Scorched Wordscapes: A Multidisciplinary Study of the Transformations in Russophone Poetry Before and After the 2022 Invasion of Ukraine

Citation

Permanent link
https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37376431

Terms of Use
This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story
The Harvard community has made this article openly available. Please share how this access benefits you. Submit a story.

Accessibility
SCORCHED WORDSCAPES:
A MULTIDISCIPLINARY STUDY OF THE TRANSFORMATIONS
IN RUSSOPHONE POETRY BEFORE AND AFTER THE 2022 INVASION OF UKRAINE

A THESIS PRESENTED BY
PAIGE S. LEE

TO
THE DEPARTMENTS OF COMPUTER SCIENCE & SLAVIC LANGUAGES AND LITERATURES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
BACHELOR OF ARTS
IN COMPUTER SCIENCE & SLAVIC LANGUAGES AND LITERATURES

ADVISOR: PROFESSOR MARTIN WATTENBERG
HARVARD UNIVERSITY

ADVISOR: PROFESSOR AMELIA GLASER
UNIVERSITY OF CALIFORNIA, SAN DIEGO

MARCH 31, 2023
Dedication
My thesis is dedicated to the enduring resilience of the Ukrainian people and the swift arrival of freedom and peace.

Acknowledgements
I would like to express my sincere gratitude to all the individuals, institutions, and experiences that have played a crucial part in bringing this thesis to life.

First, thank you to my thesis advisors, Martin Wattenberg and Amelia Glaser, for their invaluable guidance, feedback and support throughout the thesis writing process.

Thank you to Steven Clancy for introducing me to research at Harvard and being my longest-running research mentor.

Thank you to my other teachers, mentors, role models from Leverett House, the Harvard Slavic Department, the Harvard Computer Science Department, and beyond for their wisdom and encouragement.

Thank you to Anna Ivanov, Jeffrey Schnapp, Teagan Stedman, Alyssa Chen, and Kyra Duran for helping me review and revise this monstrously long work.

Thank you to my family (shout out to Mom, Dad, Owen, and Johann) and my amazing friends for always supporting my dreams, no matter how crazy.

Lastly, thank you to the Barker Center Cafe, Cafe Gato Rojo, Mt. Auburn Hospital, Leverett Dining Hall, Remy the Cat, and the MBTA/M2 for keeping me going.

Note on Translations:
Unless otherwise specified, all translations are my own. The translated texts include excerpts from poetry, websites, and articles.
Abstract
This thesis explores the changes in Russophone poetry before and after the Russian invasion of Ukraine on February 24th, 2022, by developing and analyzing a novel database of 3,222 Russian-language poems posted to the internet from 2017 to 2022. The interdisciplinary study employs a combination of quantitative and qualitative methods, such as computational semantic analyses and close readings, to uncover the shifting relationships between language and politics in Russophone poetry. The findings reveal authors' deepening estrangement from the Russian language and a changing thematic landscape. Post-invasion poetry exhibited an increased focus on themes of war, morality, and mortality and a diminished focus on nature, memory, and everyday human experiences. The research highlights the potential of natural language processing for detecting larger-scale thematic shifts in text. This study also emphasizes the value of poetry as a subject that can enhance our understanding of the political and historical moment.
# Table of Contents

1. Introduction 4  
   1.1. Overview 4  
   1.2. Research Question 8  
   1.3. Motivation 9  
   1.4. Introduction of Approach 14  
   1.5. Contributions 16  
2. Historical Background 18  
   2.1. Maidan Revolution and Donbas War 18  
   2.2. Russia's 2022 Invasion and War in Ukraine 20  
   2.3. Situation on the Russian Home Front 22  
3. Related Work 27  
   3.1. Violence and Poetic Language 27  
   3.2. Digital Literary Analysis 31  
4. Methodology 36  
   4.1. Data Collection 36  
   4.2. Frequency Calculations 39  
   4.3. Word Embedding Vectors 42  
   4.4. Principal Component Analysis 44  
5. Results 46  
   5.1. Overview of the Metadata 46  
   5.2. Significant Word Frequency Changes 53  
      5.2.1. Words with Significantly Increased Post-Invasion Frequency 56  
      5.2.2. Words with Significantly Decreased Post-Invasion Frequency 58  
   5.3. Web Interface 61  
6. Discussion 64  
   6.1. Questions on Morality and Mortality 64  
   6.2. Nature's Vanishing Voice 68  
   6.3. War's Degradation of Russian 71  
   6.4. Turning Away From Memory 76  
7. Conclusion 83  
   7.1. Overview of Methodology and Results 83  
   7.2. Limitations and Future Work 85  
8. Appendix 87
1. Introduction
   1.1. Overview

How Darwin sailed to the Galapagos so we escape into the space of our language.

