowe czy skipgramy (por. Nesselhauf 2005: 12; Greaves i Warren 2010: 213).

Najczęstszym obiektem badań w podejściu frekwencyjnym są **zbitki wielowyrzowe** (ang. *lexical bundles*), zwane również w literaturze polskojęzycznej zbitkami leksykalnymi, wiązkami leksykalnymi, jednostkami wielowyrzowymi czy n-gramami⁴.

W niniejszej pracy przyjmujemy za Ł. Grabowskim (2015b) nazwę *zbitki wielowyrzowe*, gdyż lepiej oddaje ona istotę tego zjawiska niż częstszy choć odwołujący się wyłącznie do aspektu leksykalnego termin *zbitki leksykalne*. Zbitki wielowyrzowe identyfikuje się na podstawie kryterium częstości występowania (Biber i Barbieri 2007: 264; Hyland 2008: 6) jako powtarzające się ciągi wielowyrzowe, które nie muszą być kompletnymi jednostkami semantycznymi czy gramatycznymi

4. Również na gruncie anglojęzycznym zbitki mają nieustabilizowane nazewnictwo i znane są pod różnymi synonimicznymi nazwami: *lexical bundles*, *clusters*, *n-grams*, *chunks*, *lexical phrases*, *multi-word units* (MWU).

(Biber i in. 1999: 58–59) – mogą stanowić fragmenty zdań, np. *nie później niż w, o którym mowa w, Parlamencie Europejskiego i Rady*. O ile tradycyjne frazeologizmy pełnią głównie funkcję ornamentacyjną i stylistyczną (por. Grabowski 2015a: 82), o tyle nowe typy frazeologizmów pełnią ważne funkcje dyskursywne. Badania wskazują, iż zbitki leksykalne odgrywają rolę „budulca” w dyskursie – stanowią rozpoznawalne, stabilne ramy dla dłuższych wyrażen, są szybko aktywowane z pamięci i uzupełniane nowymi informacjami (Biber i Barbieri 2007: 270). Badania dowodzą ponadto, iż zbitki wielowyrazowe są swoistymi „odciskami palców” gatunku czy rejestru (Handford 2010: 122) – mogą być zatem narzędziem do oceny zróżnicowania gatunkowego, gdyż ich układ jest na ogół specyficzny dla danego gatunku (por. Biber i in. 1999; Hyland 2008: 7). Zbitki wielowyrazowe można poddać dalszej kategoryzacji za pomocą kryteriów strukturalnych (struktura gramatyczna zbitek identyfikowana na podstawie dominującej części mowy, np. zbitki rzeczownikowe, czasownikowe czy przyimkowe) (Biber i in. 2004: 381) bądź funkcjonalnych, które uwzględniają funkcję zbitki w tekście. W popularnej klasyfikacji – opracowanej dla najczęściej badanego tą metodą dyskursu akademickiego – wyróżnia się zbitki **referencyjne** (ang. *referential bundles*), które oznaczają uczestników i obiekty, zbitki **modalne** (ang. *stance bundles*), służące wyrażaniu postaw i opinii, oraz zbitki **metatekstowe** (ang. *discourse organizers*) strukturyzujące i porządkujące dyskurs (Biber i Barbieri 2007: 265).

Zbitki wielowyrazowe badano przede wszystkim w dyskursie akademickim, głównie anglojęzycznym (zob. szerzej Biber i Barbieri 2007) i w ograniczonym zakresie w innych dyskursach: np. dyskurs farmaceutyczny (Grabowski 2013; 2014; 2015a), prasowy (Lee 2011, 2013), gatunek stron internetowych (Santini 2005), gatunek spotkań biznesowych (Handford 2010) i konwersacje (Conrad i Biber 2004). Zbitki wielowyrazowe badano również w tłumaczeniach (Tirkkonen-Condit 2005; Kemppanen 2004; Dayrell 2008; Balaskó 2008; Fiedler 2017; Lee 2011, 2013).

W szczególności niewiele badań przeprowadzono dla będącego przedmiotem niniejszego artykułu dyskursu prawnego i prawniczego czy ogólnie dyskursu administracyjnego, a ściśle – unijnego. Wśród pierwszych prac dla prawniczego języka angielskiego należy wymienić artykuły R. Jablonkai (2010) i R. Breeze (2013) oraz książkę S. Goździa-Roszkowskiego (2011), natomiast dla języka polskiego – Ł. Biel (2017). Badanie R. Jablonkai (2010) zostało przeprowadzone na relatywnie niewielkim korpusie anglojęzycznych tekstów unijnych z różnych gatunków (dotyczy więc ogólnie dyskursu unijnego) na tle korpusu narodowego British National Corpus (BNC)⁵. Wyniki badania wskazują na wysoką formułiczność tekstów unijnych na tle korpusu referencyjnego (zwłaszcza w porównaniu z podkorpusem powieści, artykułów prasowych czy tekstów ogólnych BNC), o czym świadczy bardzo duża liczba typów zbitek wielowyrazowych i łączna liczba ich wystąpień (Jablonkai 2010: 258).

Badania S. Goździa-Roszkowskiego (2011) i R. Breeze (2013) nie dotyczą co prawda dyskursu unijnego ani tekstów tłumaczonych, ale są relewantne dla niniejszego opracowania, gdyż stosują metodę zbitek wielowyrazowych do zbadania zróżnicowania kilku angielskich gatunków prawniczych. W przeciwieństwie do badania Jablonkai analizy przeprowadzono odrębnie dla każdego gatunku, a następnie skontrastowano wyniki. S. Goźdz-Roszkowski (2011) całościowo zbadał dystrybucję zbitek wielowyrazowych w siedmiu amerykańskich gatunkach prawnych i prawniczych (czasopisma naukowe, akta spraw, umowy, akty prawne, opinie, literatura fachowa, podręczniki akademickie). Zbliżone badanie R. Breeze (2013) zostało przeprowadzone na czterech brytyjskich gatunkach prawnych i prawniczych (artykuły naukowe, akty prawne, orzecznictwo i dokumenty prawnicze/umowy), przy czym badaczka kontrolowała zmienną tematyczną przez ograniczenie zakresu

5. Warto odnotować, iż bardziej adekwatnym korpusem referencyjnym byłby korpus brytyjskich tekstów administracyjnych zbudowany według porównywalnych kryteriów, zwłaszcza uwzględniający podobne gatunki tekstów.

korpusu wyłącznie do prawa handlowego. Obie prace wskazują, iż najbardziej formulicznymi gatunkami są akty prawne i umowy, gdyż posiadają największy zakres i liczbę zbitek oraz największą proporcję słów w zbitkach.

W pierwszym badaniu zbitek wielowyrazowych w polskich tekstach administracyjnych Ł. Biel (2017) przeprowadziła analizę zbitek w przetłumaczonych na język polski unijnych aktach prawnych na tle analogicznych krajowych aktów normatywnych. Badanie wykazało wyższy stopień formuliczności tłumaczeń unijnych, zarówno jeżeli chodzi o zakres, jak i łączną liczbę zbitek, zwłaszcza w przypadku zbitek mających co najmniej pięć słów. Ponadto analiza wykazała, że tłumaczenia mają bardzo mało wspólnych zbitek z referencyjnym korpusem krajowym (3% 4-gramów), co może sugerować, że tłumaczenia bazują na swoich własnych n-gramach, tworzonych pod wpływem tekstu źródłowego. Niniejsze badanie stanowi kontynuację powyższego badania, rozszerzając jego zakres o dwa dodatkowe gatunki unijne – orzeczenia i sprawozdania administracyjne – oraz testując różne progi odcięcia dla zbitek.

4. Frazeologia w tłumaczeniu

Od tłumaczeń z reguły oczekuje się, iż w warstwie frazeologicznej będą zbieżne z frazeologią języka docelowego charakterystyczną dla danego gatunku (Gouadec 2007: 23). Zasada ta dotyczy również tłumaczeń prawnych i prawniczych (zob. np. Monzó Nebot 2008: 224), choć ten typ tłumaczenia może charakteryzować się większą tendencją do nie-naturalnej czy nietypowej frazeologii z racji częściowej nieprzystawalności terminologii między systemami prawnymi (por. Biel 2014: 182). Ponadto z racji skomplikowanego dwujęzycznego przetwarzania treści przy częstym przełączaniu kodów językowych proces tłumaczenia może prowadzić do zaburzenia spontanicznego przetwarzania treści typowego dla sytuacji jednojęzycznych (Mauranen 2007: 44). W rezultacie proces

tłumaczenia pozostawia „ślady” w warstwie frazeologicznej tekstów docelowych. Mogą one mieć postać zniekształceń frazeologii w wyniku interferencji z języka źródłowego (Toury 1995: 275), połączonych z blokowaniem „wywoływania” czy „aktywowania” naturalnych frazeologizmów. Innym typem zniekształceń frazeologicznych w tłumaczeniach jest występowanie nietypowych kolokacji polegające na wyższej częstotliwości kolokacji rzadkich (choć prawidłowych) i rzadszym występowaniu kolokacji częstych w języku docelowym, większym zróżnicowaniu kolokacji i ich mniejszej stabilności (Mauranen 2006: 97). Przypuszcza się, iż tłumaczenia charakteryzują się wyższą formulicnością niż nie-tłumaczenia, co wiąże się z prawem rosnącej standaryzacji Toury’ego (1995: 268) oraz hipotezą normalizacji, tj. skłonnością tłumaczy do nadużywania typowych cech języka docelowego (Baker 1996: 183).

