

MICHAŁ SZCZYSZEK

Uniwersytet im. Adama Mickiewicza w Poznaniu

ORCID: 0000-0002-0253-7296

e-mail: szczysze@amu.edu.pl

O PROJEKCIE SŁOWNIKA POLSKIEGO PARLAMENTARYZMU
XX WIEKU (LATA 1918-2018):
ETAP WSTĘPNY – KORPUS PARLAMENTARNY

Wstęp

Trwające od kilkunastu miesięcy w Zespole Inżynierii Lingwistycznej Instytutu Podstaw Informatyki Polskiej Akademii Nauk prace nad korpusem lingwistycznym zawierającym dane językowe wyabstrahowane ze stenogramów polskiego parlamentu (Sejmu i Senatu) z lat 1918-2018 zaowocowały powstaniem już właściwie ukończonego korpusu (o danych liczbowych wspomnianego korpusu informuje Tabela 1.). Na tej podstawie zostały podjęte już wstępne językoznawcze prace analityczne nad tak zgromadzonym, niezwykle obszernym materiałem językowym. Podejmowane są też w zespole starania związane z powstaniem na bazie tych materiałów słownika polskiego parlamentaryzmu obejmującego w zamierzeniach okres właśnie ostatnich stu lat – od odzyskania przez Polskę niepodległości do dziś.

W niniejszym przyczynku omówione zostaną zatem wstępne wyniki analiz językoznawczych wykorzystujących do opracowania narzędzia cyfrowe stworzone w IPI PAN, jak i przez konsorcjum CLARIN¹. (Skupiłem się na potrzeby niniejszego tekstu na przedstawieniu danych statystycznych dotyczących skończonego korpusu stenogramów za lata 1918-2018: Korpus Dyskursu Parlamentarnego (<http://clip.ipipan.waw.pl/PSC>, http://sejm.nlp.ipipan.waw.pl/query_corpus/; Ogrodniczuk 2012: 2219-2223, Ogrodniczuk 2018: 15-19).

Dane statystyczne obrazujące bogactwo materiałowe korpusu znajdują się w Tabeli 1.

Z tabeli powyższej wynika przede wszystkim to, że łączna liczba segmentów (jednostek językowych, słowoform) pochodzących zarówno ze stenogramów polskiego Sejmu, jak i stenogramów polskiego Senatu z lat 1918-2018 wynosi

¹ Pierwsze wstępne wyniki zostały już opisane w artykule Michała Szczyszka (zob. Bibliografia).

Tabela 1. Korpus stenogramów parlamentarnych 1918-2018

Okres	Lata	Sejm	Posiedzenia	Segmenty	Senat	Posiedzenia	Segmenty	
II RP	1919-1922	Sejm	342	208 524	–	–	–	
	1922-1927	1. kadencja	340	533 109	1. kadencja	157	153 980	
	1928-1930	2. kadencja	86	111 894	2. kadencja	31	5 790	
	1930-1935	3. kadencja	148	139 166	3. kadencja	80	123 651	
	1935-1938	4. kadencja	90	141 387	4. kadencja	54	56 191	
1938-1939	5. kadencja	31	44 592	5. kadencja	20	38 776		
1943-1947	Państwowa Rada Narodowa	11	11 671	–	–	–		
PRL	1947-1952	Sejm	108	514 572	–	–	–	
	1952-1956	1. kadencja	39	76 244	–	–	–	
	1957-1961	2. kadencja	59	115 563	–	–	–	
	1961-1965	3. kadencja	32	62 799	–	–	–	
	1965-1969	4. kadencja	23	45 233	–	–	–	
	1969-1972	5. kadencja	19	33 492	–	–	–	
	1972-1976	6. kadencja	32	63 155	–	–	–	
	1976-1980	7. kadencja	29	57 352	–	–	–	
	1980-1985	8. kadencja	79	132 845	–	–	–	
	1985-1989	9. kadencja	50	89 436	–	–	–	
	1989-1991	10. kadencja	79	157 225	1. kadencja	61	111 450	
	III RP	1991-1993	1. kadencja	45	7 803 935	2. kadencja	38	1 461 165
		1993-1997	2. kadencja	115	22 299 861	3. kadencja	102	5 057 468
		1997-2001	3. kadencja	119	24 313 939	4. kadencja	90	8 261 548
		2001-2005	4. kadencja	109	28 986 555	5. kadencja	88	6 489 812
		2005-2007	5. kadencja	48	11 833 471	6. kadencja	39	3 573 955
2007-2011		6. kadencja	100	22 682 341	7. kadencja	83	8 827 024	
2011-2015		7. kadencja	102	22 587 764	8. kadencja	82	7 110 114	
2015-		8. kadencja	54	5 905 461	9. kadencja	53	3 504 637	
SUMA:			148 951 586		44 775 561	193 727 147		

193 727 137 (w tym wyekscerpowanych ze stenogramów Sejmu: 148 951 586, a Senatu: 44 775 561). Dane liczbowe dotyczące poszczególnych okresów historii Polski (II RP, PRL III RP) oraz dane za poszczególne kadencje parlamentarne znajdują się w kolumnach „Segmenty). W kolumnach „Posiedzenia” znajdują się dane dotyczące liczby posiedzeń parlamentarnych (Sejmu lub Senatu), a więc de facto – liczby stenogramów sejmowych pochodzących z danego okresu, z danej kadencji.

Narzędzia

Na potrzeby analizy tak wielkiego korpusu konieczne było skorzystanie z narzędzi cyfrowych (cyfrowej humanistyki). Zatem zarówno do wstępnego opracowania materiału językowego na potrzeby niniejszego przyczynku, jak i na potrzeby opracowywania całościowego tak bogatego materiału wykorzystywane były (i będą) narzędzia takie jak *Korpusomat* (<http://korpusomat.pl/>; Kieraś, Kobyliński, Ogrodniczuk 2018: 21-27) opracowany przez zespół Instytutu Podstaw Informatyki PAN oraz *Analizator wydziwisku* opracowany przez konsorcjum CLARIN-PL (<http://ws.clarin-pl.eu/sentymen.shtml>; Walkowiak 2017: 464-473; Walkowiak, Pol 2017: 167-174; Walkowiak, Piasecki 2015: 61-63; Piasecki 2014)².

