

Michał Szczyszek
Uniwersytet im. Adama Mickiewicza, Poznań
ORCID: <https://orcid.org/0000-0002-0253-7296>
e-mail: szczysze@amu.edu.pl

Emocje w parlamencie – parlament w emocjach: ujęcie statystyczne. O projekcie słownika polskiego parlamentaryzmu XX wieku (lata 1918–2018)

Emotions in the parliament – parliament in emotions: a statistical approach. The project of the dictionary of the Polish parliamentary system of the 20th century (1918–2018)

This contribution discusses the preliminary results of research on linguistic material included in the parliamentary transcripts of the Polish parliament in 1918–2018. Parliamentary transcripts became the basis for the Polish Parliamentary Corpus (<http://clip.ipipan.waw.pl/PSC>) created in the Linguistic Engineering Team of the Institute of Computer Science of the Polish Academy of Sciences. This contribution presents the results of the preliminary linguistic work on this corpus: emotional tensions of randomly chosen transcripts from particular decades of the twentieth century were discussed. The summary contains information on the convergence occurring in different decades and in different periods in the history of Poland (the history of Polish parliamentarism) in the construction of parliamentary discourse - its emotionality calculated by the level of language expression.

Słowa kluczowe: ekspresja językowa, leksykografia, semantyka, korpus językowy
Key words: language expression, lexicography, semantics, language corpus

1. 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 tab. 1. poniżej). Na tej podstawie zostały już podjęte 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 artykule omówione zostaną wstępne wyniki analiz językoznawczych wykorzystujących do opracowania narzędzia cyfrowe stworzone w IPI PAN, jak i przez konsorcjum CLARIN. Na potrzeby niniejszego tekstu skupiliśmy się na przedstawieniu danych statystycznych dotyczących skończonego korpusu stenogramów za lata 1918–2018. Ponadto została podjęta próba omówienia jednego z wielu potencjalnie możliwych językoznawczych aspektów analizy i interpretacji tak bogatego materiału: na potrzeby tego tekstu opracowane zostały językowe sposoby wyrażania ekspresji zawarte w stenogramach sejmowych badanego okresu, a więc pojawiające się na sali plenarnej Sejmu czy Senatu. Poprzez te analizy, poprzez opracowanie tego – wąskiego, ale bardzo istotnego – aspektu językoznawczego, jakim jest ekspresja językowa (w myśl np. koncepcji Stanisława Grabiasa o ekspresji językowej; Grabias 1980), można pokazać zarówno złożoność zjawiska w tych stenogramach parlamentarnych, jak i ogromną potencję badawczo-naukową, która tkwi w przyjętych za podstawę materiałową niniejszych badań (i niniejszego przyczynku) tekstach wystąpień sejmowych i senackich.

Z tabeli powyższej (tab. 1) 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 **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.

2. 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 wydźwięku* opracowany przez konsorcjum CLARIN-PL (<http://ws.clarin-pl.eu/sentiment.shtml>; por. także opracowania autorów *Analizator wydźwięku*: Walkowiak 2017; Walkowiak, Pol 2017; Walkowiak, Piasecki 2015; Piasecki 2014).

W dalszej części przyczynku zostanie zacytowany przykładowy stenogram sejmowy (we fragmentach) dla ukazania struktury tego rodzaju tekstów – w pewnej mierze schematycznych, realizujących wymagane prawem układy kompozycyjne (realizowane wszakże podczas posiedzeń parlamentarnych), a także – w warstwie wypowiedzi posłów czy senatorów – realizujące mniej czy bardziej swobodne vs. przygotowane wypowiedzi z mównicy sejmowej.

Następnie zostaną przedstawione wyniki pracy Korpusomatu – w postaci danych statystycznych dotyczących danego analizowanego szczegółowo na potrzeby niniejszego przyczynku stenogramu, a także wyniki pracy *Analizatora wydźwięku* – nad grupą tekstów stenogramów pochodzących z poszczególnych dziesięcioleci (tj. z lat 20., 30., 40. 50., 60., 70., 80. i 90. XX w.). *Analizator wydźwięku* pozwoli oszacować interesujący z punktu widzenia tematu niniejszego przyczynku poziom nasycenia emocjami (wyrazami ekspresywnymi) stenogramów, a więc posiedzeń sejmowych w poszczególnych okresach.