5. Materiał badawczy – gatunkowe korpusy porównawcze

Badanie wykonano na wybranych sekcjach korpusu PL Eurolect (zob. Biel 2016), obejmujących korpusy porównawcze dla trzech badanych gatunków: 1) aktów normatywnych, 2) wyroków i 3) sprawozdań. Przez korpus porównawczy rozumiemy korpus jednojęzyczny zawierający porównywalne zbiory tekstów zebranych na podstawie wspólnych ram doboru próby (ang. *sampling frame*) (McEnery i Hardie 2012: 24). W przypadku korpusów wykorzystanych w badaniu główny element ramy stanowił wspólny okres publikacji tekstów, czyli lata 2011–2015.

Tabela 1 pokazuje korpusy zastosowane w badaniu. Głównym korpusem dla każdego gatunku jest korpus unijny (oznaczony jako UE PL), natomiast korpus krajowy (PL) pełni funkcję korpusu referencyjnego. W celu kontrolowania określonych zmiennych korpus aktów normatywnych i korpus wyroków dzieli się na korpusy składowe (odpowiednio rozporządzenia i dyrektywy oraz wyroki Trybunału Sprawiedliwości).

	Nazwa korpusu	Liczba tekstów	Liczba wyrazów tekstowych (<i>tokens</i>)	Liczba wyrazów słownikowych (<i>types</i>)
	Akty normatywne UE PL			
Akty normatywne	Rozporządzenia	925	1 899 403	39 409
	Dyrektywy	92	768 187	24 288
	<i>Ogółem</i>	<i>1 017</i>	<i>2 667 590</i>	<i>45 249</i>
	Akty normatywne PL			
	Ustawy	144	1 859 880	35 142
	Wyroki UE PL			
Orzeczenia	Wyroki Trybunału	897	4 693 968	72 310
	Wyroki Sądu	384	4 515 866	65 318
	<i>Ogółem</i>	<i>1 281</i>	<i>9 209 834</i>	<i>97 335</i>
	Wyroki PL			
	Wyroki Sądu Najwyższego RP	2 564	6 833 997	76 783
	Sprawozdania UE PL			
Sprawozdania	Sprawozdania instytucji unijnych	498	2 329 498	60 466
	Sprawozdania PL			
	Sprawozdania polskich ministerstw	110	2 556 114	77 417

Tabela 1. Statystyki dla korpusu PL Eurolekt. Źródło: opracowanie własne.

Korpus unijnych **aktów prawnych** składa się z dwóch korpusów składowych zawierających rozporządzenia i dyrektywy pobrane z portalu EUR-Lex⁶. W celu zapewnienia lepszej porównywalności z korpusem krajowym z tekstów unijnych usunięto rozbudowane preambuły i załączniki, pozostawiając wyłącznie część normatywną. W korpusie krajowych

6. <http://eur-lex.europa.eu/homepage.html?locale=pl>.

aktów prawnych znalazły się ustawy uchwalone⁷ w latach 2011–2015 w wersji obowiązującej na dzień 31 grudnia 2015 roku, zaczerpnięte z bazy Lex spółki Wolters Kluwer SA⁸.

W skład korpusu **orzeczeń** wchodziły wyroki wydane przez dwa organy Trybunału Sprawiedliwości Unii Europejskiej (TSUE), tj. Trybunał Sprawiedliwości oraz Sąd, natomiast porównawczy korpus krajowy zawiera wyroki wydane przez Izbę Cywilną Sądu Najwyższego Rzeczypospolitej Polskiej (SN). Teksty pobrano pod koniec 2015 r. z portalu InfoCuria (w przypadku TSUE) oraz bazy orzeczeń SN⁹.

W ramach gatunku **sprawozdań** porównane zostaną dwa korpusy: sprawozdań unijnych sporządzonych przez dyrekcje generalne Komisji Europejskiej (Sprawozdania UE), pobranych z rejestru dokumentów Komisji, oraz sprawozdań krajowych sporządzonych przez polskie ministerstwa (Sprawozdania PL) zaczerpniętych ze stron internetowych ministerstw. W celu zapewnienia porównywalności korpusów korpus sprawozdań krajowych został ograniczony do tekstów ministerstw odpowiadających zakresem kompetencji dyrekcjom generalnym Komisji, tj. nie zostały do niego włączone raporty Ministerstwa Obrony Narodowej czy Skarbu Państwa.

Badanie wykonano w programie Wordsmith Tools 7.0 (Scott 2016). Z powodu różnej wielkości korpusów wyniki znormalizowano do 1 miliona słów.

7. Korpus krajowych aktów normatywnych zastosowany w niniejszym badaniu różni się nieznacznie od korpusu z badania Biel (2017) – obejmuje wyłącznie akty prawne z lat 2011–2015.

8. <https://sip.lex.pl/#/>.

9. http://www.sn.pl/orzecznictwo/SitePages/Baza_orzeczen.aspx.

6. Metodologia

Zbitki wielowyrazowe identyfikuje się metodami ilościowymi (Biber i Barbieri 2007: 264, Hyland 2008: 6), uwzględniając następujące kryteria dotyczące długości i dystrybucji:

1. długość zbitki (w zależności od liczby wyrazów identyfikowane zbitki to 3-gramy [ciągi trzywyrazowe], 4-gramy [ciągi czterowyrzowe] itd.),
2. wartość progowa częstości występowania zbitek,
3. wartość progowa rozproszenia (zakresu dystrybucji) zbitek w różnych tekstach w korpusie w celu wyeliminowania użyć idiosynkratycznych (Biber i Barbieri 2007: 268).

Stosowane w badaniach progi cechują się pewnym stopniem arbitralności i zależą w szczególności od długości zbitek, wielkości korpusu oraz stopnia formuliczności gatunku. W przypadku wartości progowej rozproszenia najczęściej przyjmuje się próg wystąpienia zbitki w co najmniej pięciu tekstach w korpusie (Biber i Barbieri 2007: 267), chociaż w niektórych badaniach przyjmuje się znacznie wyższe progi (np. 10% tekstów w korpusie). Bardziej dyskusyjną i zróżnicowaną kwestią jest próg częstości występowania zbitki. Za tradycyjną i rygorystyczną (tzw. konserwatywną) wielkość progową przyjmuje się 40 wystąpień na milion słów (Biber i Barbieri 2007: 267), jednakże stosowane są również niższe progi, np. 20 czy 10 wystąpień na milion słów. Rozbieżności wynikają również z niedostatecznego opracowania metodologii badań w zakresie frazeologii korpusowej (Römer 2010; Grabowski 2015a: 131; 2015b: 35)¹⁰.

Większość badań wykorzystujących metodę zbitek wielowyrzowych przeprowadzono na języku angielskim, trudno zatem ocenić, jak powinny kształtować się wartości progowe dla typologicznie

10. Przyjmowane progi są w dużej mierze uzależnione od gatunku (zwłaszcza jego schematyczności), wielkości korpusu i typu badanych zbitek.

odmiennych języków (por. Gray i Biber 2015: 144). Dotyczy to w szczególności języków fleksyjnych, w tym będącego przedmiotem niniejszego badania języka polskiego. Wydaje się, że ze względu na liczne warianty fleksyjne należałoby przyjąć niższy próg częstości.

Kolejną kwestią dyskusyjną jest długość badanych zbitek. Dla języka angielskiego najczęściej bada się ciągi czterowyrazowe (4-gramy), co podyktowane jest bardziej względami praktycznymi niż teoretycznymi (por. Gries 2010; Jablonkai 2010; Gray i Biber 2015: 136; Greaves i Warren 2010: 214; Grabowski 2015a). W przypadku porównywania zbitek na potrzeby badań przekładoznawczych należy postawić pytanie, czy zbitki czterowyrazowe rzeczywiście odzwierciedlają ten sam poziom formuliczności w różnych językach (Biel 2017), zwłaszcza czy angielskie 4-gramy odpowiadają 4-gramom w języku polskim.

