

Corpus study of lexicon dynamics in Ukrainian-Polish economic and political relations

Nina Rizun^{1,†}, Olena Palatovska^{2,†}, Yevgenii Sova^{3,†}, Oleksandr Tomilin^{4,†}, Maria Shvedova^{4,5,†}

¹ Gdańsk University of Technology, 11/12 Gabriela Narutowicza, Gdańsk, 80-233, Poland

² Kyiv National Linguistic University, 73 Velyka Vasylkivska, Kyiv, 03680, Ukraine

³ National University of «Kyiv-Mohyla Academy», 2 Skovorody, Kyiv, 04070, Ukraine

⁴ National Technical University "Kharkiv Polytechnic Institute", 2 Kyrpychova, Kharkiv, 61002, Ukraine

⁵ Friedrich-Schiller-Universität Jena, 1 Fürstengraben, Jena, 07743, Germany

Abstract

The article examines the dynamics of the lexicon in the language of the Polish mass media in the context of economic and political relations between Poland and Ukraine during 2022-2023. The source of the language material is the monitoring corpus of Polish news MoncoPL. With the help of the MoncoPL corpus, the work identifies thematic groups of economic and political vocabulary, provides quantitative and qualitative characteristics of the use of the most frequent lexical units in the specified period, clarifies the temporal stratification of the use of certain thematic groups of vocabulary with social events, as well as monitors changes in economic and political narratives in Polish news texts for the last two years.

Keywords

lexicon dynamics, lexical unit, monitoring text corpus, MoncoPL, the language of the mass media, frequency, collocation, Ukrainian-Polish economic and political relations

1. Introduction

In modern research, it has already become an axiomatic statement that dynamism, variability, and mobility are the natural state of a language [1; 2; 3]. Vocabulary is the most mobile layer of a language system, and it constantly changes, reflecting the events and moods of society, trends in its development, multi-vector relations between countries and peoples. Therefore, the study of the dynamics of the lexicon is always relevant and attractive for a researcher.

World history knows many periods that are decisive for further life and, even, the very existence of both an individual country and humanity as a whole. These are the times Ukraine is going through right now, heroically defending its freedom, independence and right to exist.

CLW-2024: Computational Linguistics Workshop at 8th International Conference on Computational Linguistics and Intelligent Systems (CoLInS-2024), April 12–13, 2024, Lviv, Ukraine

* Corresponding author.

† These authors contributed equally.

✉ nina.rizun@pg.edu.pl (N. Rizun); olenapalatovska@gmail.com (O. Palatovska); ye.sova@ukma.edu.ua (Y. Sova); omegakharkov100@gmail.com (O. Tomilin); mariia.shvedova@khi.edu.ua (M. Shvedova)

ORCID 0000-0002-4343-9713 (N. Rizun); 0000-0002-6084-6505 (O. Palatovska); 0000-0001-8158-7031 (Y. Sova); 0000-0001-6168-4862 (O. Tomilin); 0000-0002-0759-1689 (M. Shvedova)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

Since the very beginning of the full-scale Russian military aggression against Ukraine, the majority of countries in the world have been supporting our state. The diversified assistance from Poland since the first days of the war was particularly important because neighbouring economic, political, cultural and linguistic contacts between our countries have continued for many centuries. For two recent years, events in Ukraine have been broadly covered every day in the columns of Polish periodicals; relations between Ukraine and Poland have intensified in all spheres of social life and have undergone significant changes, therefore we consider the study of the dynamics of the lexicon regarding Ukrainian-Polish relations in the language of the Polish mass media to be very relevant. With the advent of language electronic corpora, researchers have a unique opportunity to obtain valid quantitative and qualitative results regarding the use, variability and distribution of lexical units of the specified subject.

The goal of this article is to investigate with the help of MoncoPL monitoring corpus resources the dynamics of lexical units' usage in the language of Polish mass media in the context of economic and political relations between Ukraine and Poland. Two full years were chosen for the analysis - from January 2022 to December 2023. It is known that media discourse reacts to social changes the fastest; therefore, it is indicative for the study of dynamic processes in the modern Polish lexicon.

The implementation of the specified goal necessitated the following tasks: 1) to demonstrate the possibilities and advantages of the MoncoPL monitoring corpus for studying the dynamics of the lexical units' usage in a clearly defined period of time; 2) to outline the thematic groups of vocabulary characterizing the economic and political relations between Ukraine and Poland in the language of the Polish mass media in the recent two years compared to pre-war period (2021); 3) to provide quantitative and qualitative characteristics of the most frequent lexical units' usage in the specified period; 4) to find out the temporal stratification of the use of certain thematic groups of vocabulary with social and political events.

The study of the lexicon dynamics in the context of economic and political relations between Ukraine and Poland was carried out on the materials of MoncoPL, the daily updated corpus search system of the Polish language (monco.frazeo.pl) [4].

2. Related Works

The problems of dynamic processes occurring in the vocabulary during a certain period of time have repeatedly been in focus of researchers' attention [5, 6, 7, 8, 9, 10, and many others]. The vast majority of works on the mentioned topic cover significant historical periods of time (from several centuries to several decades) and describe long-term diachronic processes of lexicon formation. They are based on the material of lexical units selected, as a rule, manually from dictionaries and texts of various functional styles and genres.

The study of the dynamics of lexical units' usage in a relatively short recent period is practically impossible to carry out engaging traditional methods, because it is difficult to make a representative sample for analysis in the absence of specialized dictionaries and a sufficient number of texts.

The advent of language corpora has significantly changed the situation and the ways and methods of researching language units [11]. Today, representative large language corpora have become the main resource not only in fundamental language research, they are used for applied

purposes by translators, editors and students to check the frequency of use, stylistic correctness, and communicative appropriateness of using language units [12, 13, 14].

Massive modern language corpora of the open type, which are constantly updated and can quickly fulfil requests, allow monitoring research on the state, changes and functioning of language units in various functional styles and genres, as well as individual styles of chosen authors. This makes it possible to study not only the current state of a modern language, but also to observe the processes that continuously take place in the language [15, 16].

These corpora include MoncoPL, the only daily updated search system in Poland, which consists mainly of Internet media texts by leading Polish portals on various thematic profiles (e.g. onet.pl, interia.pl, Naszamiasto.pl, kulturonline.pl, sejm.gov.pl, egospodarka.pl, plotek.pl) or press, television and radio websites (e.g. newspaper ta.pl, Fakt.pl, tvn24.pl, tvp.info, rmf24.pl).