Analizy

W dalszej części przyczynku omówione zostaną wstępnie dwa zagadnienia wynikające z przeprowadzonych dotychczas (nadal wstępnych, początkowych) analiz materiału językowego pochodzącego ze stenogramów parlamentarnych: 1) analiza losowo wybranych jednostek leksykalnych funkcjonujących w postaci wyrazów tekstowych w korpusie stenogramów; 2) analiza językowych sposobów wyrażania ekspresji pojawiających się w stenogramach. Poprzez te dwa aspekty można pokazać zarówno złożoność i bogactwo materiału językowego pochodzącego z obrad polskiego parlamentu, jak i ogromną potencję badawczo-naukową, która tkwi w przyjętych za podstawę materiałową niniejszych badań (i niniejszego przyczynku) tekstach wystąpień sejmowych i senackich.

Analiza losowo wybranych jednostek leksykalnych funkcjonujących w korpusie stenogramów

Poniżej zostaną przedstawione wyniki pracy Korpusomatu (na całym korpusie liczącym ponad 193 mln jednostek) – w postaci danych statystycznych dotyczących analizowanych szczegółowo na potrzeby niniejszego przyczynku leksemów wyekscerpowanych ze stenogramów, a także danych dotyczących semantyki tych leksemów i jej potencjalnych przemian zachodzących w stenogramach parlamentarnych badanego okresu. Badanie to przeprowadzono w przekrojach: każdy przekrój obejmował poszczególne dziesięciolecia XX wieku (tj. badano stenogramy pochodzące z lat 20., 30., 40., 50., 60., 70., 80. i 90. XX wieku). Na tej podstawie i na

² Informacje te zostały już opisane w artykule Szczyszek.

podstawie kontekstów odnalezionych przy użyciu Korpusomatu zostaną przeprowadzone analizy semantyczne dwóch leksemów: *niezawisłość*, *niepodległość*.

Niezawisłość

W słownikach wyraz ten jest definiowany następująco:

WSJP: '1. kraju: cecha czegoś, co nie pozostaje w zależności od kogoś lub czegoś; 2. poglądów: cecha czegoś, co jest właściwe ludziom, którzy oceniają rzeczy samodzielnie i nie ulegają wpływom ani naciskom innych ludzi; 3. o sądach, sędziach: niepodleganie naciskom administracyjnym i kierowanie się wyłącznie zasadami prawa'.

SJPDor. 'niezależność, niepodległość, samodzielność'.

Na podstawie analiz użycia tego leksemu w stenogramach parlamentarnych badanego okresu widać wyraźnie, że leksem ten najczęściej występuje w znaczeniu słownikowym podawanym zarówno w SJPDor, jak i w WSJP. Zauważyć należy, że według precyzyjniej wyodrębniającego znaczenia definicyjne WSJP najczęściej leksem ten jest używany w znaczeniu 3. (o sądach) oraz – choć nieco rzadziej – w znaczeniu 1. (w odniesieniu do kraju, państwa). Ponadto – w okresie II RP leksem dotyczy także szkolnictwa.

W badanym materiale dominuje neutralne użycie tego leksemu, choć zdarzają się konteksty pozytywne, z szacunkiem; nie odnotowano użyc o charakterze negatywnym.

W badanych okresie leksem *niezawisłość* został łącznie użyty 603 razy. W II RP leksem ten był często używany, natomiast w PRL liczba jego użyc zmalała o połowę. W III RP leksem ten był bardzo ważny, ponieważ jego użycia w tym okresie są największe. Zarówno w II jak i III RP omawiany leksem dotyczy nie tylko Polski, ale i innych państw m.in. Cypru i Łotwy. W III RP pojawiają się nowe konteksty – leksem *niezawisłość* dotyczy mediów, pracowników. Pojawia się też w latach 90. XX wieku zwrot *niezawisłość zawodowa*. W ten oto sposób parlamentarzyści (parlamentaryzm polski) profilują (profiluje) w dyskursie społeczno-politycznym w sejmie (i w senacie) analizowane tu pojęcie. Pomimo tego, w okresie III RP najczęściej leksem dotyczy sądów i trybunałów. Poniżej dane statystycznie i przykłady użycia wyekscerpowane z korpusu stenogramów za pomocą narzędzia Korpusomatu.

Dodać można, że na podstawie statystyk, które poniżej, można by sformułować wniosek, że w okresie PRL pojęcie *niezawisłość* nie było zbyt istotne dla ówczesnych parlamentarzystów.

a. II RP – łącznie 217 użyc leksemu *niezawisłość*

1918-1919: 16 użyc leksemu *niezawisłość*

„[...] troska to, ażeby i nazewnątrz *niezawisłość* wyborów była dostatecznie silnie zaznaczona [...]” – *neut.* znaczenie słownikowe; „[...] wskazywały, że bezstronność i *niezawisłość* sądów przysięgłych jest zagrożona.” – *neut.* znaczenie słownikowe; 3. „[...]że niepodległość państwa warunkuje się *niezawisłością* i niepodległością ekonomiczną, ze [...]” – *neut.* znaczenie słownikowe; „[...] może, zostać niezawisłym. *Niezawisłość* nie polega na frazesach, [...]” – podniośle.

1920-1929: 134 użycia leksemu *niezawistość*

„[...] Panom chodzi właśnie o podkopanie *niezawistości* sądownictwa i to znalazło swój [...]” – neutralnie; „[...] ten akt, to koniec *niezawistości* Niemiec.” – neutralnie; „[...]że wyrazem niepodległości, *niezawistości* Państwa Polskiego, wyrazem tego [...]” – neutralnie; „[...]żeby one miały istotnie Znamię *niezawistości* [...]” – podniosłe; „[...]Rząd obecny w kwestji utrzymania *niezawistości* szkolnictwa poszedł tak daleko [...]” – neutralnie.