W części ilościowej zbadano dystrybucję n-gramów w trzech gatunkach, zestawiając korpusy tłumaczeń unijnych z analogicznymi korpusami, zarówno pod kątem typów zbitek, jak i ich łącznej liczby wystąpień. W badaniu dokonano identyfikacji stopnia formuliczności analizowanych gatunków na podstawie zbitek wielowyrazowych, począwszy od 2- do 8-gramów, testując trzy progi częstości (40, 20 i 10 wystąpień na milion słów) oraz trzy progi dystrybucji zbitek w tekstach (5 tekstów, 5% i 10% tekstów)¹¹. Wybrane progi przedstawia tabela 2.

Podczas wyszukiwania zbitek program do analizy korpusowej WordSmith omijał zbitki zawierające liczby i daty, ramy frazowe, a także nie tworzył zbitek ze słowami występującymi przed znakami kończącymi zdanie i po nich.

11. W celu zapewnienia przejrzystości prezentacji danych liczbowych w niniejszym artykule zamieszczono wyłącznie dane dla progów częstości 40 i 10 wystąpień na milion słów (pomijając próg 20 wystąpień) oraz progów rozproszenia – 5 tekstów i 5% tekstów.

	Nazwa korpusu	Liczba wyr- zów tek- stowych (tokens)	Próg czę- stości 40 na 1 mln	Próg czę- stości 10 na 1 mln	Liczba tekstów	Próg dystry- bucji 5% tekstów	Średnia dl. tekstu w wyrazach tekstowych
Akty normatywne UE PL							
Akty normatywne	Rozporządzenia	1 899 403	76	19	925	46	2 053
	Dyrektywy	768 187	31	8	92	5	8 350
Akty normatywne PL							
	Ustawy	1 859 880	74	19	144	7	12 916
Wyroki UE PL							
Orzeczenia	Wyroki Trybunału	4 693 968	188	47	897	45	5 233
	Wyroki Sądu	4 515 866	181	45	384	19	11 760
	Wyroki PL						
	Wyroki Sądu Najwyższego RP	6 833 997	273	68	2 564	128	2 665
Sprawozdania UE PL							
Sprawozdania	Sprawozda- nia instytucji unijnych	2 329 498	93	23	498	25	4 678
	Sprawozdania PL						
	Sprawozda- nia polskich ministerstw	2 556 114	102	26	110	6	23 237

Tabela 2. Wybrane progi częstości i rozproszenia zastosowane przy identyfikacji zbitek. Źródło: opracowanie własne.

7. Globalna dystrybucja n-gramów w korpusach krajowych i unijnych

Dzięki przetestowaniu trzech progów częstości i dwóch progów dystrybucji zbitek uzyskano złożony i niejednoznaczny obraz formułczości. Badanie pokazuje, iż ocena formułczości jest ściśle powiązana z przyjętymi progami, a uwzględnienie tylko jednego progu może zniekształcić jej obraz.

Analizę rozpoczniemy od dystrybucji zbitek w korpusie referencyjnym gatunków krajowych, które stanowią tło dla gatunków unijnych. Tabela 3 pokazuje sumaryczne wyniki dla 2- do 8-gramów przy wysokim (40) i niskim (10) progu częstości oraz dwóch progach dystrybucji – 5 tekstów i 5% tekstów.

	Progi		Akty normatywne PL	Orzeczenia PL	Sprawozdania PL
2-8 gramy	1. wyst. na 1 mln	Teksty (min.)			
	40	5 tekstów	428 331	392 371	133 686
Łączna liczba wystąpień	40	5%	418 824	383 830	132 019
	10	5 tekstów	638 749	613 958	268 613
	10	5%	594 244	452 943	260 111
Łączna liczba typów	40	5 tekstów	2 877	2 957	1 284
	40	5%	2 736	2 812	1 264
	10	5 tekstów	14 347	15 615	8 785
	10	5%	12 035	5 200	8 315

Tabela 3. Dystrybucja n-gramów (od 2- do 8-wyrazowych) w gatunkach krajowych, częstość znormalizowana do 1 miliona słów¹². Źródło: opracowanie własne.

12. Z racji zastosowania przy identyfikacji zbitek znormalizowanych progów częstości nie dokonywano dodatkowej normalizacji typów zbitek, natomiast normalizacji poddano łączną liczbę wystąpień.

Jeżeli chodzi o łączną liczbę wystąpień wszystkich zbitek, zauważalny jest wyraźny trend: najwyższą formuliczością cechują się – w przypadku wszystkich progów – akty normatywne, jednakże orzeczenia krajowe, które mogłyby wydawać się mniej schematyczne i standaryzowane, mają niewiele mniejszą łączną liczbę zbitek od aktów normatywnych. Na tle aktów normatywnych i orzeczeń sprawozdania mają trzykrotnie mniej zbitek występujących co najmniej 40 razy na milion słów i dwukrotnie mniej zbitek o znormalizowanej częstości 10 wystąpień, czyli można je uznać za znacznie mniej formuliczne pod kątem łącznej liczby zbitek.

Inną badaną miarą formuliczości jest liczba typów zbitek, a więc ich różnorodność. W ujęciu zbiorczym od 2- do 8-gramów, dla wszystkich progów oprócz 10 wystąpień na milion słów i 5% tekstów, występuje podobna zależność – liczba typów zbitek jest niemal identyczna dla aktów prawnych i orzeczeń oraz o połowę mniejsza w sprawozdaniach. Tak wysoka zbieżność stopnia formuliczości w krajowych aktach prawnych i orzeczeniach jest zaskakująca z racji zakładanej większej idiosynkratyczności orzeczeń (w przypadku aktów normatywnych autor jest znacznie bardziej kolektywny) i wskazuje na istotne podobieństwo między tymi dwoma gatunkami w porównaniu do gatunku sprawozdań, które najwyraźniej podlegają mniejszej standaryzacji. Rozbieżności, które pojawiają się dla progu 10 wystąpień i 5% tekstów (tj. najwyższa formuliczość aktów prawnych, a najmniejsza orzeczeń), należałoby interpretować w kontekście niepełnej porównywalności korpusów dla progu 5%, który dla korpusu orzeczeń krajowych oznacza aż 128 tekstów, natomiast dla korpusu sprawozdań zaledwie 6 tekstów.

W następnej części zbadamy, czy korpus unijny wykazuje podobne zależności między gatunkami jak korpus krajowy. Tabela 4 pokazuje analogiczne sumaryczne wyniki dla zbitek unijnych od 2- do 8-gramów.

	Progi		Akty normatywne UE		Orzeczenia UE		Sprawozdania UE
	l. wyst. na 1 mln	Teksty (min.)	Rozpo- rządze- nia	Dyrek- tywy	Trybunał	Sąd	
	40	5	574 234	455 243	393 364	350 147	188 867
		tekstów					
Łączna liczba	40	5%	518 121	455 243	386 215	347 656	182 119
wystą- pień	10	5	797 807	913 447	617 876	595 510	359 018
		tekstów					
	10	5%	529 734	913 447	532 983	572 393	277 279
		tekstów					
Łączna liczba	40	5%	2 840	3 740	3 255	2 763	1 585
typów	10	5	16 435	30 956	15 957	16 759	11 523
		tekstów					
	10	5%	3 213	30 956	10 306	15 372	5 999

Tabela 4. Dystrybucja n-gramów (od 2- do 8-wyrazowych) w gatunkach unijnych, częstość znormalizowana do 1 miliona słów. Źródło: opracowanie własne.

Jeżeli chodzi o łączną liczbę wystąpień zbitok, gatunki unijne co do zasady ukazują podobne wzajemne zależności jak gatunki krajowe, tj. najbardziej formuliczne są unijne akty normatywne, następnie orzeczenia, a najmniej formuliczne są sprawozdania. Odmienna jest jednak skala różnic – formuliczność orzeczeń unijnych w stosunku do aktów normatywnych jest znacznie mniejsza niż w przypadku gatunków krajowych.