MoncoPL is a unique monitoring corpus of an open structure, with the help of which you can make representative samples of the topical use of words, phrases as well as lexical and grammatical patterns in authentic media texts in the modern Polish language. Today, the MoncoPL search engine index contains more than 8 billion words and is constantly updated. P. Pezik, describing in detail the structure, possibilities and features of MoncoPL application, writes that the main advantage of using this corpus is the possibility of creating a large sample of data for the purpose of studying the appearance of neologisms, neosemanticisms and borrowings in the language, studying the distributional properties of rare words or the variability of word collocations in online media [17]. B. Duda and K. Lisczyk write about the use of the MoncoPL search engine for didactic purposes, emphasizing that "when collecting material <...>, concerning the latest vocabulary, it is a better source from the point of view of research than, for example, Google, Internet searching system" [18].

Our research is focused on studying the dynamics of the lexicon in the language of Polish mass media "here and now". The main linguistic characteristics of modern online media texts are the rapid enrichment of the vocabulary and the penetration of spoken language into written speech [19, 20]. Researchers of media communication have repeatedly raised the question of "to what extent the media shape communicative behaviour, and with the help of which verbal means they influence the conceptualization of reality in society" [20]. The MoncoPL corpus search system provides a unique opportunity to track the dynamics of the lexicon of the Polish mass media in the context of economic and political relations between Ukraine and Poland over the past two years and to find out the temporal stratification of the use of certain thematic groups of vocabulary with corresponding events in society. The research results presented in the article are obtained on the basis of corpus data, relate to the quantitative and qualitative analysis of the lexical units' usage within the specified thematic groups and allow to determine the vectors of their dynamics and impact on society.

3. Methods and Material

To study the dynamics of the use of lexical units in the context of economic and political relations between Ukraine and Poland, the corpus analysis toolkit of the MoncoPL search engine was used, because it, unlike other corpora, consists of daily updated Polish news materials. On the one hand, this makes MoncoPL an indispensable resource for monitoring studies of the use of lexical units in a certain period of time, on the other hand, a huge increase

in the amount of accumulated information, its processing, systematization and classification require the use of a complex of highly effective methods.

The research fully reflects the modern interdisciplinary state of the study and description of linguistic units, which led to a combination of different methods of researching actual material: both the latest (corpus and quantitative methods, the use of the Copilot application from Microsoft, which is based on large language models of artificial intelligence), and traditional ones, including general scientific method, method of expert evaluations, descriptive method, etc.

The headlines and texts of publications about Polish-Ukrainian relations from news portals popular in Poland served as the actual material for the research of the topics of news and the corresponding vocabulary in the Polish mass media. The advantage of the selected resources is that they contain separate sections devoted to Polish-Ukrainian relations, which contributes to the practical implementation of the research goals. Another important characteristic of some of the selected information portals is that they publish news about Ukraine on a daily basis, ensuring sufficient coverage of news materials of various thematic orientations:

- Section "Polish-Ukrainian relations" on the Wirtualna Polska web portal
- Section "Polish-Ukrainian relations" on the Wyborcza.pl web portal
- Dedicated to Ukraine selection of daily news on the Wirtualna Polska web portal;
- A selection of daily news devoted to the economic activity and life of Ukrainian refugees in Poland.

As mentioned above, media publications can best serve as a source for studying the dynamics of the lexicon in short periods of time, and the vocabulary of mass media texts is indicative for determining the development of economic and political changes and moods in society. Therefore, it was decided to monitor the frequency of use of a set of lexical units formed on the basis of the analysis of thematic publications on popular Polish news portals of the MoncoPL corpus dedicated to Ukraine, which appeared during the last two years (January 2022 - December 2023). The obtained results were compared with the data of the previous year (January - December 2021).

So, for example, the number of mentions of different word forms of the lexeme Ukraine in the monitoring corpus of Polish news MoncoPL during the analyzed period is presented in the graph below (Figure 1). The graph shows that in 2021 there were on average less than 8,000 mentions of Ukraine in the texts of Polish news every month during January-November 2021 and about 13,000 mentions in December 2021. Quite expectedly, a jump in mentions is observed in February-April 2022 (with the peak in March 2022 – 217,023 mentions), it stabilizes at a high-level during May 2022 - March 2023, then gradually declines from April 2023, but remains quite high in the Polish media space compared to 2021.

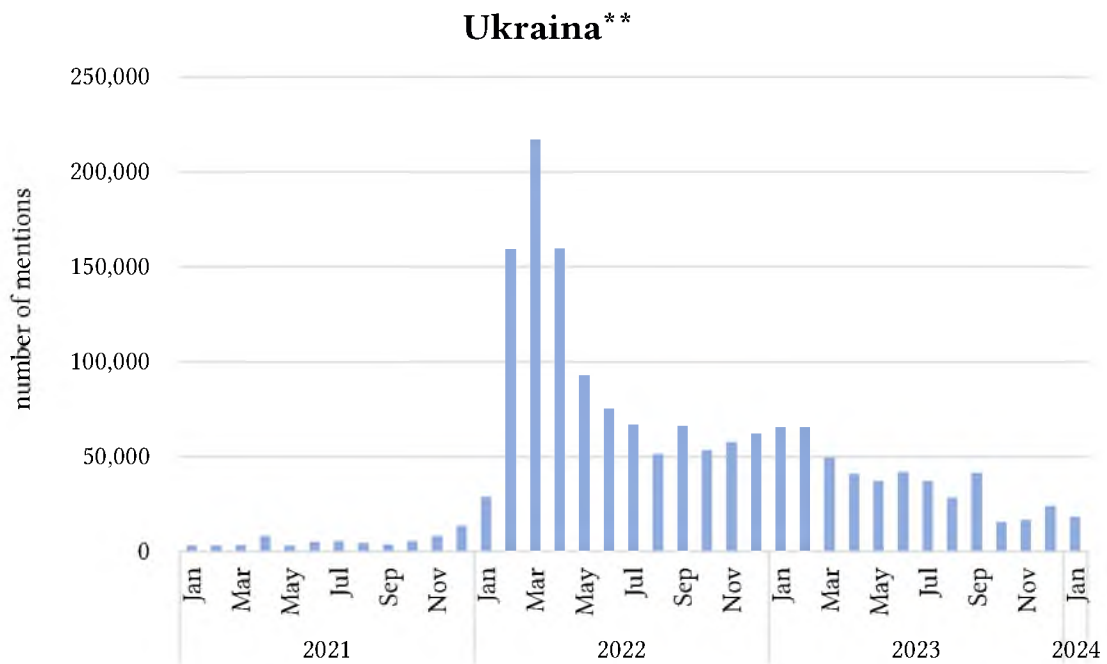


Figure 1: Dynamics of mentioning the lexeme *Ukraina* (Ukraine) in the news materials of the MoncoPL monitoring corpus, where the number of uses in the news texts of the corpus is represented on the vertical axis.