1930-1939 – 52 użycie leksemu *niezawistość*

„[...] to najistotniejszy element naszej *niezawistości* gospodarczej i politycznej.” – neutralnie; „[...] lub przenoszenia sędziego, zasada *niezawistości* sędziowskiej polega na stworzeniu takich [...]” – neutralnie; „[...] będziemy ze wszystkich sił bronić *niezawistości* szkół, wychowania narodowego [...]” – neutralnie; „[...] które tworzą siły obronności i *niezawistości* Państwa. Artystów w Polsce [...]” – neutralnie.

b. PRL – łącznie 99 użycie leksemu *niezawistość*

1950-1959 – 22 użycia leksemu *niezawistość*

„[...] zwana doktryna Eisenhowera stanowi zagrożenie *niezawistości* krajów arabskich i nosi w [...]” – neutralnie; „[...] za godność, wolność i *niezawistość* Polski płacili cenę krwi i [...]” – podniosłe; „[...] to potężna ostoja wolności i *niezawistości* naszego narodu.” – podniosłe; „[...] wolności ludu, całości i *niezawistości* Polskiej Rzeczypospolitej Ludowej.” – neutralnie; „[...] odcinku wymiaru sprawiedliwości, gdzie *niezawistość* sędziów i podporządkowanie ich tylko [...]” – neutralnie.

1960-1969 – 17 użycie leksemu *niezawistość*

„[...] w ustrój, bezpieczeństwo i *niezawistość* Polskiej Rzeczypospolitej Ludowej.” – neutralnie; „[...] w ustawie, mówiący o *niezawistości* ławnika, znalazł się na [...]” – neutralnie; „[...] wielokroć krwią serdeczną broniły *niezawistości* naszej ojczyzny. Wybraliśmy [...]” – neutralnie; „[...] jednolitością orzecznictwa i za sprzyjaniem *niezawistości* sędziów, ale postulowały również [...]” – neutralnie.

1970-1979 – 6 użycie leksemu *niezawistość*

„[...] te, którym nie odpowiada *niezawistość* Cypru. [...]” – neutralnie; „[...] użycia; wzajemne poszanowanie *niezawistości* i suwerenności; wyrzeczenie się [...]” – neutralnie; „*Niezawistość* sędziów sądów wojskowych znajduje wyraz [...]” – neutralnie;

1980-1989 – 54 użycia leksemu *niezawistość*

„[...] pracy sędziów i umacniania ich *niezawistości* jest przewidywane rozszerzenie kompetencji organów [...]” – neutralnie; „Z punktu widzenia gwarancji *niezawistości* sędziowskiej, istotne znaczenie mają [...]” – neutralnie; „Umożliwiła to *niezawistość* państwa i rozwijająca się demokracja [...]” – podniosłe.

1990-1999 – 287 użyć leksemu *niezawisłość*

„[...] KRRiT, a w sumie *niezawisłości* mediów, w kontekście przyjęcia [...]” – neutralnie; „Jest to zbyt słabe zabezpieczenie *niezawisłości* pracownika medycyny pracy wobec zleceniodawcy [...]” – neutralnie; „[...] ochrony grup szczególnego ryzyka; *niezawisłości* zawodowej osób wykonujących zadania służby [...]” – neutralnie; „[...] kontroli, które nie naruszają *niezawisłości* służby medycyny pracy.” – neutralnie; „[...] zawarty jest rozdział mówiący o *niezawisłości* sądów [...]” – neutralnie.

Niepodległość

W słownikach leksem ten jest definiowany następująco:

WSJP: ‘1. państwa: cecha państwa lub narodu, które nie zależą pod względem politycznym od innego państwa lub narodu’; 2. myśli: niezależność od zewnętrznych czynników i okoliczności’.

SJPDor. ‘1. Niezależność jednego państwa (narodu) od innych państw w sprawach wewnętrznych i stosunkach zewnętrznych; niezawisłość, suwerenność, wolność; 2. przestarzałe: swoboda działania, rozporządzania sobą, niezależność osobista, samodzielność’.

Na podstawie analiz użycia tego leksemu w stenogramach parlamentarnych badanego okresu widać wyraźnie, że leksem ten najczęściej występuje w znaczeniu słownikowym podawanym zarówno w SJPDor, jak i w WSJP. Zauważyć należy, że z obu przywołanych tu słowników leksem ten najczęściej jest używany w znaczeniu 1. (w odniesieniu do kraju, państwa) – zarówno wg SJPDor., jak i WSJP. Leksem *niepodległość* odnosi się zatem do takich pojęć, jak *tożsamość narodowa*, *patriotyzm*.

W badanym materiale dominuje neutralne użycie tego leksemu, choć zdarzają się konteksty pozytywne, z szacunkiem; nie odnotowano użyć o charakterze negatywnym. Jest to zatem podstawowy profil, w jakim występuje to pojęcie w dyskursie parlamentarnym w polskim sejmie i senacie.

W badanym okresie XX wieku wyraz ten w stenogramach wystąpił łącznie 2189. Na podstawie statystyk, które poniżej, można wysnuć wniosek, że w okresie PRL pojęcie *niepodległość* nie było istotne – podobnie jak wyżej omówione pojęcie *niezawisłość* – dla ówczesnych parlamentarzystów.

W II RP –1084 użyć leksemu *niepodległość*; leksem ten najczęściej występuje w postaci neutralnej, ale zdarzają się konteksty pozytywne, z szacunkiem; nie odnotowano użyć negatywnych.

W okresie PRL – użyto 266 razy leksem *niepodległość*, np.: 1950-1959: „[...] tego Sejmu w dziele umacniania *niepodległości* naszej Ludowej Ojczyzny, dalszego [...]” – neutralnie; „[...] wzmacnianie międzynarodowych przesłanek *niepodległości* i suwerenności narodowej, nienaruszalności [...]” – neutralnie; „[...] socjalistycznej i praworządności, umacniania *niepodległości* kraju. Zakaz broni masowego [...]”.

W III RP – 839 użyć leksemu *niepodległość*, np.: 1990-1999: „[...] Radzieckim, w republikach uzyskujących *niepodległość* i w Jugosławii. [...]” – neutralnie; „[...] Litwę, Białoruś i Ukrainę *niepodległości* na drodze pokojowej. Wiemy [...]” – neu-

tralnie; „[...] Uchwała pierwsza dotyczy proklamowania *niepodległości* przez Łotwę i Estonię oraz [...]” – neutralnie.

