

Magdalena Majdak (Kraków)

Słowa ważniejsze niż inne – metoda słów kluczy w badaniu polszczyzny dawnej¹

Celem artykułu jest próba zastosowania metody słów kluczy do badania polszczyzny dawnej. Podstawą prezentowanych analiz jest materiał z *Elektronicznego korpusu tekstów polskich XVII i XVIII w. (do 1772 r.)*, który powstaje w Pracowni Historii Języka Polskiego XVII-XVIII wieku Instytutu Języka Polskiego PAN. Materiał podzielono na cztery okresy (I 1601–1650, II 1651–1700, III 1701–1750, IV 1751–1772) i porównano wyniki dla tekstów z zakresu polityki. Rezultaty pokazały zasadność użytej metody. Wskazały także na silny związek listy wyrazów z doбором utworów oraz na konieczność stałej rozbudowy źródeł korpusowych dla polszczyzny wszystkich epok.

Słowa kluczowe: metoda słów kluczy, korpus tekstów, polszczyzna XVII-XVIII wieku

Words more important than others: The use of the keyword method in studying old Polish

The purpose of this article is to try using the keyword method in studying old Polish. The presented analyses are based on the material from the Electronic Corpus of 17th and 18th century Polish texts (up to 1772) currently being compiled at the History of the Polish Language of the 17th and 18th Century Section at the Institute of the Polish Language of the Polish Academy of Sciences. The material has been divided into four periods (1. 1601–1650, 2. 1651–1700, 3. 1701–1750, 4. 1751–1772) and the results have been compared for political texts. The results have shown the effectiveness of method used. They have also shown a connection between the list of words and the texts selection and the necessity to constantly expand corpus sources for the Polish language across centuries.

Key words: keyword method, text corpus, Polish language of the 17th and 18th century

¹ Tekst niniejszy powstał w ramach projektu badawczego „Elektroniczny korpus tekstów polskich XVII i XVIII w. (do 1772 r.)” realizowanego przez Pracownię Historii Języka Polskiego XVII i XVIII w. Instytutu Języka Polskiego PAN we współpracy z Zespołem Inżynierii Lingwistycznej Instytutu Podstaw Informatyki PAN. Grant finansowany jest ze środków Narodowego Programu Rozwoju Humanistyki na lata 2013–2018 (nr projektu: 0036/NPRH2/H11/81/2012). Artykuł jest poszerzoną wersją referatu pt. *Słowa ważniejsze niż inne* wygłoszonego na międzynarodowej konferencji „Słowa klucze 2016”, która odbyła się w dniach 20–21 października 2016 r. na Uniwersytecie Warszawskim.

Wörter, die wichtiger sind als andere – die Anwendung der Schlüsselwortanalyse auf das Altpolnische

In diesem Beitrag soll versucht werden, die Schlüsselwortanalyse auf das Altpolnische anzuwenden. Grundlage der hier dargestellten Analysen ist das Textmaterial aus dem Elektronischen Korpus polnischer Texte des 17. und 18. Jahrhunderts (bis 1772), der im Labor für Geschichte der Polnischen Sprache des 17.-18. Jahrhunderts des Instituts für Polnische Sprache an der Polnischen Akademie der Wissenschaften entsteht. Das Material wurde in vier Zeitphasen geteilt (I 1601–1650, II 1651–1700, III 1701–1750, IV 1751–1772) und es wurden Texte aus dem Bereich Politik miteinander verglichen. Wie die Ergebnisse zeigen, war die Anwendung dieser Methode begründet. Darüber hinaus konnte belegt werden, dass ein starker Zusammenhang zwischen der Wörterliste und der Auswahl der Werke besteht und dass es erforderlich ist, die Korpusquellen für Polnisch aller Epochen kontinuierlich auszubauen.

Schlüsselwörter: Analyse der Schlüsselwörter, Korpus, Polnische Sprache des 17.-18. Jahrhunderts

1. Wprowadzenie

Narzędzia elektroniczne umożliwiają wyszukiwanie w ogromnych zbiorach tekstów słów istotnych z punktu widzenia danej kultury, czasu, odbioru otaczającej rzeczywistości i sposobu jej przeżywania. Dla historyka języka interesującą perspektywę otwiera możliwość zastosowania metody wydobywania słów kluczy w odniesieniu do tekstów dawnych. Analiza leksyki statystycznie istotnej może być pomocna w odkrywaniu ważkich historycznie tematów społecznych, politycznych, a nawet w próbach odtworzenia wizji świata. Otwiera to pole do rozważań na temat kryteriów wyboru słów istotniejszych niż inne. Tytułowe słowo „ważniejsze” można również określić jako „słowa na czasie”, będzie to jednak czas siedemnastego i osiemnastego wieku.

Podstawą prezentowanych analiz jest przede wszystkim materiał z *Elektronicznego korpusu tekstów polskich XVII i XVIII w. (do 1772 r.)*, zawierającego utwory z lat 1601–1772. Korpus *KorBa* powstaje w Pracowni Historii Języka Polskiego XVII-XVIII wieku Instytutu Języka Polskiego PAN przy udziale Zespołu Inżynierii Lingwistycznej Instytutu Podstaw Informatyki PAN pod kierunkiem Włodzimierza Gruszczyńskiego. Docelowo będzie liczył 12 mln segmentów, zakończenie projektu planowane jest na rok 2018. Narzędzie tego rodzaju pozwala przeprowadzać badania na skalę dotąd niemożliwą, w czasie, jaki też dotąd nie był możliwy.

