

ALBERT LEŚNIAK

Instytut Języka Polskiego PAN, Kraków, Polska

<https://orcid.org/0000-0002-7141-576X>

ZBIGNIEW PASEK

AGH w Krakowie, Polska

<https://orcid.org/0000-0003-2580-4366>

Copyright and License: Copyright by Instytut Języka Polskiego PAN, Kraków 2020. This article is published under the terms of the Creative Commons Attribution – NoDerivatives 4.0 International (CC BY-ND 4.0) License (<https://creativecommons.org/licenses/by-nd/4.0/legalcode.pl>).

ŚWIADECTWA EWANGELIKALNE I KATOLICKIE W PERSPEKTYWIE KORPUSOWEJ ANALIZY DYSKURSU

Słowa kluczowe: językoznawstwo korpusowe, chrześcijańskie świadectwa, polska kultura religijna, korpusowa analiza dyskursu.

STRESZCZENIE

Artykuł przedstawia wyniki analizy porównawczej ewangelizacyjnych świadectw rzymskokatolickich (z portalu Deon.pl) i protestanckich – zielonoświątkowych (z pisma „Chrześcijanin” i in.). Autorzy, korzystając z narzędzi językoznawstwa korpusowego, ukazują odmienności obu korpusów. Interesujące było zrekonstruowanie „obszarów grzechu” – zła, które składający świadectwa przeżyli dzięki bożej pomocy. Są to dwie odmiennie sfery: w przypadku świadectw katolickich obszar ten tworzy kompleks znaczeniowy zbudowany wokół pojęć miłość – seks – asceza. Świadectwa zielonoświątkowe skupione są na kompleksie używki – uzależnienia – patologie. Składająca się z dwu etapów analiza (frekwencja i zakorzenienie w korpusie słów kluczowych i rozpoznanie wizualne najistotniejszych kolokacji) pozwoliła sformułować wnioski o dominujących obszarach znaczeniowych, które określają mapę mentalną życia religijnego wyznawców obu nurtów chrześcijaństwa.

We współczesnym języku religijnym coraz ważniejsze miejsce zajmuje tzw. świadectwo, definiowane przez Małgorzatę D. Nowak jako „relacjonowanie osobistego doświadczenia religijnego («spotkania» z Bogiem)” (2005, 22). Autorka ta zaznacza, że zwykle świadectwa konstruowane są na podstawie schematu: „przedtem (przed tym doświadczeniem) – wtedy (doświadczenie) – teraz (po doświadczeniu)” (2005, 22). Grzegorz Pełczyński zauważa, że świadectwo stało się „osobnym gatunkiem epicznym, funkcjonującym w obiegu oralnym (jakby folklorystycznym) i literackim” (2006, 43). W zróżnicowanym gatunkowo języku religijnym chrześcijaństwa można odnotować postępujący od drugiej połowy XX wieku wzrost znaczenia tego rodzaju wypowiedzi religijnej. Ten wywodzący się ze świata protestanckiego termin (jak i samą praktykę

składania świadectwa) przejęły liczne współczesne ruchy odnowy katolickiej, nieco odmiennie profilując rolę świadectwa w swojej duchowości. Dla ewangelikalnego protestantyzmu świadectwo pozostaje jedną z głównych form ewangelizacji, dla religijności katolickiej jest ono jedną z wielu rodzajów praktyk. Głównym celem zielonoświątkowego świadectwa jest skłonienie słuchaczy do radykalnej przemiany życia lub wzmocnienie wiary już wierzących. Silniej obecny jest schematyzm polegający na kontrastowym przedstawieniu sytuacji sprzed nawrócenia i po nim. W praktyce katolicyzmu świadectwo jest wyrazem nie tylko duchowego przełomu, ale i codziennych interwencji Boga w życie człowieka. Analizując chrześcijańskie świadectwa, należy zaznaczyć, że ich dominującą funkcją językową jest perswazja.

Przedstawione studium stawia sobie za cel dokonanie analizy porównawczej świadectw katolickich i protestanckich – zielonoświątkowych – jako fragmentów obu dyskursów religijnych – katolickiego i zielonoświątkowego. Głównym celem naszych badań było określenie wybranych cech charakterystycznych dla obu typów dyskursu wyznaniowego, który za Aleksandrą Pawlikowską rozumiemy jako:

zbiór tekstów i wypowiedzi danej wspólnoty wyznaniowej odzwierciedlający wykreowany przez tę wspólnotę sposób orzekania o świecie z perspektywy religii. Zawiera on określony system reprezentacji rzeczywistości oraz swoistą ontologię, aksjologię, a także zestaw strategii i reguł komunikacyjnych (2015, 20).

W prezentowanym artykule wnioski oparliśmy na analizie podobieństw i różnic tak w zakresie słów kluczowych, jak i kolokacji. Chcieliśmy, stosując narzędzia językoznawstwa korpusowego, pokazać, jak w języku religijnym polskich rzymskich katolików i charyzmatycznych protestantów (zielonoświątkowców) ujawniają się idee religijne. Skupiliśmy się zwłaszcza na problemie źródeł grzechu i grzeszności, w toku prowadzonych badań okazało się bowiem, że oba korpusy odmiennie profilują „obszary grzechu”.

METODOLOGIA BADAŃ

Wnioski naszej analizy formułowaliśmy po zbadaniu korpusu świadectw wyznawców kościołów rzymskokatolickiego i zielonoświątkowego w zakresie powtarzających się słów kluczowych wraz z ich kolokacjami, które zostały precyzyjnie wyodrębnione za pomocą narzędzi programistycznych oraz miar statycznych wypracowanych na gruncie językoznawstwa korpusowego i lingwistyki kwantytatywnej. Tego rodzaju korpusy językowe, czyli reprezentatywne, zapisane w formie cyfrowej dane tekstowe o skończonej wielkości (McEnery i Wilson 2001), stanowią podstawę do formułowania twierdzeń na temat jakiejś odmiany języka (np. religijnego). Dodatkowo służą odkrywaniu rekurentnych wzorców (uzusów) językowych, umożliwiających wnioskowanie o sposobach użycia tego języka do konstruowania świata dyskursywnego użytkowników (Baker 2006, 1). Nie wchodząc w szczegółowe rozważania na temat samego pojęcia dyskursu, akcentujemy tu jego utylitarny charakter, wynikający poniekąd z definicji samego Michela Foucaulta, dla którego dyskurs to „praktyki, które systematycznie

formułują przedmiot, do którego się odnoszą” (1972, 49), jest to zatem „język w działaniu” – koncept, który – jak zauważają Paul Baker i Tony McEnery – jest idealnie skrojony dla językoznawstwa korpusowego. Ten typ lingwistyki podejmuje się analizy dużych zbiorów tekstów danego języka, a wręcz „[w] swoim najgłębszym znaczeniu całe językoznawstwo korpusowe jest analizą dyskursu” (2015, 4).

