

Łucja Biel
Uniwersytet Warszawski

TRANSLATORYKA KORPUSOWA

DOI: <http://dx.doi.org/10.12775/RP.2015.001>

Zarys treści: Celem artykułu jest omówienie założeń translatoryki korpusowej jako nowej metodologii badań nad przekładem. Artykuł pozycjonuje translatorykę korpusową na tle najnowszych kierunków badań nad przekładem („zwrot technologiczny”) i omawia jej teoretyczne podstawy wywodzące się z językoznawstwa korpusowego i opisowych badań nad przekładem. Następnie przedstawiono główne typy korpusów (korpus porównawczy i równoległy), oprogramowanie służące do ich analizy oraz podstawowe metody analizy. Omówiono w nim również stan badań, zastosowania teoretyczne i praktyczne badań korpusowych, a także perspektywy rozwoju i ograniczenia.

Słowa kluczowe: translatoryka korpusowa, badania nad przekładem, analiza korpusów

Najnowsze kierunki badań nad przekładem

W dobrze znanym polskiemu odbiorcy podręczniku *Współczesne tendencje przekładoznawcze* autorki – prof. Pisarska i prof. Tomaszkiwicz – wyróżniły następujące najważniejsze kierunki badań przekładoznawczych (1996: 11–59):

- początki okresu językoznawczego (Jakobson, Nida, Mounin),
- szkoła Übersetzungswissenschaft (Wilss, Koller) i szkoła lipska (Kade, Jäger, Neubert),
- socjolingwistyczne badania francuskie (Cary, Delisle, Pergnier, szkoła paryska),
- podejście tekstowe (Ladmiral, Margot, Meschonnic),
- hermeneutyka w badaniach niemieckich (Paepcke, Stolze),
- pragmatyczne reguły Newmarka,
- manipulści (Toury, Hermans, Holmes, Bassnett, Lefevere),
- model integracyjny Snell-Hornby,

- badania psycholingwistyczne dotyczące procesów przekładu (Kusmaul, Höinig, Lörscher),
- przekład feministyczny.

Od czasu publikacji podręcznika w 1996 r. minęło niemal 20 lat. W tym okresie translatoryka dynamicznie się rozwinęła. Stała się odrębną, autonomiczną dyscypliną naukową¹ – jak podkreśla Snell-Hornby, „wymancypowała się” i nie jest już jedynie specjalnością w obrębie lingwistyki stosowanej czy literatury porównawczej (Snell-Hornby 2012: 371)². Do wzrostu popularności translatoryki bez wątpienia przyczyniła się **globalizacja**, skutkująca coraz większą współzależnością kultur i gospodarek oraz rosnącym popytem na usługi tłumaczeniowe i kształcenie tłumaczy. Na gruncie polskim istotnymi czynnikami zwiększającymi popyt na usługi tłumaczeniowe, a w rezultacie na kształcenie tłumaczy i działalność naukową były: otwarcie Polski na świat po upadku komunizmu, przejście z gospodarki planowanej na gospodarkę rynkową, napływ bezpośrednich inwestycji zagranicznych oraz przystąpienie do Unii Europejskiej i uzyskanie przez język polski statusu unijnego języka urzędowego. Przemiany społeczno-polityczne na świecie pociągnęły za sobą zmiany na **mapie geograficznej** dyscypliny – coraz więcej ośrodków badawczych powstaje w Europie Środkowo-Wschodniej, na Bliskim Wschodzie, w Indiach, Chinach, a także w Stanach Zjednoczonych.

W 2005 r. w pozytywnym artykule pt. *Trajectories of research in Translation Studies* Maria Tymoczko za centralne i przyszłe obszary badań przekładoznawczych uznała:

- 1) zdefiniowanie przekładu: badanie językoznawczych aspektów tłumaczeń, badania korpusowe, historyczne badania opisowe, badania nad procesem tłumaczenia,
- 2) internacjonalizację przekładu i zakwestionowanie popularnych na Zachodzie poglądów dotyczących specyfiki przekładu,
- 3) nowe technologie i globalizację oraz ich wpływ na teorię i praktykę przekładu,

¹ W niektórych krajach, przykładowo w Polsce, autonomia translatoryki jako dyscypliny nie została jeszcze usankcjonowana instytucjonalnie (np. brak możliwości doktoryzowania się z przekładoznawstwa).

² Na marginesie warto wspomnieć, że translatoryka nadal walczy o swój status w świecie naukowym. Jedną z bolączek jest niski ranking czasopism translatorycznych, jak i ich pomijanie w JCR, ERIH czy SJR (zob. apel Europejskiego Stowarzyszenia Badań nad Przekładem (EST) w sprawie niskiej punktacji czasopism translatorycznych http://www.est-translationstudies.org/resources/journal_rankings_statement.html; dostęp: 1 grudnia 2013 r.).

- 4) ramy interpretacyjne pochodzące z innych dyscyplin naukowych,
- 5) kognitywistykę,
- 6) neurofizjologię.

Punkt drugi wynika ze zmian na mapie geograficznej dyscypliny i otwarcia się na dorobek translatorski innych kręgów kulturowych (zob. też Munday 2010: 425). Natomiast punktem wspólnym pozostałych obszarów jest czerpanie z osiągnięć innych dyscyplin naukowych i adaptowanie ich do swoich potrzeb.

Na dynamiczny rozwój translatoryki istotny wpływ miał zatem **rozwój sąsiednich dyscyplin**, z którymi powiązana jest translatoryka – określana niekiedy mianem interdyscypliny czy polidyscypliny (zob. Gentzler 2003; Brems i in. 2012; Bassnett 2012). Dzięki kontaktom z pokrewnymi dyscyplinami otworzyły się nowe perspektywy badawcze, wprowadzono nowe metodologie badań i rozwinęto aparat pojęciowy przekładoznawstwa. Oprócz długoterminowych powiązań z literaturoznawstwem i językoznawstwem translatoryka czerpie m.in. z socjologii, psychologii, kulturoznawstwa, antropologii czy informatyki. Coraz więcej publikacji dotyczy kulturowych, socjologicznych, kognitywnych, neurologicznych, etycznych czy technologicznych aspektów przekładu.

O ile w latach 60. i 70. mówiło się o „zwrocie językoznawczym”, o tyle w latach 80. i 90. nastąpiła deprecjacja metod językoznawczych pod wpływem „zwrotu kulturowego”. Obecnie, jak zauważa Cronin, można mówić o **zwrocie technologicznym** (2010), który jest wynikiem powszechności komputerów i pojawienia się nowych form praktyki tłumaczeniowej (Cronin 2010; zob. też Gentzler 2003: 11; Munday 2010: 425–426; Snell-Hornby 2012; Brems i in. 2012: 4). W tym miejscu należałoby wspomnieć o programach wspomagających tłumaczenie, tzw. CAT-ach, tłumaczeniach maszynowych, bazach terminologicznych czy będącej przedmiotem artykułu translatoryce korpusowej (*corpus-based translation studies*), która również wpisuje się w nurt technologiczny.

Jedną z sąsiednich dyscyplin, które są szczególnie interesujące dla autorki artykułu, jest **językoznawstwo**, gdyż translatoryka korpusowa to owoc językoznawczych zainteresowań przekładoznawców. W ostatnich latach nastąpiła rehabilitacja metod językoznawczych w badaniach nad przekładem, przy czym należy podkreślić, że jest to powrót do przeżywającego rozkwit językoznawstwa funkcjonalnego. Translatoryka korzysta m.in. z narzędzi językoznawstwa korpusowego, językoznawstwa kognitywnego, analizy dyskursu, gramatyki systemowo-funkcjonalnej Hallidaya, etnolingwistyki czy neurolingwistyki. W artykule z 2013 r. Juliane House, nawiązując do „zwrotów”

w translatoryce, postuluje konieczność wprowadzenia do teorii przekładu orientacji „językoznawczo-kognitywnej” (ang. *linguistic-cognitive orientation*), polegającej na integracji wyników badań neurolingwistycznych, lingwistyki tekstu i gramatyki systemowo-funkcjonalnej.

Kolejnymi aspektami zmiany są **poszerzenie pola badawczego** translatoryki i przesunięcie jej centrum z tłumaczeń literackich, które przestały być głównym kierunkiem badań i stały się jednym z wielu nurtów, na tłumaczenia specjalistyczne. Samo pojęcie tłumaczeń specjalistycznych trudno jest zdefiniować tradycyjnymi kryteriami, o czym trafnie pisał Asensio (2007), a nowe kierunki badań dodatkowo rozmywiają to pojęcie. Na podstawie lektury najnowszych publikacji przekrojowych (np. Tymoczko 2005; Pym 2010; Munday 2010; Flynn, Gambier 2011; Brems i in. 2012; Snell-Hornby 2012; Piotrowska, Dybiec-Gajer 2012), analizy najnowszych zagranicznych publikacji książkowych oraz artykułów w czasopismach translatorycznych można zauważyć wyodrębnienie się i intensywny rozwój następujących **specjalności tematycznych** w obrębie translatoryki:

- lokalizacja stron WWW i gier,
- tłumaczenie audiowizualne oraz dostępność mediów (w tym takie zagadnienia jak audiodeskrypcja, dubbing, napisy do filmów, wielomodalność, media społecznościowe, tłumaczenia amatorskie – ang. *crowdsourcing, funsubbing*),
- tłumaczenie środowiskowe,
- tłumaczenie tekstów kultury popularnej,
- redakcja i weryfikacja tłumaczeń, a także *postediting*, czyli redakcja po tłumaczeniu maszynowym.

Warto ponadto odnotować kontynuację badań nad:

- procesem tłumaczenia (ang. *process research*) z zastosowaniem protokołów głośnego mówienia (TAP-ów), oprogramowania śledzącego ruchu na klawiaturze (Translog), badań okولوجraficznych,
- dydaktyką przekładu.