Analiza zmian semantycznych dokonujących się na przestrzeni wieków, obserwacja pojawiania się i zanikania wyrazów i znaczeń należy do głównych zadań historyków języka. Być może zastosowanie metody słów kluczy do wielkich

zbiorów tekstów, a nie tylko do konkretnych utworów, byłoby pomocne w dociekaniach leksykologicznych nad dawną polszczyzną, a także wzbogaciłoby badania z semantyki diachronicznej o wymiar społeczny, historyczny, kulturowy. Przedmiotem poszukiwania byłyby tu jednak nie „słowa dnia”, ale połowy stulecia, stulecia czy epoki.

Za próbę urzeczywistnienia tej myśli można uznać powstający pod kierunkiem prof. Adama Pawłowskiego Portal Tekstów Prasowych *ChronoPress*², czyli narzędzie służące m.in. do analizy ilościowej materiału, ilustrujące m.in. frekwencję wyrazów obecnych w polskiej prasie z lat 1945–1957. Jednym z przykładów praktycznego wykorzystania tego narzędzia jest eksperyment zainspirowany pytaniem, jakie postawił prof. Włodzimierz Gruszczyński na zjeździe PTJ w Kielcach w roku 2015, a dotyczącym dystrybucji w badanym materiale nazw *Stalinogród* i *Katowice*. Wyszukiwanie pokazało, że po 1953 nazwa *Katowice* w zasadzie (poza marginalnymi użyciami w znaczeniu historycznymi) znika i ustępuje miejsca *Stalinogrodowi* (por. też Pawłowski 2016: 68). Pytanie brzmi, czy takie „stalinogrody” istniały również w XVII-XVIII wieku bądź w innym interesującym okresie.

Przykładem pochodzącym z badanej tu epoki może być krąg tematyczny związany ze słowem *filiżanka* (pierwsza notacja w formie *fizanka* pochodzi z 1638 r.). Cytaty zawarte w artykule hasłowym *Elektronicznego słownika języka polskiego XVII i XVIII wieku* wskazują, z jakiego materiału ją wytwarzano, kiedy zaczęto pić z filiżanki, co w niej pito – była to głównie herbata i kawa (*herbathe* i *kaffa*), ale także rosół – oraz – kiedy pojawiły się *kaffenhauzy*. Takim słowem kluczem (przy udziale *Korby*) można by otworzyć wrota do języka i świata dawnych epok.

Ewolucji może podlegać oczywiście nie tylko forma, ale i treść słowa. Spójrzmy na znaczenia wyodrębnione w artykule hasłowym *aktor* pochodzącym z e-SXVII³: 1. *prawn.* »ten, kto wnosi sprawę do sądu, strona powodowa, powód«, 2. *prawn.* »obwiniony, oskarżony«, 3. »właściciel, posiadacz«, 4. *prawn.* »opiekun, pełnomocnik nieletniego«, 5. »człowiek grający rolę w teatrze lub jakimś widowisku«. Być może dzięki korpusowi tekstów XVII-XVIII wieku będzie można w przyszłości automatycznie sprawdzić, dla jakich tekstów kluczowe jest konkretne znaczenie słowa. Zabiegów związanych z wydobyciem słów kluczy z dawnej materii językowej należałoby dokonywać nie tylko w materiale anotowanym morfologicznie (co wymaga m.in. opracowania gramatyki badanego okresu), ale także poddanym dezambiguacji.

² <http://chronopress.clarin-pl.eu>. Korpus i wyszukiwarka *ChronoPress* powstały w ramach prac konsorcjum naukowego CLARIN – *Polskie wspólne zasoby językowe i infrastruktura technologiczna*.

³ http://xvii-wiek.ijp-pan.krakow.pl/pan_klient/index.php?strona=haslo&id_hasla=5163&forma=AKTOR#5235

Sztandarowe pojęcie *słowo klucz* (*słowo-klucz*) ma już swoją historię, wiele sposobów rozumienia i algorytmów wyliczania. W odniesieniu do polszczyzny dawnej było stosowane np. przy analizach języka artystycznego – mam tu na myśli prace J. Sambor (1969), E. Stachurskiego (1989, 1998, 2005) czy M. Edera (2016). O związkach słów i wartości kulturowych pisała m.in. A. Wierzbicka (2007), wprost o kulturomach A. Nagórko (2004), w przeglądowych artykułach zaś J. Bartmiński (2016) i M. Rak (2015)⁴. Relację ze stanu wiedzy w tym zakresie zdają najnowsze opracowania, m.in. J. Bartmińskiego, M. Łazińskiego, A. Pawłowskiego, W. Pisarka, zawarte w numerze 3 „Przeglądu Humanistycznego” z 2016, a także artykuły ze wspomnianej na początku konferencji, znajdujące się w tym tomie.

Ze świadomością istnienia bogatej literatury przedmiotu i rozmaitych znaczeń i kontekstów, w jakich się pojawiają, w artykule będę się odnosiła ogólnie do tytułowych „słów ważniejszych niż inne”. Jakie mogą być kryteria owej „ważności”? Czy istnieją słowa nieważne? Jak można określić ich charakter? Czy ważniejsze są te zapisane na sztandarach, jak *wolność* lub *niepodległość*, czy te pozbawione samodzielnego znaczenia, jak *w czy się?* Czy istotniejsze będzie słowo *on* czy *ja* albo *ty*, *mój* czy *nasz*, *sobie* czy *tobie*? Słowa powszechne czy indywidualne, piękne czy wulgarne, brzmiące ładnie czy pospolicie, krótkie czy długie, najnowsze czy najstarsze, te, które wyszły z użycia czy te, które zostały, rodzime czy zapożyczone, poprawne czy niepoprawne, familijne czy ogólne, podstawowe czy derywowane, określające czy określane, te, które się powtarzają, czy te trudne do zastąpienia, autosemantyczne, czy te wskazujące na zależności składniowe, o frekwencji najwyższej czy najniższej (*hapax legomena*) w porównaniu z korpusem referencyjnym. Czy rangę słowu może nadawać temat lub kontekst, czy też ma ono wartość samoistną? Czy ważniejsze jest słowo *ojczyzna* czy *chałupa*? Słowo *kocham* może wydawać się istotniejsze niż np. *deska*, choć inaczej będzie na to patrzył konstruktor kładki lub ktoś, dla kogo byłaby to ostatnia deska ratunku. Nie bez znaczenia jest też kwestia tabu językowego – czasami ważniejsze jest to, czego się nie mówi. Podobnych pytań można oczywiście zadać wiele, te nie wyczerpują tematu „słów ważniejszych niż inne”.