Poniżej przytoczone są fragmenty przykładowego stenogramu sejmowego – tu: z lat 30. XX w. (z 1935 r.):

SEJM RZECZYPOSPOLITEJ POLSKIEJ.

Okres III 1930–1935. Sesja nadzwyczajna 1935 r.

Sprawozdanie Stenograficzne

ze 147 posiedzenia w dniu 27 czerwca 1935 r.

(Początek posiedzenia o godz. 16 min. 27.)

Przewodniczący Marszałek Sejmu dr. Kazimierz Świtalski.

Marszałek: Otwieram posiedzenie. Protokół 145 posiedzenia uważam za przyjęty, gdyż nie wniesiono przeciw niemu zarzutów. Protokół 146 posiedzenia znajduje się do przejrzenia w Biurze Sejmu. Jako sekretarze zasiadają pp. Borecki i Wojtowicz. Listę mówców prowadzi p. Wojtowicz.

Usprawiedliwiają nieobecność posłowie: Adamowicz, dr. Stefan Zygmunt Dąbrowski, Jaźwiński i Pac.

Udzieliłem urlopu posłom: Rząsie na 2 dni, Pacowi i Winiarskiemu na 3 dni, Nowodworskiemu na 4 dni, Bryle, Grünbaumowi, Witkowskiemu i Tadeuszowi Wróblowi na 6 dni, Jakubowskiemu i Kulisiewiczowi na 1 tydzień.

Proszą Izbę o udzielenie urlopów do końca bieżącej sesji posłowie Lewicki i Pużak. Nie słyszę protestu, uważam więc, że Izba zgodziła się na udzielenie tych urlopów. Przystępujemy do porządku dziennego: Sprawozdanie Komisji Konstytucyjnej o projekcie ustawy o wyborze Prezydenta Rzeczypospolitej (druki nr 1168 i 1174).

Głos ma sprawozdawca p. Car.

P. Car: Wysoki Sejmie! Zanim przystąpię do referatu, chciałbym zwrócić uwagę Wysokiej Izby, że w druku nr 1168, który Panom rozesłano, znalazły się dwie omyłki, które sprostowano w rozdanych Panom odbitkach. W tej chwili chciałbym zwrócić

uwagę na jeszcze jedno sprostowanie. Mianowicie w art. 52 ust. 2 ostatnie dwa wyrazy: w poszczególnym obwodzie” należy zamienić przez wyrazy: w poszczególnych obwodach”. Wnoszę to także jako poprawkę drukarską. [...]

Sekretarz p. Wawrzynowski (czyta): W ust. 3 wyrazy: „Przewodniczącego Państwowej Komisji Wyborczej” zastąpić przez „najmłodszego członka Komisji Wyborczej”.

Marszałek: Kto jest za tą poprawką, zechce wstać. Mniejszość. Poprawka upadła. A zatem art. 57 przechodzi w brzmieniu komisyjnym, tak samo, jak wszystkie artykuły od art. 58 aż do końca, ponieważ nie zgłoszono do nich żadnych poprawek.

Przystępujemy do głosowania nad całością ustawy. Kto jest za całością ustawy, zechce wstać. Stoi większość, ustawa została w drugim czytaniu przyjęta.

Przystępujemy do trzeciego czytania. Kto jest za całością ustawy, zechce wstać. Większość. Stwierdzam, że ustawa została przez Sejm przyjęta.

Proponuję odbyć następne posiedzenie jutro o godz. 10 rano z następującym porządkiem dziennym: trzecie czytanie projektu ustawy o ordynacji wyborczej do Sejmu.

Nie słyszę sprzeciwu, uważam propozycję moją za przyjętą.

Zamykam posiedzenie.

(Koniec posiedzenia o godz. 17 min. 52.).

3. Wyniki prac *Analizatora wydźwięku* dla poszczególnych dziesięcioleci XX w. oraz dla przykładowych stenogramów z poszczególnych okresów