This language is the whole world, it contains countries and their borders, peaks of communism and failures of militaristic rhetoric, under the burial mounds slumber letters that have lost their sound, through the forests and swamps walk sounds devoid of letters and signs.

In this world "I love you" is a random set of three words, and "somewhere there is a war" is a statement of obvious facts.

And now, rising to that shifting ledge, I see how in the light of the comet of 1812 words and meanings perish, and text corpora fall apart in the heat of battle.

—Maksim Semyonov, 2022

In the middle of Leo Tolstoy's 1,225-page novel War and Peace, the character Pierre witnesses an "enormous and brilliant comet [...] which was said to portend all kinds of woes and the end of the world."² This is the comet of 1812, a real astronomical phenomenon which, at the time, was popularly

---

thought to have foreshadowed Napoleon Bonaparte's 1812 failed invasion of Russia (known in Russia as the Patriotic War of 1812). In the poem above from August 2022, the contemporary Russian poet Maksim Semyonov references 1812 in relation to how the Russian language often collides catastrophically with war and imperialism. In the 210 years following Napoleon's 1812 invasion of Russia, Russia has shamelessly embarked on a number of its own imperialist pursuits, and in 2022, it appalled and outraged the world by launching a full-scale military invasion of Ukraine. As opposed to Tolstoy, who was inspired by the Patriotic War of 1812 to write what some consider one of the greatest works of world literature, Semyonov does not find war to be a fitting context for conveying meaning. In Semyonov's opinion, war renders language unintelligible. Reflecting on Russia's contemporary actions in light of 1812, Semyonov stresses how war, imperialism, and revanchism transform and distort language. The violence of war alters or destroys the significance of words, their meanings, and the broader context in which they exist (text corpora).

Semyonov's reflection on the disintegration of language due to war serves as a powerful reminder of the complex relationship between language and the world in which it is used. As language constantly evolves, it becomes necessary to study these changes to understand how meaning is constructed and transformed. In order to study these changes, one must find a medium and a method to assess change. Literature is one resource for analyzing such evolutions in language. Poetry in particular is an especially fruitful medium for language change because it is emotionally and stylistically condensed, has experimental tendencies, and can react quickly to the events of the day. Poetic language is not necessarily representative of the language of everyday life, so changes occurring in the poetic domain reflect changes in a language which puts a high value on intense, connotative, and experimental speech patterns. In an ideal sense, poetic language is more human than the language of the news, politics, and historians. Poets, because of their ability to register sensitive changes in reality through their work, could be seen as "barometers" of society. While socio-political events are rapidly changing, one would expect poetry to rapidly change as well. Utilizing computational tools for large-scale analyses of these texts allows researchers to uncover patterns and trends in subtle language changes across a vast corpora. The wide-scale analysis of poetic texts, under certain historical circumstances, can serve as
worthwhile anthropological sources because poets strongly link language with reality. For those reasons, poetry can be an excellent source of raw language data for studying language change as it reflects societal conditions.

I have chosen to compare the situation of Russian-language (Russophone) poetry before and after the February 24th, 2022, Russian invasion of Ukraine. With this case study, I analyze changes in the language through poetry and examine the situation of poetry written in the language of the aggressor. The contemporary Russian-language case study is particularly promising for studying the evolution of language because the relationship between the Russian language and political reality has undergone significant changes in recent history. Exploring this subject offers further insights into the international perception and reception of Russian cultural artifacts, as well as how these processes are reflected in the language used within the works themselves.

This thesis is built around a collection of Russian-language poetry posted to the internet from 2017-2022 which I have assembled as a novel database. Using this time frame has allowed me to work with a substantial number of poems from both before and after the February 24th, 2022, invasion of Ukraine. The poetry is drawn from a variety of sources, including oppositional poetry blogs, Telegram channels for sharing poetry, and public Facebook profiles of individual authors. Because I am interested in the reflective and humanitarian roles of poetry during war, I have ruled out propagandistic poetry blogs which support Putin and the Kremlin. The sources in the database feature Russian-language poetry by authors from Russia and a variety of other countries such as Ukraine, Belarus, and Kazakhstan. I gathered poems from authors with a variety of national,

---

ethnic, and cultural backgrounds with the goal of understanding what happens to a subset of literature when the common language is under duress.