Analiza językowych sposobów wyrażania ekspresji pojawiających się w stenogramach

Wyniki tu przedstawione zostały uzyskane za pomocą narzędzia *Analizatora wydźwięku* – dzięki któremu możliwe było zbadanie obszernej grupy tekstów (ale jeszcze nie wszystkich) wchodzących w skład Korpusu Dyskursu Parlamentarnego. Stenogramy (losowo dobrane) przebadano (podobnie jak powyżej leksemy; powtórzmy więc tu tylko tę samą informację) w przekrojach – każdy przekrój obejmował poszczególne dziesięciolecia XX wieku (tj. badano stenogramy pochodzące z lat 20., 30., 40., 50., 60., 70., 80. i 90. XX wieku; z każdego wskazanego tu dziesięciolecia wybrano losowo po 10 stenogramów – na potrzeby tych analiz stworzono mniejszy korpus liczący 3 334 054 jednostek (segmentów, słowoform)). *Analizator wydźwięku* pozwolił oszacować poziom nasycenia emocjami (wyrazami ekspresywnymi) stenogramów. Mimo że nie wszystkie stenogramy zostały poddane badaniu, to wybrana tu próbka może stanowić już podstawę do orzekania o tendencjach tkwiących w poszczególnych dziesięcioleciach XX funkcjonowania polskiego parlamentu.

Dane dotyczące nasycenia ekspresywizmami stenogramów z poszczególnych 10-leci XX wieku przedstawia Tabela 2. Zawiera ona informacje w ujęciu procentowym pokazujące stosunek liczbowy jednostek leksykalnych określonych a pomocą narzędzia *Analizator wydźwięku* jako ekspresywizmy funkcjonujące w wybranych tu stenogramach z poszczególnych 10-leci do ogólnej liczby jednostek leksykalnych wyliczonych dzięki narzędziu Korpusomat w wybranych tu stenogramach z poszczególnych 10-leci³. Są to dane pokazujące zbiorczo udział emocji pozytywnych, jak i negatywnych ujawniających się w tych tekstach

Tabela 2. Dane procentowe pokazujące stosunek ekspresywizmów do ogólnej liczby jednostek leksykalnych w danym dziesięcioleciu XX wieku

dziesięciolecia XX wieku	procentowy udział ekspresywizmów w stenogramach z poszczególnych dziesięcioleci XX wieku
lata 20.	21,57%
lata 30.	24,75%
lata 40.	28,26%
lata 50.	27,78%
lata 60.	23,93%
lata 70.	25,12%
lata 80.	24,70%
lata 90.	20,18%
średnia:	24,32%

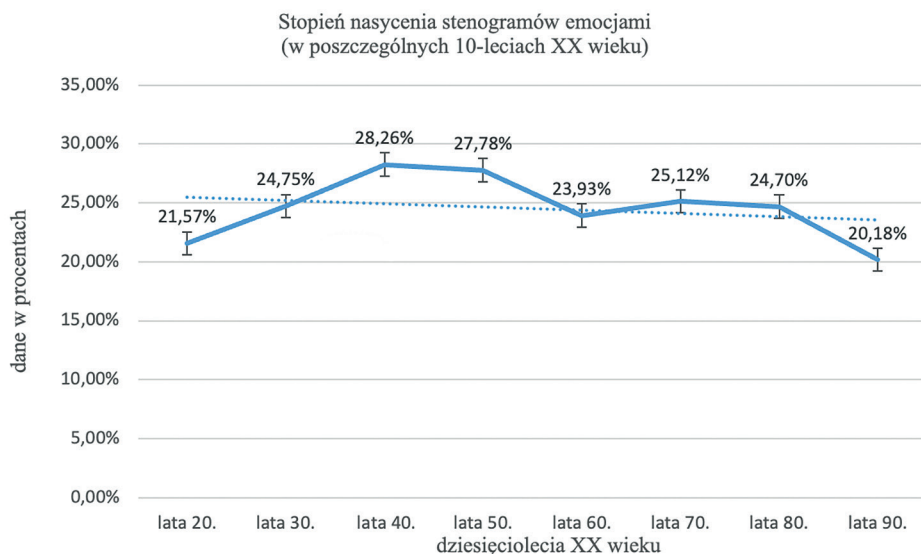
³ Pełna informacja o danych statystycznych, na podstawie których opracowano tę część artykułu, zawiera tabela znajdująca się na końcu artykułu, w Aneksie.

Jak wynika z powyższej Tabeli 2. średnia nasycenia emocjami (ekspresywizmami) stenogramów parlamentarnych w XX wieku (na podstawie danych z losowo wybranych stenogramów) wynosi 24,32%.

Powyżej średniej sytuują się stenogramy z większości badanych okresów. Znacząco powyżej średniej są dane za lata 40. i 50. wieku; natomiast nieznacznie powyżej średniej sytuują się dane z lat 30., 70. i 80. XX wieku.

Poniżej średniej sytuują się dane za lata 20., 60. oraz 90. XX wieku.

Niemniej na uwagę zasługuje fakt (dostrzegalny już i w Tabeli 2.), że od połowy XX wieku poziom nasycenia ekspresywizmami (emocjami) stenogramów polskiego parlamentu systematycznie maleje. Wyraźnie widać to na poniższym Wykresie 1., gdzie linia trendu (na wykresie: linia przerywana) ma charakter opadający.



Wykres 1.

Należy tu dodać, że dane za lata 90. XX wieku pochodzą z wyższej izby polskiego parlamentu – z senatu. Taki dobór danych miał na celu skonfrontowanie danych sejmowych z danymi senackimi. Działanie to ukazało, że senat, uznawany za izbę refleksji, także nasycony jest emocjami, w stopniu nie małym (nieco ponad 20%) – chociaż dane te sytuują się znacząco poniżej średniej z XX wieku. Oczywiście – zdaję sobie sprawę, że przytoczenie tu danych senackich (niemalże „pojedynczych”) może wypaczać, a nawet wypacza ogólny obraz danych uzyskiwanych tu głównie na podstawie danych sejmowych. Jednakże trzeba pamiętać, że to jest w jakimś sensie pokłosie historii polskiego parlamentaryzmu w XX wieku, kiedy to przez blisko półwiecze senatu nie było. Ponadto – niniejszy tekst stanowi jeden z pierwszych w planowanym cyklu artykułów, które piszę w oparciu o tak zebrany

korpus, za pomocą tak dobranych narzędzi. Poruszane w tym akapicie zagadnienie, tj. porównania danych językowych pochodzących ze stenogramów sejmu z danymi językowymi pochodzącymi ze stenogramów senatu, wymaga oczywiście dalszych pogłębionych badań i będzie przedstawiane w opracowaniach z tego zakresu.