The work on the research took place in the following stages:

1. Qualitative expert analysis by the authors of the work was used to determine the thematic groups of news and vocabulary that was collected from the specified sources at different time intervals. Traditionally, in linguistics, the method of expert evaluations allows obtaining sufficiently objective results based on the totality of individual opinions of the researchers.
2. Additionally (to confirm the results of expert analysis), the Copilot tool from Microsoft (GPT- 4 model from OpenAI) was used, which, based on artificial intelligence algorithms trained on large data sets, allows, among other things, to analyze and summarize the content of large amounts of textual information. In the research process, the Copilot tool allowed the authors to synthesize the main topics of the news on the basis of automated analysis of headlines and brief descriptions of news from the selected resources.
3. Based on the results of the research at the previous stages, a sample of more than 100 words and phrases, which were most often used in the context of the six identified thematic groups, was formed.
4. The toolkit of the MoncoPL corpus was used to objectively and systematically check the frequency of use of each set of lexical units from the formed sample. Search queries for the corpus database were formed iteratively taking into account lexical units' collocations, which made it possible to exclude inappropriate examples of their use (for example, similar proper names or random word collocations that do not correspond to

the context of the study). This approach made it possible to significantly improve the quality of requests, which is important in frequency analysis to ensure sufficient accuracy of the results.

At the same time, it is necessary to pay attention to certain limitations regarding the completeness and objectivity of the obtained results, because the process of selection, analysis and synthesis of political and economic topics of news in the Polish media space was not continuous, but was carried out only on the above-mentioned news portals; the application of an expert approach to the definition of thematic groups of the vocabulary by the authors of the study is also somewhat subjective.

The results of the analysis of lexical units' usage may be slightly biased due to the technical complexity of processing a large volume of material; therefore, it is advisable to perceive the indicators of the frequency of word usage as indicative values of the relative frequency of mentions, but not as absolute values.

By the *indicative number*, we mean the number of times a particular phrase is used in different combinations, but this number should not be taken as an exact value. We use the term "indicative number" to represent general trends, and not to reflect the absolute frequency of a certain lexical unit. This is due to the specifics of using the news monitoring corpus, in which, unlike reference text corpora, the number of uses increases every day, the vocabulary is less standardized, more variable, so it is difficult to take into account all the criteria, restrictions and cases of word usage when forming a search query. There may be hundreds, thousands or tens of thousands of mentions for a certain word or phrase. Accordingly, it is difficult to check whether there are some examples of homonyms, proper names, etc., that do not correspond to the sought context. After all, when forming a query, it is necessary to achieve a balance between narrowing it so that unnecessary words do not enter the analysed quantity, and expanding it in order to cover as many variants of mentions of a certain word as part of word collocations in different grammatical forms and at different distances from each other (for example: help for refugees; refugees' help; assistance aimed at protecting refugees, etc.). This also applies to analyzing themes and changing narratives into mentions of words across different time periods. For this, the frequency dynamics was analyzed. And then, during periods of sudden changes in the number of uses of a certain lexical unit, the contexts of its use were selectively reviewed.

Despite the indicated limitations, we believe that the proposed method of analyzing factual material allows us to generalize the key topics and track the dynamics of the lexicon of economic and political news related to Ukraine on popular media resources in Poland during a certain period of time.

Working with such a monitoring corpus as MoncoPL allows you to track the natural course of changes in the language vocabulary and draw conclusions about the textual functioning of lexical units and their temporal stratification with social and political events, taking into account objective quantitative indicators.

4. Results and Discussion

Since the beginning of the full-scale invasion by Russia, Poland has become one of the most important strategic partners of Ukraine. Polish support and assistance were exceptional and took place in all spheres of life. Bilateral relations between the countries experienced an

unprecedented rise, which, of course, was reflected in the language of the mass media. In the news publications of the first half of 2022, the prevailing vocabulary denoted:

- words that have acquired symbolic semantics to indicate the heroism and national identity of Ukrainians (*Azowstal*; *czzerwona kalina*; *Dobry wieczór, jesteśmy z Ukrainy*; *pies Patron*; *Sily Zbrojne Ukrainy (ZSU)*; *palanyca*; *oborona terytorialna*);
- names and types of weapons (*Bajraktar*, *Kindzal*, *HIMARS*, *Stuhna*, *NASAMS*, *Patriot*, *szahed*, *dron kamikadze*);
- the newest realities of wartime (*brudna bomba*, *gauleiter*, *filtracja*, *denazifikacja*, *Lend-Lease*, *Mordor*, *orkowie*, *ruskij wojennyj korabl*);
- Ukrainian toponyms primarily from the places of battlefield (*Awdijiwka*, *Bachmut*, *Balaklija*, *Borodzianka*, *Buczna*, *Hostomel*, *Kachowka*, *Melitopol*);
- ironic depictions of Russian propaganda rhetoric (*biolaboratoria*, *bojowe komary*, *denazifikacja*, *gest dobrej woli*, *nie majacy analogow*, *czmobiki*, *druga armia swiata*, *zombirowani*).

Most of these lexemes represent neologisms for the modern Polish lexicon, but the high frequency of their use in the Polish mass media language gives reasons to suggest the possibility of further fixation of some of the indicated words and expressions in the Polish linguistic picture of the world.