Wraz ze wzrostem skali działalności naukowo-badawczej specjalności tematyczne – zarówno nowe, jak i tradycyjne (np. tłumaczenia prawne i prawnicze, tłumaczenia techniczne, tłumaczenia ekonomiczne, tłumaczenia medyczne) – stają się coraz bardziej autonomiczne. Negatywnym aspektem tak dynamicznego rozwoju skutkującego mnogością specjalności, tematów, zagadnień i metodologii jest zagrożenie nadmierną fragmentacją dziedziny (Brems i in. 2012: 3).

Zastosowanie w translatoryce metod językoznawstwa korpusowego i narzędzi informatycznych doprowadziło do wyłonienia się odrębnej interdyscy-

scypinarnej specjalności o nazwie translatoryka korpusowa. Na mapie przekładoznawstwa Holmesa (1972/2004: 184) translatorykę korpusową można umieścić wśród badań podstawowych (*pure translation studies*) w tzw. nurcie opisowym, czyli empirycznym (*descriptive translation studies*), ale także wśród badań stosowanych (np. zastosowanie korpusów w dydaktyce przekładu). W klasyfikacji głównych nurtów badawczych zaproponowanej przez Marco, która obejmuje model tekstowo-deskryptywny, model kognitywny, model kulturowy i model socjologiczny, translatoryka korpusowa wpisuje się w model tekstowo-deskryptywny (2009: 15). Hatim natomiast klasyfikuje translatorykę korpusową jako badania empiryczne i badania nad procesem tłumaczenia (2001: 151).

Geneza translatoryki korpusowej

Translatoryka korpusowa została zapoczątkowana w połowie lat 90. serią pionierskich artykułów Mony Baker, która zaproponowała zastosowanie metodologii językoznawstwa korpusowego do badania cech tłumaczeń, tzw. uniwersaliów translatorycznych, na tle tekstów nieprzetłumaczonych (1993, 1994, 1996). Inspirację teoretyczną do badań nad językiem tłumaczeń stanowiły opisowe badania nad przekładem Toury'ego, teoria polisystemów oraz *skopostheorie*, dzięki którym zaczęto postrzegać tłumaczenia jako niezależne teksty funkcjonujące w kulturze docelowej. Pomysł Mony Baker podchwyciło wielu przekładoznawców. Obecnie badania korpusowe prowadzone są w licznych ośrodkach na świecie, jednak w Polsce są stosunkowo mało popularne, co może wiązać się z preferowaniem w polskich ośrodkach „lingwistyki gabinetowej” i ograniczoną popularnością empirycznych nurtów językoznawstwa, w tym lingwistyki korpusowej (przykładowo – chociaż na Uniwersytecie Łódzkim zespół PELCRA zaczął tworzyć korpus już w latach 90., to Narodowy Korpus Języka Polskiego powstał dopiero w 2012 r.).

Założenia językoznawstwa korpusowego jako podstawy metodologiczne translatoryki korpusowej

Językoznawstwo korpusowe jest jedną z bardzo prężnie rozwijających się metodologii w językoznawstwie. W jego obrębie można wyróżnić dwa nurty, które odmiennie postrzegają jego status jako teorii i jako metodologii (Hardie, McEnery 2010; Tognini-Bonelli 2001). Bardziej rozpowszechnione jest

drugie stanowisko, które uznaje językoznawstwo korpusowe za zbiór systematycznych metod i zasad stosowanych w wielu dziedzinach językoznawstwa przy wykorzystaniu różnych ram teoretycznych do interpretacji danych korpusowych (Hardie, McEnery 2010: 384–386). W translatoryce korpusy stosuje się jako metodologię (Olohan 2004: 9).

Badania korpusowe opierają się na dwóch zasadach sformułowanych przez Michaela Stubbsa (2004: 111):

- 1) obserwatorowi nie wolno ingerować w to, co obserwuje,
- 2) powtarzające się zdarzenia są istotne.

Pierwsza zasada oznacza konieczność pracy z autentycznymi danymi językowymi, natomiast druga jest kluczowa dla podejścia do znaczenia. Istotnymi informacjami dla językoznawcy korpusowego są częstość i dystrybucja danych językowych (zob. szerzej Lewandowska-Tomaszczyk 2005: 134). Jak podkreśla Stubbs, korpusy pokazują to, co jest „centralne i typowe, normalne i oczekiwane” w użyciu języka, dowodzą również, że język dostarcza nam licznych „prefabrykatów” (2004: 111). W centrum zainteresowania znajdują się nie słowa, ale powtarzające się jednostki (wzorce) leksykalno-gramatyczne. Prowadzi to do stawiania pytań dotyczących minimalnej jednostki znaczeniowej (zob. szerzej de Beaugrande 1996; Teubert 2002; Stubbs 2004; Lewandowska-Tomaszczyk 2005).

Językoznawstwo korpusowe jest podejściem empirycznym i indukcyjnym dzięki formułowaniu twierdzeń teoretycznych na podstawie obserwacji rzeczywistego użycia języka (zob. szerzej de Beaugrande 1996: 514; Tonini-Bonelli 2001: 2; Lewandowska-Tomaszczyk 2005: 11; McEnery i in. 2006: 7). Jest to metoda ilościowa, co jest przedmiotem krytyki, jednakże językoznawstwo korpusowe stara się również integrować metody jakościowe w znacznie większym stopniu niż językoznawstwo komputerowe. Główną zaletą korpusów jest ograniczenie spekulacji i subiektywności, autentyczność danych językowych, możliwość weryfikowania hipotez w sposób systematyczny, bardziej wiarygodny i na dużym materiale językowym. Ograniczenia metodologii wiążą się z dostępnością korpusów, w szczególności dla języków o ograniczonym zasięgu, reprezentatywnością i porównywalnością korpusów, a także z możliwościami technicznymi oprogramowania. O ile metody analiz korpusowych bardzo dobrze sprawdzają się przy badaniu bezpośredniego otoczenia słów, współwystępowania wyrazów i zbitek leksykalnych, frekwencyjności i dystrybucji jednostek językowych, o tyle znacznie gorzej radzą sobie ze zjawiskami tekstowymi czy z aspektami pragmatycznymi i dyskursywnymi, zwłaszcza jeżeli korpus nie jest znakowany (anotowany). Dlatego też często wymagają triangulacji metodologicznej, o czym będzie mowa w końcowej części artykułu.

Czym jest korpus?

Należy odróżnić potoczne (nad)użycie terminu „korpus” jako dowolnego zbioru tekstów od korpusu w formie elektronicznej stosowanego w językoznawstwie korpusowym. Korpus jest reprezentatywnym i zrównoważonym zbiorem tekstów zgromadzonych dla celu komputerowej analizy językowej (por. McEnery i in. 2006: 4–5). Wyróżnia się następujące rodzaje korpusów:

- 1) **korpus jednojęzyczny** – zawierający nieprzetłumaczone teksty w jednym języku (np. Narodowy Korpus Języka Polskiego). Jest to najbardziej typowy korpus stosowany przez językoznawców;
- 2) **korpus porównawczy** (ang. *comparable corpus*)
 - a) korpus dwujęzyczny/wielojęzyczny – zawierający teksty w co najmniej dwóch językach zgromadzone według podobnych kryteriów. Jest to korpus stosowany przez językoznawców kontrastywnych do badania różnic i podobieństw między językami,
 - b) korpus jednojęzyczny – zawierający korpus tłumaczeń i korpus spontanicznie tworzonych (tj. nieprzetłumaczonych) tekstów w języku docelowym. Tego typu korpusy są stosowane przez przekładoznawców do badania, w jaki sposób język tłumaczeń odbiega od języka docelowego. Metodę badań opracowała i rozpropagowała Mona Baker (1993, 1994, 1996);
- 3) **korpus równoległy** (ang. *parallel corpus, translation corpus*) – jest to korpus tłumaczeniowy w największym tego słowa znaczeniu. Zawiera zintegrowany korpus tłumaczeń i korpus tekstów źródłowych (jest zatem korpusem dwu- lub wielojęzycznym), w których zdania przetłumaczone są powiązane z odpowiednimi zdaniami w tekstach źródłowych. Tworzenie i analizowanie korpusów równoległych jest czasochłonne, dlatego są dużo mniejsze niż korpusy porównawcze. Jednym z największych korpusów równoległych jest korpus wspólnotowego dorobku prawnego JRC Acquis opracowany przez Joint Research Centre przy Komisji Europejskiej (<http://optima.jrc.it/Acquis>). Inne korpusy równoległe, zawierające również teksty w języku polskim, dostępne są na stronie Opus (<http://opus.lingfil.uu.se/>) oraz na stronie projektu Pelcra (<http://pelcra.pl>);
- 4) **korpus porównawczo-równoległy** (ang. *multilingually comparable corpus* Hansen-Schirra, Teich 2009, *composite bilingual corpus* Bernardini, Ferraresi 2011: 231) – jest to najnowszy rodzaj korpusu, łączący korpus równoległy z porównawczym, tj. zawiera co najmniej trzy zbiory tekstów: tłumaczenia, teksty źródłowe i teksty nieprze-

tłumaczone w języku docelowym. Jego powstanie wynika z głosów krytycznych kwestionujących prawidłowość badania tłumaczeń wyłącznie na korpusach porównawczych, bez odniesienia do tekstów źródłowych (np. Bernardini, Zanettin 2004: 59; Pym 2010: 82; Bernardini, Ferraresi 2011: 228).

W translatoryce korpusowej stosuje się zatem złożone korpusy o specjalnej konstrukcji: korpusy porównawcze (2b), korpusy równoległe (3) i porównawczo-równoległe (4). Bez względu na typ korpusu jedną z części składowych jest zawsze podkorpus tłumaczeń.

Oprogramowanie do analizy korpusowej

Korpusy porównawcze i równoległe wymagają innego rodzaju oprogramowania. Metody analizy korpusów porównawczych są szybsze, znacznie bardziej zaawansowane i wieloaspektowe. W tabeli 1 zestawiono popularne programy do analizy korpusów jednojęzycznych (a więc i porównawczych)³.