Na marginesie tych rozważań należałoby wspomnieć – na zasadach niemalże studium przypadku – o stenogramie 24_00018_000050413. Jest to stenogram z 18 posiedzenia Sejmu Polskiej Rzeczypospolitej Ludowej, które odbyło się w dniach 25 i 26 stycznia 1982 roku. Poziom nasycenia emocjami wynosi w nim 28,08%. Znacząco przekracza średnią z całego badanego okresu w XX wieku, jak i z średnią wyliczoną dla lat 80. XX wieku. Dodać należy jedynie, że jest to stenogram z pierwszego (albo jednego z pierwszych) posiedzenia sejmu po wprowadzeniu stanu wojennego w Polsce. Dominują (choć raczej nieznacznie) w nim emocje negatywne, które – w sumie – osiągnęły poziom 53,72% ogólnego stanu ekspresywizmów tego stenogramu (24_00018_000050413), przy poziomie emocji pozytywnych sięgających – w sumie – 46,28%. Na uwagę, nawet zaskoczenie, zasługują dane dotyczące poszczególnych emocji wyodrębnionych w tym stenogramie. Najliczniej występują, najczęściej się werbalizowały dwie pojedyncze emocje pozytywne, tj. najliczniej występuje emocja pozytywna Radość (3429 tego rodzaju ekspresywizmów, co stanowi 11,44% ogólnego stanu ekspresywizmów tego stenogramu) oraz Użyteczność (3540 ekspresywizmów; 11,82%). Pozostałe emocje występujące w tym stenogramie, to:

- emocje pozytywne: Zaufanie (2258 ekspresywizmów), Cieszenie się na coś oczekiwanego (589), Prawda (468), Wiedza (771), Piękno (381), Szczęście (2417);
- emocje negatywne: Smutek (2801), Złość (2533), Strach (962), Wstręt (812), Zaskoczenie czymś nieprzewidywanym (189), Nieużyteczność (2313), Krzywda (2327), Niewiedza (385), Błąd (1468), Brzydota (114), Nieszczęście (2188).

Przykładowe cytaty nacechowane emocjonalne z wybranych losowo stenogramów z XX wieku:

– z lat 20. XX:

[...] istotne nasze prawo. (*Oklaski wśród postów ukraińskich. Głos: Nie dawno byliście Rosjanami. Pan jest Rosjaninem*). Nie, jestem Ukraińiec z rodu, z dziada pradziada i kocham tak moją ojczyznę i moją narodowość, jak i Panowie waszą. (*P. Dębski: Czy śmiałybyś Pan w Dumie rosyjskiej po ukraińsku przemówić?*). Jesteśmy w państwie i w republice demokratycznej, a nie w ciemnym, rosyjskim państwie. My, przedstawiciele narodu ukraińskiego Wołynia, Chełmszczyzny (*Głos: Jakiej Chełmszczyzny? Wrzawa*), Podlasia i Polesia, obrani na zasadzie prawa ogólnego głosowania, wstępując po raz pierwszy do Sejmu Rzeczypospolitej Polskiej (*P. Dębski: Carską Rosję przypomina! Wrzawa*), uważamy za swój [...]

(10 posiedzenie Sejmu Rzeczypospolitej z dn. 25 stycznia 1923 r., s. 32.)

– z lat 30. XX:

[...] i może być krytykowany, (*Wrzawa na ławach B. B. i lewicy. Okrzyki na ławach B. B.: Napraszacie się na bicie w gębę! Głosy na ławach P. P. S.: Jak to z tą gębą? Głos na ławach B. B.: Nie będę mówił, tylko dam.*)

(SEJM RZECZYPOSPOLITEJ POLSKIEJ. OKRES III. Sesja zwyczajna r. 1932/33. Sprawozdanie Stenograficzne z 82 posiedzenia w dniu 8 lutego 1933 roku., s. 37)

– z lat 40. XX:

Tym gwałtowniej to oszczerstwo zwraca się przeciwko niemu.

Nie inny charakter Imają też pozostałe zarzuty pana Mikołajczyka pod adresem naszej partii, a w szczególności w związku z procesem wLesznie, który jest wynikiem atmosfery, do której niemało przyczynili się PSL-owcy, (Poseł Mikołajczyk: Dlaczego?) Dlatego z tej trybuny najostrzej piętnujemy oszczercze metody pana Mikołajczyka, które są wyrazem rozkładu i upadku jego partii. (Oklański. Głosy: Hańba. Różne głosy na ławach PSL. Głos na ławach PPR: Milczcie, łobuzy.)

Wicemarszałek Barcikowski: Może Panowie raczą odłożyć te rozmowy.

(Sejm Ustawodawczy Rzeczypospolitej Polskiej Sesja zwyczajna r. 1947. Sprawozdanie Stenograficzne z 23 posiedzenia Sejmu Ustawodawczego w dniu 27 czerwca 1947 r., s. 2)

– z lat 50. XX:

Daliśmy wyraz naszym dążeniom do niepodległości za wszelką cenę, do suwerenności politycznej i dlatego dobrze będzie, jeśli w tej sali, w której siedzi ta Partia, której VIII Plenum odbywało się w październiku, nie będzie się powtarzało kiepskich żartów. Głosować będziemy nie z konieczności, bo nie ma innego preliminarza budżetowego i nie ma innych form, tak jak o tym mówił poseł Kisielewski, ale dlatego, że w tej sytuacji, w jakiej się znajdujemy, jest to jedyny budżet, którego złe strony znamy, i że jest naszym obowiązkiem świętym ten budżet – jeśli jest złym, niedoskonałym – doprowadzić do doskonałości. Na co mamy czekać? Na 3 miliardy dolarów pożyczki, żeby stać się kolonią kapitalistycznych Stanów Zjednoczonych? Za darmo nie daje się takich pieniędzy żadnemu państwu. My już dwa miesiące walczyliśmy o pożyczkę dużo niniejszą w Stanach Zjednoczonych, ale Stany Zjednoczone powinny wiedzieć, że jeżeli jej nie dadzą, to Polska dążąca do socjalizmu szczerze, nie tylko w słowach, musieć sobie sarna radę, i da sobie radę, bo za tą Polską stoją elementy i czynniki, które przeżywały już nieraz ciężkie czasy, a nie poddały się. Czy nie pamiętamy, jak w 1947 r. ambasador Stanów Zjednoczonych Bliss Lane powiedział: „Nie pomożemy wam, nawet UNRRY nie damy”.