Na potrzeby niniejszego przyczynku przeanalizowano kilka losowo wybranych stenogramów – po jednym z każdego dziesięciolecia XX w. Taki dobór materiałów miał na celu sprawdzenie i ukazanie potencji badawczej zebranych tekstów w postaci korpusu stenogramów sejmowych (Polski Korpus Sejmowy: <http://clip.ipipan.waw.pl/PSC>; Ogrodniczuk 2012: 2219–2223; Ogrodniczuk 2018: 15–19). Jednakże należy zauważyć, że analizy pojedynczych przykładów stenogramów parlamentarnych tu omówionych pokazują tendencję tkwiącą w poszczególnych dziesięcioleciach XX w. funkcjonowania polskiego parlamentu – zostało to przebadane na szerszej grupie materiałowej, większej liczbie stenogramów z poszczególnych wskazanych tu przedziałów czasowych, natomiast tu, w tym przyczynku, pokazane są szczegółowe dane pochodzące właśnie z pojedynczych, losowo dobranych stenogramów. Innymi słowy, niniejszy tekst (będący wynikiem wstępnych badań nad korpusem polskich stenogramów parlamentarnych) jest swoistym zbiorem studiów przypadku – analizy konkretnych, przykładowych stenogramów sejmowych obrazujących tendencje w zakresie ekspresji językowej dającej się zaobserwować w obszerniejszej bazie materiałowej (stenogramowej).

3.1. Lata 20. XX w.

Wyniki pracy *Analizatora wydźwięku* dla stenogramu z 1923 r. (44 strony)

- 18 600 segmentów – wszystkie słowoformy wynotowane za pomocą Korpusomatu;
- 1085 segmentów – oszacowanych jako emocjonalnie nacechowane słowoformy (za pomocą *Analizatora wydźwięku*);
- to stanowi ok. 5,8% całości tekstu – czyli emocje wystąpiły w tym tekście w takim właśnie natężeniu;
- ujawniły się emocje:
 - pozytywne: Radość (1136), Zaufanie (673), Cieszenie się na coś oczekiwanego (187), Użyteczność (1033), Prawda (86), Wiedza (164), Piękno (127), Szczęście (810),
 - negatywne: Smutek (935), Złość (957), Strach (434), Wstręt (507), Zaskoczenie czymś nieprzewidywanym (74), Nieużyteczność (815), Krzywda (931), Niewiedza (162), Błąd (589), Brzydota (67), Nieszczęście (830).

Przykładowy cytat nacechowany emocjonalnie ze stenogramów z lat 20. XX w.:

„... istotne nasze prawo. (Oklaski wśród posłów ukraińskich. Głos: Nie dawno byliście Rosjanami. Pan jest Rosjaninem). Nie, jestem Ukrainiec 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.

Nota bene: w całym tym cytacie wyraźnie przebijają się emocje, rzutujące na cały fragment poprzez sformułowania typu, np. „Pan jest Rosjaninem” czy informacja o sytuacji komunikacyjnej: „(Wrzawa)”.

3.2. Lata 30. XX w.

Wyniki pracy *Analizatora wydźwięku* dla stenogramu z 1933 r. (12 stron)

- 4757 segmentów – wszystkie słowoformy wynotowane za pomocą Korpusomatu;
- 234 segmenty – oszacowane jako emocjonalnie nacechowane słowoformy (za pomocą *Analizatora wydźwięku*);

- to stanowi ok. 5% całości tekstu – czyli emocje wystąpiły w tym tekście w takim właśnie natężeniu;
- ujawniły się emocje:
 - pozytywne: Radość (949), Zaufanie (631), Cieszenie się na coś oczekiwanego (170), Użyteczność (934), Prawda (110), Wiedza (174), Piękno (84), Szczęście (716),
 - negatywne: Smutek (960), Złość (1002), Strach (415), Wstręt (545), Zaskoczenie czymś nieprzewidywanym (91), Nieużyteczność (861), Krzywda (844), Niewiedza (165), Błąd (660), Brzydota (65), Nieszczęście (820).

Przykładowy cytat nacechowany emocjonalnie ze stenogramów z lat 30. XX w.:

„...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

3.3. Lata 40. XX w.

Wyniki pracy *Analizatora wydźwięku* dla stenogramu z 1947 r. (38 stron)

- 12 007 segmentów – wszystkie słowoformy wynotowane za pomocą Korpusomatu;
- 894 segmenty – oszacowane jako emocjonalnie nacechowane słowoformy (za pomocą *Analizatora wydźwięku*);
- to stanowi ok. 7,45% całości tekstu – czyli emocje wystąpiły w tym tekście w takim właśnie natężeniu;
- ujawniły się emocje:
 - pozytywne: Radość (589), Zaufanie (375), Cieszenie się na coś oczekiwanego (118), Użyteczność (592), Prawda (58), Wiedza (117), Piękno (55), Szczęście (443),
 - negatywne: Smutek (413), Złość (446), Strach (159), Wstręt (175), Zaskoczenie czymś nieprzewidywanym (21), Nieużyteczność (386), Krzywda (328), Niewiedza (64), Błąd (292), Brzydota (35), Nieszczęście (352).