W prezentowanej analizie posługujemy się dwiema technikami opartymi na frekwencji występowania i współwystępowania leksemów w tekście, czyli na ekstrakcji słów kluczowych oraz kolokacji. W badaniach korpusowych zagadnieniu kluczowości słów towarzyszą dwa główne założenia: po pierwsze, znaczenie w tekście nie jest dystrybuowane równomiernie, tzn. istnieją słowa, których potencjał semantyczny jest większy, znaczą one *bardziej*. Po drugie, dyferencjacja słów o zwiększonej sile znaczenia jest mierzalna, co stanowi podstawę do tego, by za kluczowe uznawać te słowa, których wysoka frekwencja w korpusie jest statystycznie istotna, gdy porównać ją do korpusu odniesienia (tzw. korpusu referencyjnego). Podczas gdy lista frekwencyjna słów informuje o (względnej lub bezwzględnej) *częstości* występowania poszczególnych słów w korpusie, lista słów kluczowych mówi, jak *istotne* są to słowa dla badanego korpusu (Baker 2006, 125). Mike Scott definiuje słowa kluczowe jako „słowa, które występują z nieoczekiwaną częstością w danym tekście” (1997, 236), wspólnie z Christopherem Tribble’em dodaje, że:

kluczowość jest własnością, którą słowa mogą przejawiać w obrębie danego tekstu, bądź zbioru tekstów, sugerującą, że są one ważne, odzwierciedlają to, czego naprawdę dotyczy tekst, pomijając drobiazgi i nieznaczące detale (Scott i Tribble 2006, 55–56).

Z perspektywy statystyki języka są to zatem leksemy, których frekwencja w danym korpusie jest niespodziewanie wyższa (pozytywne słowa kluczowe) lub niższa (negatywne słowa kluczowe), posiadają przy tym potencjał semantyczny, który może dostarczać ważnych informacji o samym tekście lub korpusie.

Dla analizy dyskursu słowa kluczowe stanowią swoiste instancje językowe, na bazie których konstruowana jest rzeczywistość społeczna i kulturowa. Prekursorem podejścia, które uwzględni kulturotwórczy potencjał słów kluczowych, był William Raymond. W słynnej już pracy *Keywords* zwracał uwagę, iż kulturowe słowa kluczowe nie tylko opisują kulturę swoich czasów, lecz także są używane przez osoby zaangażowane w dyskusję na temat tej kultury. Według Williamsa są one: „wspólnym zbiorem słów i znaczeń w naszych najbardziej ogólnych dyskusjach w języku angielskim, na temat praktyk i instytucji, które określamy zbiorczo jako kulturę i społeczeństwo” (1983, 15). Williams nie stosował statystycznych metod ekstrakcji słów mających kulturowo istotny potencjał semantyczny głównie ze względu na brak metod umożliwiających zbadanie tekstów w odpowiedniej skali. Alan Durant twierdził, że:

rozwój możliwości elektronicznego przeszukiwania stosowany do dużych korpusów językowych (techniki zwane zbiorczo językoznawstwem korpusowym) zachęca do ponownego zwrócenia uwagi na słowa kluczowe związane z kulturą (2006, 19).

Podkreślmy, że wygenerowanie listy słów kluczowych musi być w przypadku nauk o kulturze jedynie etapem wstępnym, prowadzącym do pogłębionej analizy kontekstu, w którym występują. Jak píše Victoria Kamasa:

słowa kluczowe stanowią jednak istotną wskazówkę kierującą uwagę badaczy na zjawiska charakterystyczne dla analizowanego przez nich dyskursu. Mogą być więc swego rodzaju punktem wejścia do zebranych danych (2014, 107).

Badania korpusowe mają najczęściej charakter emergentny i dystrybucyjny (Lewandowska-Tomaszczyk 2005). Według Anny Bączkowskiej emergentność, która polega na formułowaniu wniosków wyłącznie na podstawie obserwacji wynikających z analizy korpusu (bez wcześniejszych założeń), istotnie dowartościowuje rolę kontekstu, ponieważ zakłada, że znaczenie nie tkwi tylko w statycznej definicji pojedynczych słów, ale także w ich kontekście, obejmującym grupy słów je otaczających (Bączkowska 2016, 8–9). Według Bączkowskiej kontekst współtworzy znaczenie słów kluczowych, profilując je, precyzując i wskazując na ich preferencje semantyczne. Warto tu zauważyć, że pojęcie kontekstu można rozumieć znacznie szerzej niż tylko jako zespół jednostek leksykalnych otaczających dane słowo kluczowe, odnosząc ów termin także do sytuacji komunikacyjnej, w której dane wypowiedzi padają (por. Boniecka 1994). W tym sensie będziemy rozumieli kontekst również jako „strukturę, obejmującą wszystkie właściwości sytuacji społecznej, które są istotne dla wytwarzania i odbierania dyskursu” (Skowronek 2006, 26–27).

W węższym znaczeniu kontekst może być rekonstruowany poprzez identyfikację tzw. kolokacji, czyli słów, które mają tendencję do współwystępowania obok siebie znacznie częściej niż wynikałoby to z teorii prawdopodobieństwa przy założeniu, że rozkład słów w korpusie byłby losowy. John R. Firth również uważał, że część znaczeń słowa można ustalić właśnie na podstawie jego kolokacji, np. jedno ze znaczeń słowa *noc* (ang. *night*) może być sformułowane ze względu na jego łączliwość ze słowem *ciemna* (ang. *dark*). Streścił ów postulat w słynnej sentencji: „możesz poznać słowo na podstawie towarzystwa, z którym przestaje” (1957, 179). Dla Bakera natomiast analiza kolokacji to „sposób rozumienia znaczeń oraz powiązań między słowami, które w innym przypadku byłyby trudne do ustalenia na podstawie analizy pojedynczego tekstu na małą skalę” (2006, 96). W analizie dyskursu tego rodzaju profile kolokacyjne wykorzystuje się do identyfikacji bardziej złożonych konstrukcji dyskursywnych:

najczęściej służą one uzyskaniu bardziej szczegółowych informacji na temat funkcjonowania określonych słów w badanych tekstach. Na podstawie takiej informacji badacz identyfikuje później struktury dyskursywne (Kamasa 2014, 108).

Pod wieloma względami pomiar istotności liczbowej współwystąpień dwóch wyrazów może być podobny do pomiaru odrębności leksykalnej korpusów na podstawie ich słów kluczowych, a wręcz często korzysta się w tym celu z tych samych miar asocjacji. Jednostki leksykalne (zarówno słowa kluczowe, jak i kolokacje), których wysoka frekwencja wystąpień w korpusie jest statystycznie istotna, tzn. znacząco wyższa niż wynikałoby

to z przypadku, zostały w prezentowanym badaniu wyodrębnione za pomocą miary *logarytmu wskaźnika wiarygodności* (ang. *log-likelihood*), zaproponowanej przez Teda Dunninga (1993), zgodnie z procedurą opisaną szczegółowo w podręczniku Vaclava Breziny (2018b, 66–101). Miara ta jest jedną z najczęściej używanych do określenia, czy wystąpienie zdarzenia (czyli wystąpienia słowa w korpusie A w porównaniu do korpusu B lub współwystąpienia dwóch słów w danym korpusie) jest istotne statystycznie, czy też wynika z przypadku. Wszystkie obliczenia zostały wykonane na formach zlematyzowanych, przy pomocy środowiska programistycznego Python 3.7 opublikowanego na wolnej licencji. Są one dostępne do wglądu (wraz z wygenerowanymi danymi) na stronie <http://github.com/alblesniak/swiadectwa>.