(Sejm Polskiej Rzeczypospolitej Ludowej. Kadencja II – Sesja I. Sprawozdanie Scenograficzne z 6 posiedzenia Sejmu Polskiej Rzeczypospolitej Ludowej w dniach od 24 do 28 kwietnia 1957 r.)

– z lat 60. XX:

Jest to jednakże a razem wielki problem, który staje przed dalszym naszym rozwojem. Stare konflikty klasowe, związane z przedwojenną strukturą społeczną, należą pod każdym względem do przeszłości. Konflikty współczesne, to konflikty nowego społeczeństwa, w których chodzi nie o to, czy, ale o to – jak kształtować oblicze naszej rzeczywistości w warunkach socjalizmu, jak uczynić ją najlepszą i najefektywniejszą. Trzeba przy tym mieć na uwadze dwa zjawiska, które odzwierciedlają sytuację naszego społeczeństwa. Po pierwsze – że rewolucja kulturalna rozbudziła potrzeby ludzkie, wyprzedzając zdecydowanie możliwości gospodarcze kraju, mimo ich wielokrotnego wzrostu, oraz stan organizacyjny w zakresie zaspokajania tych potrzeb.

(Sejm Polskiej Rzeczypospolitej Ludowej. Kadencja III – Sesja IX.
Sprawozdanie Stenograficzne 30 posiedzenia Sejmu Polskiej Rzeczypospolitej Ludowej
w dniach 7 i 8 grudnia 1964 r.)

– z lat 70. XX:

Uwidoczniała równocześnie skalę społecznych rezerw, inicjatywy i gospodarności, a więc zasobów, które można skutecznie wyzwać przy pomocy odpowiedniego oddziaływania moralno-politycznego i zainteresowania materialnego. W rezultacie otrzymaliśmy wszyscy niezbita dowody utrwalania się bardzo korzystnej atmosfery społeczno-politycznej, która po VI Zjeździe partii wyraziła się powszechną działalnością na rzecz przyspieszenia wzrostu gospodarczego. Materialne efekty tej działalności wyraża się – w samym tylko przemyśle – przekroczeniem tegorocznego planu sprzedaży produkcji i usług o około 53 mld zł, tj. o około 4%, z czego około 33 mld zł, tj. 2,4% zadań planu będzie rezultatem realizacji zobowiązań podjętych w ramach akcji „20 miliardów”. Efekty bieżącego roku dowodzą, że realizacja wielkiego i ambitnego programu osiągnięcia dobrobytu społecznego możliwa jest jedynie w warunkach dynamicznego wzrostu gospodarki. I odwrotnie – wysokie tempo wzrostu gospodarczego uwarunkowane jest systematycznym i odczuwalnym postępem w poziomie życia ludzi pracy. Jak zaś wskazują doświadczenia 1971 i 1972 roku, ogniwem łączącym szybki rozwój kraju i osiąganie dobrobytu społecznego jest stałe utrzymywanie w polu widzenia podstawowego celu socjalistycznej ekonomiki, jakim jest zaspokajanie materialnych, socjalnych i kulturalnych potrzeb społeczeństwa.”

(Sejm Polskiej Rzeczypospolitej Ludowej. Kadencja VI – Sesja II.
Sprawozdanie Stenograficzne z 7 posiedzenia Sejmu Polskiej Rzeczypospolitej Ludowej
w dniu 29 listopada 1972 r.)

– z lat 80. XX:

Dekret bowiem zredagowany został przed 20 laty, rzeczywiście na stan wojenny z wrogiem zewnętrznym, i w tej nie poprawionej formie została chyba zbyt pośpiesznie podpisany w nocy z 12 na 13 grudnia. Tym głębokim nieporozumieniem, które wywołuje w kraju i w świecie nasz termin: „stan wojenny” wydaje mi się, czas położyć kres odpowiednią inicjatywą Sejmu. Już Komisja Prac

Ustawodawczych sugerowała, aby dekret o stanie wojennym przygotować w nowej redakcji w formie ustawy sejmowej. Proponuje niemylącą nazwę: „Ustawa o stanie nadzwyczajnym w przypadku wojny lub zagrożenia bezpieczeństwa wewnętrznego kraju”. Ustawa wyraźnie rozróżniałaby restrykcje rzeczywistego stanu wojny od ograniczeń stanu wyjątkowego, wynikających z zagrożenia wewnętrznego.

(Sejm Polskiej Rzeczypospolitej Ludowej. Sprawozdanie Stenograficzne z 21 posiedzenia Sejmu Polskiej Rzeczypospolitej Ludowej w dniach 3 i 4 maja 1982 r., s. 2)

– z lat 90. XX:

Wysoki Senacie! Przed porządkiem dziennym chciałbym wrócić do faktu, który nas głęboko poruszył, mianowicie już po naszym poprzednim posiedzeniu, w dniu 7 marca doszło do tragicznego wypadku w kopalni „Halemba”, w wyniku którego pięciu górników poniosło śmierć. Sądzę, że naszą moralną powinnością jest oddanie hołdu ofiaram i uczczenie minutą ciszy. (Wszyscy wstają, minuta ciszy).”

(Senat Rzeczypospolitej Polskiej. 1 kadencja 45 posiedzenie Senatu Rzeczypospolitej Polskiej w dniu 15 marca 1991 r.)

Podsumowanie

Powyższe informacje dotyczące stenogramów parlamentu polskiego za lata 1918-2018 nie stanowią, bo stanowić nie mogą w tym krótkim przyczynku, pełnej palety danych i ich interpretacji. Jest to raczej niewielka próbka możliwości tkwiących zarówno w samym materiale językowym, jak i narzędzi cyfrowych, za pomocą których przeprowadzone zostały te wstępne analizy zaprezentowane w niniejszym tekście. Jednakże już ta próbka badań pokazuje – w moim przekonaniu – przeogromny potencjał badawczy tkwiący w samym materiale językowym, który – jako niemały korpus danych – może być z powodzeniem przeszukiwany i opracowywany m. in. za pomocą tych narzędzi z zakresu humanistyki cyfrowej (jak i innych narzędzi, które można by testować też i na tym materiale). Stąd też już krótka droga do zrealizowania projektu słownika parlamentaryzmu polskiego XX wieku, tj. dokładniej – słownika obejmującego okres 1918-2018. Dane zgromadzone w postaci korpusu (w IPI PAN) obejmują ten okres i są nieustannie pomnażane – do korpusu dodawane są dane językowe pochodzące zarówno z posiedzeń plenarnych sejmu i senatu, jak i z posiedzeń komisji parlamentarnych.