Przykładowy cytat nacechowany emocjonalnie ze stenogramów z lat 40. XX w.:

„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 w Lesznie, który jest wynikiem atmosfery, do której niemało przyczynili się PSL-owcy, (Poseł Mikołajczyk: Dlaczego?) Dlatego z tej trybuny najostrożniej piętnujemy oszczercze metody pana Mikołajczyka, które są wyrazem rozkładu i upadku jego partii. (Oklaski. 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

3.4. Lata 50. XX w.

Wyniki pracy *Analizatora wydźwięku* dla stenogramu z 1957 r. (118 stron)

- 138 933 segmenty – wszystkie słowoformy wynotowane za pomocą Korpusomatu;
- 10 595 segmentów – oszacowanych jako emocjonalnie nacechowane słowoformy (za pomocą *Analizatora wydźwięku*);
- to stanowi ok. 7,6% całości tekstu – czyli emocje wystąpiły w tym tekście w takim właśnie natężeniu;
- ujawniły się emocje:
 - pozytywne: Radość (176), Zaufanie (105), Cieszenie się na coś oczekiwanego (36), Użyteczność (185), Prawda (14), Wiedza (30), Piękno (14), Szczęście (148),
 - negatywne: Smutek (116), Złość (108), Strach (53), Wstręt (69), Zaskoczenie czymś nieprzewidywanym (10), Nieużyteczność (98), Krzywda (114), Niewiedza (40), Błąd (59), Brzydota (4), Nieszczęście (122).

Przykładowy cytat nacechowany emocjonalnie ze stenogramów 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, musimy 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.

3.5. Lata 60. XX w.

Wyniki pracy *Analizatora wydźwięku* dla stenogramu z 1964 r. (38 stron)

- 54 571 segmentów – wszystkie słowoformy wynotowane za pomocą Korpusomatu;
- 4089 segmentów – oszacowanych jako emocjonalnie nacechowane słowoformy (za pomocą *Analizatora wydźwięku*);
- to stanowi ok. 7,5% całości tekstu – czyli emocje wystąpiły w tym tekście w takim właśnie natężeniu;
- ujawniły się emocje:
 - pozytywne: Radość (2703), Zaufanie (1908), Cieszenie się na coś oczekiwanego (751), Użyteczność (2940), Prawda (461), Wiedza (926), Piękno (141), Szczęście (1689),
 - negatywne: Smutek (1575), Złość (1449), Strach (437), Wstręt (488), Zaskoczenie czymś nieprzewidywanym (136), Nieużyteczność (1617), Krzywda (1044), Niewiedza (257), Błąd (881), Brzydota (117), Nieszczęście (1305).

Przykładowy cytat nacechowany emocjonalnie ze stenogramów z lat 60. XX w.:

„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
z 30 posiedzenia Sejmu Polskiej Rzeczypospolitej Ludowej
w dniach 7 i 8 grudnia 1964 r.

3.6. Lata 70. XX w.

Wyniki pracy *Analizatora wydźwięku* dla stenogramu z 1972 r. (24 strony)

- 26 670 segmentów – wszystkie słowoformy wynotowane za pomocą Korpusomatu;
- 2064 segmenty – oszacowane jako emocjonalnie nacechowane słowoformy (za pomocą *Analizatora wydźwięku*);
- to stanowi ok. 7,7% całości tekstu – czyli emocje wystąpiły w tym tekście w takim właśnie natężeniu;
- ujawniły się emocje:
 - pozytywne: Radość (993), Zaufanie (685), Cieszenie się na coś oczekiwanego (263), Użyteczność (1122), Prawda (133), Wiedza (305), Piękno (40), Szczęście (640),
 - negatywne: Smutek (521), Złość (474), Strach (257), Wstręt (169), Zaskoczenie czymś nieprzewidywanym (84), Nieużyteczność (531), Krzywda (456), Niewiedza (76), Błąd (255), Brzydota (26), Nieszczęście (462).

Przykładowy cytat nacechowany emocjonalnie ze stenogramów z lat 70. XX w.:

„Uwidoczniła 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 niezbite 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 wyrażą 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.