Podobne zależności zauważalne są również w odniesieniu do liczby typów zbitok – najbardziej zróżnicowane zbitki występują w aktach normatywnych i orzeczeniach, natomiast w sprawozdaniach unijnych jest ich niemal dwukrotnie mniej. Rozbieżności ponownie pojawiają się przy progu 10 wystąpień na milion słów i 5% tekstów, tj. można by wnioskować, iż najbardziej formuliczne są dyrektywy, a najmniej rozporządzenia, jednakże należy uwzględnić fakt, iż korpus dyrektyw

jest najmniejszy¹³, a zatem zbitki mają najniższe znormalizowane progi, natomiast korpus rozporządzeń ma dużo tekstów i jego próg 5% wynosi 46 tekstów, podczas gdy dla dyrektyw 5. Zatem próg 5% przy 10 wystąpieniach na milion słów wydaje się być najbardziej wrażliwym na wielkość i skład korpusu. Co ciekawe, ten problem neutralizuje się przy progu 40 wystąpień na milion słów, co oznacza, że zbitki o wysokiej częstości są z reguły na tyle dobrze rozproszone w korpusie, iż odcięcie na poziomie 5% tekstów ma na nie niewielki wpływ.

Reasumując, korpusy krajowe i unijne wykazują zbliżone relacje między gatunkami w zakresie formuliczności.

Jeżeli chodzi o formuliczność tłumaczeń na tle analogicznych tekstów nieprzetłumaczonych, tj. tekstów unijnych na tle tekstów krajowych (zob. tabele 3 i 4), hipotezę o większej formuliczności tłumaczeń można potwierdzić – z wyjątkiem problematycznego progu – dla gatunku aktów prawnych i sprawozdań, zarówno pod kątem łącznej liczby zbitek, jak i typów zbitek. Ten trend jest szczególnie widoczny w odniesieniu do typów zbitek przy progu 40 wystąpień na milion słów (a więc częstszych zbitek) i dla progu 10 wystąpień na milion słów i rozproszenia w co najmniej 5 tekstach. W przypadku natomiast orzeczeń formuliczność tłumaczeń jest bardzo zbliżona do tekstów nieprzetłumaczonych. Jeżeli chodzi o łączną liczbę zbitek, większa formuliczność tłumaczeń jest statystycznie istotna¹⁴ w przypadku sprawozdań (wszystkie progi; współczynnik Bayesa (BIC) wynosi od 1332 do 31908) oraz aktów normatywnych (BIC na poziomie od 6949 do 40013), natomiast ponownie w przypadku orzeczeń nie zaobserwowano większych różnic pomiędzy korpusem unijnym a krajowym (BIC wynosi od -9 do 53). Uzyskane wyniki potwierdzają wielowymiarowy i niejednoznaczny charakter formuliczności.

13. Można przypuszczać, iż normalizacja nie neutralizuje w pełni różnic w wielkości korpusów, a zbitki nie przyrastają proporcjonalnie do wzrostu liczby słów w korpusie.

14. Do oceny istotności statystycznej zastosowano wskaźnik LL oraz współczynnik Bayesa.

8. Analiza 3-gramów w dyskursie administracyjnym: dystrybucja i funkcje

Następnie przejdziemy do analizy formułiczności poszczególnych gatunków na poziomie zbitok trzywyrazowych (3-gramy) (zob. tabela 5). W toku badania do dalszej analizy wybrano 3-gramy¹⁵, uznając je za najbardziej zbliżone do angielskich 4-gramów z racji kodowania w języku polskim informacji fleksyjnych w afiksach i braku rodzajników (*a, the*).

3-gramy	Progi		Akty normatywne UE		Akty normatywne PL
	l. wyst. na 1 mln	Teksty (min.)	Rozporządzenia	Dyrektywy	
	40	5	111 044	98 632	95 149
Łączna liczba wystąpień	40	5%	99 824	98 632	92 915
	10	5	162 894	191 155	143 450
	10	5%	102 052	191 155	133 130
	40	5	758	866	605
Łączna liczba typów	40	5%	573	866	574
	10	5	3 748	6 138	3 259
	10	5%	643	6 138	2 729
3-gramy	Progi		Orzeczenia UE		Orzeczenia PL
	l. wyst. na 1 mln	Teksty (min.)	Trybunał	Sąd	
	40	5	73 507	68 330	77 275
Łączna liczba wystąpień	40	5%	72 273	67 930	75 220
	10	5	130 984	130 732	130 077
	10	5%	108 399	125 005	91 558
	40	5	660	611	615
Łączna liczba typów	40	5%	637	604	580
	10	5	3 803	4 204	3 673
	10	5%	2 398	3 831	1 157

15. Do identyfikacji 3-gramów zastosowano program Wordsmith Tools 7.0 bez eliminowania nachodzących na siebie zbitok (ang. *overlapping bundles*) i zbitok zawierających się w dłuższych ciągach (ang. *subsumption*) (szerzej zob. Chen i Baker 2010: 33; Biel 2017).

3-gramy	Progi		Sprawozdania UE	Sprawozdania PL
	l. wyst. na 1 mln	Teksty (min.)		
	40	5	61 301	39 408
Łączna liczba wystąpień	40	5%	58 894	38 445
	10	5	141 199	107 551
	10	5%	102 045	101 407
	40	5	288	182
Łączna liczba typów	40	5%	269	177
	10	5	2 315	1 691
	10	5%	1 139	1 560
	40	5	288	182

Tabela 5. Dystrybucja 3-gramów w dyskursie administracyjnym.

Źródło: opracowanie własne.

Analiza 3-gramów potwierdza tendencje zauważone dla globalnej dystrybucji 2-8 gramów. Tłumaczenia aktów normatywnych i sprawozdań cechują się większą formulicnością niż teksty krajowe pod względem typów zbitek (ich zróżnicowania), natomiast w orzeczeniach liczba typów jest bardzo zbliżona. We wszystkich korpusach (oprócz dyrektyw) następuje odwrócenie trendu przy problematycznym progu 10 wystąpień na milion słów i 5% tekstów, co należałoby tłumaczyć różnicami w strukturze korpusów¹⁶. Jeżeli chodzi o łączną liczbę zbitek, największe różnice co do stopnia formulicznego zauważalne są w przypadku sprawozdań – w sprawozdaniach unijnych liczba ta jest znacznie większa niż w sprawozdaniach krajowych, co może implikować, iż gatunek ten podlega w większym stopniu prawu rosnącej standaryzacji Toury’ego (1995: 275)¹⁷. Również unijne akty normatywne mają większą łączną liczbę 3-gramów niż akty krajowe. Natomiast w przypadku orzeczeń nie

16. Możliwa jest również interpretacja, iż gatunek sprawozdań i aktów normatywnych cechuje się większą liczbą typów zbitek o relatywnie niskiej częstotliwości i znacznym rozproszeniu w korpusie w odmianie krajowej, natomiast w przypadku orzeczeń – w odmianie unijnej.

17. Z racji niemal identycznej wielkości krajowego i unijnego korpusu sprawozdań ten gatunek cechuje się najwyższym stopniem porównywalności.

odnotowano istotnych różnic pomiędzy korpusem krajowym a unijnym, z wyjątkiem progu 10 wystąpień i 5% tekstów, przy którym to orzeczenia krajowe są mniej formułiczne, choć różnice nie są duże. Zbliżona formułiczność orzeczeń krajowych i unijnych wymaga pogłębionych badań, zwłaszcza na korpusach o innej strukturze – wstępnie można ją próbować wyjaśnić faktem, iż ten typ tekstów tłumaczą prawnicy-lingwiści posiadający wykształcenie prawnicze i większą wrażliwość na konwencje docelowego języka prawniczego, choć nie można wykluczyć wpływu stopnia porównywalności korpusów pod względem wielkości i liczby tekstów.

W ostatniej części przejdziemy do analizy funkcji najczęstszych piętnastu 3-gramów. Przyjmujemy za D. Biberem i F. Barbieri (2007: 265) podział na zbitki pełniące funkcje **referencyjne** (oznaczające uczestników i obiekty), **metatekstowe** (strukturyzujące i porządkujące dyskurs) i **modalne** (wyrażające postawy i opinie). Ze względu na specyfikę dyskursu administracyjnego konieczne było wyróżnienie w obrębie częstych zbitek metatekstowych dalszych podkategorii, bazując częściowo na klasyfikacji S. Goździa-Roszkowskiego (2011: 113–142), w tym zbitek **wyjaśniająco-precyzujących**, zbitek **przyczynowo-skutkowych** (zawierających zbitki wyrażające cel i warunek) oraz zbitek **nawigacyjnych**, które realizują odesłania intra- i intertekstualne. Warto podkreślić, iż klasyfikacja zbitek jest do pewnego stopnia arbitralna, gdyż zbitka może: (1) równocześnie pełnić więcej niż jedną funkcję w zdaniu, np. zbitka *dyrektywy w sprawie* tworzy obiekt w zdaniu *Poniżej przedstawiono główne ustalenia wynikające z oceny dyrektywy w sprawie przyrządów pomiarowych*, jednocześnie pełniąc funkcję odesłania intertekstualnego do innego dokumentu, a więc pełni funkcję nawigacyjną; (2) zbitka może pełnić różne funkcje w zależności od kontekstu, np. zbitka *ze względu na* może pełnić funkcję precyzującą (zawężającą znaczenie) – *ze względu na płeć* lub wyrażać relacje przyczynowo-skutkowe – *ze względu na fakt*, *ze względu na brak*; w takiej sytuacji uwzględniano dominującą funkcję zbitki na podstawie analizy otoczenia

kolokacyjnego. Podobnym przykładem jest przyimek złożony w związku z, który wyraża głównie relację przyczynowo-skutkową w sprawozdaniach i orzeczeniach (*w związku z tym, ~ czym, ~ opóźnieniami, ~ brakiem*), jednakże w orzeczeniach pełni również funkcję nawigacyjną (*w związku z art., ~ umową*), w których wykształcił również formę skrótową pełniącą wyłącznie funkcję odsyłającą *w zw. z art.*

Tabela 6 pokazuje koncentrację najczęstszych piętnastu 3-gramów w poszczególnych korpusach.