Let us investigate the temporal dynamics of Ukrainian neologisms' usage by the example of the words and expressions most used in news publications, which symbolize heroism and national identity of Ukrainians (Figure 2):

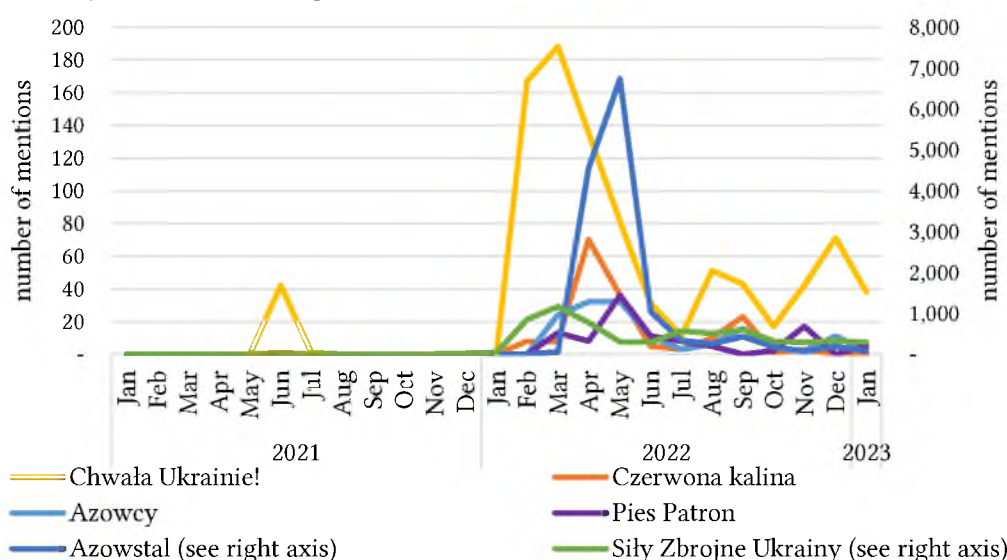


Figure 2: The frequency of use of words and expressions that have acquired symbolic semantics to denote the heroism and national identity of Ukrainians in Polish news during February 2022–January 2023.

Note. The right vertical axis is created for lexical units, the frequency of which significantly exceeds the use of most analyzed words. In the description of the schedule, the following words are indicated separately: *Azowstal*, *Sily Zbrojne Ukrainy*.

As depicted in Figure 2, the use of this vocabulary sharply rises in the first half of 2022, then gradually declines, but does not completely disappear from the Polish mass media even in January 2023.

However, gradually the relations between Ukraine and Poland underwent significant changes, a political crisis began. International expert S. Zhelikhovsky notes: "In general, relations between Ukraine and Poland can be divided into strategic and tactical levels: strategically, Warsaw has never stopped supporting Ukraine; however, there is a crisis at the tactical level that needs to be resolved" [21]. Accordingly, the focus of Polish mass media is beginning to shift to the plane of economic and political topics.

Monitoring of the Polish media texts vocabulary from the specified sources for two war years (January - December 2022 and January - December 2023) made it possible to determine the following economic and political trends in publications during this period:

- **January - December 2022:** financial assistance to Ukrainian refugees; possibility to exchange hryvnas in Polish banks; special preferential conditions for train travel for refugees from Ukraine; impact of refugees from Ukraine on the Polish labour market; health care funding for the refugees; education for Ukrainian children in Polish schools; a potential union between Ukraine, Poland, Great Britain, as well as the USA and the Baltic states; receiving funding from the EU to help refugees from Ukraine by Poland; gradual increase in the number of refugees from Ukraine and their social integration.
- **January - December 2023:** growth in the number of registered Ukrainian enterprises in Poland; plans of many refugees to return to Ukraine; transit of grain from Ukraine through the territory of Poland; features of the integration of refugees from Ukraine and their impact on the labour market; survey of Polish citizens regarding the attitude towards refugees in the context of their impact on the labour market; relations between the countries under the influence of the election campaign in Poland; introduction of restrictions on the import of agricultural products from Ukraine; the assurance of the continuation of the support provided by Poland to Ukraine and emphasizing its important role in maintaining peace in Europe; blocking the trucks on the Ukrainian-Polish border due to farmers' protests; declaration on unblocking the borders; analysis of relations between countries based on the results of the year; agreement on re-launching and further improvement of the relationship between the two countries.

In the process of analyzing Polish news texts, a sample of the most frequent words and word collocations used in the context of economic and political relations between Ukraine and Poland was formed. The selected lexical units, in turn, were distributed according to the following thematic groups:

- **Group 1. Refugee issues:** *refugees (uchodźcy), refugee relocation (relokacja uchodźców), humanitarian aid (pomoc humanitarna), aid package (pakiet wsparcia), system aid (pomoc systemowa), immigration reform (reformacja imigracyjna).*
- **Group 2. Economic issues:** *impose sanctions (nałożyć sankcje), embargo (embargo), hryvna (hrywna), increase in interest rates (wzrost stóp procentowych), risk of recession (ryzyko recesji), import restrictions (ograniczenia importu), export restrictions (ograniczenia eksportu), commercial permit (zezwolenie komercyjne), trade liberalization*

(liberalizacja handlu), economic interest (interes ekonomiczny), Ukrainian capital (kapitał ukraiński).

- **Group 3. Polish-Ukrainian relations:** Polish-Ukrainian union (sojusz polsko-ukraiński); Polish-Ukrainian solidarity (solidarność polsko-ukraińska); Polish-Ukrainian relations (stosunki polsko-ukraińskie); Polish-Ukrainian Union (unia polsko-ukraińska).
- **Group 4. Agricultural issues:** agricultural products (produkty rolne), grain import/export (import/eksport zboża), grain crisis (kryzys zbożowy), grain influx (napływ zboża), grain transit (tranzyt zboża), transit of agricultural products (tranzyt produktów rolnych), trade in agricultural products (handel produktami rolnymi), agricultural infrastructure (infrastruktura rolnicza), grain conflict (konflikt zbożowy).
- **Group 5. Logistics and transport:** Baltic Pipe, transporters/farmers' protests (protesty przewoźników/rolników), transport industry (branża transportowa), truck queues (kolejki ciężarówek), electronic queue (elektroniczna kolejka), kilometer queue (kilometrowa kolejka), logistics hub (hub logistyczny), sea corridor (korytarz morski), blocking the border (blokada granicy), unblocking the border (odblokowanie granicy).
- **Group 6. Reconstruction and development of Ukraine:** reconstruction of Ukraine (odbudowa Ukrainy), energy infrastructure (infrastruktura energetyczna), billions [amounts of funds] for Ukraine (miliardy [środków] dla Ukrainy), military and financial aid (pomoc wojskowa i finansowa), process entry (proces akcesyjny), conditions of competition (warunki konkurencji).