Tabela 1. Przykładowe programy do analizy korpusów jednojęzycznych

Nazwa programu	Opis	Koszt licencji
Wordsmith Tools	Kompleksowy program autorstwa Mike'a Scotta do tworzenia korpusów i ich analizy; jeden z najczęściej stosowanych program do badań naukowych http://www.lexically.net/wordsmith/	50 GBP
AntConc	Program autorstwa Laurence'a Anthony'ego, uznawany za najlepszą darmową alternatywę dla Wordsmitha http://www.antlab.sci.waseda.ac.jp/software.html	program bezpłatny
The Sketch Engine	Program autorstwa Adama Kilgarriffa. Automatycznie tworzy „mapy” zachowania gramatycznego i kolokacyjnego słów (ang. <i>word sketch</i>). Ma również wbudowane narzędzia analizy korpusów równoległych http://www.sketchengine.co.uk/	50 GBP rocznie
ConcGram	Program oferowany przez wydawnictwo John Benjamins, służący do analizy współwystępowania słów (ang. <i>concordgram</i>) i tworzenia profili frazeologicznych http://benjamins.com/#catalog/software/cls.1/main	45 EUR
KfNgram	Program do analizy zbitek leksykalnych, tzw. n-gramów http://kwicfinder.com/kfNgram/kfNgramHelp.html	program bezpłatny

³ Praktyczne zestawienie programów do analizy korpusowej znajduje się na stronie <http://www.uow.edu.au/~dlee/software.htm> (dostęp: 1 grudnia 2013 r.).

Warto pamiętać, że programy korpusowe powstają głównie z myślą o języku angielskim i znacznie trudniej pracuje się w nich na językach fleksyjnych, takich jak język polski (zob. szerzej Piotrowski 2007). Zespół IPI PAN opracował darmowy zestaw narzędzi Poliqarp do przeszukiwania korpusów języka polskiego (<http://korpus.pl/index.php?page=poliqarp>), jednakże ich funkcjonalność w porównaniu z Wordsmithem jest ograniczona.

Programy do analizy korpusów równoległych, zawierających teksty w dwóch językach, nie oferują tak dobrych narzędzi analizy ilościowej jak programy do korpusów jednojęzycznych. Analiza korpusów równoległych jest znacznie bardziej pracochłonna i technicznie ograniczona. Najbardziej znanym programem obsługującym korpusy równoległe jest ParaConc (koszt licencji 95 USD, <http://www.athel.com/para.html>). Zrzut ekranu (ilustracja 1) pokazuje zapytanie dla zwrotu z *zastrzeżeniem* w polsko-angielskim korpusie statutów spółek. W górnym oknie ParaConc pojawia się szukany zwrot w tekstach źródłowych. Jest on zaznaczony na niebiesko, a na czerwono – jego kolokacje z prawej i lewej strony. W dolnym oknie pojawiają się odpowiednie zdania z angielskich tłumaczeń.

ParaConc nie rozpoznaje automatycznie i nie wyróżnia kolorem angielskich ekwiwalentów zwrotu z *zastrzeżeniem*, nie jest w stanie ich posortować i przedstawić danych statystycznych dla poszczególnych wariantów. Dlatego też analizowanie tłumaczeń wymaga żmudnej ręcznej pracy. W podobny sposób działa funkcja konkordancji w narzędziach wspomagających tłumaczenia CAT (np. w Tradosie). Ograniczenia techniczne związane z analizą korpusów równoległych istotnie redukują możliwości zastosowania tej metody, w tym pracę na dużych korpusach.

Szczegółowe informacje praktyczne na temat budowania korpusów i ich znakowania można znaleźć w książkach Olohan (2004) i Zanettina (2012) oraz w języku polskim na stronie <http://korpusy.net> pod redakcją Błażeja Gałkowskiego.

25 matches	Polish - Original text order	Strings matching: z zastrzeżeniem
English (United States)	20 659 / 24 649	

ParConc - Statyngrupp - Parallel Concordance - [z zastrzeżeniem]	
File	Search
...	... o pocią głośowi na Walnym Zgromadzeniu, z zastrzeżeniem postanowień § 17. 4. Akcjonari A ...
...	... e wymagalą zgody Rady Nadzorczej, 3. z zastrzeżeniem ust. 4 w granicach kapitału doposażonego ...
...	... Klasyfikacja akcji i udziałów w kapitale zakładowym, 3. z zastrzeżeniem postanowień niniejszego paragrafu § 11 ...
...	... 2. Właściciel akcji i udziałów w kapitale zakładowym niniejszego paragrafu § 11 ...
...	... niniejszego paragrafu § 11 składowi, 3. z zastrzeżeniem postanowień niniejszego paragrafu, zania ...
...	... ch, przepisach innych ustaw i statucie, 3. z zastrzeżeniem postanowień ust. 2. Nabywca i zby ...
...	... dywanego wykonywana nauczyci, 3. z zastrzeżeniem postanowień ust. 4. Bez względu na większość 3/4 (p ...
...	... ch i sposobach: członkami Zarządu, 3. z zastrzeżeniem ust. 2, do 4 niniejszego paragrafu, wyła ...
...	... onkowie Rady Nadzorczej są powoływani - z zastrzeżeniem wyjątków opisanych w lit. a) i b) niniejsz ...
...	... głosie, Zarząd lub Rada Nadzorcza, 2. z zastrzeżeniem wyjątków przewidzianych w ust. 4. Zasadny ...
...	... zego paragrafu od § 20 niniejszego paragrafu, 3. z zastrzeżeniem postanowień niniejszego paragrafu i uz ...
...	... wyrażony jest przez Walne Zgromadzenie, 3. z zastrzeżeniem postanowień ust. 3. Niniejszego paragraf ...
...	... teni ust. 3 niniejszego paragrafu, 2. z zastrzeżeniem postanowień niniejszego paragrafu, Zarząd sk ...
...	... co najmniej 50% akcji plus jedna akcja, z zastrzeżeniem bezwzględnie obowiązujących przepisów p ...
...	... omy, w tym w szczególności: 1) z zastrzeżeniem postanowień niniejszego paragrafu, 2) w zakresie ...
...	... niosa funkcja członka Rady Nadzorczej, z zastrzeżeniem ust. 3, 2) w razie resignacji czł ...
...	... tuzza i wymaga uprzedniej zgody Spółki z zastrzeżeniem, iż Spółka nie odwołuje uzbelenia zgody ...
...	... em Europejskiego Urzędu Gospodarczego, z zastrzeżeniem, dochowania warunków ustanowionych w wy ...
...	... onka Rady Nadzorczej lub Zarządu, 6. z zastrzeżeniem postanowień ust. 7 niniejszego paragraf ...
...	... 6) Spółki lub jej podmiotów zależnych, z zastrzeżeniem, że niniejsze postanowienie nie stosuje ...
...	... Phees Zarządu kieruje pracami Zarządu, z zastrzeżeniem ust. 2, niniejszego paragrafu. Zarząd pod ...
...	3. Subject to the provisions of § 17 herein, series A and C shares shall be preferred shares in that each such share shall entitle its holder to five votes at the General Meeting of the Shareholders.
...	3. Subject to Section 4 within the limits of the authorized capital, the Management Board shall have the right to waive or to limit the preemptive rights upon consent of the Supervisory Board.
...	3. Redemption of the Company's shares requires a resolution of the General Meeting of Shareholders, subject to the provisions of Art. 363 paragraph 5 of the Commercial Companies Code.
...	3. Conversion of registered series A, B and C shares to bearer shares shall be made within 30 days from the date of filing an application by the shareholder holding such shares, subject to this paragraph and § 11 of the Statutes.
...	1. The General Meeting of the Shareholders shall have competence in matters reserved under the Commercial Companies Code, provisions of other laws and as provided herein, subject to section 2.
...	1. Subject to section 4, the absolute majority of 3/4 (three quarters) votes cast when the Shareholders representing at least 50% of the Company's share capital are present, shall be required for the resolution on the removal of matters from the agenda of the General Meeting of the Share
...	b) subject to Sections 2, 3 and 4 of this paragraph, granting consent for the Company to enter into or to amend an agreement with an affiliate of the Company within the meaning of the Law on the Public Trading of Securities, and enforcement regulations issued on the basis thereof (Pills
...	o) in the event that the nomination of the candidates is not made in accordance with the above guidelines and the provisions of § 20 Sections 4 and 6, the Management Board of the Supervisory Board shall nominate the candidates for members of the Supervisory Board.
...	3. Subject to the exceptions provided in section 5, the principles of making nominations for members and appointment of members of the Supervisory Board as provided in Section 2 of this paragraph and § 20, Section 2, shall apply to the newly appointed members in case of dismissal, ev
...	§ 22.
...	1. The Management Board is elected by the General Meeting of the Shareholders, except for provisions of section 3 of this paragraph.
...	4. The General Meeting of Shareholders referred to in Section 3 should be held on the date falling - not later than within eight weeks, after the General Meeting of Shareholders which has not adopted the resolutions for the lack of quorum.
...	§ 11
...	2) In the event of resignation of a member of the Supervisory Board from his position.
...	§ 9.
...	The Agreement transferring title to the buyer named by the Management Board and payment of the price by the buyer will take place within 14 days of naming such a buyer by the Management Board, unless the buyer and the selling Shareholder agree to a different date. If the buyer nam
...	2. The General Shareholders' Meeting is authorized to set the date as of which the list of shareholders entitled to dividend for a given financial year shall be determined ("the Dividend Date") and the date as of which the dividend shall be paid.
...	4. To the extent permitted by the law, the members of the Supervisory Board may participate in addition of the Supervisory Board's resolutions by casting their votes in writing through another member of the Supervisory Board
...	4. Elaborating publications of annual reports and financial statements of the Company prepared pursuant to the International Financial Reporting Standards (IFRS).