Jedną z możliwych form prezentacji wyników kolokacji jest stworzenie grafów i sieci kolokacyjnych (Brezina 2018a). Ten rodzaj wizualizacji umożliwi analizę skomplikowanych zależności w zestawie danych językowych, które byłyby znacznie trudniejsze do eksploracji za pomocą tradycyjnych form tabelarycznych. Podczas gdy grafy kolokacyjne pokazują najistotniejsze kolokaty słowa węzłowego (przy zdefiniowanym *zasięgu*, czyli liczbie słów poprzedzających to słowo i następujących po nim, które będą uwzględniane w obliczeniach, oraz określonej minimalnej wartości użytej miary asocjacji), sieci kolokacyjne idą o krok dalej, włączając w poczet węzłów sieci więcej słów kluczowych (w naszym przypadku 50), a także uwzględniając kolokacje drugiego rzędu (kolokaty kolokatów). Szerokość krawędzi łączącej dwa węzły odpowiada ich sile kolokacyjnej, natomiast wielkość samych węzłów obrazuje kluczowość, które mają te słowa w porównaniu do korpusu referencyjnego.

Podczas pracy chcieliśmy także pokazać, jak narzędzia językoznawstwa korpusowego otwierają drogę do cyfrowej humanistyki i nie tylko usprawniają pracę z ogromnymi rozmiarowo zbiorami tekstów (danych), ale także, dzięki dostarczaniu dowodów na istnienie (bądź brak) pewnych stałych tendencji, umożliwiają potwierdzanie lub falsyfikację obserwacji jakościowych; a co za tym idzie pozwalają na formułowanie daleko pewniejszych wniosków również w badaniach nad współczesnym językiem religijnym i religijną kulturą danego środowiska.

KORPUSY

Analizowany korpus zawiera łącznie 281 746 tokenów (373 dokumenty). Składają się na niego dwa mniejsze podkorpusy: świadectwa spisane przez wyznawców Kościoła Zielonościwkowego (73 554 tokeny, 78 dokumentów) oraz świadectwa wyznawców Kościoła Rzymskokatolickiego (208 192 tokeny, 295 dokumentów). Teksty, z których składał się podkorpus świadectw ewangelikalnych, zostały zaczerpnięte z dwu źródeł, tzn. z czasopisma „Chrześcijanin” – organu prasowego Kościoła Zielonościwkowego – oraz z oficjalnych stron internetowych zborów tego Kościoła. Świadectwa katolickie, stanowiące podstawę korpusu referencyjnego, zostały pobrane ze strony internetowej katolickiego serwisu Deon.pl. Wykorzystano metodę automatycznej ekstrakcji danych tekstowych przy pomocy robota internetowego stworzonego przy użyciu struktury

biblioteki *Scrapy*, czyli tzw. *web scrapingu* (zob. Mitchell 2019). Zdecydowana większość tekstów, które składają się na podkorpus ewangelikalny, pochodzi z przełomu XX i XXI wieku, natomiast czas powstania świadectw katolickich to lata 2010–2019.

Mimo że teksty źródłowe pochodzą z różnych rodzajów mediów, są one jednolite ze względu na kanał przekazu, są to bowiem teksty pisane. Trudna do rozstrzygnięcia jest kwestia istotności wkładu redaktorskiego w ich edycję, ale można z pewną dozą ostrożności przyjąć, że mają one charakter autorski, a czas ich powstania jest zbliżony, wykazują więc jednorodność pod względem terytorialnym, chronologicznym i socjologiczno-kulturowym (Sambor 1972, 24). Dysproporcja rozmiaru badanych korpusów spowodowana jest łatwiejszym dostępem do katolickich tekstów źródłowych w wersji cyfrowej. Pozwoliło to na ich ekstrakcję w sposób zautomatyzowany. Ponadto zauważyliśmy, że liczba wiernych Kościoła Zielonoświątkowego (ok. 25 tysięcy wyznawców) w porównaniu do katolickiego nie pozwala na stworzenie zupełnie adekwatnych korpusów.

Choć podkorpusy różnią się liczbą dokumentów, to ich średnia długość jest zbliżona, a ewentualne problemy wynikające z różnicy w ich wielkości zostaną zminimalizowane poprzez posługiwanie się w analizie względną liczbą wystąpień tokenów (tzn. znormalizowaną przez wielkość korpusu) oraz zastosowanie odpowiednio dobranej miary statystycznej. Uznajemy zatem, że korpusy są reprezentatywne i porównywalne.

Tabela 1

Statystyczna charakterystyka badanych korpusów

Świadectwa	PROTESTANCKIE	KATOLICKIE
Liczba tekstów	78	295
Liczba tokenów	73 554	208 192
Liczba lematów	6209	11 448
Średnia liczba tokenów w dokumencie	943	705,8
Minimalna liczba tokenów w dokumencie	193	28
Maksymalna liczba tokenów w dokumencie	5201	3299

Źródło: oprac. aut.

Wszystkie dokumenty wchodzące w skład korpusów podzielone zostały na segmenty (nazywane także tokenami), które następnie (po usunięciu interpunkcji) poddane zostały lematyzacji, czyli sprowadzeniu leksemów do ich gramatycznych form bazowych, tzw. lematów. Proces ten umożliwia kwantyfikację wystąpień wszystkich leksemów niezależnie od tego, w jakich formach fleksyjnych były reprezentowane w tekście. Tak przygotowany korpus umożliwia przeprowadzanie szeregu obliczeń statystycznych z wykorzystaniem tzw. miar asocjacji, które stosuje się do badania korelacji między dwoma elementami, np. frekwencją leksemu w dwóch różnych korpusach lub współwystępowaniem dwóch leksemów w tym samym korpusie. Pierwszy przypadek pozwala

na wyodrębnienie leksemów charakteryzujących dany korpus tzw. słów kluczowych, drugi zaś umożliwia ekstrakcję związków wyrazowych mających tendencję do współwystępowania, czyli tzw. kolokacji.

WYNIKI BADAŃ 1 – SŁOWA KLUCZOWE

W prezentowanych tabelach nr 2 i 3 pokazujemy w zestawieniu wyniki uzyskane podczas analizy. Na potrzeby tej publikacji ograniczono się do pierwszych 50 słów kluczowych dla obu korpusów. Wnioski wyciągamy także z danych, które znajdują się w dalszych częściach obu tabel, udostępnionych przez nas do wglądu na profilu w portalu Github.