⁴ Wydaje się, że w szerokim rozumieniu „kluczowości” słowa rozumianej jako swego rodzaju *signum temporis* do badań można by dołączyć także analizę konstrukcji gramatycznych (por. sprostowanie Marka Łazińskiego dotyczące angielskiego słowa *because*, które w części kontekstów zmieniło funkcję gramatyczną i może teraz np. rozpoczynać zdanie (2016: 90). Idąc tym tropem – w odniesieniu do czasów współczesnych – za symptomatyczne można by uznać nawet pojawianie się nowych realizacji fonetycznych (np. sybilantów) rejestrowanych w korpusach języka mówionego.

2. Cel

Głównym celem tego artykułu jest przedstawienie wyników próby zastosowania metody słów kluczy do zbioru tekstów dawnych oraz sprawdzenie, czy taki pomysł może przynieść rezultaty już na obecnym etapie zawansowania prac. Poniżej zostaną zaprezentowane trzy takie eksperymenty, każdy z zastosowaniem innych kryteriów doboru tekstów.

3. Materiał

Podstawą materiałową rozważań jest wspomniany korpus tekstów barokowych *KorBa*. Na potrzeby tego artykułu korzystam z wersji liczącej ponad 10 mln słów, obejmującej 692 teksty w transliteracji. Zbiór ten pełni tutaj funkcję korpusu referencyjnego. O istotnym ograniczeniu, jakim jest stosunek zawartości korpusów do realnej materii językowej i wynikających stąd zniekształceń nie trzeba tu przypominać. Jak we wszystkich tego rodzaju przedsięwzięciach należy stale pamiętać, że za zawartością korpusu stoją konkretne decyzje jego konstruktorów. Mają one wpływ m.in. na dobór tekstów, który i tak jest już pewnym wyborem z tych, jakie do naszych czasów dotrwały, a te z kolei reprezentacją tych w odmianie pisanej, odzwierciedlających preferowane wówczas style czy typy.

Jednym ze sposobów przeszukiwania i pozyskiwania materiału w korpusie *KorBa* jest możliwość wyboru tekstu określonego gatunku (jednego z 54), pochodzącego z konkretnego regionu (jednego z 11), o zadanej tematyce (jednej z 38). Repertuar kategorii obejmuje też m.in. konkretny przedział czasowy, rodzaj mowy i inne. Aby podjąć próbę odnalezienia słów kluczowych dla poszczególnych połówek XVII i XVIII stulecia, wybrano teksty z zakresu „polityka”. Chodziło o to, by uzyskać grupę jednolitą tematycznie, a jednocześnie podejrzaną o rejestrowanie aktualnych wydarzeń. W zamierzeniu utwory miały dotyczyć w miarę możliwości bieżących problemów Rzeczypospolitej, zrezygnowano zatem z przekładów np. politycznych tekstów starożytnych (np. *Mów przeciwko Katylinie* Cycerona w tłumaczeniu Ignacego Nagurczewskiego (Warszawa 1763)). Następnie zawężono wybór do takich gatunków, jak akta sejmikowe, diariusze sejmowe, konstytucje sejmowe, mowy polityczne i pisma polityczne. Uzyskano w ten sposób rezultat w postaci 89 tekstów. Podzielono je według okresów I 1601–1650, II 1651–1700, III 1701–1750, IV 1751–1772 i dobrano materiał tak, ażeby zbiory były porównywalne nie tylko tematycznie, ale i pod względem objętości. Liczba znajdująca się po dacie wydania oznacza liczbę znaków ze spacjami w utworze, liczba pogrubiona zaś sumę ogólną znaków, jakie znajdują się w każdej części materiału. Ostatecznie do czterech podkorpusów weszły następujące utwory z poszczególnych części stuleci:

I (1601–1650): Stanisław Witkowski, *Złota wolność koronna*, Kraków 1609; 79217, Mikołaj Chabielski, *Sposób rządu koronnego i gotowości. Obrona Rzeczypospolitej, jakaby miała być* Kraków 1615; 36843, Adam Grodziecki, *Przestroga o tytułach i dygnitarstwach*, nieznane 1634; 35110; Jeremiasz Bielejowski, *Obrona tytułów książęcych od Rzeczypospolitej uchwałą sejmową pozwoionych*, nieznane 1641; 31229, Anonim, *Dyskurs o zawziętych terażniejszych zaciągach, skąd in Republica urosły, i o postępku sejmu 1606*, nieznane 1606; 81226; Anonim, *Sposób podania drogi do ś. sprawiedliwości i pokoju pospolitego w Koronie polskiej*, nieznane 1606; 20350, Anonim, *Libera respublica — absolutum dominium — rokosz*, nieznane 1606; 24259, Anonim, *Zdanie szlachcica polskiego o rokoszu*, nieznane 1606; 14708, Anonim, *Przestroga i sposób na czasy przyszłe naprawy Rzpltej*, nieznane 1606; 21904; Anonim, *Collatio tego wszystkiego, co na rokoszu i w Wiślicy zawarto*, nieznane 1606; 42124, Anonim, *Przestroga Rzpltej potrzebna, którą kanclerz on Zamoyski dwiema szlachcicom godnym wiary, ukazawszy się niedawno, in publicum podać i komu należy, odnieść kazał*, nieznane 1607; 23924, Anonim, *Rewersał listu szlachcica jednego do drugiego pisany*, nieznane 1606; 38818 Mikołaj Zebrzydowski, *Fundament rokoszu, pod Jeziorną 22 Iunii ustnie oświadczony*, nieznane 1606; 110575 – **560 287**