The Russian war in Ukraine has wreaked immense casualties, destruction, and displacement. Out of Ukraine's population of 41 million, 8.1 million people fled the country as of March 20th, 2023, and 5.4 million people were estimated to be internally displaced in Ukraine at the time of January 23, 2023.\(^4\)\(^5\) There has been a catastrophic loss of life, and the damage to Ukraine's infrastructure and environment may linger for generations.\(^6\) Later on, I provide a more extensive historical background on the Russo-Ukrainian War.

Living comfortably outside the war zone, Russians face different kinds of day-to-day worries. Putin's repression of freedom of speech and freedom of the media have made it dangerous for Russians to publicly oppose the war. Russian laws instituted in March 2022 punish those who "disseminate unreliable information" and "discredit" the Russian military. The dissemination of "unreliable information" about the Russian Armed Forces and the war in Ukraine may lead to imprisonment of up to 15 years. Nonetheless, some Russians have voiced their opposition to the war. Much of the content in the project is sourced from two significant blogs for oppositional Russian-language poetry: No War Poetry by the KriK Publishing House and Russian Oppositional Arts Review (ROAR) led by Linor Goralik, a Russian-speaking author who was born in the Ukrainian Soviet Socialist Republic and has mainly lived in Israel since 1989. No War Poetry and ROAR include many submissions from authors who are not Russian or do not live in Russia, since Russophone authors are connected by their language, not their nationality. Among the authors studied in this project, 71.25% were born in Russia, 16.56% were born in Ukraine, and 12.19% were born in other countries (as reflected in Figure 1).\(^7\)

---

\(^7\) I identify the cities on the map by their current names and borders. Many of the authors were born in the Soviet Union, which is not reflected in these statistics or on the map.
1.2. Research Question

This thesis responds to the following broad question: how do political events change language, particularly when the language of literature is the language of the aggressor?

Haleta asserted that in times of political violence, the metamorphosis of the literary language made it "impossible to predict the meanings and associations of words used by writers." On the Maidan in 2014, Haleta recognized that the violent struggle between protestors and the state set the stage for Ukrainian poetry to change in stylistic and thematic ways. Violent reality and poetry were tightly linked. In 2022, the geopolitics of Eastern Europe again rendered reality unrecognizable to those living through it, albeit in a different time and under a different scale of violence. In contrast to the 2014 Maidan protests, the Russian language has assumed an important role in the conflict (more in the Historical Background section). One would expect huge changes to the Russian literary language, changes which are "impossible to predict." The main focus of my study is not prediction but

---

8 More zoomed-in maps can be found in the Results section or at the web link: https://eelegiap.github.io/thesis-code/authorMap/index.html
9 Not shown is one author born in Pittsburgh, USA.
rather a characterization of changes which occurred in the linguistic landscape. I use a mixture of quantitative and qualitative methods to better understand how shifts in reality are reflected in semantic and associative shifts in the literary language.

This thesis approaches the task of understanding shifts in poetic language on the level of the word: how do words change in significance, and how do the associations between words change? I divide the dataset into pre-invasion and post-invasion subsets which are the experimental sets for language shift. The quantitative analyses serve as proxies for analyzing different aspects of language shift, generating high-level statistics and results about the textual dataset which I use as jumping-off points for close reading.

1.3. Motivation

I am interested in how the language of poetry changes in a time of social and political stress. Those who speak Russian face a collective recognition that their language is associated with the horrors of Russia's war in Ukraine. As Russian society becomes more and more totalitarian, Putin and his allies have abused the language to serve their needs. Russia's war has been terrorizing Ukrainians since 2014, but the scale of Russian aggression and propaganda has severely escalated in 2022. From the start of the 2022 invasion of Ukraine, the Russian government forbade the public use of words such as "invasion," and it frequently censures artists for "discrediting" or spreading "unreliable information" about the Russian Armed Forces and other Russian state bodies.\(^\text{11}\) The use of the Russian language in literature is particularly fraught as authors seek outlets for opposition. Russophone poets who oppose the war face a reckoning with the moral and ethical ramifications of their art, which raises one of the central questions about contemporary Russophone literature: what happens when the language of the aggressor is also the language of art?

I am motivated to study the Russophone literary community primarily because of how their common language binds them. All Russophone writing is

---

necessarily governed by the same systemic constraints of grammar and vocabulary. This project looks at how words become more or less significant in the language and how their connotations may change due to political events. These words belong to an overarching Russian language, with existing regional and generational differences in mind. Russophone authors have disparate views and backgrounds and write in various styles, but they are all linked by the fact that they are more or less perceived as part of a community with Russian cultural or linguistic heritage. If we can better understand the ongoing transformation of the Russian language, we can shed more light on how Russian cultural artifacts are perceived and received internationally, and how these processes manifest in the words themselves.