3.7. Lata 80. XX w.

- Wyniki pracy *Analizatora wydźwięku* dla stenogramu z 1982 r. (63 strony)
- 61 329 segmentów – wszystkie słowoformy wynotowane za pomocą Korpusomatu;
 - 5769 segmentów – oszacowanych jako emocjonalnie nacechowane słowoformy (za pomocą *Analizatora wydźwięku*);
 - to stanowi ok. 9,4% całości tekstu – czyli emocje wystąpiły w tym tekście w takim właśnie natężeniu;
 - ujawniły się emocje:
 - pozytywne: Radość (1773), Zaufanie (1280), Cieszenie się na coś oczekiwanego (360), Użyteczność (2019), Prawda (310), Wiedza (519), Piękno (166), Szczęście (1204),
 - negatywne: Smutek (1864), Złość (1787), Strach (715), Wstręt (639), Zaskoczenie czymś nieprzewidywanym (145), Nieużyteczność (1520), Krzywda (1453), Niewiedza (266), Błąd (986), Brzydota (88), Nieszczęście (1433).

Przykładowy cytat nacechowany emocjonalnie ze stenogramów z lat 80. XX w.:

„Dekret bowiem zredaowany został przed 20 laty, rzeczywiście na stan wojenny z wrogiem zewnętrznym, i w tej nie poprawionej formie został 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.

3.8. Lata 90. XX w.

Wyniki pracy *Analizatora wydźwięku* dla stenogramu z 1991 r. (66 stron)

- 26 419 segmentów – wszystkie słowoformy wynotowane za pomocą Korpusomatu;
- 1348 segmentów – oszacowanych jako emocjonalnie nacechowane słowoformy (za pomocą *Analizatora wydźwięku*);
- to stanowi ok. 5,1% całości tekstu – czyli emocje wystąpiły w tym tekście w takim właśnie natężeniu;
- ujawniły się emocje:
 - pozytywne: Radość (905), Zaufanie (662), Cieszenie się na coś oczekiwanego (202), Użyteczność (973), Prawda (176), Wiedza (270), Piękno (32), Szczęście (589), Radość (1810), Zaufanie (1324), Cieszenie się na coś oczekiwanego (404), Użyteczność (1946), Prawda (352), Wiedza (540), Piękno (64), Szczęście (1178),
 - negatywne: Smutek (923), Złość (835), Strach (242), Wstręt (421), Zaskoczenie czymś nieprzewidywanym (61), Nieużyteczność (609), Krzywda (538), Niewiedza (84), Błąd (411), Brzydota (9), Nieszczęście (1534), Smutek (1846), Złość (1670) Strach (484), Wstręt (842), Zaskoczenie czymś nieprzewidywanym (122), Nieużyteczność (1218), Krzywda (1076), Niewiedza (168), Błąd (822), Brzydota (18), Nieszczęście (1138).

Przykładowy cytat nacechowany emocjonalnie ze stenogramów z lat 90. XX w.:

„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 ofiarom i uczenie minutą ciszy. (Wszyscy wstają, minuta ciszy).”

Senat

Rzeczypospolitej Polskiej

1 kadencja

45 posiedzenie Senatu Rzeczypospolitej Polskiej

w dniu 15 marca 1991 r.

4. Podsumowanie

Powyższe informacje dotyczące stenogramów Sejmu 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, obejmującym ponad 193 mln

jednostek pochodzących z zapisów posiedzeń polskiego parlamentu w ostatnich stu latach, jak i narzędzi cyfrowych, za pomocą których przeprowadzone zostały 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 narzędzi z zakresu humanistyki cyfrowej (jak i innych narzędzi, które można by testować też i na tym materiale).

Stąd już krótka droga do zrealizowania projektu słownika parlamentaryzmu polskiego XX w., tj. dokładniej – w stulecie odzyskania niepodległości 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 z posiedzeń komisji parlamentarnych (prace trwające od kilkunastu miesięcy w zespole skupionym w IPI PAN zaowocowały powstaniem roboczego (i nadal rozwijanego) jeszcze korpusu polskiego parlamentaryzmu XX w.).