Planowany słownik miałby postać tezaury, czyli opisywałby jednostki leksykalne pogrupowane w pola tematyczne według swoich znaczeń (na podobnej koncepcji, wzbogaconej dodatkowo o zestaw relacji semantycznych, są oparte wordnety, czyli bazy danych relacji semantycznych). W pracach nad słownikiem wykorzystane zostaną ponadto: tradycyjna koncepcja pól leksykalno-semantycznych oraz wywodząca się z nurtu kognitywizmu teoria ram interpretacyjnych (Fillmore 1982). W budowie hasła

słownikowego zespół odwoła się do przyjętego w polskiej leksykografii sposobu konstruowania haseł, opartego na pracach Witolda Doroszewskiego (por. SJPDor.). W zadaniach dotyczących przetwarzania języka naturalnego stosowana będzie metodologia lingwistyczno-informatyczna przejęta z projektu Narodowego Korpusu Języka Polskiego.

Na podstawie analiz przeprowadzonych na potrzeby niniejszego tekstu – analiz na większej próbie materiałowej niż w przywołanym we wstępie moim wcześniejszym artykule (będącym w druku) – należy zauważyć, że poziom emocji w polskim parlamencie jest jednak stosunkowo niemały. Średni poziom nasycenia emocjami stenogramów, a więc – dyskursu parlamentarnego – wynosi nieco ponad 24% każdego tekstu notującego posiedzenie parlamentu. Warto tu dodać, że istotny wkład w tę i każdą następną interpretację emocjonalności stenogramów wnoszą prace Stanisława Grabiasa (Grabias 1980; Grabias 1994).

Drugim składnikiem niniejszego tekstu było zaprezentowanie szczegółowych analiz semantycznych dwóch losowo wybranych, ale ważnych leksemów pojawiających się w stenogramach na przestrzeni badanego tu okresu. Dzięki danym z korpusu widać wyraźnie, czy i jak zmienia się pod względem użycia w tym dyskursie znaczenie leksemu, a zatem – jakiej kategoryzacji językowej poddawane są pojęcia (tu: te dwa analizowane) w dyskursie parlamentarnym (o kategoryzacji – w myśl prac powstających w nurcie kognitywizmu, dotyczących językowego obrazu świata, np. Bartmiński 1999; Bartmiński 2006).

Bibliografia

- Bartmiński J., 2006, *Językowe podstawy obrazu świata*, Lublin.
- Fillmore Ch., 1982, *Frame Semantics. Linguistics in the Morning Calm*, The Linguistics Society of Korea (red.), Seoul, s. 111-137.
- Grabias S., 1980, *O ekspresywności języka. Ekspresja a słowotwórstwo*, Lublin.
- Grabias S., 1994, *Język w zachowaniach społecznych*, Lublin.
- Językowy obraz świata*, 1999, J. Bartmiński (red.), Lublin.
- Kieraś W., Kobyliński Ł., Ogrodniczuk M., 2018, *Korpusomat – a tool for creating searchable morphosyntactically tagged corpora*, „Computational Methods in Science and Technology” 24(1), s. 21-27.
- Ogrodniczuk M., 2012, *The Polish Sejm Corpus*, w: *Proceedings of the Eighth International Conference on Language Resources and Evaluation, LREC*, Istanbul, Turkey, ELRA, p. 2219-2223.
- Ogrodniczuk M., 2018, *Polish Parliamentary Corpus*. In: *Proceedings of the LREC 2018 Workshop ParlaCLARIN: Creating and Using Parliamentary Corpora*, Darja Fišer, Maria Eskevich, and Franciska de Jong (edit.), European Language Resources Association (ELRA), Paris, p. 15–19.
- Piasecki M., 2014, *User-driven language technology infrastructure – the case of CLARIN-PL. Proceedings of the Ninth Language Technologies Conference. Ljubljana, Slovenia*.
- Szczyszek M., 2019, *Emocje w parlamencie – parlament w emocjach: ujęcie statystyczne. O projekcie słownika polskiego parlamentaryzmu XX wieku (lata 1918–2018)*, „Prace Językoznawcze”, tom XXI, nr 3, s. 203-218.

- Walkowiak T., 2017, *Language Processing Modelling Notation – Orchestration of NLP Micro-services*, w: *Advances in Dependability Engineering of Complex Systems: Proceedings of the Twelfth International Conference on Dependability and Complex Systems DepCoS-RELCOMEX*, s. 464-473.
- Walkowiak T., Pol M., 2017, *The impact of administrator working hours on the reliability of the Centre of Language Technology*, „Journal of Polish Safety and Reliability Association Summer Safety and Reliability Seminars” Volume 8, Number 1, s. 167-174.
- Walkowiak T., Piasecki M., 2015, *Web-based natural language processing workflows for the research infrastructure in humanities*, w: „5th Conference of the Japanese Association for Digital Humanities”, JADH, s. 61-63.
- Żmigrodzki P., Bańko M., Batko-Tokarz B., Bobrowski J., Czelakowska A., Grochowski M., Przybylska R., Waniakowa J., Węgrzynek K. (red.): *Wielki słownik języka polskiego PAN. Geneza, koncepcja, zasady opracowania*, Instytut Języka Polskiego PAN, Kraków 2018.

<http://clip.ipipan.waw.pl/PSC>

<http://korpusomat.pl/>

http://sejm.nlp.ipipan.waw.pl/query_corpus/

<http://ws.clarin-pl.eu/sentymet.shtml>

SJPDor. – *Słownik języka polskiego*, red. W. Doroszewski, t. I-XI, Warszawa 1958-1969.

WSJP – *Wielki słownik języka polskiego*, red. P. Żmigrodzki (<http://www.wsjp.pl/>).