Tabela 2

50 pierwszych słów kluczowych w korpusie świadectw katolickich, obliczonych na podstawie logarytmu wskaźnika wiarygodności

L.p.	Słowo kluczowe	Współczynnik <i>log-likelihood</i>	Liczba wystąpień na 1000 słów	L.p.	Słowo kluczowe	Współczynnik <i>log-likelihood</i>	Liczba wystąpień na 1000 słów
1	rekolekcja	101,5917	1,1240	21	miłosierdzie	24,3482	0,3266
2	spowiedź	80,5576	0,8598	22	cierpienie	23,3219	0,3795
3	świadek	71,0707	1,4266	23	kolejny	23,2697	1,3305
4	chłopak	70,6767	1,0231	24	siebie	23,2233	5,2404
5	związek	65,8248	0,9126	25	miesiąc	21,8457	1,0951
6	czystość	56,9192	0,6148	26	operacja	20,9935	0,3266
7	różaniec	53,0052	0,5380	27	poród	19,7190	0,2113
8	ciąża	50,2416	0,5572	28	przypadek	19,6632	0,4227
9	odmawiać	46,1537	0,5956	29	dług	19,1594	0,2065
10	Maryja	43,5417	0,6052	30	dziewczyna	18,7765	0,5091
11	modlitwa	36,3880	2,6802	31	uzdrowienie	18,2710	0,3554
12	ślub	36,2593	0,8358	32	walczyć	18,1164	0,3266
13	badanie	35,5301	0,3891	33	pojawiać	17,1941	0,3170
14	sakrament	34,9741	0,3843	34	dobra	16,7309	0,2546
15	intencja	34,4187	0,3795	35	konkretny	16,3781	0,1825
16	tydzień	34,1442	1,1096	36	choroba	16,3686	0,5236
17	matka	33,8066	0,7829	37	zaufanie	16,0594	0,2161
18	małżeństwo	29,4044	0,8021	38	przyczyna	15,8257	0,1777
19	ksiądz	27,0809	0,8406	39	pierwsza	15,6810	1,1192
20	ciężki	24,4665	0,3602	40	walka	15,6145	0,3266

L.p.	Słowo kluczowe	Współczynnik <i>log-likelihood</i>	Liczba wystąpień na 1000 słów	L.p.	Słowo kluczowe	Współczynnik <i>log-likelihood</i>	Liczba wystąpień na 1000 słów
41	seks	15,6145	0,3266	46	obój	14,7251	0,1681
42	warto	15,2374	0,2690	47	wspólny	14,5784	0,4995
43	lekarz	14,9790	0,6100	48	kryzys	14,1772	0,1633
44	oglądać	14,7715	0,2642	49	znak	13,8904	0,3074
45	owoc	14,7681	0,2354	50	doświadczenie	13,5202	0,6677

Źródło: oprac. aut.

Tabela 3

50 pierwszych słów kluczowych w korpusie świadectw protestanckich, obliczonych na podstawie logarytmu wskaźnika wiary

L.p.	Słowo kluczowe	Współczynnik <i>log-likelihood</i>	Liczba wystąpień na 1000 słów	L.p.	Słowo kluczowe	Współczynnik <i>log-likelihood</i>	Liczba wystąpień na 1000 słów
1	bóg	189,9654	12,5894	19	zbawienie	54,1362	0,5302
2	Jezus	173,5969	6,3627	20	słowo	52,7294	2,8414
3	Biblia	164,6221	1,2780	21	chrześcijański	52,0894	0,5030
4	życie	161,3172	9,8295	22	zbawić	50,7596	0,4351
5	zbór	154,2113	0,8293	23	testament	49,8966	0,3399
6	alkohol	133,1858	1,3595	24	wieczny	48,5263	0,4215
7	Chrystus	112,4584	1,7402	25	człowiek	46,4807	4,1330
8	czytać	96,0239	1,6450	26	Basia	43,8891	0,4079
9	chrzest	86,2822	0,6798	27	uczęszczać	43,1411	0,3399
10	społeczność	85,4751	0,4758	28	pustka	42,3944	0,5030
11	pastor	77,0850	0,4622	29	pieśń	39,5986	0,3807
12	kościół	73,5060	2,5831	30	wodny	38,5425	0,2311
13	nabożeństwo	72,8846	0,7613	31	dom	38,3874	2,4880
14	papieros	67,4022	0,5166	32	pismo	37,2055	0,5846
15	ośrodek	64,5074	0,6526	33	wieczność	35,3232	0,2583
16	więzienie	57,3104	0,3807	34	narkotyk	33,6500	0,4622
17	przyjąć	56,7842	1,0876	35	zbawiciel	33,3054	0,4486
18	pić	54,6864	0,6254	36	żyć	32,7220	1,7266

37	religia	31,4581	0,3535	44	wołać	27,3432	0,3127
38	boży	30,1777	2,2432	45	duch	27,2701	1,4547
39	grzesznik	29,2594	0,2719	46	palić	27,0840	0,2855
40	serce	28,6687	2,8686	47	śpiewać	23,7291	0,3943
41	swój	28,4267	2,2432	48	syn	23,7145	1,1284
42	Wojtek	28,3178	0,1767	49	kolega	23,6056	0,5982
43	werset	28,3178	0,1767	50	przyjść	23,6048	0,9789

Źródło: oprac. aut.

Analiza danych zawartych w obu tabelach (1 i 2) pozwala na sformułowanie następujących wniosków. Przede wszystkim należy odnotować silnie obecny w korpusie katolickim kompleks, nazwijmy go tymczasowo, seksualno-ascetyczny (miłość – seks – asceza). Analiza pierwszego 1000 słów kluczowych listy wykazała, iż kompleks ten budują wysokie frekwencje takich leksemów jak *chłopak*, *związek*, *czystość*, *cięża*, *małżeństwo*, *poród*, *dziewczyna*, *seks*, *małżeński*, *ukochana*, *łóżko*, *przytulic*, *zakochać*, *rodzinny*, *kobieta*. W korpusie protestanckim (zielonoświątkowym) pojawia się natomiast kompleks, który można określić jako używki – uzależnienia – patologie. Budują go takie słowa jak *alkohol*, *papieros*, *pić*, *więzienie* i podobne. Należy zaznaczyć, że świadomie pominęliśmy oczywiste zjawiska, jak np. dominacja w korpusie protestanckim: *czytać*, *Biblia*, *pismo*, *testament*, *pastor*, a w korpusie katolickim: *różaniec*, *Maryja*, *sakrament*, *ksiądz*, *eucharystia*.

W prowadzonej analizie skierowaliśmy uwagę na słowa mniej lub bardziej wiążące się z wyróżnionymi domenami znaczeniowymi. Dane liczbowe opisujące odnoszące się do bezwzględnej oraz względnej częstości leksemów konstruujących obie domeny zawarte są w tabeli 4. Po przeprowadzeniu analizy porównawczej uzyskaliśmy obraz dysproporcji w frekwencji słów, które wyróżniliśmy jako znaczące.