Akty normatywne		Orzeczenia		Sprawozdania			
UE	PL	UE	PL	UE	PL		
Rozporządzenia	Dyrektywy	Trybunał	Sąd				
27 829	25 319	42 863	12 799	12 980	16 361	6 655	4 108

Tabela 6. Koncentracja najczęstszych piętnastu 3-gramów w dyskursie administracyjnym (wielkości znormalizowane do 1 miliona słów).

Źródło: opracowanie własne.

Największą koncentracją 3-gramów charakteryzują się akty normatywne – w korpusie krajowym najczęstsze 3-gramy mają łączną dystrybucję rzędu aż 42 863 wystąpień na milion słów, co oznacza, że są niezwykle często powtarzane. W rezultacie tak wysoka koncentracja może potęgować wrażenie schematyczności i szablonowości tekstów w obrębie tego gatunku. Koncentracja 3-gramów jest dwukrotnie niższa w orzeczeniach, a najniższa w sprawozdaniach – dla porównania najczęstsze 3-gramy mają częstość zaledwie 4 108 wystąpień na milion słów w korpusie sprawozdań krajowych. W przypadku aktów prawnych i orzeczeń koncentracja jest wyższa w tekstach przetłumaczonych, natomiast sprawozdania wykazują odwrotny trend.

W następnej tabeli (tabela 7) przedstawiono najczęstsze 3-gramy w badanych korpusach z uwzględnieniem ich funkcji, znormalizowanej częstości do 1 miliona słów i ich rozproszenia w korpusie.

Zbitki (typy)	Akty normatywne UE			Akty normatywne PL		
	Rozporządzenia			Dyrektywy		
	I. wyst. na bucja 1 mln	Dystry- Funkcja Zbitki (typy)	I. wyst. na bucja 1 mln	Dystry- Funkcja Zbitki (typy)	I. wyst. na bucja 1 mln	Dystry- Funkcja Zbitki (typy)
	(%) tekstów)	(%) tekstów)	(%) tekstów)	(%) tekstów)	(%) tekstów)	(%) tekstów)
o których mowa, których mowa w mowa w art.	2952 (2930) 2776	37 MT-NAV o których mowa, których mowa w art.	2868 (2844) 2153	88 MT-NAV mowa w art.	5323	94 MT-NAV
zgodnie z art.	2611	35 MT-NAV w przypadku gdy	2134	88 MT-PS mowa w ust.	4191	92 MT-NAV
w odniesieniu do	2493	78 MT-PR zgodnie z art.	1997	84 MT-NAV o którym mowa, którym mowa w	3958 (3948)	93 MT-NAV
mowa w ust.	1635	30 MT-NAV państwa członkowskie	1764	86 M o której mowa, której mowa w	2160 (2155)	89 MT-NAV
o którym mowa, w którym mowa w w przypadku gdy	1428 (1423) 1248	74 MT-NAV w odniesieniu do zapewniania	1721	89 MT-PR ustawy z dnia	2137	94 MT-NAV
o której mowa, której mowa w rozporządzenia UE nr	1105 (1104) 966	23 MT-PS mowa w ust.	1474	89 MT-NAV minister właściwy do, właściwy do spraw	1156 (1155)	79 R
w stosownych przypadkach	750	25 MT-NAV którym mowa w; o którym mowa	1121 (1120)	89 MT-NAV z późn.zm.	1022	92 MT-NAV
do niniejszego rozporządzenia	730	21 MT-NAV rozporządzenia UE nr	1044	51 MT-NAV ustawie z dnia, w ustawie z	956 (931)	86 MT-NAV
rozporządzenia WE nr	727	22 R państwa członkowskie mogą	1010	80 M w przypadku gdy	836	81 MT-NAV
w dzienniku urzędowym, dzienniku urzędowym, Unii, urzędowym Unii Europejskiej	627 (626, 625)	73 R w stosownych przypadkach	788	75 R terytorium Rzeczypospolitej Polskiej, na terytorium Rzeczypospolitej	785 (712)	51 R
wchodzi w życie	555	34 MT-NAV członkowskie zapewniania aby	773	54 M na podstawie art.	673	76 MT-NAV
zgodnie z procedurą	518	99 R bez uszczerbku dla	668	93 MT-NAV wejścia w życie	672	83 R
Łącznie	27829	Łącznie	25319	Łącznie	42863	

Sprawozdania UE			Sprawozdania PL				
Zbitki (typy)	I. wyst. na I mln	Dystrybucja (% tekstów)	Funkcja	Zbitki (typy)	I. wyst. na I mln	Dystrybucja (% tekstów)	Funkcja
w odniesieniu do	1673	93	MT PR	w związku z	380	83	MT PS
w związku z	687	80	MT PS	w stosunku do	369	85	MT PR
przez państwa członkowskie	440	58	R	ze względu na	326	75	MT PR
ze względu na	426	68	MT PR	narodowych i etnicznych	318	7	R
zgodnie z art.	342	55	MT NAV	r. w sprawie	306	55	MT NAV
w państwach członkowskich	341	56	R	w odniesieniu do	286	79	MT PR
Parlamentu Europejskiego i	335	70 (68)	R	mniejszości narodowych i	271	7	R
Europejskiego i Rady	(327)						
związku z tym	316	61	MTPS	Powiatowy Urząd Pracy	262	5	R
					(258)		
dyrektywy w sprawie	300	22	R/MT NAV	z późn.zm.	203	47	R/MT
w oparciu o	291	56	MT PR	na podstawie danych	198	54	MT NAV
w porównaniu z	261	44	MT PR	w porównaniu z	194	64	MT PR
jeżeli chodzi o	259	39	MT PR	Administracji i Cyfryzacji	191	23	R
sprawozdanie Komisji dla	227	86	R	w ramach programu	185	48	MT PR
w szczególności w	215	48	MT PR	przemocy w rodzinie	182	12	R
dla Parlamentu Europejskiego	215	76	R	w ramach projektu	179	43	MT PR
Łącznie	6655			Łącznie	4108		

Tabela 7. Piętnaście najczęstszych 3-gramów – funkcje, dystrybucja i rozproszenie (oznaczenie skrótów: R – zbitka referencyjna, M – zbitka modalna, MT-NAV – zbitka metatekstowa nawigacyjna, MT-PR – zbitka metatekstowa precyzująco-wyjaśniająca, MT-PS zbitka metatekstowa precyzująca). Źródło: opracowanie własne.

We wszystkich gatunkach wśród najczęstszych 3-gramów dominują zbitki metatekstowe¹⁸ – o ile w aktach normatywnych i orzeczeniach zdecydowanie najwięcej jest zbitek nawigacyjnych służących odesłaniom intra- i intertekstualnym (*o których mowa w, na podstawie art., ww. w pkt., w rozumieniu art.*), w sprawozdaniach najwięcej jest zbitek metatekstowych, które pełnią funkcję precyzująco-wyjaśniającą (*w odniesieniu do, ze względu na, w szczególności w, w stosunku do, w ramach programu*) i pojawiają się również w niewielkim stopniu w orzeczeniach unijnych. Dominacja zbitek nawigacyjnych w aktach normatywnych i orzeczeniach wynika z wysokiej intertekstualności tych gatunków, a także konieczności zapewnienia precyzji odesłań. Z kolei przewaga zbitek precyzująco-wyjaśniających w sprawozdaniach jest powiązana z ich podstawową funkcją komunikowania wiedzy. Następnie występuje relatywnie dużo zbitek referencyjnych, zwłaszcza w sprawozdaniach (*w państwach członkowskich, przez państwa członkowskie, Powiatowy Urząd Pracy*). W większości korpusów pojawia się przynajmniej jedna zbitka metatekstowa przyczynowo-skutkowa (*w przypadku gdy, ze względu na, w związku z*). Zbitki modalne występują we wszystkich korpusach aktów normatywnych, w których realizują modalność deontyczną, i w jednym korpusie orzeczeń unijnych dla wyrażenia modalności epistemicznej (*wchodzi w życie, państwa członkowskie mogą, należy interpretować w*), co jest zgodne z dominującą cechą tych gatunków.