Ilustracja 1

Podstawowe metody analizy korpusowej

W tej części zaprezentuję podstawowe metody analizy korpusowej na przykładzie funkcji programu Wordsmith: listy frekwencyjne, słowa kluczowe i konkordancje.

Listy frekwencyjne (WordList)

Jest to podstawowe narzędzie umożliwiające uzyskanie globalnych informacji o tekstach w korpusie. Prezentuje listę słów pojawiających się w korpusie według częstości występowania. Zrzut ekranu (ilustracja 2) pokazuje górną część listy słów dla zbudowanego przez autorkę Korpusu polskiego prawa (zob. Biel 2014).

Wordsmith tworzy również statystykę dla całego korpusu i dla poszczególnych tekstów w korpusie, w tym: liczbę słów i zdań, średnią długość zdania, odchylenia standardowe, słowa według liczby znaków, co pokazuje zrzut ekranu (ilustracja 3).

Słowa kluczowe (KeyWords)

Funkcja ta jest bardzo przydatna szczególnie dla przekładoznawców. Umożliwia porównanie list słów z dwóch korpusów, pokazując, które słowa są nadreprezentowane lub niedoreprezentowane, tj. występują znacznie częściej lub rzadziej niż w korpusie referencyjnym. Zrzut ekranu (ilustracja 4) pokazuje część listy słów kluczowych dla korpusu tłumaczeń (dyrektyw unijnych), dla którego korpusem referencyjnym jest Korpus polskiego prawa, a więc korpus tekstów nieprzetłumaczonych w języku docelowym.

Wśród słów występujących znacznie częściej w korpusie tłumaczeń znajdują się: *niniejszej*, *musi/muszą*, *zgodnie*, *aby*, *jeśli*, *typu*, *dla*. W trzeciej kolumnie oznaczonej jako „Freq.” wyświetla się liczba wystąpień danego słowa w korpusie podstawowym, czyli w tym przypadku w korpusie dyrektyw. W piątej kolumnie („FC Freq.”), zaznaczona na czerwono, pojawia się frekwencyjność danego słowa w korpusie referencyjnym, tutaj w Korpusie polskiego prawa. Oba korpusy mają podobną wielkość, tj. ok. 6 mln słów. Przykładowo *musi* pojawia się aż 10 928 razy w dyrektywach, natomiast w korpusie referencyjnym tylko 213 razy. Dla badacza jest to wskazówka do dalszej analizy w celu ustalenia przyczyn nadreprezentacji poszczególnych jednostek (np. interferencja z języka źródłowego).

Istnieje również możliwość odwrócenia list słów, dzięki czemu uzyskamy wykaz słów niedoreprezentowanych w tłumaczeniu.

PLC-HAB.txt

File Edit View Compute Settings Windows Help

case sensitive search

N	Word	Freq.	%	Texts	%	Lemmas	Set
1	#	790 806	11,56	755	100,00		
2	W	307 767	4,50	755	100,00		
3	I	140 658	2,06	745	98,68		
4	Z	139 703	2,04	755	100,00		
5	O	120 972	1,77	755	100,00		
6	ART	115 279	1,68	755	100,00		
7	DO	107 090	1,57	734	97,22		
8	NA	90 198	1,32	741	98,15		
9	LUB	83 762	1,22	703	93,11		
10	NR	82 317	1,20	692	91,66		
11	POZ	78 282	1,14	687	90,99		
12	UST	71 324	1,04	699	92,58		
13	SIE	56 277	0,82	745	98,68		
14	ORAZ	54 535	0,80	720	95,36		
15	MOWA	54 001	0,79	689	91,26		
16	R	50 865	0,74	755	100,00		
17	PRZEZ	38 649	0,56	720	95,36		
18	NIE	37 539	0,55	709	93,91		
19	KTÓRYCH	31 778	0,46	701	92,85		
20	DNIA	30 169	0,44	755	100,00		
21	OD	25 862	0,38	719	95,23		
22	JEST	24 995	0,37	693	91,79		
23	KTÓRYM	23 613	0,35	665	88,08		
24	PKT	22 860	0,33	641	84,90		
25	MOŻE	22 426	0,33	673	89,14		
26	A	21 958	0,32	689	91,26		
27	JEŻELI	21 706	0,32	644	85,30		
28	USTAWY	20 813	0,30	709	93,91		
29	ZA	19 648	0,29	688	91,13		

frequency alphabetical statistics filenames notes

Ilustracja 2

	1	2	3	4	5	6	7
Overall							
text file	Konstyytucja_RP.txt	ustawa_budzetowa_2011.txt	Kodeks_cywilnyy.txt	Kodeks_karny.txt	Kodeks_karny_skarbowy.txt	Kodeks_morski.txt	Codeks_pracy_1974_rodzini
file size	216 258	40 106	888 250	455 398	370 694	404 910	568 118
tokens (running words) in text	14 009	3 066	30 829	30 829	25 145	26 500	37 932
tokens used for word list	12 981	2 455	54 506	27 914	22 159	24 727	34 655
sum of entries							
types (distinct words)	2 810	555	6 157	3 845	3 207	4 147	4 524
type/token ratio (TTR)	21.65	22.61	11.30	13.77	14.47	16.77	13.05
standardised TTR	35.26	42.11	26.40	41.06	34.80	41.91	37.15
standardised TTR std.dev.	63.40	50.20	56.63	60.93	60.69	54.99	60.47
standardised TTR basis	1 000	1 000	1 000	1 000	1 000	1 000	1 000
mean word length (in characters)	5.82	5.11	6.18	5.93	5.90	6.22	6.05
word length std.dev.	3.91	3.72	3.60	3.67	3.80	3.84	3.86
sentences	309 410	944	3 590	1 264	1 123	1 366	1 681
mean (in words)	19.56	12.09	15.18	22.08	19.73	18.10	20.62
std.dev.	33.49	24.89	15.73	18.82	19.15	18.02	21.56
paragraphs	93 661	87	115	137	159	204	454
mean (in words)	149.21	144.41	473.97	203.75	139.36	121.21	76.33
std.dev.	181.54	274.86	1 128.13	615.16	369.56	314.15	164.78
headings							
mean (in words)							
std.dev.							
sections	755						
mean (in words)	8 014.67	2 455.00					
std.dev.	10 949.04						
numbers removed	1 028	611	2 694	2 915	2 986	1 773	3 277
stoptlist tokens removed							
stoptlist types removed							
1-letter words	1 185 951	537	5 553	3 865	3 741	3 435	6 112
2-letter words	609 844	427	4 323	2 390	1 975	2 187	2 650
3-letter words	757 487	488	6 825	4 519	3 432	2 539	3 635
4-letter words	501 156	254	4 244	2 090	1 630	1 685	2 503

Ilustracja 3

The screenshot shows the KeyWords application window. The main area displays a list of words with columns for N, Key word, Freq, % RC, Freq, RC, %, Keyness, and Lemmas. The words are sorted by frequency, with 'ARTYKUŁ' having the highest frequency (26,047) and 'ZŁĄCZNIKA' having the lowest (6,749).

N	Key word	Freq	% RC	Freq	RC	%	Keyness	Lemmas
1	ARTYKUŁ	26 047	0.36	13			34 591.09	0.0000000000
2	CZŁONKOWSKIE	28 076	0.39	381			34 013.91	0.0000000000
3	DYREKTYWY	30 946	0.43	983	0.01		33 983.64	0.0000000000
4	EWG	27 063	0.38	956	0.01		29 190.60	0.0000000000
5	DYREKTYWA	13 509	0.19	56			17 389.05	0.0000000000
6	NINIEJSZEJ	17 266	0.24	2 057	0.03		12 897.77	0.0000000000
7	MUSI	10 928	0.15	213			12 786.23	0.0000000000
8	STR	15 446	0.21	1 538	0.02		12 514.24	0.0000000000
9	ZGODNIE	23 406	0.32	4 838	0.07		12 347.85	0.0000000000
10	C	18 931	0.26	3 243	0.05		11 486.67	0.0000000000
11	MUSZA	9 538	0.13	257			10 850.63	0.0000000000
12	L	15 349	0.21	2 085	0.03		10 725.72	0.0000000000
13	DYREKTYWA	8 074	0.11	46			10 274.70	0.0000000000
14	PPKT	7 667	0.11	1			10 214.34	0.0000000000
15	ZŁĄCZNIKU	10 566	0.15	654			10 057.82	0.0000000000
16	ABY	10 221	0.14	697	0.01		9 461.90	0.0000000000
17	JESLI	7 989	0.11	215			8 981.78	0.0000000000
18	PANSTWA	29 640	0.41	10 255	0.15		8 857.25	0.0000000000
19	II	9 815	0.14	813	0.01		8 528.74	0.0000000000
20	TYPU	8 440	0.12	445			8 373.10	0.0000000000
21	ZŁĄCZNIK	7 943	0.11	439			7 788.79	0.0000000000
22	DLA	32 561	0.45	13 124	0.19		7 567.39	0.0000000000
23	CZŁONKOWSKICH	9 963	0.14	1 226	0.02		7 328.06	0.0000000000
24	HOMOLOGACJI	5 897	0.08	70			7 208.93	0.0000000000
25	BYĆ	28 885	0.40	11 244	0.16		7 141.68	0.0000000000
26	MG	5 752	0.08	60			7 094.34	0.0000000000
27	WE	18 100	0.25	5 108	0.07		7 050.91	0.0000000000
28	KG	5 672	0.08	85			6 806.33	0.0000000000
29	DYREKTYWIE	5 187	0.07	24			6 650.36	0.0000000000
30	MALEZY	14 626	0.20	3 600	0.05		6 589.81	0.0000000000
31	ZŁĄCZNIKA	6 749	0.09	417			6 426.03	0.0000000000