II (1651–1700) Jerzy Sebastian Lubomirski, *Jawnej niewinności manifest*, nieznane 1666; 298 209, Anonim, *Perspektywa na objaśnienie niewinności*, nieznane 1666; 160 156 – **458 365**

III (1701–1750) Stanisław Konarski, *Rozmowa pewnego ziemianina ze swoim sąsiadem*, Warszawa 1733; 252 286, Anonim, *Przestroga generalna stanów Rzpltej...*, nieznane 1713; 19 184, Wojciech Bystrzonowski, *Polak sensat ...*, Lublin 1730; 8580, S. Dunin-Karwicki, *Egzorbitancje we wszystkich trzech stanach Rzeczypospolitej ...*, Lublin 1703; 21 748, Stanisław Leszczyński, *Głos wolny wolność ubezpieczający*, Wilno 1733; 330 480 – **632 278**

IV (1751–1772) Stanisław Konarski, *O skutecznym rad sposobie*, Warszawa 1760; 320 454, Stefan Garczyński, *Anatomia Rzeczypospolitej Polskiej*, Warszawa 1753; 220962 – **541 416**

4. Analiza

Do analiz wykorzystano rozszerzoną wersję programu AntConc z 2015 roku⁵. Jedną z funkcji, jakie oferuje, jest możliwość stworzenia listy słów kluczy (*Keywords*), czyli ważniejszych niż inne w porównaniu z korpusem referencyjnym, niekoniecznie o najwyższej frekwencji. Trzeba pamiętać, że analizowany materiał nie jest lematyzowany, co w tym przypadku oznacza, że program liczy wystąpienia poszczególnych form paradygmatu. Planowana anotacja Korpusu, wiążąca się ze szczegółową analizą kontekstów, pozwoli także ustalić znaczenia słów i odpowiedzieć na pytanie, czy np. *miłość* występuje w funkcji nazwy uczucia, czy też jest elementem zwrotu grzecznościowego.

Kolejność postępowania badawczego była następująca. Dla każdej z czterech części stuleci XVII i XVIII (do 1772 r.) wygenerowano listę *Keywords*. Pod uwagę wzięto pierwsze 200 pozycji rangowych, a następnie pozostawiono rzeczowniki i zastępujące je skróty w języku polskim. Kwestią wymagającą rozstrzygnięcia był status słów łacińskich. Pozostawiono je, jeśli wystąpiła rodzima forma odmiany (np. *script*, *scriptu*) lub w przypadkach wątpliwych. Na listach z poszczególnych przedziałów czasowych znalazły się następujące formy:

	I (1601–1650)		II (1651–1700)		III (1701–1750)		IV (1751–1772)	
1.	rzpltej	983.306	mości	2565.282	prawo	892.930	sejmów	1607.344
2.	kjmci	650.484	pána	1415.362	marszałka	861.183	sejmy	1215.038
3.	kjm	476.706	wm	1226.474	seymie	770.984	rzpltej	669.740
4.	tytuły	300.591	rzeczyposp	1022.065	sejmy	679.428	sejm	466.627
5.	tytułow	239.604	mość	1015.703	sąsiad	669.530	ojczyzny	443.202
6.	rzplta	204.141	márszałek	729.566	ziemianin	635.504	rady	404.854
7.	wolnosc	202.135	márszałká	649.295	seymow	610.920	utrzymywa- nia	356.351
8.	rokoszanie	174.486	elekcyey	626.533	buławy	599.745	sposób	328.283
9.	wm	173.595	oyczyzny	535.438	seym	576.146	sejmach	324.752
10.	rzplta	141.970	woysko	510.562	prawa	522.004	rwania	313.521
11.	pp	128.772	oyczyznie	491.045	mći	516.552	dzieci	278.490
12.	rokoszu	114.676	rzeczypospolitey	488.852	seymu	514.133	rzpltej	239.440
13.	sejmu	113.123	instigator	483.999	buław	512.678	oyczyźnie	236.442
14.	praktyk	110.897	pánu	460.726	elekcyi	442.251	radach	234.771
15.	sejmie	105.685	niewinności	396.149	pan	339.271	wakansów	217.659
16.	wolności	102.389	pan	372.444	panie	320.940	elekcyje	213.612
17.	stan	91.536	pánów	368.663	elekcyą	320.385	sejmu	188.353
18.	rokosz	89.400	woyská	361.761	rzpтей	301.989	rzeczypospolitey	175.072