Tabela 4

Leksemy charakteryzujące kompleks miłość – asceza – seks

Słowo kluczowe	Liczba wystąpień w korpusie katolickim	Liczba wystąpień w korpusie katolickim w przeliczeniu na 1000 słów	Liczba wystąpień w korpusie protestanckim	Liczba wystąpień w korpusie protestanckim w przeliczeniu na 1000 słów	Różnica
chłopak	213	1,02	11	0,15	0,87
związek	190	0,91	9	0,12	0,79
czystość	128	0,61	3	0,04	0,57
małżeństwo	167	0,8	19	0,26	0,54

Słowo kluczowe	Liczba wystąpień w korpusie katolickim	Liczba wystąpień w korpusie katolickim w przeliczeniu na 1000 słów	Liczba wystąpień w korpusie protestanckim	Liczba wystąpień w korpusie protestanckim w przeliczeniu na 1000 słów	Różnica
ciąża	116	0,56	3	0,04	0,52
dziewczyna	106	0,51	12	0,16	0,35
seks	68	0,33	6	0,08	0,25
poród	44	0,21	1	0,01	0,2
małżeński	44	0,21	5	0,07	0,14
łóżko	54	0,26	9	0,12	0,14
kobieta	102	0,49	26	0,35	0,14
rodzinny	47	0,23	9	0,12	0,1
ukochana	21	0,1	1	0,01	0,09
zakochać	20	0,1	2	0,03	0,07
prztyulić	15	0,07	1	0,01	0,06

Źródło: oprac. aut.

Tabela 5

Leksemy charakteryzujące kompleks używki – uzależnienia – patologie

Słowo kluczowe	Liczba wystąpień w korpusie katolickim	Liczba wystąpień w korpusie katolickim w przeliczeniu na 1000 słów	Liczba wystąpień w korpusie protestanckim	Liczba wystąpień w korpusie protestanckim w przeliczeniu na 1000 słów	Różnica
alkohol	36	0,17	100	1,36	1,19
pić	20	0,1	46	0,63	0,53
papieros	7	0,03	38	0,52	0,48
narkotyk	19	0,09	34	0,46	0,37
więzienie	3	0,01	28	0,38	0,37
palenie	4	0,02	15	0,2	0,18
picie	4	0,02	15	0,2	0,18
wyrok	4	0,02	15	0,2	0,18
alkoholik	8	0,04	14	0,19	0,15
karny	1	0	8	0,11	0,1
kac	1	0	6	0,08	0,08

Źródło: oprac. aut.

Uwzględnienie liczby kluczowych słów w przeliczeniu na 1000 słów pozwoliło potwierdzić zasadność wyróżnienia w korpusie katolickim domeny znaczeniowej obejmującej sferę relacji między mężczyzną a kobietą, dziewczyną a chłopcem z silnym akcentem seksualnym i ascetycznym (w tym zakresie – *vide* wysoka pozycja słowa *czystość* odnoszącego się do wstrzemięźliwości seksualnej).

Zatem, rozszerzając rozważania, można sformułować pytanie, czy należy te statystyczne tendencje interpretować jako ilustrację tezy, że podstawowym problemem, z jakim mierzą się współcześni polscy katolicy, są napięcia związane z erotycznym wymiarem relacji (miłości/zakochania) między partnerami? Analizowane świadectwa ukazują „moc działania Boga” w uregulowaniu przede wszystkim właśnie tych relacji. Ta sfera wyróżnia się jako główne źródło religijnego niepokoju, problemów czy – mówiąc językiem chrześcijaństwa – grzechu. Przeprowadzona analiza ukazuje siłę i zakres obecności kompleksu seksualnego w języku tych świadectw, co dowodzi tezy, że sfera ta jest jedną z dominujących w codziennych duchowych zmaganiach katolików. Rozważając tę problematykę, warto przywołać wielokrotnie podkreślany przez socjologów religii rozziw między nauczaniem Kościoła i postawami i praktykami polskich katolików kontestujących naukę Kościoła w tym zakresie (np. zakaz pożycia przed małżeństwem, stosowanie środków antykoncepcyjnych). Socjologowie religii podkreślają, iż zasady etyki głoszone przez Kościół są akceptowane przez co piątego Polaka (m.in. Mariański 2015, 12–13).

Autorzy portalu Deon.pl (podobnie redaktorzy „Chrześcijanina”) selekcjonują przed publikacją otrzymywane świadectwa. Przyjmujemy, że upubliczniając je, prawdopodobnie biorą pod uwagę ich aktualność, czyli związek z bieżącymi problemami wiernych czy potencjalną ewangelizacyjną skutecznością. Publikowane świadectwa odpowiadają na problemy, jakimi żyją wierni, dlatego wydaje się, że dokonywana na ich podstawie diagnoza jest prawomocna. Tylko częściowo można wyjaśnić obecność semantycznego kompleksu związanego z seksualnym wymiarem miłości odmienną grupą docelową, jaką zakłada portal Deon.pl, adresowany przede wszystkim do młodzieży (choć w korpusie znajdują się liczne świadectwa pochodzące od starszych osób, to określenie wieku wszystkich autorów jest niestety niemożliwe). Zatem, zawiązując nasz wniosek, można sformułować tezę, że dla młodych polskich katolików problem uzgodnienia nauczania Kościoła ze swoimi potrzebami w zakresie życia intymnego można uznać za doniosły, albo nawet kluczowy.

Warto także zaznaczyć, że w korpusie katolickim ponad 65% świadectw pochodzi od kobiet (tab. 6). Kwalifikacji tej dokonano automatycznie na podstawie frekwencji form aglutynacyjnych czasowników w czasie przeszłym (liczby mnogiej lub pojedynczej), które posiadały odpowiedni znacznik określający rodzaj (męski lub żeński) czasownika. Ustalenie tego faktu częściowo wyjaśnia np. dwukrotnie wyższą liczbą wystąpień słowa *chłopak* niż *dziewczyna*.

Tabela 6

Rozkład świadectw pod względem płci autora

Świadectwa	Zakwalifikowane jako męskie	Zakwalifikowane jako żeńskie	Nierozpoznane
Katolickie	89 (30,07%)	194 (65,54%)	13 (4,39%)
Protestanckie	37 (47,44%)	41 (52,56%)	0 (0,0%)

Źródło: oprac. aut.

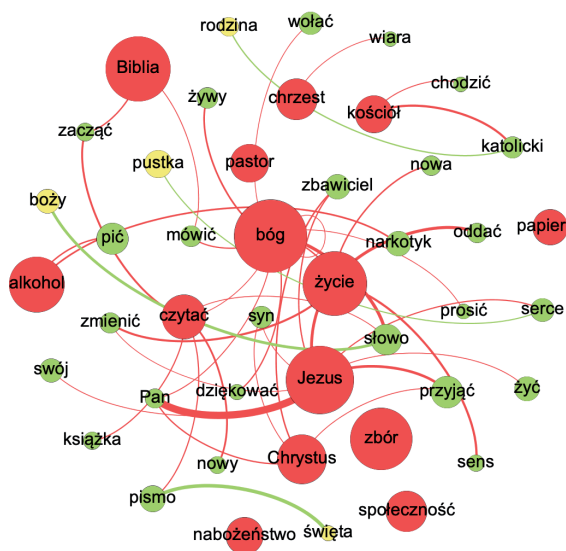
Zgromadzone wyniki związane z obecnością kompleksu erotycznego w korpusie katolickim mogą również stanowić ilustrację trendów bardziej ogólnych, takich jak przypadający na ostatnie dekady wzrost popularności poradnictwa duchowego (wydawnictw książkowych, filmów w serwisach społecznościowych, seminariów, warsztatów i rekolekcji), których tematyka coraz częściej dotyka zagadnień oraz problemów natury intymnej w perspektywie chrześcijańskiej. Wystarczy przytoczyć w tym kontekście katolicką działalność publikacyjną, np. kapucyna Ksawerego Knotza (Knotz i Strączek 2010) oraz medialną dominikanina Adama Szustaka, którego konto w serwisie YouTube obserwuje ponad 500 000 użytkowników.