Żadna zbitka z listy najczęstszych 3-gramów nie powtarza się we wszystkich korpusach. Wspólne zbitki, tj. *zgodnie z art.* oraz *w odniesieniu do*, mają jedynie gatunki unijne, natomiast w korpusach krajowych nie zidentyfikowano wspólnych zbitek. Więcej zbieżności można znaleźć w obrębie dwóch gatunków – np. orzeczenia i sprawozdania mają wspólną zbitkę metatekstową przyczynowo-skutkową *w związku z*,

18. Nie potwierdza to wyników Jablonkai, która wykazała, że w dyskursie unijnym najwięcej jest zbitek referencyjnych, które denotują ilość, jakość, cel, czas i miejsce, a także zbitek tematycznych odnoszących się do Unii Europejskiej, natomiast mało jest zbitek modalnych i metatekstowych (2010: 261). Nie jest jednak jasne, czy zbitki nawigacyjne zostały zaklasyfikowane do grupy zbitek referencyjnych czy metatekstowych.

a ich warianty unijne dodatkową zbitkę o tej samej funkcji – *ze względu na*. Z kolei krajowe akty normatywne i orzeczenia mają wspólne zbitki nawigacyjne: *ustawy z dnia i na podstawie art.* Wspólne zbitki występują również w niewielkim stopniu w obrębie poszczególnych gatunków:

- krajowe i unijne akty prawne: *o których mowa, mowa w art., w przypadku gdy*;
- krajowe i unijne orzeczenia: *w związku z, na podstawie art.*;
- krajowe i unijne sprawozdania: *w odniesieniu do, w związku z, ze względu na, w porównaniu z*.

Ogólnie rzecz biorąc, można uznać za R. Breeze (2013: 248, 250), iż zbitki odgrywają istotną rolę w badanych gatunkach – o ile zbitki referencyjne rozwijają tekst tematycznie i wypełniają go treścią, zbitki metatekstowe (głównie przyimkowe) tworzą ważne ramy dla tych treści.

9. Wnioski

W niniejszej pracy zbadano formuliczność trzech gatunków w obrębie dyskursu administracyjnego, wskazując na największą formuliczność gatunku aktów normatywnych, porównywalną formuliczność orzeczeń i najmniejszą formuliczność sprawozdań, zarówno jeżeli chodzi o liczbę typów zbitek, ich łączną liczbę wystąpień, jak i koncentrację najczęstszych 3-gramów. Dla gatunku aktów normatywnych i sprawozdań wykazano również co do zasady większą formuliczność tłumaczeń (ten trend nie został potwierdzony dla orzeczeń).

Badanie wykazało, iż użycie zbitek jest specyficzne dla poszczególnych gatunków i ich odmian, można więc przyjąć, iż zbitki dobrze profilują specyfikę gatunku i realizują jego cele komunikacyjne. Znajduje to również potwierdzenie w tym, iż niewiele zbitek jest wspólnych dla wszystkich gatunków. W aktach normatywnych i orzeczeniach najczęstsze 3-gramy pełnią głównie funkcję metatekstową nawigacyjną, natomiast w sprawozdaniach najczęstsze są zbitki przyczynowo-skutkowe i referencyjne.

Dzięki przetestowaniu różnych progów badanie wykazało, iż formuliczność jest zjawiskiem niejednoznacznym i wielopłaszczyznowym. Jednocześnie analiza uwypukliła problemy metodologiczne związane z identyfikacją zbitek w różnych korpusach – identyfikacja wydaje się być szczególnie czuła na wielkość i skład korpusu, tj. liczbę tekstów i liczbę wyrazów w korpusie, które mogą ograniczyć porównywalność wyników. Ponadto z analizy wynika, iż normalizacja wyników nie neutralizuje w pełni różnic w zakresie wielkości korpusów. Na porównywalność wyników z pewnością ma również wpływ spójność tematyczna (dziedzinowa) poszczególnych korpusów lub jej brak. Potwierdza to konieczność dopracowania aspektów metodologicznych identyfikacji zbitek.

Bibliografia

- Aijmer, Karin i Diana Lewis (red.)
2017 *Contrastive Analysis of Discourse-pragmatic Aspects of Linguistic Genres*. Yearbook of Corpus Linguistics and Pragmatics vol. 5. Springer.
- Anderman, Gunilla M. i Margaret Rogers (red.)
2007 *Incorporating Corpora: The Linguist and the Translator*. Clevedon: Multilingual Matters.
- Bachtin, Michał
1986 *Estetyka twórczości słownej*. Warszawa: PIW.
- Baker, Mona
1996 „Corpus-based translation studies: the challenges that lie ahead”, w: Harold Somers (red.), 175–186.
- Balaskó, Maria
2008 „What does the *figure* show? Patterns of translationese in a Hungarian comparable corpus”, *Trans-Kom* 1(1): 58–73.
- Berber Sardinha i Antonio Paulo
1997 „Patterns of lexis in original and translated business reports: Textual differences and similarities”, w: Karl Simms (red.), 147–154.

- Bhatia, Vijay K.
 1993 *Analysing Genre – Language Use in Professional Settings*. London and New York: Longman.
- 2006 „Legal Genres”, w: Keith Brown (red.), 1–7.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad i Edward Finegan
 1999 *Longman Grammar of Spoken and Written English*. Harlow: Longman.
- Biber, Douglas, Susan Conrad i Viviana Cortes
 2004 „If you look at...: Lexical bundles in university teaching and textbooks”, *Applied linguistics* 25(3): 371–405.
- Biber, Douglas i Federica Barbieri
 2007 „Lexical bundles in university spoken and written registers”, *English for Specific Purposes* 26/3: 263–286.
- Biber, Douglas i Randi Reppen (red.)
 2015 *The Cambridge Handbook of English Corpus Linguistics*. Cambridge: Cambridge University Press.
- Biel, Łucja
 2014 *Lost in the Eurofog. Textual Fit of Translated Law*. Frankfurt am Main: Peter Lang.
- 2016 „Mixed corpus design for researching the Eurolect: a genre-based comparable-parallel corpus in the PL EUROLECT project”, w: Ewa Gruszczyńska i Agnieszka Leńko-Szymańska (red.), 197–208.
- 2018 „Lexical bundles in EU law: The impact of translation process on the patterning of legal language”, w: Stanisław Goźdz-Roszkowski i Gianluca Pontrandolfo (red.), 11–26.
- Bładowski, Bogdan
 2012 *Pisma sądowe w sprawach cywilnych*. Warszawa: Wolters Kluwer.
- Breeze, Ruth
 2013 „Lexical bundles across four legal genres”, *International Journal of Corpus Linguistics* 18/2: 229–253.
- Brown, Keith (red.)
 2006 *Encyclopedia of Language and Linguistics*. Vol. 13. (wydanie 2.) Oxford: Elsevier.
- Candlin, Christopher N. i Maurizio Gotti (red.)
 2006 *Intercultural Aspects of Specialized Communication. Linguistic Insights/Studies in Language and Communication*, vol. 14. Bern: Peter Lang.