Ilustracja 4

N	Key word	Freq.	%	RC	Freq.	RC	%	Keymess	F	Lenmas
1	POZ	78 282	1,14	18	112 862,83	0,0000000000				
2	NR	82 317	1,20	5 311	85 990,27	0,0000000000	0,07			
3	O	120 972	1,77	31 740	61 055,96	0,0000000000	0,44			
4	MOWA	54 001	0,79	3 484	56 308,91	0,0000000000	0,05			
5	ART	115 279	1,68	40 071	42 521,59	0,0000000000	0,56			
6	USTAWY	20 813	0,30	179	28 176,26	0,0000000000				
7	SPRAWY	15 611	0,23	250	20 257,66	0,0000000000				
8	UST	71 324	1,04	29 772	20 002,33	0,0000000000	0,41			
9	KTORYM	23 613	0,35	4 861	14 485,09	0,0000000000	0,07			
10	MINISTER	9 758	0,14	16	13 841,28	0,0000000000				
11	ALBO	18 917	0,28	3 705	12 024,91	0,0000000000	0,05			
12	KTORYCH	31 778	0,46	11 068	11 600,61	0,0000000000	0,15			
13	POLSKIEJ	7 810	0,11	8	11 134,32	0,0000000000				
14	PKT	22 860	0,33	6 284	10 934,21	0,0000000000	0,09			
15	USTAWIE	7 394	0,11	7	10 548,20	0,0000000000				
16	RZECZYPOSPOLITEJ	7 317	0,11	2	10 505,25	0,0000000000				
17	R	50 865	0,74	25 514	10 022,95	0,0000000000	0,35			
18	URZEDU	7 090	0,10	133	9 063,50	0,0000000000				
19	POSTEPOWANIA	9 291	0,14	826	8 763,54	0,0000000000	0,01			
20	DRODZE	10 518	0,15	1 408	8 369,80	0,0000000000	0,02			
21	UCHYLONY	5 709	0,08	11	8 077,65	0,0000000000				
22	ROZPORZADZENIA	8 461	0,12	747	7 999,42	0,0000000000	0,01			
23	SLUZBY	7 410	0,11	414	7 986,06	0,0000000000				
24	KTOREJ	14 447	0,21	3 391	7 983,39	0,0000000000	0,05			
25	SAD	5 868	0,09	101	7 562,00	0,0000000000				
26	FUNDUSZU	6 511	0,10	258	7 530,13	0,0000000000				
27	WLASCIWY	10 685	0,16	1 941	7 143,95	0,0000000000	0,03			
28	SADU	5 482	0,08	117	6 914,83	0,0000000000				
29	CZYNNOSCI	8 389	0,12	1 164	6 555,36	0,0000000000	0,02			
30	POZN	4 345	0,06	1	6 240,14	0,0000000000				
31	ZM	4 342	0,06	4	6 194,65	0,0000000000				

Ilustracja 5

N	Concordance	Word#	Sent.	Sent.	Para.	Para.	Sect.	Sect.	File	Date	%	
		f	f	f	f	f	f	f				
1370	Marynarki Wojennej, Straży Granicznej i Policji, z zaszereżeniem przypadków określonych w ustawie 1a.	189	9	75%	2	75%	0	1%	ustawa_o_bezpieczenstwie_j	2012-mar-23	00	1%
1371	o udostępnienie informacji publicznej stosuje się, z zaszereżeniem przypadku , o którym mowa w art. 22.	2.352	81	47%	33	90%	0	80%	ustawa_o_dostepie_do_infor	2012-mar-23	00	83%
1372	lub niewykonanie rozkazu lub innego polecenia, z zaszereżeniem przypadku określonego w art. 63 ust. 2	26.308	910	25%	323	95%	0	82%	ustawa_o_strazy_granicznej	2012-mar-23	00	84%
1373	na dopuszczenie do obrotu produktu leczniczego, z zaszereżeniem spełnienia przez podmiot	15.694	315	16%	266	49%	0	34%	Prawo_farmaceutyczne_bt	2012-mar-23	00	35%
1374	wprowadzonych na terytorium kraju olejów smarowych, z zaszereżeniem spełnienia warunków określonych w	1.587	38	82%	33	44%	0	21%	ustawa_o_obowiazkach_prze	2012-mar-23	00	21%
1375	wprowadzonych na terytorium kraju olejów smarowych, z zaszereżeniem spełnienia warunków określonych w	1.457	33	84%	33	7%	0	9%	ustawa_o_obowiazkach_prze	2012-mar-23	00	20%
1376	zastępcę lub zastępców pełnomocnika ochrony, z zaszereżeniem spełnienia przez te osoby warunków, o	3.694	101	54%	94	52%	0	9%	ustawa_o_ochronie_informac	2012-mar-23	00	10%
1377	na dopuszczenie do obrotu produktu leczniczego, z zaszereżeniem spełnienia przez podmiot	6.188	156	11%	104	20%	0	44%	ustawa_o_urzedzie_rejestrac	2012-mar-23	00	47%
1378	wyroczona w pierwszej instancji orzeka sąd rejonowy, z zaszereżeniem praw określonych w art. 10. § 2. Sąd	740	32	75%	15	89%	0	6%	Kodeks_postepowania_w_sp	2012-mar-23	00	6%
1379	czyżności w zakresie ich merytorycznego działania z zaszereżeniem praw objętych wydawaniem decyzji	41.217	113	19%	810	43%	0	90%	Prawo_farmaceutyczne_bt	2012-mar-23	00	91%
1380	14 dni od doręczenia zawiadomienia o jej terminie, z zaszereżeniem sprawy określonej w art. 2 ust. 3, do	4.332	236	58%	51	75%	0	61%	ustawa_o_tribunale_konstytj	2012-mar-23	00	65%
1381	uznania, jak też przystępować do tych organizacji, z zaszereżeniem stosowania się do ich statutow. 2.	44	4	79%	0	14%	0	3%	ustawa_o_organizacjach_pra	2012-mar-23	00	3%
1382	osoby po uprzednim, osobistym jej zbadaniu, z zaszereżeniem sytuacji określonych w odrębnych	25.635	692	71%	433	35%	0	90%	ustawa_o_zawodach_lekarza	2012-mar-23	00	91%
1383	na podstawie osobiste wykonanych badań i ustaleń, z zaszereżeniem sytuacji określonych w odrębnych	25.656	694	71%	433	40%	0	90%	ustawa_o_zawodach_lekarza	2012-mar-23	00	91%
1384	w udzielaniu tych świadczeń. 3. Piętniąrka, pobozna, z zaszereżeniem sytuacji , o których mowa w ust. 2 pkt	9.974	304	23%	215	5%	0	87%	ustawa_o_zawodach_pielegn	2012-mar-23	00	88%
1385	nastąpić wyłącznie w niezbędnym zakresie. 3. Lekarz, z zaszereżeniem sytuacji , o których mowa w ust. 2 pkt	25.550	686	19%	433	12%	0	90%	ustawa_o_zawodach_lekarza	2012-mar-23	00	91%
1386	powinno zakończyć się nie później niż w ciągu 210 dni, z zaszereżeniem terminów określonych w art. 16a i 19.	10.582	203	71%	164	88%	0	23%	Prawo_farmaceutyczne_bt	2012-mar-23	00	23%
1387	nie dłuższym niż 3 miesiące od dnia złożenia wniosku, z zaszereżeniem terminu określonego w ust. 1. 3.	971	29	73%	28	55%	0	5%	ustawa_Karta_Nauczyciela_	2012-mar-23	00	5%
1388	nie może być przeniesiona pod warunkiem ani z zaszereżeniem terminu § 2. Jeżeli umowa	8.506	546	88%	24	33%	0	15%	Kodeks_cywilny_bt	2012-mar-23	00	14%
1389	i bezwarunkowy, jak również nie może być zbżony z zaszereżeniem terminu . Art. 27. 1. Towarzystwo	4.100	120	91%	119	31%	0	7%	ustawa_o_funduszach_inwes	2012-mar-23	00	7%
1390	nieruchomości została zawarta pod warunkiem lub z zaszereżeniem terminu , do przeniesienia własności	8.523	547	47%	24	33%	0	15%	Kodeks_cywilny_bt	2012-mar-23	00	14%
1391	lub poleceniem albo uchyleniu pod warunkiem lub z zaszereżeniem terminu . Art. 999. Jeżeli spadkobierca	55.011	335	94%	102	81%	0	94%	Kodeks_cywilny_bt	2012-mar-23	00	94%
1392	Art. 975. Zapis może być uchyleniu pod warunkiem lub z zaszereżeniem terminu . Art. 976. W braku odmienej	54.076	329	83%	102	37%	0	93%	Kodeks_cywilny_bt	2012-mar-23	00	93%
1393	§ 4. Zapis na akcje dokonany pod warunkiem lub z zaszereżeniem terminu jest nieważny. § 5. Nieważne	36.973	199	73%	320	9%	0	64%	Kodeks_spolek_handlowych	2012-mar-23	00	65%
1394	lub o odruczeniu spadku złożone pod warunkiem lub z zaszereżeniem terminu jest nieważne. § 2.	55.785	341	80%	105	80%	0	96%	Kodeks_cywilny_bt	2012-mar-23	00	96%
1395	Funduszu. 2) gospodarowanie środkami Funduszu, z zaszereżeniem uprawnień Rady Funduszu: 3)	2.879	86	34%	56	56%	0	25%	ustawa_o_Bankowym_Fundu	2012-mar-23	00	26%
1396	Narodowego Funduszu i wojewódzkich funduszy, z zaszereżeniem uprawnień Rady Nadzorczej	47.015	164	50%	978	45%	0	82%	Prawo_ochrony_srodowiska_i	2012-mar-23	00	83%
1397	ust. 2. (określony) Art. 3. 1. Przekazaniu, z zaszereżeniem ust. 1a , mogą podlegać nieruchomości,	345	24	5%	4	29%	0	29%	ustawa_o_zasadach_przekaz	2012-mar-23	00	27%
1398	Art. 31. 1. Ilistalania kontrolni, inkumentowane sa, z zaszereżeniem ust. 10 , w formie wrotokolu. 2	4.641	157	57%	99	83%	0	28%	ustawa_o_Panstwowej_Ispec	2012-mar-23	00	30%