Autorzy świadectw zielonoświątkowych pochodzą z różnych grup społecznych, a ich przekaz adresowany jest możliwie szeroko. Szukając podobnej sfery „winy i grzeszności” w tym korpusie, należy odnotować nadreprezentację (silniejszą pod względem statystycznym) słów, które pozwalają na wyróżnienie kompleksu skupionego wokół sfer: używki – uzależnienia – patologie. Wysoki stopień schematyzacji (mniejsze zróżnicowanie słów/tematów) świadectw pentekostalnych sytuuje pojawienie się tego kompleksu w opisach okresu sprzed nawrócenia, przełomowego momentu życia dla protestantyzmu ewangelikalnego, do którego należy ruch zielonoświątkowy. Aby ukazać wszechmoc Boga przemieniającego dogłębnie jednostkę, w wymowny sposób opisuje się wcześniejsze okresy życia. Sfera, jakiej dotyczą te świadectwa, w zasadzie pokrywa się z mapą współczesnych polskich patologii (gł. alkoholizm i narkomania).

WYNIKI BADAŃ 2 – KOŁOKACJE I SIECI KOŁOKACJI

Kolejne etapy analizy pozwoliły na pozyskanie nowych danych. Kierując się sformułowanymi dotąd wnioskami, sprawdziliśmy kolokacje kluczowych słów dla obu korpusów, jak zaznaczono bowiem w części metodologicznej tekstu, to kontekst wypowiedzi wyznacza znaczenie słów, a ten etap analizy wydaje się niezbędny dla potwierdzenia wstępnych hipotez.

Humanistyka cyfrowa posługuje się wizualizacjami jako elementem analizy i prezentacji wyników. Przedstawiamy dwie sieci i dwa grafy pokazujące wizualizację sieci kolokacji dla słów kluczowych obu korpusów. Rozmiar kół oznacza wagę dyskursywną słowa kluczowego, ustaloną na podstawie stosunku liczby wystąpień do rozmiaru korpusu. Kolor czerwony oznacza słowa kluczowe, które wydobyto z korpusów, zielony to kolokacje tych słów kluczowych, a żółty to kolejny rząd związków kolokacyjnych.



Rysunek 2. Wizualizacja sieci kolokacji dla pierwszych piętnastu słów kluczowych korpusu protestanckiego
Źródło: oprac. aut.

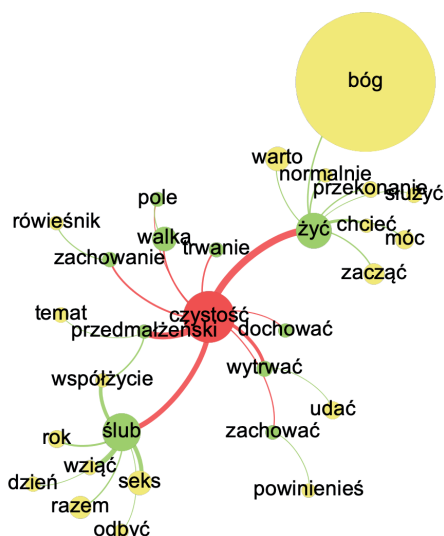
Jak wspomniano, wyróżniającą się pod względem statystycznym grupą świadectw są te o czystości przedmałżeńskiej czy o tym, iż boża interwencja pomaga w znalezieniu życiowego partnera i utrzymaniu go (301 i in.). Zwróćmy uwagę, że np. słowo *cięża* jest bardzo rzadko obecne w korpusie protestanckim (*log-likelihood* na poziomie 10 i mniej), natomiast wchodzi w rozliczne kolokacje w korpusie katolickim, a *być w ciąży* okazuje się jednym z kluczowych zwrotów tego korpusu.

W korpusie katolickim zwraca też uwagę wysoka frekwencja związku dwóch słów: *nowenna pompejańska*. Ten rodzaj modlitwy powstał w 1884 roku we Włoszech w wyniku religijnych objawień Fortunaty Agrelli. Nowenna została rozpropagowana przez dominikańskiego tercjarza Bartolo Longa. Analiza świadectw z nią związanych pokazuje, że ci, którzy ją przywołują, głęboko wierzą w jej skuteczność, określają ją jako „nie do odparcia” (*Nowenna pompejańska*).

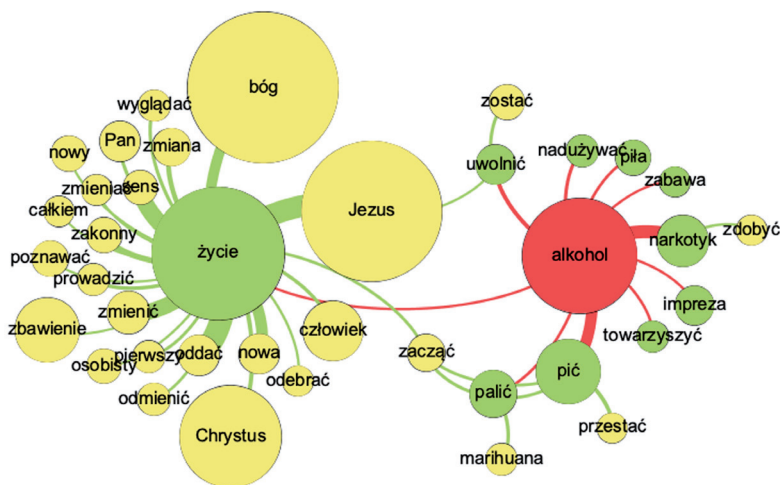
Kolejnym etapem było sprawdzenie słów kluczowych i wydobycie sieci kolokacji, aby zweryfikować dotychczasowe wnioski. Wybraliśmy na potrzeby prezentowanego tekstu dwa wyrazy: *czystość* (korpus katolicki) – rys. 3 – i *alkohol* (korpus zielonoświątkowy), rys. 4. Wizualizacja ta miała pokazać kolejne poziomy naszej penetracji „w głąb” tekstu i ujawnić dalsze struktury znaczenia kluczowych słów. Przykładowo na rysunku 3 widzimy silny związek czystości z frazemem *żyć z Bogiem*, a samo to słowo wykazuje silne związki ze słowami *przedmałżeński*, *współżycie*, *seks*, jak również *dochować*, *walka trwanie/wytrwać*, *zachować/zachowanie*. Zleksykalizowany w dyskursie katolickim zwrot *dochować* czy *zachować czystość* jawi się jako

pomost między dwoma silnymi „gniazdami” znaczeń – pomiędzy ‘życiem z Bogiem’ a przedmażeńskim życiem seksualnym.