- Cap, Piotr i Urszula Okulska
2013 „Analyzing genres in political communication: An introduction”, w: Piotr Cap i Urszula Okulska (red.), 1–26.
- Cap, Piotr i Urszula Okulska (red.)
2013 *Analyzing Genres in Political Communication*. Amsterdam: John Benjamins.
- Charteris-Black, Jonathan i Timothy Ennis
2001 „A comparative study of metaphor in Spanish and English financial reporting”, *English for Specific Purposes* 20: 249–266.
- Charteris-Black, Jonathan i Andreas Musolff
2003 „Battered hero’ or ‘innocent victim’? A comparative study of metaphors for euro trading in British and German financial reporting”, *English for Specific Purposes* 22: 153–176.
- Chen, Yu-Hua i Paul Baker
2010 „Lexical bundles in L1 and L2 academic writing”, *Language Learning & Technology* 4/2: 30–49.
- Cheng, Winnie i Le Cheng
2014 „Epistemic modality in court judgments: A corpus-driven comparison of civil cases in Hong Kong and Scotland”, *English for Specific Purposes* 33: 15–26.
- Choduń, Agnieszka
2007 *Słownictwo tekstów aktów prawnych w zasobie leksykalnym współczesnej polszczyzny*. Warszawa: Trio.
- Conrad, Susan M. i Douglas Biber
2004 „The frequency and use of lexical bundles in conversation and academic prose”, w: Wolfgang Teubert i Michaela Mahlberg (red.), 56–71.
- Crosthwaite, Peter, Lisa Cheung i Feng (Kevin) Jiang
2017 „Writing with attitude: Stance expression in learner and professional dentistry research reports”, *English for Specific Purposes* 46: 107–123.
- Davies, Alan i Catherine Elder (red.)
2004 *The Handbook of Applied Linguistics*. Oxford: Blackwell.
- Dayrell, Carmen
2008 „Investigating the preference of translators for recurrent lexical patterns: a corpus-based study”, *Trans-Kom* 1(1): 36–57.

- Dijk, Teun A. van
 1988a *News as Discourse*. Hillsdale, NJ: Lawrence Erlbaum Associates.
 1988b *News Analysis: Case Studies of International News in the Press*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ditlevsen, Marianne Grove i Jan Engberg
 2007 „Translating an auditor’s report: An instance of specialised legal translation”, *JoSTrans* 7: 190–194.
- Duszak, Anna
 1998 *Tekst, dyskurs, komunikacja międzykulturowa*. Warszawa: Wydawnictwo Naukowe PWN.
- Duszak, Anna, Elżbieta Gajek i Urszula Okulska (red.)
 2006 *Korpusy w angielsko-polskim językoznawstwie kontrastywnym: Teoria i praktyka*. Kraków: Universitas.
- Duszak, Anna, Juliane House i Łukasz Kumięga (red.)
 2010 *Globalization, Discourse, Media: In a Critical Perspective*. Warsaw: WUP.
- Duszak, Anna, Arkadiusz Jabłoński i Agnieszka Leńko-Szymańska (red.)
 2017 *East-Asian and Central-European Encounters in Discourse Analysis, Translation and Language Teaching*. Warszawa: Instytut Lingwistyki Stosowanej.
- Duszak, Anna i Grzegorz Kowalski (red.)
 2015 *Academic (Inter)genres: Between Texts, Contexts and Identities*. Frankfurt am Main: Peter Lang.
- Duszak, Anna i Urszula Okulska (red.)
 2004 *Speaking from the Margin. Global English from a European Perspective*. Frankfurt am Main: Peter Lang.
- Engberg, Jan
 1999 „Übersetzen von Gerichtsurteilen: Der Einfluss der Perspektive”, w: Peter Sandrini (red.), 83–102.
- Fairclough, Norman
 2003 *Analysing Discourse. Textual Analysis for Social Research*. London: Routledge.
- Fiedler, Sabine
 2017 „Phraseological borrowing from English into German: Cultural and pragmatic implications”, *Journal of Pragmatics* 113: 89–102.

- Flowerdew, John i Maurizio Gotti (red.)
2006 *Studies in Specialized Discourse. Linguistic Insights/ Studies in Language and Communication* (Vol. 41). Bern: Peter Lang.
- Flowerdew, John i Alina Wan
2010 „The linguistic and the contextual in applied genre analysis: The case of the company audit report”, *English for Specific Purposes* 29: 78–93.
- Friginal, Eric
2013 „Developing research report writing skills using corpora”, *English for Specific Purposes* 32: 208–220.
- Gapska, Edyta
2009 *Wady orzeczeń sądowych w postępowaniu cywilnym*. Warszawa: Oficyna Wolters Kluwer business.
- Garzone, Giuliana
2006a „The use of discursive features expressing causal relations in annual company reports”, w: John Flowerdew i Maurizio Gotti (red.), 81–108.
2006b „Annual company reports and CEOs’ letters: Discoursal features and cultural markedness”, w: Christopher N. Candlin i Maurizio Gotti (red.), 311–341.
- Gizbert-Studnicki, Tomasz
1986 *Język prawny z perspektywy socjolingwistycznej*. Warszawa: PWN.
- Gotti, Maurizio i Christopher John Williams (red.)
2010 *Legal Discourse Across Languages and Cultures*. Bern: Peter Lang.
- Gouadec, Daniel
2007 *Translation as a Profession*. Amsterdam: John Benjamins.
- Goźdz-Roszkowski, Stanisław
2011 *Patterns of Linguistic Variation in American Legal English. A Corpus-Based Study*. Frankfurt am Main: Peter Lang.
2013 „Evaluative patterns in judicial discourse: A corpus-based phraseological perspective on American and Italian criminal judgments”, *International Journal of Law, Language and Discourse* 3/2: 9–69.
- Goźdz-Roszkowski, Stanisław i Gianluca Pontrandolfo (red.)
2018 *Phraseology in Legal and Institutional Settings: A Corpus-based Interdisciplinary Perspective*. London: Routledge.

Grabowski, Łukasz

2013 „Register variation across English pharmaceutical texts: A corpus-driven study of keywords, lexical bundles and phrase frames in patient information leaflets and summaries of product characteristics”, w: *Procedia – Social and Behavioral Sciences* 95: 391–401.

2014 „On lexical bundles in Polish patient information leaflets: A corpus-driven study”, w: *Studies in Polish Linguistics* 9/1: 21–43.

2015a *Phraseology in English Pharmaceutical Discourse. A Corpus-Driven Study of Register Variation*. Opole: Uniwersytet Opolski.

2015b „O frazeologii z perspektywy językoznawstwa korpusowego”, w: *Problemy frazeologii europejskiej X*: 23–48.

Granger, Sylviane i Magali Paquot

2008 „Disentangling the phraseological web”, w: Sylviane Granger i Fanny Meunier (red.), 27–49.

Granger Sylviane i Fanny Meunier (red.)

2008 *Phraseology. An Interdisciplinary Perspective*. Amsterdam: John Benjamins.

Gray, Bethany i Douglas Biber

2015 „Phraseology”, w: Douglas Biber i Randi Reppen (red.), 125–145.

Greaves, Chris i Martin Warren

2010 „What can a corpus tell us about multi-word units?”, w: Anne O’Keeffe i Michael McCarthy (red.), 212–226.

Gries, Stefan Th.

2010 „Corpus linguistics and theoretical linguistics. A love-hate relationship? Not necessarily...”, *International Journal of Corpus Linguistics* 15/3: 327–343.

Grochowski, Mateusz, Michał Raczkowski i Sławomir Żółtek (red.)

2015 *Studia i analizy Sądu Najwyższego. Materiały naukowe*. Warszawa: Biuro Studiów i Analiz Sądu Najwyższego.

Gruszczyńska, Ewa i Agnieszka Leńko-Szymańska (red.)

2016 *Polskojęzyczne korpusy równoległe. Polish-language Parallel Corpora*. Warszawa: Instytut Lingwistyki Stosowanej.

- Guangrong, Dai
2009 „The applications of keywords and collocations to Translation Studies and teaching: A tentative research on the parallel corpus of the 17th NCCPC Report”, *Translation Journal* 13/3. Wydanie online na: <http://www.bokorlang.com/journal/49corpus.htm> [data dostępu: maj 2017].
- Halmari, Helena i Tuija Virtanen (red.)
2005 *Persuasion Across Genres. A Linguistic Approach*. Amsterdam: John Benjamins.
- Hałas, Bożena
2005 *Terminologia języka prawnego*. Zielona Góra: WSP.
- Handford, Michael
2010 *The Language of Business Meetings*. Cambridge: Cambridge University Press.
- Heine, Bernd i Heiko Narrog (red.)
2009 *The Oxford Handbook of Linguistic Analysis*. Wydanie online na: <http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199544004.001.0001/oxfordhb-9780199544004> [data dostępu: maj 2017].
- Hyland, Ken
2008 „As can be seen: Lexical bundles and disciplinary variation”, *English for Specific Purposes* 27: 4–21.
- Hyon, Sunny
2008 „Convention and inventiveness in an occluded academic genre: A case study of retention–promotion–tenure reports”, *English for Specific Purposes* 27: 175–192.
- Jablonkai, Reka
2010 „English in the context of European integration: A corpus-driven analysis of lexical bundles in English EU documents”, *English for Specific Purposes* 29/2: 253–267.
- Jadacka, Hanna
1996 „Styl tekstów prawnych na tle innych odmian polszczyzny”, *Przegląd Legislacyjny* 3: 112–117.
- Jopek-Bosiacka, Anna
2006 *Przekład prawny i sądowy*. Warszawa: Wydawnictwo Naukowe PWN.
- Károly, Krisztina i Ágota Fóris (red.)
2005 *New trends in Translation Studies: In Honour of Kinga Klaudy*. Budapest: Akadémiai Kiadó.