The complete list of lexical units selected from the MoncoPL news monitoring corpus and sorted by decreasing number of mentions is shown in Figure 3. In addition, the frequency of their use is presented in percentage terms and marked with a certain colour to visually represent the share of mentions in each of the analysed periods. The percentage of lexical units whose number of mentions did not exceed 10% in a given period is marked in red; shades of yellow mark the share up to 40%; shades of green mark the share above 40%.

The seven most frequently mentioned lexical units in the news in 2022-2023 (top 7) are shown in Table 1 in comparison with their usage in 2021, and the dynamics of their usage during the analysed period can be seen in Figure 4. These are the following words and phrases: *refugees* (uchodźcy), *humanitarian aid* (pomoc humanitarna), *impose sanctions* (nałożyć sankcje), *embargo* (embargo), *agricultural products* (produkty rolne), *Baltic Pipe* (Baltic Pipe), *reconstruction of Ukraine* (odbudowa Ukrainy).

№	Lexical units	Total number of mentions	Distribution by period, %		
			January - December 2021	January - December 2022	January - December 2023
1	Refugees / Uchodźcy	167.966	14%	71%	15%
2	Humanitarian aid / Pomoc humanitarna	26.428	14%	56%	30%
3	Impose sanctions / Nalożyć sankcje	19.024	20%	62%	18%
4	Embargo / Embargo	17.550	4%	66%	30%
5	Agricultural products / Produkty rolne	14.058	23%	28%	48%
6	Baltic Pipe / Baltic Pipe	9.326	27%	65%	9%
7	Reconstruction of Ukraine / Odbudowa Ukrainy	7.989	0%	51%	49%
8	Energy infrastructure / Infrastruktura energetyczna	7.069	13%	59%	28%
9	Hryvnia [Ukrainian currency] / Hrywna	6.065	11%	70%	20%
10	Grain import/export / Import/eksport zboża	4.713	1%	46%	53%
11	Protests by carriers/farmers / Protesty przewoźników/rolników	2.509	21%	18%	61%
12	Aid package / Pakiet wsparcia	1.890	14%	46%	41%
13	Transport industry / Branża transportowa	1.558	38%	26%	36%
14	Increase in interest rates / Wzrost stóp procentowych	1.248	6%	80%	14%
15	World Trade Organization / Światowa Organizacja Handlu	1.190	26%	17%	57%
16	Truck queues / Kolejki ciężarówek	1.073	26%	59%	15%
17	Refugee relocation / Relokacja uchodźców	901	8%	44%	48%
18	Grain crisis / Kryzys zbożowy	804	0%	5%	95%
19	Accession process / Proces akcesyjny	798	5%	57%	38%
20	Influx of grain / Napływ zboża	643	0%	4%	96%
21	Grain transit / Tranzyt zboża	618	0%	6%	94%
22	Border blockade / Blokada granicy	466	14%	26%	60%
23	Maritime corridor / Korytarz morski	428	0%	68%	32%
24	Polish-Ukrainian alliance / Sojusz polsko-ukraiński	427	3%	54%	43%
25	Risk of recession / Ryzyko recesji	415	1%	72%	27%
26	Billions [of funds] for Ukraine / Miliardy [środków] dla Ukrainy	413	0%	65%	35%
27	Systemic assistance / Pomoc systemowa	374	24%	64%	13%
28	Import restrictions / Ograniczenia importu	334	8%	53%	38%
29	Electronic queue / Elektroniczna kolejka	324	1%	2%	98%
30	Kilometer-long queue / Kilometrowa kolejka	322	21%	57%	22%
31	Export restrictions / Ograniczenia eksportu	302	20%	50%	30%
32	Polish-Ukrainian solidarity / Solidarność polsko-ukraińska	285	5%	81%	14%
33	Polish-Ukrainian relations / Stosunki polsko-ukraińskie	279	4%	48%	49%
34	Commercial permit / Zezwolenie komercyjne	275	0%	0%	99%
35	Trade liberalization / Liberalizacja handlu	234	6%	35%	59%
36	Logistics hub / Hub logistyczny	195	22%	46%	32%
37	Conditions of competition / Warunki konkurencji	191	35%	32%	34%
38	Transit of agricultural products / Tranzyt produktów rolnych	179	0%	1%	99%
39	Military and financial aid / Pomoc wojskowa i finansowa	179	1%	59%	40%
40	Border unlocking / Odblokowanie granicy	172	19%	7%	74%
41	Trade in agricultural products / Handel produktami rolnymi	147	17%	53%	30%
42	Polish-Ukrainian union / Unia polsko-ukraińska	130	0%	73%	27%
43	Economic interest / Interes ekonomiczny	128	37%	41%	22%
44	Ukrainian capital / Kapital ukraiński	109	2%	22%	76%
45	Agricultural infrastructure / Infrastruktura rolnicza	54	4%	54%	43%
46	Immigration reform / Reforma imigracyjna	39	13%	18%	69%
47	Grain conflict / Konflikt zbożowy	37	0%	14%	86%

Figure 3: The list of lexical units included in the sample with an indicative number of occurrences in the MoncoPL news monitoring corpus.

Table 1

Top 7 lexical units according to the number of their mentions in 2022-2023 compared to 2021 in MoncoPL

Lexical units	January-December 2021	January-December 2022	January-December 2023
<i>Refugees / Uchodźcy</i>	22,763	120,094	25,109
<i>Humanitarian aid / Pomoc humanitarna</i>	3,745	14,879	7,804
<i>Impose sanctions / Nałożyć sankcje</i>	3,791	11,752	3,481
<i>Embargo / Embargo</i>	683	11,623	5,244
<i>Agricultural products / Produkty rolne</i>	3,296	3,997	6,765
<i>Baltic Pipe / Baltic Pipe</i>	2,494	6,028	804
<i>Reconstruction of Ukraine / Odbudowa Ukrainy</i>	1	4,074	3,914