Tak obszerna podstawa materiałowa (już średniej wielkości światowych korpusów językowych typu *big data*), poddawana takiemu opracowaniu (choćby za pomocą Korpusomatu IPI PAN) daje – w przekonaniu twórców korpusu, jak i osób chcących zaangażować się w prace projektowe nad słownikiem – wystarczająco mocne argumenty za opracowaniem takiego słownika. Byłby on przydatny zarówno filologom (językoznawcom i literaturoznawcom), jak i politologom, historykom, prawnikom, socjologom, kulturoznawcom. Wydaje się bowiem uzasadnione twierdzenie, że w parlamencie, w przemówieniach parlamentarzystów zapisanych w stenogramach, skupiają się jak w soczewce wszelkie, a przynajmniej znaczna większość, procesy kulturowe, społeczne, polityczne (politologiczne), historyczne – jak i językowe (filologiczne) danego czasu, okresu historyczno-społeczno-polityczno-kulturowo-językowego.

Planowany słownik ma mieć postać tezauryusa, czyli opisywać 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 przyczynku należy zauważyć, że poziom emocji w polskim parlamencie nie jest jakoś szczególnie duży. Zawiera się w przedziale od 5% objętości danego analizowanego tekstu (zwłaszcza w stenogramach przedwojennych) do ponad 9% (tu w latach 80. XX w.). Najczęściej pojawiały się dane świadczące o poziomach 5- i 7-procentowych (tu zwłaszcza w stenogramach powojennego ćwierćwiecza). Trudno jednak dziś ze stuprocentową pewnością stwierdzić, że są to poziomy albo znacząco niskie, albo znacząco wysokie, albo uśrednione. Aby któreś z tych twierdzeń mogło być sformułowane i udowodnione, należałoby przeprowadzić na szeroką skalę badania innych (nieparlamentarnych) tekstów z wykorzystaniem metod i narzędzi tu zaprezentowanych, w celu uzyskania danych mogących służyć za porównawcze w tej materii (takich badań systemowych i systematycznych trudno by szukać poza odległymi już nieco w czasie pracami np. Stanisława Grabiasa; Grabias 1980). Otwiera się więc – jak widać – kolejne pole eksploracji językoznawczej – właśnie dzięki analizom stenogramów za pomocą narzędzi cyfrowych.

Na uwagę zasługują jeszcze inne aspekty emocjonalności stenogramów sejmowych. Nawet jeżeli – tylko roboczo – przyjąć, że poziom nasycenia emocjami tych tekstów nie jest zbyt wysoki – choć to jedynie hipoteza wymagająca weryfikacji badawczej opisanej powyżej, to ze względu na rangę tych tekstów, miejsce ich powstania i/lub wygłoszenia – swoiste oddziaływanie społeczno-polityczno-kulturowo-historyczno-językowe jest (a przynajmniej może być większe) niż oddziaływanie w tych zakresach innych tekstów (co też stanowi obecnie raczej hipotezę badawczą, którą należałoby zweryfikować naukowo, materiałowo). W tym kontekście wyraźnie trzeba by zaznaczyć, że – co prawda stenogramy przedwojenne nie wykazują nawet w kontekście całego zbioru stenogramów wysokiego poziomu emocjonalnego (ok. 5%), to dosadność sformułowań ekspresywnych wydaje się – właśnie pod względem jakościowym – znacząco bardziej wyrazista niż dosadność wyrażania emocji w stenogramach powojennych, i to zarówno w tych z pierwszego powojennego ćwierćwiecza, jak i z ostatniego okresu XX w. (co jest widoczne w przywołanych powyżej przykładowych cytatach ze stenogramów z różnych dziesięcioleci XX w.).

Na oddzielną uwagę zasługują dane pochodzące z *Analizatora wydźwięku*, a dokładniej – to, jakie emocje zostały ujawnione w przeanalizowanych przez analizator stenogramach pochodzących z różnych okresów. Powielanie się tych samych lub bardzo zbliżonych emocji – zarówno pozytywnych, jak i negatywnych (co zostało opisane powyżej) – można interpretować na dwa różne sposoby wzajemnie się niewykluczające, a może nawet komplementarne względem siebie. Po pierwsze – jako spodziewane, bo narzucone przez „gatunek”