Kemppanen, Hannu

- 2004 „Keywords and ideology in translated history texts: A corpus-based analysis”, *Across Languages and Cultures* 5(1): 89–106.

Kjær, Anne Lise

- 2000 „On the structure of legal knowledge: The importance of knowing legal rules for understanding legal texts”, w: Lita Lundquist i Robert J. Jarvella (red.), 127–161.

Kruger, Alet

- 1997 „The translator as agent of reconciliation: translating an eye-witness report in a historical context”, w: Karl Simms (red.), 77–86.

Łapa, Romana

- 2008 „Znaczenia gramatyczne znominalizowanych grup imiennych we współczesnych tekstach prawnych”, *Język Polski* 88/4–5: 280–288.

Lee, Chang-Soo

- 2011 „N-gram analysis of phrasal differences between translational and original Korean news texts”, *Journal of Interpretation & Translation Research Institute* 15/1: 317–340.
- 2013 „Using lexical bundle analysis as discovery tool for corpus-based translation research”, *Perspectives* 21 (3): 378–395.

Lundquist, Lita i Robert J. Jarvella (red.)

- 2000 *Language, Text, and Knowledge. Mental Models of Expert Communication*. Berlin: de Gruyter.

Malinowski, Andrzej

- 2006 *Polski język prawny. Wybrane zagadnienia*. Warszawa: LexisNexis.
- 2010 *Teksty prawne Unii Europejskiej. Opracowanie treściowe i redakcyjne oraz zasady ich publikacji*. Warszawa: LexisNexis.
- 2012 *Polski tekst prawny. Opracowanie treściowe i redakcyjne*. (wydanie 3.) Warszawa: LexisNexis.

Matulewska, Aleksandra

- 2007 *Lingua Legis in Translation*. Frankfurt am Main: Peter Lang.

Mauranen, Anna

- 2006 „Translation universals”, w: Keith Brown (red.), 93–100.
- 2007 „Universal tendencies in translation”, w: Gunilla M. Anderman i Margaret Rogers (red.), 32–48.

Mazzi, Davide

- 2007 „The construction of argumentation in judicial texts: Combining a genre and a corpus perspective”, *Argumentation* 14 (1): 21–38.
- 2010 „«This argument fails for two reasons»: A linguistic analysis of judicial evaluation strategies in US Supreme Court judgments”, *International Journal For The Semiotics Of Law* 23: 373–385.
- 2015 „«It must be obvious that this line of argument is utterly inconsistent...»: On attitudinal qualification in English judicial discourse across legal systems”, *Bulletin Suisse de Linguistique Appliquée* 2: 51–67.

McEnery, Tony i Andrew Hardie

- 2012 *Corpus Linguistics – Method, Theory and Practice*. Cambridge: Cambridge University Press.

Miller, Carolyn R.

- 1984 „Genre as social action”, *Quarterly Journal of Speech* 70: 151–167.

Monzó Nebot, Esther

- 2008 „Corpus-based activities in legal translator training”, *The Interpreter and Translator Trainer* 2/2: 221–251.

Nathan, Philip

- 2016 „Analysing options in pedagogical business case reports: Genre, process and language”, *English for Specific Purposes* 44: 1–15.

Nesselhauf, Nadja

- 2005 *Collocations in a Learner Corpus*. Amsterdam: John Benjamins.

O’Keeffe, Anne i Michael McCarthy (red.)

- 2010 *The Routledge Handbook of Corpus Linguistics*. London: Routledge.

Parkinson, Jean

- 2017 „The student laboratory report genre: A genre analysis”, *English for Specific Purposes* 45: 1–13.

Pontrandolfo, Gianluca

- 2016 *Fraseología y lenguaje judicial. Las sentencias penales desde una perspectiva contrastive*. (Dialogoi Ispanistica 17). Roma: Aracne.

Przetak, Marzena

2015 *Struktura tekstu prawnego na przykładzie kodeksu karnego*. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.

Renkema, Jan

1993 *Discourse Studies. An Introductory Textbook*. Amsterdam/Philadelphia: John Benjamins.

Römer, Ute

2010 „Establishing the phraseological profile of a text type: The construction of meaning in academic book reviews”, *English Text Construction* 3/1: 95–119.

Rzucidło-Grochowska, Iwona

2015 „Uzasadnienie orzeczenia sądowego w świetle jednolitości orzecznictwa. Zarys problematyki”, w: Mateusz Grochowski, Michał Raczkowski i Sławomir Żółtek (red.), 62–72.

Salmi-Tolonen, Tarja

2005 „Persuasion in judicial argumentation: The Opinions of the Advocates General at the European Court of Justice”, w: Helena Halmari i Tuija Virtanen (red.), 59–101.

Sandrini, Peter (red.)

1999 *Übersetzen von Rechtstexten. Fachkommunikation im Spannungsfeld zwischen Rechtsordnung und Sprache*. Tübingen: Narr Francke Attempto Verlag.

Santini, Marina

2005 „Genres in formation? An exploratory study of web pages using cluster analysis”, w: *Proceedings of the 8th annual colloquium for the UK special interest group for computational linguistics (CLUK05)*, 11 stycznia 2005, Manchester, UK. Wydanie online na: <http://www.nltg.brighton.ac.uk/home/Marina.Santini/ITRI-05-01.pdf> [data dostępu: maj 2017].

Šarčević, Susan

1997 *New Approach to Legal Translation*. The Hague: Kluwer Law International.

Scott, Mike

2016 *WordSmith Tools 7*. Stroud: Lexical Analysis Software.

Simms, Karl (red.)

1997 *Translating Sensitive Texts: Linguistic Aspects*. Amsterdam, Atlanta: Rodopi.

- Somers, Harold (red.)
1996 *Terminology, LSP and Translation. Studies in Language Engineering in Honour of Juan C. Sager*. Amsterdam/Philadelphia: John Benjamins.
- Spence, Paul i Liu Gi-Zen
2013 „Engineering English and the high-tech industry: A case study of an English needs analysis of process integration engineers at a semiconductor manufacturing company in Taiwan”, *English for Specific Purposes* 32: 97–109.
- Stubbs, Michael
2004 „Language corpora”, w: Alan Davies i Catherine Elder (red.), 106–132.
- Swales, John M.
1990 *Genre Analysis. English in academic and research settings*. Cambridge: Cambridge University Press.
- Szczyrbak, Magdalena
2009 „Genre-based analysis of the realisation of Concession in judicial discourse”, *Studia Linguistica Universitatis Iagellonicae Cracoviensis* 126(2009): 127–147.
2014 *The Realisation of Concession in the Discourse of Judges: A Genre Perspective*. Kraków: WUJ.
2017 „Modal adverbs of certainty in EU legal discourse: A parallel corpus approach”, w: Karin Aijmer i Diana Lewis (red.), 91–115.
- Teubert, Wolfgang i Michaela Mahlberg (red.)
2004 „The Corpus Approach to Lexicography”, *Lexicographica. Internationales Jahrbuch für Lexicographie* 20: 1–19.
- Tirkkonen-Condit, Sonja
2005 „Do unique items make themselves scarce in translated Finnish?”, w: Kristina Károly i Ágota Fóris (red.), 177–189.
- Tománková, Veronika
2016 „Lexical bundles in legal texts corpora – Selection, classification and pedagogical implications”, *Discourse and Interaction* 9/2: 75–94.
- Toury, Gideon
1995 *Descriptive Translation Studies and Beyond*. Amsterdam: John Benjamins.

Vazquez-Orta, Ignacio

- 2010 „A genre-based view of judgments of appellate courts in the common law system: Inter-subjective positioning, intertextuality and interdiscursivity in the reasoning of judges”, w: Maurizio Gotti i Christopher John Williams (red.), 263–284.

Wharton, Sue

- 2007 „Social identity and parallel text dynamics in the reporting of educational action research”, *English for Specific Purposes* 26: 485–501.

Wronkowska, Sławomira i Maciej Zieliński

- 1993 *Problemy i zasady redagowania tekstów prawnych*. Warszawa: Urząd Rady Ministrów.
2012 *Komentarz do zasad techniki prawodawczej*. (wydanie 2.) Warszawa: Wydawnictwo Sejmowe.

Yeung, Lorrita

- 2007 „In search of commonalities: Some linguistic and rhetorical features of business reports as a genre”, *English for Specific Purposes* 26: 156–179.