Analiza kolokacji ujawnia również niezwykle interesujące metafory pojęciowe, których funkcją jest konceptualizacja wewnętrznego doświadczenia trudności związanych z przedmażeńską abstynencją seksualną katolików. Oprócz wymienionych wcześniej (częściowo spodziewanych) uzusów, jak *zachowanie czystości*, mamy tu także mniej oczywiste, a wiele mówiące o skali podejmowanego wysiłku wyrażenia metaforyczne czerpiące z domeny źródłowej WALKA. Zaczerpniętymi z korpusu przykładami zdań odwołujących się do tej domeny konceptualnej mogą być: „Nie bójcie się *walczyć o czystość*, bo chociaż czekanie bywa gorzkie, jego owoce są przesłodkie” (114), „[...] prowadzę normalne życie młodej kobiety. Codziennie *walczę o czystość*” (159), „Do dziś mam z tym problem jednak *walczę* – skutecznie. Już długo trwam w czystości” (183), „My sami *walczyliśmy* ze sobą, aby mimo wszystko nie przekraczać barier” (123), „Potrzeba więcej takich ludzi, którym chce się *walczyć o czystość*” (286) itd. Konceptualizacja wstrzemięźliwości seksualnej jako podejmowania nieustającej *walki* tworzy językowy obraz postawy pełnej ciągłego napięcia, wymagającego bardzo aktywnego działania, którego celem jest zapanowanie nad popędem seksualnym w imię dochowania wierności wyznawanym wartościom.



Rysunek 3. Wizualizacja sieci kolokacji dla słowa kluczowego czystość w korpusie katolickim
Źródło: oprac. aut.



Rysunek 4. Wizualizacja sieci kolokacji dla słowa kluczowego alkohol w korpusie protestanckim
Źródło: oprac. aut.

Wizualizacje sieci słowa *alkohol* dla korpusu protestanckiego ukazują z jednej strony silny związek tego określenia ze sferą różnych używek i patologii: *pić/piła*, *palić*, *narkotyk*, *nadużywać*. Z drugiej strony bardzo silny związek z słowem *życie* i kompleksem: *zmienić*, *odmienić*, *zmiana*, *nowe*, skupionym wokół słów arcykluczowych (vide rozmiar kółka na grafie) takich jak: *bóg*, *Jezus*, *Chrystus*. Diagram ilustruje z jednej strony chrystocentryzm zielonoświątkowego świadectwa (to Jezus jest agensem przemiany życia), z drugiej ujawnia główne obszary zła i zagrożenia, z którymi mierzą się ewangelikalni charyzmatycy.

WNIOSKI

Przeprowadzona analiza umożliwiła wgląd w pewien fragment katolickiego i zielonoświątkowego dyskursu wyznaniowego, jakim są świadectwa religijne. Różnicowanie obu korpusów jest tylko częściowo uzasadnione tym, że zacerpnięto je z odmiennych mediów (portal internetowy i czasopismo). W obu odmianach dyskursów nadawcy inaczej traktują samo chrześcijańskie świadectwo. Dla zielonoświątkowców jest ono bardziej relacją ze zmieniającego całe życie nawrócenia, dla katolików to opowieści o stałej pomocy i obecności Boga w codziennym życiu. Model ewangelikalnego życia przewiduje gwałtowne nawrócenie, wzór katolicki jest bliższy stałemu i stopniowemu wysiłkowi doskonalenia się. Zatem świadectwa opowiadają o zróżnicowanych formach bożej pomocy w tym zakresie.

Konsekwencją tej różnicy jest większy schematyzm świadectw zielonoświątkowych, operujących mniejszą liczbą słów, które są mocniej akcentowane w tym podkorpusie.

Opisywane sytuacje są bardziej zestandaryzowane, bo skupiają się na kontraście ciemnego, złego życia przed nawróceniem (najczęściej związanego z jakimś uzależnieniem) i pozytywnej, całkowitej przemiany dokonanej za sprawą duchowego przełomu. Nawrócenie, zgodnie z protestancką perspektywą teologiczną, jest darem łaski, darem Boga. Katolickie świadectwa są daleko bardziej zróżnicowane leksykalnie, bo opowiadają o większej różnorodności w skutkach bożej interwencji i pomocy. Poza nielicznymi wyjątkami w zasadzie brak w nich elementu przełomowego, duchowego nawrócenia. W dyskursie katolickim zapisana jest wiara w stałą asystencję i pomoc, jaką Bóg niesie w codziennym życiu człowieka. W dyskursie protestanckim Bóg działa niejako „punktowo”, zmieniając dogłębnie całe życie jednostki.

Analiza ujawniła, że elementem wysoce różnicującym oba korpusy jest problem zła. W tekstach katolickich semantyczny „kompleks zła” pojawia się w kontekście relacji chłopak–dziewczyna, mężczyzna–kobieta i bardzo często dotyczy erotyki. Analiza statystyczna frekwencji słów ukazuje tę sferę jako stale obecną w świadomości młodych ludzi, którzy budują swoje partnerstwo. Można sformułować hipotezę, że przyczyną tak mocnego akcentowania seksualnego (cielesnego) wymiaru tych relacji są rozterki, jakie targają młodymi katolikami, którzy chcą pozostać wierni wymaganiom stawianym im przez ich Kościół (zakaz seksu przedmałżeńskiego, regulacje dotyczące pożycia), a jednocześnie konfrontują je z liberalnymi zasadami współczesnej zachodniej kultury. Katolickie świadectwa są zapisem napięć związanych z konfliktem tych dwu światów wartości.

W naszej trójstopniowej (od leksyki po kolokację) interpretacji kilkuset świadectw można zobaczyć, że głównym obszarem bożej obecności w życiu polskich katolików jest „uporządkowanie” sfery erotyki i uczuć. Dla protestantów interwencje Boga to przede wszystkim przemiana nałogowca (tu bardzo szeroka paleta uzależnień od alkoholu, przez seks, po narkotyki) w nawróconego chrześcijanina.

Analiza niepokojów sumienia, które targają młodymi katolikami, ujawnia dylematy, jakie przeżywają, pracując nad swoją tożsamością religijną. Badania potwierdzają tezę, że świat językowy analizowanych świadectw skupia się wokół problematyki poprawnego życia seksualnego, tak jakby bycie katolikiem polegało na określonych regulach, w których moralność nierozłącznie jest związana z „właściwym” seksem. Zło dla dyskursu zielonoświątkowego jest właściwie sprawą przeszłości. Nawrócenie „uwalnia od zła” i lokuje je w przeszłości.

Zaprezentowany rekonesans badawczy pokazuje możliwości, jakie daje połączenie metod ilościowych i jakościowych w humanistyce, w tym przypadku w korpusowej analizie dyskursu. Miała za zadanie pokazać, jak narzędzia językoznawstwa korpusowego (traktowanego jako jeden z elementów humanistyki cyfrowej) mogą być pomocne w interpretacji dużych rozmiarowo danych tekstowych. Chcieliśmy pokazać, jak narzędzia te mogą uzupełniać naszą wiedzę o współczesnym życiu religijnym i polskiej kulturze religijnej.