MICHAŁ SZCZYSZEK

The project of a dictionary of the Polish parliamentarism of the 20th century (1918-2018): the preliminary stage – the parliamentary corps

Abstract

The article discusses the preliminary results of research into the language corpus of the Polish parliamentary transcripts. Emphasis has been placed on a study of the emotional level of randomly selected transcripts. Two important lexemes (concept): *niezawistość* and *niepodległość* have been studied by academics in terms of their use in transcripts. The frequency of these words has been estimated. Moreover, the semantic transformations of these words depending on the use of these lexemes in this discourse have been subjected to an academic study.

Keywords: language expression, lexicography, semantics, language corpus, language categorization.

Aneks

Tabela 3.

nazwa pliku	rok	liczba segmentów wg <i>Korpusomatu</i> (Korpusomat.pl)	liczba segmentów emocjonalnych wg <i>Analizatora</i> <i>Wydźwięku</i> (https:// ws.clarin-pl.eu/ sentyment.shtml)	procentowy udział ekspresywizmów w danym stenogramie
4_035_000010568	senat1990	49293	10570	21,44320695
4_036_000010906	senat1990	48550	8488	17,48300721
4_037_000010576	senat1990	37831	8055	21,29206207
4_038_000010580	senat1990	39920	6856	17,1743487
4_039_000010596	senat1990	32235	4969	15,41492167
4_040_000010600	senat1990	57886	11909	20,57319559
4_041_000010614	senat1991	88866	18548	20,87187451
4_042_000010622	senat1991	24500	4223	17,23673469
4_043_000010628	senat1991	45358	7920	17,46108735
4_044_000010642	senat1991	65654	17378	26,4690651
suma za lata 90:		490093	98916	20,1831081
24_00007_000050269	sejm1980	113702	29471	25,91950889
24_00008_000050354	sejm1980	57666	12637	21,91412617
24_00016_000050304	sejm1981	94052	20776	22,08990771
24_00017_000050312	sejm1981	77623	19996	25,76040607
24_00018_000050413	sejm1982	106683	29955	28,07851298
24_00019_000050318	sejm1982	72624	17491	24,08432474
24_00020_000050322	sejm1982	93263	23825	25,54603648
25_00061_000050818	sejm1985	62340	16460	26,4035932
25_00062_000050820	sejm1985	59433	12419	20,89579863
25_00063_000050733	sejm1985	38693	8648	22,35029592
suma za lata 80:		776079	191678	24,69825881
5_00006_000014679	sejm1970	17819	4201	23,57595825
5_00007_000014681	sejm1970	15368	5126	33,35502343
5_00009_000014685	sejm1970	21111	4863	23,0353844
6_00008_000046194	sejm1972	70182	17808	25,37402753
6_00009_000046360	sejm1973	48695	11102	22,79905534
6_00010_000046362	sejm1973	53660	15692	29,24338427
6_00015_000046354	sejm1973	57254	12568	21,95130471
suma za lata 70:		284089	71360	25,1188888
6_00020_000045305	sejm1963	74906	20451	27,30221878
6_00021_000045331	sejm1963	52295	12473	23,85122861
6_00029_000045349	sejm1964	53720	12313	22,92069993

6_00030_000045351	sejm1964	95429	23025	24,12788565
5_00010_000014571	sejm1966	95014	21569	22,70086514
5_00011_000014579	sejm1966	97934	20999	21,44199155
5_00012_000014581	sejm1967	34754	8007	23,03907464
5_00013_000014587	sejm1967	29430	7558	25,68127761
5_00018_000014567	sejm1968	81054	20285	25,02652553
5_00020_000014575	sejm1968	24741	6307	25,49209814
suma za lata 60:		639277	152987	23,93125359
5_00075_000014527	sejm1950	71161	21985	30,89473166
5_00076_000014667	sejm1950	91659	27220	29,69702921
5_00086_000013811	sejm1950	13448	4923	36,607674
5_00088_000013815	sejm1951	18009	4699	26,0925093
6_00001_000033591	sejm1957	8264	1560	18,87705712
6_00009_000030668	sejm1957	8481	1818	21,4361514
6_00002_000033597	sejm1957	29834	8375	28,07199839
6_00003_000033603	sejm1957	10993	1862	16,93805149
6_00004_000033605	sejm1957	23407	5489	23,45024993
6_00005_000034719	sejm1957	22339	4741	21,22297328
suma za lata 50:		297595	82672	27,78003663
5_00017_000013749	sejm1947	27783	9670	34,80545657
5_00018_000013751	sejm1947	18722	5879	31,40155966
5_00019_000013755	sejm1947	49107	13704	27,90640846
5_00020_000013757	sejm1947	44270	12527	28,296815
5_00021_000013779	sejm1947	31684	9177	28,96414594
5_00022_000013785	sejm1947	12784	3725	29,13798498
5_00023_000013787	sejm1947	22617	5647	24,96794447
5_00060_000014349	sejm1949	37813	9802	25,92230185
5_00062_000014505	sejm1949	18813	5322	28,28894913
5_00066_000014013	sejm1949	25486	6246	24,50757279
suma za lata 40:		289079	81699	28,26182462
18_060_000034495	sejm1932	49928	12845	25,72704695
18_065_000034500	sejm1932	42013	8507	20,24849451
18_072_000034507	sejm1932	16841	3538	21,00825367
18_031_000034618	sejm1936	12793	2092	16,35269288
18_034_000034621	sejm1937	12220	2786	22,79869067
18_035_000034622	sejm1937	8248	1966	23,83608147
18_037_000034624	sejm1937	37802	10067	26,63086609
18_040_000034627	sejm1937	61810	17481	28,28183142
18_041_000034628	sejm1937	49588	13044	26,30475115
18_072_000034659	sejm1938	58164	14153	24,33292071
suma za lata 30:		349407	86479	24,75021966

17_004_000034003	sejm1922	6796	1770	26,0447322
17_005_000034004	sejm1922	433	165	38,10623557
17_006_000034005	sejm1923	11865	2760	23,26169406
17_007_000034006	sejm1923	10182	3503	34,40384993
17_008_000034007	sejm1923	21616	5403	24,9953738
17_015_000034014	sejm1923	29169	6691	22,93873633
17_065_000034064	sejm1923	41448	7203	17,37840185
17_215_000034214	sejm1925	31678	6194	19,55300208
17_216_000034215	sejm1925	21245	4626	21,77453518
17_217_000034216	sejm1925	34003	6641	19,53062965
suma za lata 20:		208435	44956	21,56835464
SUMA całości:		3334054	810747	24,31715263