⁵ <http://www.laurenceanthony.net/software.html>

	I (1601–1650)		II (1651–1700)		III (1701–1750)		IV (1751–1772)	
19.	sejm	84.926	rzeczposp	327.685	seymowania	275.189	interessów	158.020
20.	senatorów	84.658	mánifestu	306.673	krol	245.732	ludzi	151.327
21.	senatory	83.960	punct	306.484	rzpta	209.771	kraju	148.018
22.	gotowości	76.722	kšiedzá	278.138	poniatowski	207.886	rzplitą	134.646
23.	huffy	73.118	wolności	266.130	konfederacya	207.529	duchowni	120.044
24.	stanu	72.586	márszałkowi	262.198	krola	187.269	wolność	119.090
25.	egzekucyey	72.004	związku	256.205	seymach	180.490	rodziców	118.422
26.	xiąząt	69.732	krolá	255.114	województw	180.053	dwór	116.329
27.	wolności	68.806	krolovey	221.551	zagaieniem	150.853	wolności	112.759
28.	wišliczanie	67.239	confidentá	213.419	konfederacyi	149.334	oyczyznę	110.557
29.	jkm	65.885	mci	206.570	posłów	134.129	rzepłtej	110.195
30.	oyczyźnie	65.831	práwá	206.334	instancye	133.637	prawa	104.159
31.	wišlicy	65.259	confoederáye	204.743	hetmanow	123.535	stan	103.992
32.	praktyki	64.669	czás	191.490	okolicznosciach	117.345	panowie	102.168
33.	šlachcicowi	62.893	seymie	190.111	pana	115.709	elekcyjami	101.993
34.	praw	62.751	listu	185.931	prawu	113.609	rzplita	95.194
35.	senatorowie	62.579	tátárov	177.048	wakansow	112.927	racyje	93.330
36.	wmci	62.360	posłów	176.649	panu	112.264	rzádu	89.363
37.	rzeczypospolitey	60.991	elekcyą	162.493	oyczyźnie	107.099	rzplta	88.226
38.	práwá	56.968	elekcyja	152.669	poniatowskiego	99.160	interessu	87.247
39.	republicá	56.184	krolowa	152.079	seymem	97.749	praw	84.626
40.	tumultów	55.312	woysku	151.972	porzádku	97.241	interessach	83.539
41.	fámiiom	55.283	mánifest	150.343	krolowi	94.690	rzplite	81.595
42.	senator	54.697	związek	150.143	vota	94.602	anarchiją	79.010
43.	opresyą	52.835	kšiadz	145.481	boze	91.377	elekcyjom	74.795
44.	strony	50.333	mošcią	141.704	veto	85.082	obrad	73.385
45.	praktykach	50.037	vslugi	140.146	rzeczypospolitey	84.968	panów	72.872
46.	szwecyey	50.037	dwor	139.774	rwania	83.946	szlachta	72.478
47.	prochow	49.151	márszałku	138.059	posłowie	81.629	paragrafie	72.223
48.	mierze	47.776	oyczyznę	136.625	konfederacyą	75.896	racyj	71.329
49.	nieprzyacielowi	47.703	záwziętošci	136.565	dwor	74.490	patryjotów	67.996
50.	kul	47.348	seym	133.329	dyskursu	73.425	sejmie	67.919
51.	egzekucyą	46.143	confoederacyą	130.266	uzpokoienia	72.892	sposobów	67.792
52.	prawa	46.075	senatorá	129.417	rzepłtey	71.457	cicero	67.366
53.	koronie	44.610	rády	128.243	interes	71.112	konfederacyją	66.916
54.	wolność	43.569	złošliwošć	126.226	wolności	70.304	sposobu	65.854

	I (1601–1650)		II (1651–1700)		III (1701–1750)		IV (1751–1772)	
55.	ludzie	43.252	senatorów	125.627	traktacie	68.903	rzplątą	65.705
56.	środków	42.519	pánem	122.081	prywat	68.840	elekcyjach	65.439
57.	xiążęta	41.914	pánowie	117.660	instrukcyi	68.346	interest	64.915
58.	stanowi	40.97	copia	116.075	miarą	67.160	podłości	63.470
59.	rakuszany	40.65	obiecta	110.275	alternatach	63.517	głosów	63.407
60.	senatorom	40.54	prawo	109.483	pretensye	62.671	polwsze	62.456
61.	ojczyzny	39.408	práwu	107.518	seymom	60.986	pána	61.865
62.	królestwa	39.404	falsz	106.827	rzpltej	60.357	konserwacyą	60.514
63.	majątności	39.402	list	106.249	wolność	60.291	rada	59.611
64.	kar	38.846	szláhcicá	103.655	ministerya	59.862	fakcyje	58.511
65.	nieprzyacię- łom	38.846	pvnet	101.638	woiewodz- twach	59.686	wsie	57.908
66.	rzeczy	38.381	instigatorá	98.891	hetmani	58.787	liczba	57.071
67.	potomstwu	38.246	mándat	98.028			sejmików	56.938
68.	przodkowie	38.156	książstwa	96.399			veto	56.302
69.	polwsze	37.877	poselstwa	96.399			racyjami	56.180
70.	porządku	37.772	pánom	95.657			posłów	55.967
71.	konstytucyą	37.398	mości	94.541			zdania	55.612
72.	senatora	37.339	łáski	93.253			wsiach	54.575
73.	rzeczach	37.022	commissiey	90.468			zrywania	54.397
74.	duchowni	36.635	seymu	90.351			sposoby	54.345
75.	tytvlach	36.423	práwom	89.713			słowa	54.142
76.	pokoju	36.272	fortun	89.547				
77.	obcych	36.091	stanu	89.527				
78.	osób	35.490	poddánych	88.019				
79.	tytułach	35.468	punctá	87.326				
80.	czas	35.196	thronu	86.278				
81.			praw	86.271				
82.			woyskiem	85.083				
83.			mościow	84.258				
84.			successorá	84.227				
85.			sądu	83.819				
86.			rzeczpospo- lita	83.716				
87.			wárszawy	83.655				
88.			script	83.485				
89.			polská	82.656				
90.			oyczyzná	80.368				
91.			práwem	80.198				
92.			instrucyą	80.171				
93.			izbá	79.424				
94.			seymikách	78.197				
95.			oppressyą	76.173				
96.			spráwiedli- wości	74.959				