stenogramu ograniczenie do pewnej listy emocji (raczej podstawowych); po drugie – pokazuje to, w jaki sposób różni mówcy parlamentarni działający w różnych okresach XX-wiecznej historii Polski, historii polskiego parlamentu – mówcy niezależnie prawdopodobnie od siebie – konstruują ramy emocjonalne swoich wystąpień. Wszyscy z nich (większość?) budują napięcie emocjonalne, wyrażając swoją ekspresję poprzez odwołanie się do tych samych lub bardzo zbliżonych emocji. To natomiast pokazuje, że niezależnie od mówcy i okresu dziejowego parlamentarzyści tworzą w swoich wystąpieniach, a może – odwołują się do tej samej (chciałoby się powiedzieć) językowo-emocjonalnej (ekspresywnej) wizji świata, tj. korzystają – z danej lub przez siebie samodzielnie wytwarzanej – „kategoryzacji” emocjonalnej towarzyszącej ich wystąpieniom (por. prace powstające w nurcie kognitywizmu, dotyczące językowego obrazu świata, np. Bartmiński 1999; Bartmiński 2006). Takie stwierdzenie, znów będące raczej jeszcze hipotezą, wymagałoby oczywiście dalszych badań i weryfikacji, jednakże prowadzi wprost do konieczności opisanie językowej kategoryzacji rzeczywistości widzianej oczami parlamentarzystów w różnych okresach XX w., swoistego językowego obrazu świata z parlamentu, a temu problemowi w zamierzeniach badawczo-naukowych ma być poświęcony właśnie słownik polskiego parlamentarysty XX w.

Literatura

- Bartmiński J. (2006): *Językowe podstawy obrazu świata*. Lublin.
- Fillmore Ch. (1982): *Frame Semantics. Linguistics in the Morning Calm*. The Linguistics Society of Korea (eds.). Seoul, s. 111–137.
- Grabias S. (1980): *O ekspresywności języka. Ekspresja a słowotwórstwo*. Lublin.
- Językowy obraz świata* (1999). Red. J. Bartmiński. Lublin.
- Kieraś W., Kobyliński L., 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*. [In:] *Proceedings of the Eighth International Conference on Language Resources and Evaluation, LREC*, p. 2219–2223, Istanbul. Turkey.
- Ogrodniczuk M. (2018): *Polish Parliamentary Corpus*. [In:] D. Fišer, M. Eskevich, and F. de Jong (eds). *Proceedings of the LREC 2018 Workshop ParlaCLARIN: Creating and Using Parliamentary Corpora*, p. 15–19, Paris. France.
- Piasecki M. (2014): *User-driven language technology infrastructure – the case of CLARIN-PL*. Proceedings of the Ninth Language Technologies Conference. Ljubljana.
- Słownik języka polskiego* (1958–1969). Red. W. Doroszewski. T. I–XI. Warszawa (SJPDor.).
- Walkowiak T. (2017): *Language Processing Modelling Notation – Orchestration of NLP Microservices*. [In:] *Advances in Dependability Engineering of Complex Systems: Proceedings of the Twelfth International Conference on Dependability and Complex Systems DepCoS-RELCOMEX*, pp. 464–473.
- Walkowiak T., Piasecki M. (2015): *Web-based natural language processing workflows for the research infrastructure in humanities*. [In:] 5th Conference of the Japanese Association for Digital Humanities, pp. 61–63.

Walkowiak T., Pol M. (2017): *The impact of administrator working hours on the reliability of the Centre of Language Technology*. [In:] *Journal of Polish Safety and Reliability Association Summer Safety and Reliability Seminars*, Vol. 8, Nr 1, pp. 167–174.

Źródła internetowe

<http://korpusomat.pl/>

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

Streszczenie

W artykule omówiono wstępne wyniki badań nad materiałem językowym zawartym w stenogramach sejmowych polskiego parlamentu w latach 1918–2018. Stenogramy parlamentarne stały się podstawą dla Polskiego Korpusu Sejmowego (<http://clip.ipipan.waw.pl/PSC>) stworzonego w Zespole Inżynierii Lingwistycznej Instytutu Podstaw Informatyki Polskiej Akademii Nauk. W niniejszym przyczynku przedstawiono rezultaty początkowych prac lingwistycznych nad tym korpusem: omówiono nacechowanie emocjonalne losowo wybranych stenogramów z poszczególnych dziesięcioleci XX w. W podsumowaniu zawarto informacje dotyczące zbieżności zachodzącej w różnych dziesięcioleciach i w różnych okresach historii Polski (historii polskiego parlamentaryzmu) w zakresie konstruowania dyskursu parlamentarnego – jego emocjonalności liczonej poziomem ekspresji językowej.