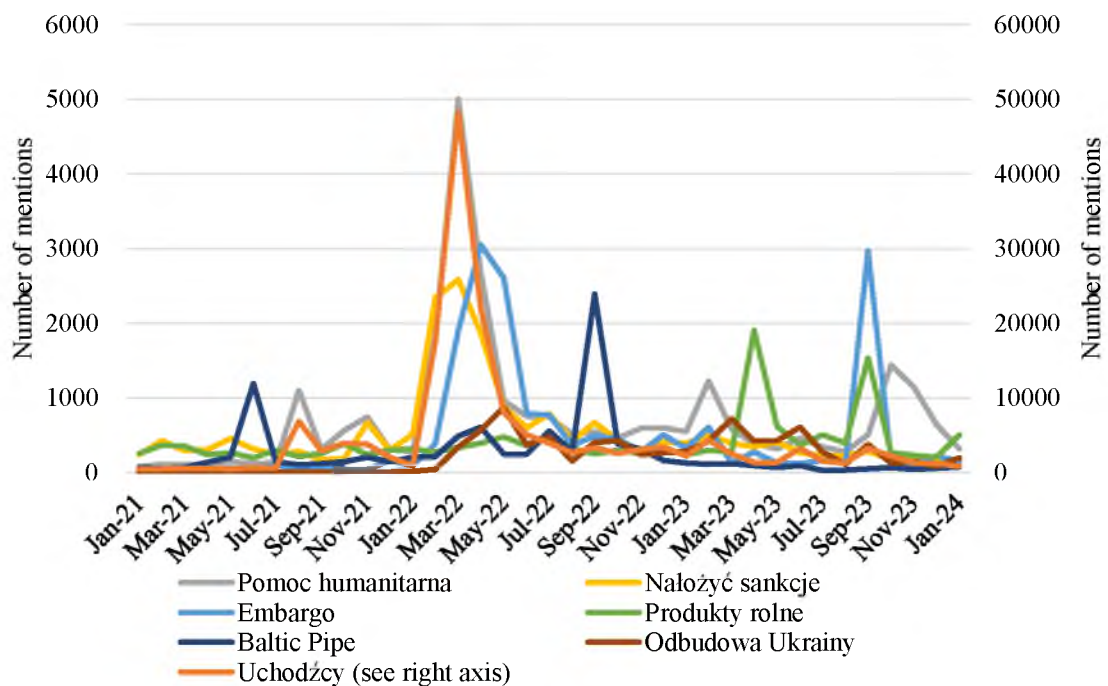


Figure 4: Dynamics of lexical units' usage (top 7) in the context of economic and political relations between Ukraine and Poland in MoncoPL.

Note. The right vertical axis is created for the lexical unit *Uchodźcy* (Refugees), the frequency of which significantly exceeds the use of most of the analyzed words

Let us consider in more detail the dynamics of the use of economic and political vocabulary by the example of the lexemes *uchodźcy* (refugees) and *embargo* (embargo).

As can be seen from Table 1 and Figure 4, the issue of refugees remains relevant in the news texts throughout the analysed period. During this period, the lexeme *uchodźcy* (refugees) was used 167,966 times. However, the contexts of its usage have change significantly since 2022.

The frequency of the lexeme *uchodźcy* usage in 2021-2023 according to MoncoPL is shown in the graph below (Figure 5).

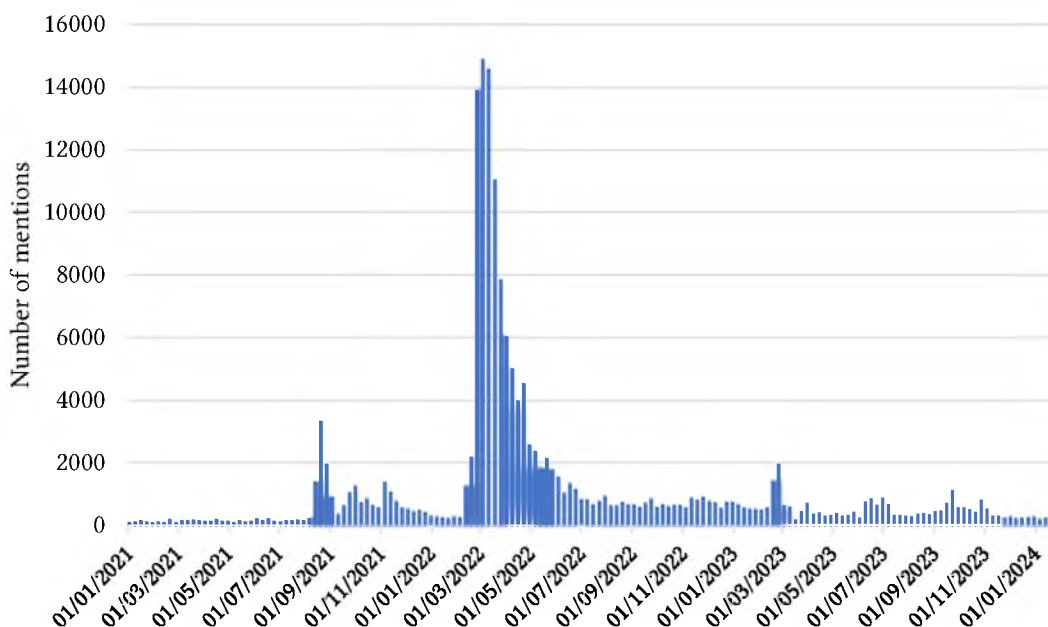


Figure 5: Frequency of the lexeme *Uchodźcy* (Refugees) usage during 2021-2023 in MoncoPL.

In 2021, the lexeme *uchodźcy* refers primarily to refugees from Belarus and Afghanistan. Hence, for example, the peak of mentions of *uchodźcy* in 2021 falls in August due to concern of the Polish community regarding new waves of refugees from the Middle East, particularly Afghanistan, after the withdrawal of US troops. Example:

- ... gdański samorząd gotów jest wspierać działania rządu, jeśli ten sprowadzi afgańskich **uchodźców** do Polski (Wiadomosti.wp.pl, 2021);
- Wysoki rangą przedstawiciel Komisji Europejskiej ostrzega tymczasem, że konflikt ten grozi pojawieniem się nowego kryzysu **uchodźczego**./.../ Według przedstawiciela Komisji Europejskiej wkrótce około pół miliona mieszkańców Afganistanu będzie próbować uciec ze swojego kraju (tvn24.pl, 2021).

Another topic of the migration crisis in Poland in 2021 is refugees on the Polish-Belarusian border. Example:

- Straż graniczna informuje o rekordowej liczbie **uchodźców** przybywających do kraju. Migranci docierają na granicę polsko-białoruską w ramach wojny hybrydowej... (gazetaprawna.pl, 2021).