	I (1601–1650)	II (1651–1700)	III (1701–1750)	IV (1751–1772)
97.		máiestatu	74.370	
98.		dobrodzieiá	70.910	
99.		šwiádectwo	70.787	
100.		przyaóioł	70.661	
101.		mándaóie	70.267	
102.		seymikow	69.368	
103.		scriptu	69.029	
104.		zwiąskowi	69.029	
105.		czármieckiego	67.428	
106.		oppressiey	66.773	
107.		vprzeymoóí	66.686	
108.		fortuny	66.564	
109.		šwiádkow	66.403	
110.		seymách	65.943	
111.		niewinnoó	65.404	
112.		kozakow	65.366	
113.		szláóhtę	64.850	
114.		deputaóí	64.440	
115.		kóšiestwo	64.440	
116.		condicye	64.433	
117.		primássá	64.433	
118.		vspokoienia	63.747	
119.		szláóhty	63.554	
120.		decret	63.347	
121.		electiey	63.113	
122.		kánclerzá	61.912	
123.		ichmoóóíow	61.439	
124.		obiáóńnienie	61.268	
125.		záwžíótoó	61.268	
126.		óátoóóí	61.163	
127.		slugá	60.999	
128.		commissiá	60.746	

5. Wnioski

Łącznie uzyskano 349 form wyrazowych o najwyższym stopniu „kluczowości”. Lista I zawiera 80 pozycji, II – 128, III – 66, IV – 75. Zestawienie i porównanie ich zawartości prowadzi do wielu wniosków. Najbardziej interesujące są tu z jednej strony wyrazy o jak największym zakresie czasowym, czyli pojawiające się w każdej części stulecia, z drugiej strony te charakterystyczne dla jednej z nich. Wśród tych pierwszych aż 23 pozycje zajmują różne formy i postaci wy-

razu *Rzeczpospolita*⁶. Szczegółowe rozpisanie form tego słowa wygląda następująco dla poszczególnych okresów:

I: *rzpltej, rzeczypospolitey, rzplta, rzpltą,*

II: *rzeczyposp, rzeczypospolitey, rzeczposp, rzeczpospolita*

III *rzptey, rzpta, rzeczypospolitey, rzepzey, rzpltej*

IV: *rzeczypospolitey: rzepltej, rzplita, rzplitą, rzplite, rzplitej, rzplitych, rzplta, rzpltą, rzpltej*

Na czele słów, które obecne są w każdym podokresie występują – obok *Rzeczpospolitej* – *pan, wolność (wolności), prawo (prawa)*, a także *sejm* (na liście I i IV *sejm*, na II i III *sejm*⁷). Na trzech listach pojawiają się także takie wyrazy, jak: *dwór* (II, III, IV), *stan* (I, II, IV), *confoederacya, konfederacja* (II, III, IV), *elekcyja* (II, III, IV). Na dwóch listach obecne są następujące: *duchowni* (I, IV), *ludzie* (I, IV), *miara* (I, III), *vspokoienie* (II, III), *instructiā, instrukcyj* (II, III), *szlachta* (II, IV), *ślachcic, szláchćic*, (II, III), *interes, interest* (III, IV), *veto* (III, IV), *wakanse* (III, IV), *rwanie* (III, IV), *porządek* (I, III), *opresya, oprresya* (I, II), *seymik, sejmik* (II, IV), *czas* (I, II). Należy stwierdzić częstą obecność wszelkiego rodzaju form adresatywnych i określających osobę, np. *kjm, jkm* (o królu), do innych osób: *wm, wmcj* (I), *mość, ichmość, mci* (II).

Warto przyjrzeć się też słowom łacińskim (które nie są tu przedmiotem analizy), ponieważ mają one znaczny udział w liście słów kluczowych⁸. Procentowy udział słownictwa łacińskiego w poszczególnych podzbiorach może być elementem analizy leksyki dawnej. Wymaga to osobnej uwagi (rozstrzygnięcie, czy mamy do czynienia z wyrazem polskim czy łacińskim utrudnia niekiedy sama pisownia, np. *republica*), wnosi jednak wiele cennych informacji do obrazu świata, w tym wypadku ograniczonego do kręgu polityki. Warto zwrócić uwagę, że niektóre znaczenia są synonimiczne względem polskich, np. obok *przykład*

⁶ Słowo *ojczyzna (oyczyzna)* pojawia się łącznie 5 razy, z czego 4 w II okresie. *Polska* występuje trzykrotnie, w każdym oprócz okresu III. Prawdopodobnie do zestawu nazw można by także dołączyć część kontekstów słów *rzecz i kraj*, a także łacińskiego (?) *republica*.

⁷ Zróźnicowanie graficzne form może być jednak częściowo wynikiem uwspółcześnionych wydań tekstów zawartych w *Korbie*.

⁸ Dla okresu I są to m.in.: *patria, privatim, publice, poena, officia, regni, aequalitatem, articulatum, ducum, nobilitatis, republica, luxus, vetustae, dignitatem, auctoritatem, libertatis, iura, pro, honoris, publicis, comites, baronum, peregrina, genere, principes, aequalitate, saeculares, iure, absolutum, ratione, titulorum, dominium, leges, res*, dla II *crimen perduellionis probationes, publice, majestatis, obiecta*, dla III *activitas, activitate, activitatem, aliquo, amovibiles, anni, auctoritate, aulae, bené, casus, dispensari, exemplo, fato, illimitatam, ius, leges, lex, libertatis, liberum, maiestatis, medium, mens, mentem, merita, ministeria, necessitate, necessitatis, paritas, passivitate, plena, praeiudicata, rationem, reipublicae, vim, vigore*, dla IV *clima, vers, vetandi, vicissim, prolis*.

pojawia się *exemplo*, co w pewnym sensie wzmacnia „kluczowość” słów w języku polskim.