BIBLIOGRAFIA

- Baker, P. 2006. *Using Corpora in Discourse Analysis*. London, New York: Continuum.
- Baker, P., i T. McEnery, red. 2015. *Corpora and Discourse Studies: Integrating Discourse and Corpora*. London: Palgrave Macmillan.
- Bączkowska, A. 2016. „Korpusowa analiza dyskursu związanego z tematyką imigracji w brytyjskiej prasie opiniotwórczej”. *Conversatoria Linguistica* 10: 7–25.
- Boniecka, B. 1994. „Tekst w kontekście (problemy metodologiczne)”. *Polonica* 16: 43–67.
- Brezina, V. 2018a. „Collocation Graphs and Networks: Selected Applications”. W *Lexical Collocation Analysis: Quantitative Methods in the Humanities and Social Sciences*, red. P. Cantos-Gómez, i M. Almela-Sánchez. Cham: Springer, 59–83.
- Brezina, V. 2018b. *Statistics in Corpus Linguistics: A Practical Guide*. New York: Cambridge University Press.
- Chlondowski OFM, W.P. 2018. „Pentekostalizacja w ujęciu kardynała J. Ratzingera oraz dokumentów episkopatu – dar czy zagrożenie?”. <https://www.deon.pl/magazyn/pentekostalizacja-w-ujeciu-kard-j-ratzingera-oraz-dokumentow-episkopatow-dar-czy-zagrozenie,481475> (19.07.2020).
- Dunning, T. 1993. „Accurate Methods for the Statistics of Surprise and Coincidence”. *Computational Linguistics* 19 (1): 61–74.
- Durant, A. 2006. „Raymond Williams’s Keywords: Investigating Meanings «Offered, Felt for, Tested, Confirmed, Asserted, Qualified, Changed»”. *Critical Quarterly* 48 (4): 1–26.
- Eder, M. 2016. „Słowa znaczące, słowa kluczowe, słowozbiory – o statystycznych metodach wyszukiwania wyrazów istotnych”. *Przegląd Humanistyczny* 60 (3, 454): 31–4.
- Firth, J.R. 1957. *Papers in Linguistics 1934–1951*. London: Oxford University Press.
- Foucault, M. 1972. *The Archaeology of Knowledge*. New York: Pantheon Books.
- Kamasa, V. 2014. „Techniki językoznawstwa korpusowego wykorzystywane w krytycznej analizie dyskursu. Przegląd”. *Przegląd Socjologii Jakościowej* 10 (2): 100–117.
- Kilgarriff, A. 2001. „Comparing Corpora”. *International Journal of Corpus Linguistics* 6 (1): 97–133.
- Knotz, K., i K. Strączek. 2010. *Seks jest boski, czyli erotyka katolika*. Kraków: Znak.
- Lewandowska-Tomaszczyk, B., red. 2005. *Podstawy językoznawstwa korpusowego*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Mariański, J. 2015. „Religia i moralność w świadomości Polaków: zależność czy autonomia?”. *Konteksty Społeczne* 3 (1, 5): 8–26.
- McEnery T., i A. Wilson. 2001. *Corpus Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.
- Migda, A. 2013. *Mistycyzm pentekostalny*. Kraków: Nomos.
- Mitchell, R. 2019. *Ekstrakcja danych z językiem Python. Pozyskiwanie danych z Internetu*. tłum. K. Sawka. Gliwice: Helion.
- Nowak, M. 2005. *Świadectwo religijne. Gatunek – język – styl*. Lublin: Towarzystwo Naukowe KUL.
- „Nowenna pompejańska”. <http://mysliborska28.pl/nawenna-pompejanska> (16.07.2020).
- Pasek, Z., i B. Sorbjan. 2017. „Świadectwa wiary członków Kościoła Zielonoświątkowego jako forma perswazji religijnej”. *Rocznik Teologiczny* 59 (1): 173–196.
- Pawlikowska, A. 2015. *Zróżnicowanie leksykalno-semantyczne dyskursów wyznaniowych na materiale polskich tekstów ewangelicznych, katolickich i prawosławnych*. Wrocław: Quaestio.
- Pelczyński, G. 2006. „«Świadectwa». Opowieści o własnym nawróceniu”. *Lud* 90: 37–52.

- Boguszewski, R. 2018. „Religijność Polaków i ocena sytuacji Kościoła katolickiego”. *Komunikat z Badań CBOS* 147: 1–6.
- Sambor, J. 1972. *Słowa i liczby. Zagadnienia językoznawstwa statystycznego*. Wrocław: Zakład Narodowy im. Ossolińskich.
- Scott, M. 1997. „PC Analysis of Key Words – And Key Key Words”. *System: An International Journal of Educational Technology and Applied Linguistics* 25 (2): 233–45.
- Scott, M., i C. Tribble. 2006. *Textual Patterns: Keyword and Corpus Analysis in Language Education*, Amsterdam: John Benjamins Publishing.
- Skowronek, K. 2006. *Między sacrum a profanum. Studium językoznawcze listów pasterskich Konferencji Episkopatu Polski (1945–2005)*. Kraków: Wydawnictwo Lexis.
- Wierzbička, A. 2007. *Słowa klucze. Różne języki – różne kultury*. tłum. A. Duraj-Nowosielska. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.
- Williams, R. 1983. *Keywords: A Vocabulary of Culture and Society*. London: Fontana.
- Woźniak, M. 2017. *Jak znaleźć igłę w stogu siana? Automatyczna ekstrakcja wielosegmentowych jednostek leksykalnych z tekstu polskiego*. Kraków: Instytut Języka Polskiego PAN.
- Zieliński, A. 2004. *Zjawiska cudowne w polskim katolicyzmie ludowym*. Kraków: Nomos.

WYKAZ ŹRÓDEŁ

- „Chrześcijanin”, miesięcznik Kościoła Zielonoświątkowego w Polsce, świadectwa z lat 1995–2010, <http://www.chn.kz.pl/>.
- „Deon.pl”, portal społecznościowo-informacyjny jezuitów, świadectwa z lat 2010–2019, <https://deon.pl/>.

Evangelical and Catholic Testimonies in the Perspective of Corpus-based Discourse Analysis

Key words: corpus linguistic, Christian testimonies, Polish religious culture, corpus-based discourse analysis.

ABSTRACT

The article presents the results of a comparative analysis of Roman Catholic (from the portal Deon.pl) and Protestant, Pentecostal (from the magazine “Chrześcijanin” et al.) Christian testimonies. Using the tools of corpus linguistics, the authors show the differences between both collections of texts. Especially interesting was the reconstruction of the „areas of sin”, the evils the testimony authors overcame thanks to God’s help. These are different spheres for both types of texts, in the case of Catholic testimonies it is formed by a semantic complex built around the terms of “love – sex – asceticism”, and Pentecostal testimonies are focused on the “stimulants – addictions – antisocial behaviour” complex. The analysis consisting of two stages (frequency, describing the role of keywords in the corpus, and visual recognition of the most important collocations) allowed us to formulate conclusions on the dominant areas of meaning which define the mental map of the religious life of the followers of both the Christian denominations.