concordance plot patterns clusters timeline filenames source text notes

Ilustracja 6

Concord
File Edit View Compute Settings Windows Help
case sensitive search

N	Word	With Relation	Texts	Total	Total	f	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5	
1	UST	z zastrzeżeniem	0,000	435	2 581	171	2 410	25	50	10	86	0	0	1 898	26	439	27	20
2	ART	z zastrzeżeniem	0,000	336	1 228	130	1 098	14	51	14	51	0	0	934	36	18	84	26
3	PRZEPISÓW	z zastrzeżeniem	0,000	95	156	38	118	4	6	8	2	18	0	109	2	1	3	3
4	PKT	z zastrzeżeniem	0,000	124	277	68	209	16	19	3	30	0	0	86	2	59	3	59
5	ŻE	z zastrzeżeniem	0,000	51	74	4	70	3	1	0	0	0	0	59	0	4	7	7
6	2	z zastrzeżeniem	0,000	266	766	184	582	33	40	34	9	68	0	55	382	30	100	15
7	PRZEPISU	z zastrzeżeniem	0,000	27	37	0	37	0	0	0	0	0	0	36	0	1	0	0
8	WYJĄTKÓW	z zastrzeżeniem	0,000	18	33	0	33	0	0	0	0	0	0	33	0	0	0	0
9	3	z zastrzeżeniem	0,000	240	525	100	425	9	16	23	21	31	0	27	189	26	168	15
10	4	z zastrzeżeniem	0,000	149	301	61	240	8	9	12	15	17	0	16	131	14	73	6
11	5	z zastrzeżeniem	0,000	139	238	51	187	5	9	13	11	13	0	12	111	11	50	3
12	LIT	z zastrzeżeniem	0,000	25	30	8	22	1	0	1	6	0	0	11	0	5	0	6
13	PRZYPADKÓW	z zastrzeżeniem	0,000	7	8	0	8	0	0	0	0	0	0	8	0	0	0	0
14	TERMINU	z zastrzeżeniem	0,000	10	15	5	10	1	2	2	0	0	0	8	0	0	0	2
15	WARUNKÓW	z zastrzeżeniem	0,000	19	24	11	13	4	2	0	4	1	0	8	2	2	0	1
16	ZACHOWANIA	z zastrzeżeniem	0,000	11	13	3	10	1	0	0	2	0	0	7	2	1	0	0
17	6	z zastrzeżeniem	0,000	106	152	33	119	7	1	6	5	14	0	5	69	9	33	3
18	OBOWIĄZKU	z zastrzeżeniem	0,000	14	17	9	8	3	0	2	3	1	0	5	0	0	1	2
19	SPEŁNIENIA	z zastrzeżeniem	0,000	5	6	1	5	0	0	1	0	0	0	5	0	0	0	0
20	ZŁĄCZNIKA	z zastrzeżeniem	0,000	2	6	1	5	0	0	1	0	0	0	5	0	0	0	0
21	7	z zastrzeżeniem	0,000	59	83	15	68	2	6	3	0	4	0	4	37	4	22	1
22	DOPEŁNIENIA	z zastrzeżeniem	0,000	2	4	0	4	0	0	0	0	0	0	4	0	0	0	0
23	SYTUACJI	z zastrzeżeniem	0,000	2	5	0	5	0	0	0	0	0	0	4	0	0	0	1
24	2A	z zastrzeżeniem	0,000	47	77	0	77	0	0	0	0	0	0	3	40	17	12	5
25	POSTANOWIEN	z zastrzeżeniem	0,000	5	5	1	4	0	0	1	0	0	0	3	0	0	0	1
26	WYNIKAJĄCYM	z zastrzeżeniem	0,000	3	3	0	3	0	0	0	0	0	0	3	0	0	0	0
27	ZMIAN	z zastrzeżeniem	0,000	6	6	1	5	0	1	0	0	0	0	3	1	1	0	0
28	2-4	z zastrzeżeniem	0,000	22	24	4	20	0	1	0	0	3	0	2	17	0	1	0
29	3A	z zastrzeżeniem	0,000	25	42	0	42	0	0	0	0	0	0	2	20	14	6	0

concordance collocates plot patterns clusters timeline filenames source text notes

Ilustracja 7

Zrzut ekranu (ilustracja 5) pokazuje słowa kluczowe w Korpusie polskiego prawa, gdy korpusem referencyjnym jest korpus dyrektyw. Są to więc słowa, które znacznie częściej występują w aktach prawnych w języku docelowym i nie są „wywoływane” w tłumaczeniu (tzw. *unique items* – jednostki unikalne). Brak jednostek unikalnych lub ich niska frekwencyjność w tłumaczeniach wywołuje u odbiorcy wrażenie „trzeciego języka”, tj. powoduje, że tłumaczenia są postrzegane jako teksty przetłumaczone, nienaturalne (zob. szerzej Tirkkonen-Condit 2004).

Konkordancje

Funkcja Concord umożliwia badanie otoczenia słowa lub wyrażenia. Zrzut ekranu (ilustracja 6) pokazuje zapytanie dla zwrotu z *zastrzeżeniem* w Korpusie polskiego prawa. Program pozwala na sortowanie według kolokacji z lewej lub prawej strony (do pięciu wyrazów w prawo i lewo od wyszukiwanego zwrotu), a także sortowanie według tekstu czy miejsca w tekście.

Wordsmith również automatycznie podsumowuje kolokacje oraz pokazuje typowe zbitki leksykalne dla danego zwrotu. Na następnym zrzucie ekranu (ilustracja 7) znajduje się podsumowanie kolokacji z pierwszego miejsca po prawej stronie dla zwrotu z *zastrzeżeniem*. I tak, najczęstszymi kolokacjami są: z *zastrzeżeniem ust.* (1898 wystąpień w korpusie), z *zastrzeżeniem art.*, z *zastrzeżeniem przepisów*, z *zastrzeżeniem pkt*, z *zastrzeżeniem że*.

Ciekawe wyniki może dać porównanie kolokacji danego zwrotu w korpusie tłumaczeń i korpusie nie-tłumaczeń.

Kierunki badań translatoryki korpusowej

W obrębie translatoryki korpusowej można wyróżnić dwa podstawowe nurty ze względu na charakter badań (Laviosa 2010: 83–86):

- **badania podstawowe**: (ang. *corpus-based descriptive studies*) – zastosowanie korpusów w celach naukowych w badaniach nad specyfiką tłumaczeń,
- **badania stosowane** (ang. *corpus-based applied studies*) – zastosowanie korpusów w celach praktycznych, głównie w kształceniu tłumaczy, praktyce tłumaczeniowej oraz przy opracowywaniu materiałów dla tłumaczy.

Celem badań podstawowych jest pogłębienie wiedzy na temat specyfiki przekładu. Badania te prowadzą do nowych pytań i hipotez. Dostarczają bogatego materiału empirycznego, który umożliwia testowanie hipotez i weryfikowanie twierdzeń teoretycznych (Laviosa 2010: 84).

Głównym nurtem badań są tzw. **uniwersalia translatoryczne**, czyli cechy charakterystyczne tłumaczeń:

- a) w stosunku do tekstów źródłowych – tzw. S-uniwersalia dotyczące relacji ekwiwalencji; badania prowadzone są na korpusach równoległych;
- b) w stosunku do nieprzetłumaczonych tekstów w języku źródłowym – tzw. T-uniwersalia dotyczące relacji dopasowania tekstowego, ang. *textual fit*; badania prowadzone głównie na korpusach porównawczych (zob. szerzej Chesterman 2004). W szczególności badaczy interesują zniekształcenia językowe i inne rozbieżności w porównaniu z językiem docelowym.

O ile relacja ekwiwalencji jest często omawianym zagadnieniem, o tyle relacja dopasowania tekstowego rzadko jest przedmiotem badań. Do wzrostu zainteresowania tą relacją przyczyniły się korpusy porównawcze, które umożliwiają badanie jej w bardziej systematyczny i kompleksowy sposób. Warto podkreślić, że badania nad dopasowaniem tekstowym zwracają uwagę na aspekty komunikacyjne przekładu, pokazując, jak tłumaczenia odbiegają od konwencji języka docelowego, a zarazem od oczekiwań odbiorcy.

Do potencjalnych uniwersaliów zalicza się (zob. Baker 1993, 1996; Chesterman 2004; Olohan 2004; Tirkkonen-Condit 2004; Mauranen 2007; Biel 2010; Zanettin 2012; Biel 2014):

- eksplicytację – tendencję tłumaczy do precyzowania znaczenia, które w tekście źródłowym jest *implicite*,
- uproszczenie przekazu tekstu źródłowego lub jego języka oraz eliminowanie wieloznaczności,
- normalizację, konwencjonalizację, standaryzację: tendencję do przesadnego odzwierciedlenia typowych cech języka docelowego,
- spłaszczenie (ang. *levelling-out*): tłumaczenia są mniej idiosynkratyczne i bardziej podobne do siebie niż oryginały,
- nietypowe i mniej stabilne kolokacje,
- niedoreprezentację jednostek unikalnych.