Jeżeli chodzi o słowa znajdujące się na drugim biegunie rozważań, czyli charakterystyczne tylko dla konkretnego podkorpusu, to na liście słów ważniejszych niż inne znalazły się następujące rzeczowniki (ułożono je w kolejności alfabetycznej, tak by frekwencja poszczególnych form paradygmatu nie sugerowała rangi całego leksemu):

I *egzekucye, familie, gotowość, huffy, kara, konstytucyą, korona, królestwo, kule, majętności, nieprzyjaciel, obcych, osób, pokój, potomstwu, praktyki, prochow, przodkowie, rakuszany, rokosz, rokoszanie, rzeczy, strony, Szwecya, środki, tumulty, tytuły, Wiślica, Wiśliczanie, xiqżęta,*

II *całość, commissia, condicya, confidentá, Czárniecki, decret, deputáci, dobrodziei, fałsz, fortuna, instygator, izbá, kanclerz, kozak, krolowa., ksiądz, księstwo, list, łaska, máiestat, mandat, manifest, niewinność, obiásnienie, poddani, poselstwa, primáss, przyjaciele, punct, sąd, script, sługá, sprawiedliwość, successor, świadectwo, świadkowie, tatarowie, thronu, vprzeyność, vsługi, Warszawa, woysko, zawziętość, złośliwość, związek,*

III *alternata, Boze, buława, dyskurs, hetman, instancye, król, okoliczności, Poniatowski, pretensye, prywaty, sąsiad, seymowanie, traktat, vota, województwo, zaganienie, ziemianin,*

IV *anarchia, Cicero, dzieci, fakcyjne, glosy, konserwacya, kraj, obrady, paragraf, patryjota, podłość, racyje, rada, rodzice, rząd, słowa, sposób, utrzymywanie, wsie, zdania, zrywania.*

W drugiej próbie, jaką przeprowadzono, aby porównać wyniki, nie ujednoczono rozmiarów podkorpusów, tzn. wydobyto słowa klucze ze wszystkich tekstów zawartych w *Korbie*, czyli 70 (zamiast 13) dla okresu I, 2 dla II, 10 (zamiast 5) dla III i 2 dla IV. Nie miało to zasadniczego wpływu na katalog słów kluczy, warto jednak przytoczyć te nieodnotowane wcześniej. Dla pierwszego okresu są to: *artykuły, Stężycza, bracia, sprawy, król, wojewoda, Janowiec, deputaty, senat, małżeństwo, swobody, zjazdy, deklaracja*; dla III zaś *szlachta, sejmik, podskarbi, Polacy, Piast, cudzoziemiec*.

W próbie trzeciej zdjęte zostało ograniczenie w postaci wyboru konkretnych gatunków tekstów na rzecz wyłącznie tych o tematyce *polityka*. Mimo że ich liczebność była znacznie większa w porównaniu z pierwszym eksperymentem – uwzględniono 71 utworów dla okresu I i 9 dla III, tylko w tym ostatnim pojawiły

się nieobecne w pierwszej i drugiej próbie słowa: *status, minister, affekt, respons, applauz, skarb*.

Zgodnie z przewidywaniami zakres słownictwa dotyczy zasadniczo spraw politycznych i społecznych dawnej Rzeczypospolitej. Wyodrębnić można tu krąg leksyki prawnej, sądowej, związanej z obradami sejmiku, takich jak *konstytucja, prawo, prawa, puncta, paragraf, dekret, deputat, niewinność instrukcje, głosy, veto*. W badanym materiale pojawiło się też wiele nazw funkcji i stanowisk – *król, marszałek, hetman, senator, poseł, prymas, podskarbi*. W analizach kultury politycznej warto wydobyć również takie słowa, jak *anarchia, rokosz, tumulty, uspokojenie*⁹. Jeżeli chodzi o nazwy uczuć czy wartości, oprócz wspomnianej *wolności*, na liście nr III pojawia się (dwukrotnie) *uprzejmość*, na II *złośliwość*, na IV *podłość*. Kolejne przeprowadzane próby pokazały, że udział słownictwa wcześniej nieobecnego nie jest znaczący, pomaga jednak znaleźć człony dla wyłaniających się par, takich jak *rodzice – dzieci, Polacy – cudzoziemcy*. Warto postawić tu pytanie o zasadność wyciągania daleko idących wniosków z obecności słownictwa w badanych podokresach. Z pewnością zaistnienie takich „słów ważniejszych” jak *Szwecja, Wiślica, wiśliczanie, Tatarowie, huffy, Kozacy* czy nazwisk Czarnieckiego i Poniatowskiego jest nieprzypadkowe (por. rokosz Zembrzydowski, który miał miejsce w 1606 r. w Wiślicy czy wojnę ze Szwecją o Inflanty w latach 1600–1611). Z drugiej strony wyciąganie wniosków z faktu, że Szwecja pojawiła się w I okresie (a nie np. w II, podczas potopu szwedzkiego) mogłoby być przedwczesne. Interpretacja tych faktów stwarza miejsce do współpracy z historykami.