Since February 2022, the lexeme *uchodźcy* (refugees) usually designates refugees from Ukraine who left their homes as a result of Russian military aggression: January-December

2022 – 120,094 mentions. As can be seen from Figure 4, the peak frequency falls on March 2022, when the operation of evacuation trains was established and people began to leave the dangerous territories en masse, primarily to Poland. Polish volunteers met them at the border and helped in every possible way. Example:

- *Uchodźcom z Ukrainy 24 godziny na dobę pomagają wolontariusze, którzy, często bez żadnego doświadczenia, przyjechali tu z całej Polski (Wiadomosci.onet.pl, 2022).*

The topic of refugees remains relevant in 2023 - 25,109 mentions. But the frequency of the lexeme *uchodźcy* returns almost to the level of 2021 and primarily concerns financial aid and the integration of Ukrainians into Polish society. Example:

- *Równoległe powinny zostać stworzone ramy formalne i finansowe dla mądrego pomagania, które pozwoli władzom lokalnym oferować wsparcie w integracji i aktywizacji uchodźców (gazetaprawna.pl, 2023).*

In 2021, the lexeme *embargo* is mentioned in the Polish media 683 times and refers to various events in the world. Further, the peak of mentions occurs in the first half of 2022 (for the year 2022 – 11,623 mentions in total) and, as we can see from the most frequent collocations (Figure 6), concerns primarily Russian products (oil, gas, etc.). Example:

- *Parlament Europejski przyjął rezolucję, w której postulowane jest pełne embargo na import ropy, węgla, paliwa jądrowego i gazu z Rosji (gazetaprawna.pl, 2022).*

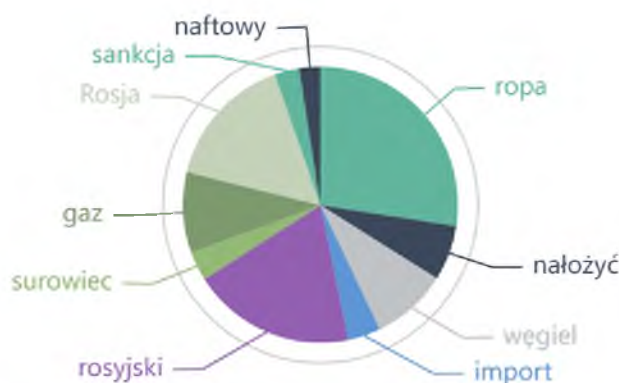


Figure 6: The most frequent collocations to the lexeme *Embargo* in 2022 in MoncoPL.

In 2023, a significant part of the lexeme *embargo* mentions refers to trade and economic restrictions on the grain import from Ukraine (Figure 7). Example:

- *Komisja Europejska nie przedłużyła embargo na ukraińskie zboże dla pięciu krajów - Polski, Węgier, Słowacji, Bułgarii i Rumunii. Po tej decyzji Warszawa, Budapeszt i Bratysława przedłużyły zakaz importu produktów rolnych z Ukrainy (wnp.pl, 2023).*

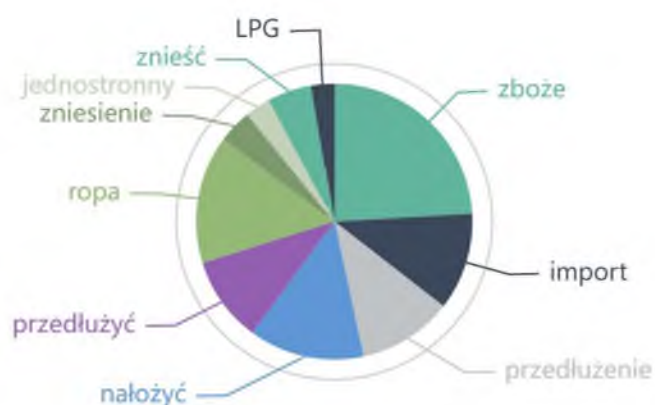


Figure 7: The most frequent collocations to the lexeme *Embargo* in 2023 in MoncoPL.

Figure 8 lists the lexical units that, according to MoncoPL, became frequent in Polish news texts only in 2023. Their frequency is presented as a percentage and marked by a certain color by period: red colour marks 0-6%, orange colour marks 14% (2021-2022), and green colour marks 86-99% (2023). As can be seen, all these phrases refer to events related to the ban on the import and transit of Ukrainian grain across the Polish border and the blockade of border checkpoints by Polish drivers and farmers: *transit of agricultural products (tranzyt produktów rolnych)*, *commercial permit (zezwolenie komercyjne)*, *electronic queue (elektroniczna kolejka)*, *influx of grain (napływ zboża)*, *grain crisis (kryzys zbożowy)*, *grain transit (tranzyt zboża)*, *grain conflict (konflikt zbożowy)*.

№	Lexical units	Total number of mentions	Distribution by period, %		
			January - December 2021	January - December 2022	January - December 2023
1	Grain crisis / Kryzys zbożowy	804	0%	5%	95%
2	Influx of grain / Napływ zboża	643	0%	4%	96%
3	Grain transit / Tranzyt zboża	618	0%	6%	94%
4	Electronic queue / Elektroniczna kolejka	324	1%	2%	98%
5	Commercial permit / Zezwolenie komercyjne	275	0%	0%	99%
6	Transit of agricultural products / Tranzyt produktów rolnych	179	0%	1%	99%
7	Grain conflict / Konflikt zbożowy	37	0%	14%	86%

Figure 8: Top-7 lexical units reflecting the crisis of economic and political relations between Ukraine and Poland, in the texts of Polish mass media in MoncoPL (in %).

The quantitative calculations shown in Figure 8 convincingly prove that the problems related to the grain crisis and the transit of Ukrainian agricultural products became more acute in 2023 and led to a political crisis between Poland and Ukraine. In 2021-2022, the given lexical units were either not mentioned at all, or their usage was sporadic.

5. Conclusions

The language of mass media righteously occupies a leading place in modern processes of language dynamics, in particular in the lexical units' usage. Journalism is the most responsive

to social events, moods, sentiments and changes. Moreover, the processes of influence in the coordinates "language - society" are two-way. On the one hand, the events taking place in the world are quickly reflected in newspaper and magazine publications, stimulating the attraction of new lexical units and forming the emergence of new trends in their use. On the other hand, the appearance of high-frequency lexical units in the columns of media publications is one of the most influential factors in the formation of public opinion, which leads to certain consequences and changes in relations between individuals, political forces, and even countries.