Hipoteza uniwersaliów translatorycznych wzbudziła (i wzbudza) liczne kontrowersje wśród badaczy, przede wszystkim ze względu na niezbyt trafnie dobraną nazwę. Obecnie w literaturze anglojęzycznej odchodzi się od terminu „uniwersalia”, który implikuje konieczność wykazania prawdziwości twierdzeń o przekładzie w każdych okolicznościach, na rzecz terminów implikujących prawidłowości, tendencje, regularności (ang. *patterns, features, tendencies, regularities, qualified generalisations*). Kolejnym punktem spornym jest radykalizm silnej wersji hipotezy zaproponowanej przez Monę Baker w 1993 r.,

która zakłada, iż uniwersalia są rezultatem ograniczeń specyficznych dla procesu tłumaczenia, nie wynikają z interferencji i są niezależne od takich czynników jak para językowa, kierunek tłumaczenia, gatunek tekstu, kultura czy normy. Badania pokazały, że większość cech tłumaczeń jest uzależniona od asymetrii między językami czy od gatunku tekstu (początkowo większość badań prowadzono na tekstach literackich). Stąd ostatnio popularność zyskała słabsza wersja hipotezy zakładająca, że specyficzne cechy tłumaczeń mogą wynikać z różnorodnych czynników, w tym np. z interferencji. Wkładem badań nad uniwersaliami jest, jak podkreśla Chesterman, odkrywanie prawidłowości i tendencji w przekładzie oraz formułowanie generalizacji na jego temat, które ułatwiają zrozumienie specyfiki przekładu (2004: 11).

Innym zagadnieniem badanym na korpusach porównawczych są style tłumaczy (literackich) – zestawia się w korpusie kilka tłumaczeń tego samego dzieła i bada różnice między nimi. Podejmowane są również próby porównania tłumaczeń wykonanych przez tłumaczy oraz tłumaczeń maszynowych.

Ponieważ w skład korpusów równoległych wchodzi teksty źródłowe, głównym przedmiotem zainteresowań jest relacja ekwiwalencji na różnych poziomach języka. Badania na korpusach równoległych dotyczą najczęściej zjawisk gramatycznych w przekładzie (np. użycie zaimków osobowych, strony biernej, konstrukcji czasownikowych), interferencji oraz technik i strategii tłumaczenia, zwłaszcza przy doborze ekwiwalentów terminologicznych. Badania dostarczają informacji deskryptywnych o praktyce tłumaczeniowej. Znaczącym wkładem badań korpusowych, w szczególności badań na korpusach równoległych, jest zdaniem Baker odejście od podejścia preskryptywnego na rzecz podejścia deskryptywnego (1995: 231).

Praktyczne zastosowania korpusów zostały ujęte w tabeli 2.

Tabela 2. Praktyczne zastosowania korpusów

Praktyka tłumaczeniowa	<p>Korpusy narodowe można wykorzystać do redukcji interferencji z języka źródłowego oraz zwiększenia naturalności tłumaczeń, zwłaszcza w przypadku tłumaczenia na język nieojczysty. Zaletą korpusów są informacje o częstotliwości określonych sformułowań/łączności wyrazów w danym kontekście, które umożliwiają wyeliminowanie połączeń możliwych, lecz rzadkich na rzecz połączeń typowych, częstych. Narodowy Korpus Języka Polskiego (http://nkjp.pl) można przeszukiwać według rodzajów tekstów, co zwiększa stopień kontroli nad tekstami referencyjnymi w porównaniu z nie zawsze wiarygodnym googlowaniem. Jeszcze większy stopień kontroli można uzyskać, budując niewielkie dziedzinowe korpusy ad hoc (ang. <i>DIY corpora</i>, <i>ad hoc corpora</i>) na potrzeby danego zlecenia czy klienta (zob. Monzó Nebot 2008, Scott 2012). Przykładowo – gdy tłumacz dostaje do przetłumaczenia umowę licen-</p>
------------------------	--

Tabela 2. Cd.

Praktyka tłumaczeniowa	<p>cyzną na język angielski, może stosunkowo szybko (w ciągu 20–30 minut) zbudować korpus angielskich umów licencyjnych, w którym będzie weryfikować prawidłowość ekwiwalentów, sposób osadzenia terminów w tekście i typowe związki wyrazowe. Tłumacze często nie są ekspertami dziedzinowymi i dzięki pracy z korpusami mogą w większym stopniu dostosować się do oczekiwań stylistycznych odbiorcy specjalistycznego, sprawniej imitując jego język.</p> <p>Aby korpusy były wykorzystywane na większą skalę w praktyce tłumaczeniowej, niezbędne jest zintegrowanie narzędzi korpusowych z narzędziami wspomagającymi tłumaczenia (CAT-ami) (por. Kübler i Aston 2010: 512).</p>
Kształcenie tłumaczy	<p>Korpusy wiążą się z indukcyjnymi metodami nauczania określanymi mianem nauczania/„odkrywania” opartego na danych językowych (ang. <i>data-driven (discovery) learning</i>; zob. Granger 2003: 24, Stubbs 2004: 107). Praca z korpusami jednojęzycznymi i porównawczymi uwrażliwia studentów na konwencje języka docelowego, natomiast praca z korpusami równoległymi umożliwia zgłębianie technik i rozwiązań stosowanych przez profesjonalnych tłumaczy. Jak podkreśla Monzó Nebot, ćwiczenia z korpusami ułatwiają studentom przyjęcie bardziej aktywnego podejścia do rozwoju kompetencji tłumaczeniowych (2008: 224; w artykule Monzó Nebot znajduje się wiele przykładowych ćwiczeń korpusowych wykorzystywanych przez autorkę na zajęciach z przekładu prawnego i prawniczego).</p>
Materiały dla tłumaczy	<p>Leksykografia i terminografia:</p> <ul style="list-style-type: none"> – słowniki jednojęzyczne (w tym słowniki kolokacji, np. <i>Słownik dobrego stylu, czyli wyrazy, które się lubią</i> prof. M. Bańko (2007) czy opracowany przez autorkę <i>Słowniczek kolokacji polskiego prawa spółek</i>, 2009) – słowniki dwujęzyczne i bazy terminologiczne – zastosowanie korpusów równoległych i oprogramowania do ekstrakcji terminologii i kolokacji umożliwia tworzenie bardziej praktycznych słowników, zawierających więcej ekwiwalentów stosowanych w praktyce. Korpusy stosuje się również do weryfikowania i udoskonalania istniejących słowników, w tym również w celu redukcji arbitralnych decyzji dotyczących doboru haseł (Teubert 2002: 204). <p>Gramatyki korpusowe pokazujące zróżnicowanie języka w zależności od gatunku tekstu (np. <i>Longman Grammar of Spoken and Written English</i> pod red. Bibera i in., 2007).</p> <p>Gatunki tekstowe (np. projekt GENTT na Uniwersytecie Jaume I mający na celu opracowanie wielojęzycznej elektronicznej encyklopedii tekstów specjalistycznych dla tłumaczy, http://www.gentt.uji.es/).</p>

Przykładowy korpusowy projekt badawczy: projekt EUROFOG

Projekt Eurofog miał na celu zbadanie dopasowania tekstowego prawa unijnego do języka prawa polskiego. Istotne dla gatunku elementy wzorców leksykalno-gramatycznych, jak np. struktury służące wyrażaniu modalności deon-

tycznej, relacji przyczynowo-skutkowej, odesłań do jednostek redakcyjnych aktu prawnego (tzw. mapowanie tekstowe), zostały najpierw zidentyfikowane, a następnie zbadane na korpusie rozporządzeń R-Acquis JRC (~15 mln słów) i korpusie dyrektyw L-Acquis JRC (~7 mln słów), a następnie na korpusie prawa polskiego PLC zbudowanym przez autorkę (~7 mln słów). Ogólnie badanie wskazuje na niski stopień dopasowania tekstowego tłumaczeń unijnych. W badaniu zidentyfikowano liczne wzorce leksykalno-gramatyczne, które są nadreprezentowane i niedoreprezentowane w tłumaczeniu lub które w ogóle w tłumaczeniu nie są „wywoływane”. Następnie zweryfikowano, w jakim stopniu tak duży napływ tłumaczeń unijnych, które podlegają transpozycji do polskiego prawa, wpływa na europeizację języka polskiego prawa. W tym celu sprawdzono frekwencyjność zidentyfikowanych nadreprezentowanych i niedoreprezentowanych wzorców leksykalno-gramatycznych w korpusie przedakcesyjnym (korpus kodeksów i ustaw typu prawo w wersji z roku 1999) i w korpusie poakcesyjnym (korpus tych samych ustaw w wersji z roku 2011). Projekt był realizowany w latach 2010–2012 w ramach grantu badawczego z Ministerstwa Nauki i Szkolnictwa Wyższego. Szczegółowe wyniki badań zostały opublikowane w książce *Lost in the Eurofog. The Textual Fit of Translated Law* (Biel 2014).

Wnioski: perspektywy i wyzwania

Translatoryka korpusowa, będąca interdyscyplinarną specjalnością korzystającą z osiągnięć językoznawstwa korpusowego, przyczyniła się do umocnienia nurtu empirycznego, metod ilościowych oraz podejścia indukcyjnego i deskryptywnego w badaniach nad przekładem. Jest, zdaniem Malmkjear, głównym motorem postępu w przekładoznawstwie (2003: 119) i trudno nie zgodzić się z tym stwierdzeniem, chociażby ze względu na samą liczbę i skalę badań korpusowych. Niewątpliwą zaletą translatoryki korpusowej jest możliwość identyfikacji zdarzeń typowych, regularnych i częstych w użyciu języka, a także możliwość systematycznego i efektywnego porównania dwu i więcej dużych zbiorów tekstów. Intensywne badania korpusowe prowadzą do nowych pytań badawczych i hipotez oraz generalizacji na temat przekładu. Ponadto badania mają praktyczne zastosowanie w dydaktyce przekładu i w opracowywaniu zasobów dla tłumaczy.

Translatoryka korpusowa umożliwia badanie przekładu z nowej perspektywy, która choć uwypukla określone aspekty, jednocześnie przysłania inne. Z powodu ograniczeń technicznych w niezadowalającym stopniu może

badać np. zjawiska dyskursywne czy pragmatyczne. Jest przede wszystkim metodą ilościową, wymagającą triangulacji danych (z korpusów porównawczych i równoległych) i metodologicznego eklektyzmu/triangulacji metodologii (metod ilościowych i jakościowych). To również metoda czasochłonna, nierzadko wymagająca wysiłku zespołowego. Do dalszego rozwoju konieczny jest rozwój oprogramowania i dopracowanie metod badania korpusów równoległych, a także tworzenie i udostępnianie zasobów korpusowych dla języków o ograniczonym zasięgu.