Grupa utworów z zakresu polityki sprzyja uchwyceniu słów specyficznych. Leksyka związana z obradami sejmiku i sejmików i omawianych tam kwestii politycznych jest stosunkowo stała. Ponadto to właśnie teksty o tej tematyce obecne są we wszystkich przedziałach czasowych. Inne – medyczne, muzyczne – nie miały takiej reprezentacji. W przebadanych, dotyczących polemik religijnych czy obyczajowości dominowały konkretne, poruszone tam kwestie, niekiedy bardzo szczegółowe, por. *Dyskursu o pijanństwie kontynuacja A Gdacjusza, Gorzka wolność młodzieńska* Andrzeja Żydowskiego, *Niebo ziemskie aniołów w ciele* (tekst anonimowy), *Spizarnia aktów rozmaitych przy zalotach, weselach, bankietach, pogrzebach* (tekst anonimowy), *Wodka albo gorzałka* (tekst anonimowy), *Barwiczka dla ozdoby twarzy panińskiej* Radopatrzeka Gładkotwarskiego. Obecność niektórych słów może tu wynikać ze specyfiki tekstu lub idiolektu autora. Przy takiej próbie, jaką dysponujemy, pochopte uogólnianie może wpływać na

⁹ Anonimowy Recenzent tego artykułu podpowiada w tym miejscu ciekawą obserwację następującej treści: „warto zaobserwować, że zakres znaczeniowy jest do dzisiejszych kulturomów podobny, zaś formy wyrażania są inne: *dziś pucz/chaos/balagan...*”.

zniekształcenie obrazu przeszłości. Przy większym zbiorze tekstów wymienione tu cechy jednostkowe przestałyby mieć znaczenie.

6. Zakończenie

Zaprezentowane tu analizy miały charakter pilotażowy, wskazały jednak na zasadność dalszych prób podejmowanych we współpracy z historykami obyczajowości, polityki, religii. Słowo może stać się kluczem do rozumienia historii i kultury, pomóc w zrozumieniu rzeczywistości pozajęzykowej. Być może, przy wszystkich ograniczeniach, będzie można z czasem odnaleźć we współczesności siedemnasto- i osiemnastowiecznej słowa równie żywo obecne, jak dzisiejsze *islamista*, *uchodźca* czy *pięćset plus*. Pozostaje mieć nadzieję, że tego rodzaju analizy będą przeprowadzane na coraz większą skalę, a ich wyniki przyczynią się do rozwoju badań zarówno językoznawczych, jak i historioznawczych.

Literatura

- Adamiec, Dorota (2015): Kryteria doboru tekstów do „Elektronicznego korpusu tekstów polskich z XVII i XVIII w. (do 1772 r.)”. W: *Prace Filologiczne*. Tom LXVII, s. 11–20.
- Bartmiński, Jerzy (2016): Słowa klucze, kulturemy, koncepty kulturowe. W: *Przegląd Humanistyczny* 3, s. 21–29.
- Bronikowska, Renata (2015): Możliwości przeszukiwania korpusu barokowego – cele i założenia. W: *Prace Filologiczne* LXVII, s. 45–56.
- Eder, Maciej (2016): Słowa znaczące, słowa kluczowe, słowozbiory – o statystycznych metodach wyszukiwania wyrazów. W: *Przegląd Humanistyczny* 3, s. 31–44.
- Gruszczyński, Włodzimierz/ Adamiec, Dorota/ Ogrodniczuk, Maciej (2014): Elektroniczny korpus tekstów polskich z XVII i XVIII w. (do 1772 r.) – prezentacja projektu badawczego. W: *Polonica* XXXIII, s. 311–318.
- Łaziński, Marek (2016): Słowa klucze – do czego? W: *Przegląd Humanistyczny* 3, s. 75–94.
- Majdak, Magdalena (2016): Słowa klucze w materiale historycznym – wyzwania i ograniczenia. W: *Przegląd Humanistyczny* 3, s. 45–55.
- Nagórko, Alicja (2004): Metody kontrastywne a etnolingwistyka (lingwistyka kulturowa). W: Fórián Evá (ed.): *Multikulturowość, tożsamość narodowa, mniejszości na Węgrzech i w Polsce*. Debreczyn, s. 23–33.
- Pawłowski, Adam (2016): Korpusy chronologiczne i leksykalne szeregów czasowe jako narzędzia wykrywania słów kluczy i neosemantyzmów. Konceptualizacja czasu w korpusach chronologicznych. W: *Przegląd Humanistyczny* 3, s. 57–74.
- Pisarek, Walery (2002): *Polskie słowa sztandarowe i ich publiczność*. Kraków.
- Pisarek, Walery (2016): Słowa ważne i ważniejsze. W: *Przegląd Humanistyczny* 3, s. 11–20.
- Rak, Maciej (2015): Co to jest kulturem? W: *LingVaria* X, s. 305–316.
- Sambor, Jadwiga (1969): *Badania statystyczne nad słownictwem (na materiale „Pana Tadeusza”)*. Wrocław.

- Stachurski, Edward (1989): Słownictwo w utworach polskich naturalistów: badania statystyczne. Kraków.
- Stachurski, Edward (1998): Słowa-klucze polskiej epiki romantycznej. Kraków.
- Stachurski, Edward (2005): Słownictwo „Pana Tadeusza” Adama Mickiewicza na tle tekstów innych poematów romantycznych. Kraków.
- Wierzbicka, Anna (2007): Słowa kluczowe. Różne języki – różne kultury. Tłum.: Izabela Duraj-Nowosielska. Warszawa.

dr Magdalena Majdak
Instytut Języka Polskiego Polskiej Akademii Nauk
al. Mickiewicza 31, 31-120 Kraków
e-mail: magdalena.majdak@gmail.com