To study the latest dynamics in the use of lexical units of economic and political orientation, we consider it absolutely appropriate and justified to turn to news texts of a daily updated monitoring corpus such as MoncoPL. The analysis of the frequency of a certain word or stable expression usage, the formation of collocations makes it possible to imagine a generalized picture of the language activity that is happening right now. It can be noted that the significance of the frequency of lexical units' usage has considerably increased recently thanks to the use of language corpora. Thanks to the huge volumes of actual material analyzed, its indicators are able to provide objective and qualitative information about the dynamics of the lexicon, real models of usage and functioning of vocabulary in authentic texts, their compatibility with other units of natural language, as well as their influence on changes in the linguistic picture of the world of a certain community.

With the help of the opportunities provided by the MoncoPL monitoring corpus, the conducted analysis made it possible to study the dynamics of the use of lexical units that characterize changes in the economic and political relations between Ukraine and Poland over the recent two years in the following way: 1) thematic groups of vocabulary in the language of the Polish mass media were outlined in 2022-2023 in comparison with pre-war 2021; 2) the temporal stratification of the use of certain thematic groups of vocabulary with social events is clarified; 3) on the basis of the quantitative and qualitative characteristics of the most frequent lexical units usage in the specified period, changes in narratives in Polish news texts were monitored.

Therefore, investigating the dynamics of the lexicon in this way makes it possible to reveal the latest trends in language conceptualization and categorization of the world of a certain language community and its quick response to changes in economic and political events on the one hand; and on the other hand, it helps to monitor the influence of media language on the formation of sentiments in society.

References

- [1] A. O. Biletsky, *Mova – dynamichna znakova systema* [Language – a Dynamic Sign System], in: A.O. Biletsky, *Vybrani pratsi* [Selected Works]. Kyiv, 2012, pp. 529–535.
- [2] S. V. Semchynsky, *Statyka i dynamika movy. Funktsionuvannya movy* [Statics and Dynamics of Language. Language Functioning], in: S. V. Semchynsky, *Zahal'ne movoznavstvo* [General Linguistics], Kyiv, 1988, pp. 251–262.
- [3] A. Martinet, *Fonction et dynamique des langues* [Function and Dynamics of Languages], Paris, 1989.
- [4] MoncoPL: Wyszukiwarka korpusowa Monco. URL: <https://monco.frazeo.pl>.
- [5] S. Grytsenko, *Statics and Dynamics of the Language: the Issue of Balance*, *Studia Linguistica*, 13 (2018): 48-70. doi: 10.17721/StudLing2018.13.47-70.

- [6] K. Kleszczowa, Tajemnice dynamiki języka [Secrets of Language Dynamics], Katowice, 2012.
- [7] N. F. Klymenko, E. A. Karpilovska, L. P. Kislyuk., Dynamichni protsesy v suchasnomu ukrayinskomu leksykoni [Dynamic Processes in the Modern Ukrainian Lexicon]: monograph, Kyiv, 2008.
- [8] M. I. Navalna, Dynamika leksyky v suchasnykh ukrayins'kykh masmedia [Dynamics of Vocabulary in Modern Ukrainian Mass Media], REICST, pp. 81–84, Nov 2021.
- [9] O. A. Styshov, Ukrayinska leksyka kintsya XX stolittya (na materialy movy zasobiv masovoyi informatsiyi) [Ukrainian Vocabulary of the Late 20th Century (based on the language of mass media)], Kyiv, 2005.
- [10] H. Jadacka, System slowotwórcy polszczyzny (1945-2000) [The Word-Former System of the Polish Language (1945-2000)], Warszawa: PWN, 2001.
- [11] Z. Kunch, L. Kharchuk & M. Hnatyuk, Research of Medical Terms-substantives in O. Zabuzhko's Artistic Works on the Basis of the Text Corpus, COLINS-2023: 7th International Conference on Computational Linguistics and Intelligent Systems, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine, pp. 359-372.
- [12] A. R. Gabidullina, O. L. Kolesnichenko, Methodology of Modern Linguistic Research: a Study Guide, Sloviansk, 2019.
- [13] O. L. Kapinus, I. I. Bailo, The Use of the Text Corpora in Linguistic Research, Transcarpathian Philological Studies, 25, 1 (2022). doi: 10.32782/tps2663-4880/2022.25.1.44.
- [14] M. Ogodniczuk, L. Gorski, M. Lazinski, P. Pezik, From the National Corpus of Polish to the Polish Corpus Infrastructure, Journal of Linguistics, 70, 2 (2019): 315 – 323. doi: 10.2478/jazcas-2019-0061.
- [15] V. Stariko, O. Synchak, O. Feminine Personal Nouns in Ukrainian: Dynamics in a Corpus. COLINS-2023: 7th International Conference on Computational Linguistics and Intelligent Systems, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine, 2023, pp. 407-425.
- [16] O. Palatovska, E. Sova, Neologisms of Ukrainian Origin in Polish News: January 2021 – January 2023 (Corpus Research), Linguistic Studies, 58 (2023): 268–279. doi: 10.34142/23127546.2023.58.20.
- [17] P. Pęzik, Budowa i zastosowania korpusu monitorującego MoncoPL [Design and Applications of MoncoPL as a Monitor Corpus of Polish]. Forum Lingwistyczne, 7, 7 (2020): 133-150, doi: <https://doi.org/10.31261/FL.2020.07.11>.
- [18] B. Duda, K. Liszyk, Narzędzia cyfrowe w polonistycznej dydaktyce akademickiej – zastosowania, możliwości, perspektywy [Digital Tools in the Polish Language Academic Didactics – Applications, Possibilities, Prospects], Forum Lingwistyczne, 5 (2018): 143–154.
- [19] J. Jarosz, R. Opiłowski, P. Staniewski, Lingwistyka mediów. Antologia tłumaczeń [Media Linguistics. Translation Anthology], Wrocław, 2015.
- [20] B. Skowronek, Mediolingwistyka. Wprowadzenie [Mediolinguistics. Introduction], Kraków, 2013.
- [21] S. Zhelikhovskiy, Novyy pol'skyy uryad i Ukrayina: shcho chekaye na vidnosyny mizh krayinamy u 2024 rotsi [The new Polish Government and Ukraine: What awaits Relations between the Countries in 2024], December 18, 2023. URL: https://24tv.ua/ukrayina-polshha-shho-chekaye-vidnosini-mizh-derzhavami-2024_n2454386.