Literatura

- Asensio, R.M., 2007, „Specialised translation: A concept in need of revision”, [w:] *Babel*, nr 53(1), s. 48–55.
- Baker, M., 1993, „Corpus Linguistics and Translation Studies. Implications and Applications”, [w:] *Text and Technology: In Honour of John Sinclair*, M. Baker i in. (red.), Amsterdam–Philadelphia, s. 233–250.
- Baker, M., 1995, „Corpora in Translation Studies: An Overview and Some Suggestions for Future Research”, [w:] *Target*, nr 7(2), s. 223–243.
- Baker, M., 1996, „Corpus-based translation studies: the challenges that lie ahead”, [w:] *Terminology, LSP and Translation. Studies in Language Engineering in Honour of Juan C. Sager*, H.L. Somers (red.), Amsterdam–Philadelphia, s. 175–186.
- Bańko, M., 2007, *Słownik dobrego stylu, czyli wyrazy, które się lubią*, Warszawa.
- Bassnett, S., 2012, „Translation studies at a cross-roads”, [w:] *Target*, nr 24(1), s. 15–25.
- Beaugrande de, R., 1996, „The «pragmatics» of doing language science: The «warrant» for large-corpus linguistics”, [w:] *Journal of Pragmatics*, nr 25, s. 503–535.
- Bernardini, S., Ferraresi, A., 2011, „Practice Description and Theory Come Together – Normalization or Interference in Italian Technical Translation?”, [w:] *Meta: Translators’ Journal*, nr 56(1), s. 226–246.
- Bernardini, S., Zanettin, F., 2004, „When is a universal not a universal? Some limits of current corpus-based methodologies for the investigation of translation universals”, [w:] *Translation Universals. Do they exist?*, A. Mauranen, P. Kujamäki (red.), Amsterdam–Philadelphia, s. 51–62.
- Biel, Ł., 2009, *Słowniczek kolokacji polskiego prawa spółek*, http://www.academia.edu/1746040/Słowniczek_kolokacji_polskiego_prawa_spolek_Glosary_of_Polish_company_law_collocations (dostęp: 1 grudnia 2013 r.).

- Biel, L., 2010, „Corpus-Based Studies of Legal Language for Translation Purposes: Methodological and Practical Potential”, [w:] *Reconceptualizing LSP. Online proceedings of the XVII European LSP Symposium 2009*, C. Heine, J. Engberg, Aarhus (red.), <http://www.asb.dk/fileexplorer/fetchfile.aspx?file=16841> (dostęp: 1 grudnia 2013 r.).
- Biel, L., 2014, *Lost in the Eurofog. The Textual Fit of Translated Law*, Frankfurt am Main.
- Biber, D., Johansson, S., Leech, G., Conrad S., Finegan, E., 2007, *Longman Grammar of Spoken and Written English*, Harlow.
- Brems, E., Meylaerts, R., van Doorslaer, L., 2012, „A discipline looking back and looking forward. An introduction”, [w:] *Target*, nr 24(1), s. 1–14.
- Chesterman, A., 2004, „Hypotheses about translation universals”, [w:] *Claims, Changes and Challenges in Translation Studies. Selected contributions from the EST Congress, Copenhagen 2001*, G. Hansen i in. (red.), Amsterdam–Philadelphia, s. 1–13.
- Cronin, M., 2010, „The Translation Crowd”, [w:] *Revista Tradumàtica*, nr 8, <http://www.fti.uab.es/tradumatica/revista/num8/articles/04/04central.htm> (dostęp: 1 grudnia 2013 r.).
- Flynn, P., Gambier, Y., 2011, „Methodology in Translation Studies”, [w:] *Handbook of Translation Studies*, t. 2, Y. Gambier, L. van Doorslaer (red.), Amsterdam–Philadelphia, s. 88–96.
- Gentzler, E., 2003, „Interdisciplinary connections”, [w:] *Perspectives: Studies in Translatology*, nr 11(1), s. 11–24.
- Granger, S., 2003, „The corpus approach: a common way forward for Contrastive Linguistics and Translation Studies?”, [w:] *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*, S. Granger i in. (red.), Amsterdam–New York, s. 17–29.
- Hansen-Schirra, S., Teich, E., 2009, „Corpora in human translation”, [w:] *Corpus Linguistics. An International Handbook*, t. 2, A. Lüdeling, M. Kytö (red.), Berlin–New York, s. 1286–1304.
- Hardie, A., McEnery, T., 2010, „On two traditions in corpus linguistics, and what they have in common”, [w:] *International Journal of Corpus Linguistics*, nr 15(3), s. 384–394.
- Hatim, B., 2001, *Teaching and Researching Translation*, Harlow.
- Holmes, J.S., 1972/2004, „The name and nature of translation studies”, [w:] *The Translation Studies Reader*, wyd. 2 (2004), L. Venuti (red.), New York–London, s. 180–192.
- House, J., 2013, „Towards a new linguistic-cognitive orientation in translation studies”, [w:] *Target*, nr 25(1), s. 46–60.

- Laviosa, S., 2010, „Corpora”, [w:] *Handbook of Translation Studies*, t. 1, Y. Gambier, L. van Doorslaer (red.), Amsterdam–Philadelphia, s. 80–86.
- Lewandowska-Tomaszczyk, B., 2005, „Analiza języka z zastosowaniem korpusu”, [w:] *Podstawy językoznawstwa korpusowego*, B. Lewandowska-Tomaszczyk (red.), Łódź, s. 133–173.
- Malmkjær, K., 2003, „On a pseudo-subversive use of corpora in translator training”, [w:] *Corpora in Translator Education*, F. Zanettin i in. (red.), Manchester, s. 119–134.
- Marco, J., 2009, „Training Translation Researchers: An Approach Based on Models and Best Practice”, [w:] *The Interpreter and Translator Trainer*, nr 3(1), s. 13–35.
- McEnery, T., Xiao, R., Tono Y., 2006, *Corpus-Based Language Studies. An Advanced Resource Book*, London–New York.
- Monzó Nebot, E., 2008, „Corpus-based Activities in Legal Translator Training”, [w:] *The Interpreter and Translator Trainer*, nr 2(2), s. 221–252.
- Munday, J., 2010, „Translation Studies”, [w:] *Handbook of Translation Studies*, t. 1, Y. Gambier, L. van Doorslaer (red.), Amsterdam–Philadelphia, s. 419–428.
- Olohan, M., 2004, *Introducing Corpora in Translation Studies*, London–New York.
- Piotrowska, M., Dybiec-Gajer, J., 2012, *Verba Volant Scripta Manent. How to Write an M.A. Thesis in Translation Studies*, Kraków.
- Piotrowski, T., 2007, „The Translator and Polish-English Corpora”, [w:] *Incorporating Corpora. The Linguist and the Translator*, G.M. Anderman, M. Rogers (red.), Clevedon, s. 117–132.
- Pym, A., 2005, „Explaining Explicitation”, [w:] *New Trends in Translation Studies. In Honour of Kinga Klauudy*, K. Krisztina, Á. Fóris (red.), Budapest, s. 29–34.
- Scott, J., 2012, „Towards professional uptake of DIY electronic corpora in legal genres”, [w:] *Salford Working Papers in Translation and Interpreting*, nr 1, http://www.salford.ac.uk/__data/assets/pdf_file/0010/229492/WorkingPapersT-and-I.Scott.pdf (dostęp: 1 grudnia 2013 r.).
- Snell-Hornby, M., 2012, „From the Fall of the Wall to Facebook. Translation Studies in Europe twenty years later”, [w:] *Perspectives: Studies in Translatology*, nr 20(3), s. 365–373.
- Stubbs, M., 2004, „Language Corpora”, [w:] *Handbook of Applied Linguistics*, A. Davies, C. Elder (red.), Oxford, s. 106–132.
- Teubert, W., 2002, „The role of parallel corpora in translation and multilin-

- gual lexicography”, [w:] *Lexis in Contrast*, B. Attenberg, S. Granger (red.), Amsterdam–Philadelphia, s. 189–214.
- Tirkkonen-Condit, S., 2004, „Unique items – over- or under-represented in translated language?”, [w:] *Translation Universals. Do They Exist?*, A. Mauranen, P. Kujamäki (red.), Amsterdam–Philadelphia, s. 177–184.
- Tognini-Bonelli, E., 2001, *Corpus Linguistics at Work*, Amsterdam–Philadelphia.
- Tymoczko, M., 2005, „Trajectories of research in Translation Studies”, [w:] *Meta: Translators’ Journal*, nr 50(4), s. 1082–1097.
- Zanettin, F., 2012, *Translation-Driven Corpora. Corpus Resources for Descriptive and Applied Translation Studies*, Manchester.

Corpus-Based Translation Studies

Summary

The aim of the paper is to discuss Corpus-Based Translation Studies as a new methodology in Translation Studies (TS): its place on the TS map, its genesis, tracks of research, methods of analysis, limitations and perspectives. Considering recent developments in TS, Corpus-Based Translation Studies has emerged owing to the technological and linguistic turns in TS and the combination of corpus linguistics and Descriptive Translation Studies. It uses two types of corpora: comparable corpora (to research textual fit, that is, features of translations against the target language) and parallel corpora (to research equivalence and other aspects of translations against source texts). Corpora also have practical applications in translation practice, translator training and the development of resources for translators. As a method, Corpus-Based Translation Studies is an empirical, inductive, descriptive and data-driven approach. It is especially well-suited for the identification of typical, regular and frequent patterns and for comparison of two or more large data sets (corpora). It also has its limitations, mainly due to technological and corpus design constraints, and it requires triangulation of data and methodologies. Overall, it is a promising methodology which contributes new knowledge about translations.

Keywords: Corpus-Based Translation Studies, Translation Studies research, corpus analysis