The Ukrainian literary scholar and poet Olena Haleta discusses the anthropological potential of poetic texts in her 2020 article "Mined Words: An Unimaginable Reality and the Search for a New Language in the Poetry of Maidan." She observed how changes in poetry from the 2014 Maidan protests in Ukraine expressed the "individual and collective experience of that time." Citing several scholars of literary and symbolic anthropology such as Wolfgang Iser, Clifford Geertz, Rudolf Behrens, and Ronald Galle, she goes on to argue that:

"...literature provides unconventional material, direct evidence of what was going on not only in reality itself but also in the relation between words and reality. The analysis of these relations can show the depth of the anthropological distress and the seriousness of the challenges faced by both society and literature, and poetry first of all, which appear to be not only a reflection but also a direct response to the ongoing events."12

Haleta emphasizes that literature is a specific type of historical evidence which is particularly useful for showing how language and reality are intertwined. In this way, poetry is an expression of humanity at a place and time, and studying it is a way of studying what makes us human, something that becomes particularly important when the events of history appear unthinkable and inhumane.

12 Haleta, “Mined Words.”
This study also sheds light on Russian speakers, a group which has been relatively de-emphasized in Western narratives about the war. For one, a muted Russophone response to the war can be attributed to draconian speech laws in Russia which limit the extent of Russian oppositional culture within Russia. For very different reasons, there is a general apprehension from the international community about promoting Russian voices at this time while many recognize a necessity to listen to Ukrainians' stories first and foremost. A data-driven study of Russian-speaking poetry from authors both inside and outside Russia could help us break through warped conceptions of the "Russian world" (Russian: русский мир; russkiy mir) which the Kremlin seeks to portray.

Yevgenia Belorusets, a Ukrainian writer and artist, has expressed ambivalence about the categorization of peoples into an amorphous "Russian cultural community." The "Russian cultural community," in her view, is a term imposed by outsiders to unnecessarily affiliate a scattered group of individuals including Russian speakers from the Ukrainian Donbas to Russian-born emigrés living abroad. On December 9th, 2022 in a work titled "Who May Speak," she wrote:

"Since the start of the war, I have seen various requests made of 'others'—to disappear from view, to remain silent. In the first weeks, such a request was made of an entity as vague as the 'Russian cultural community.' There were statements calling for those in the 'Russian cultural community' to not appear in public and to keep silent about their own misfortune in the face of Russia's genocidal crimes, violations of international law, and disregard for human rights."

Belorusets, a Ukrainian and native Russian speaker, goes on to lament the loss of nuance brought about by war:

"Through the grids of fixed aggressive perception, I see lost opportunities in many languages and voices that I miss."

---

14 Belorusets.
In these statements, she points out that the boundaries of the "Russian cultural community" are ambiguous, and she expresses hesitation about taking an oversimplified approach to who is allowed to enter the international discourse about the war. This thesis strives to step outside "grids of fixed aggressive perception" by considering the multi-stranded conversation happening among culturally distinct Russian-speaking artists.

The ethical question of working with Russophone poetry in 2023 exists in the background of this thesis, although it is not the main point of investigation. Many of the authors in the archive live in Russia, are ethnically Russian, or hold Russian citizenship, bringing them in closer proximity to the aggressor nation than someone with other national ties. These authors, who are Russian and Russophone, have faced some degree of exclusion from the international community despite expressing anti-war sentiment in their works. While I recognize the political reasons for de-prioritizing Russian poetry at this time, I view a study of Russian poetry as an opportunity to study the effects of war on the language of the aggressor state. I looked particularly at poets who display opposition to the war to varying degrees.

Cultural exclusion is a small price to pay when Russia's war of aggression has displaced at least 14 million Ukrainians, killed and wounded at least 20,000 Ukrainian civilians, and destroyed at least 339 Ukrainian cultural sites as of February 2023.1516 While I agree with Belorusets that unilateral, indiscriminate cultural exclusion should raise questions, I also understand that discussions of Russian and Russophone poetry may come across as harmful and unproductive to some audiences. Poetry is clearly not solving the problem of Russia invading a sovereign nation and terrorizing innocent people. As the Polish poet Zbigniew Herbert said in a 1987 interview, "It is vanity to think that one can influence the course of history by writing poetry.

It is not the barometer that changes the weather." But this thesis is not about how poetry can save lives and end wars, nor does it aim to venerate contemporary Russophone poetry for its aesthetic or political excellence. My aim is to study the puzzling phenomenon of wartime Russophone poetry in the context of linguistic evolution.

Part of the impetus for this project comes from digital archival work on contemporary Ukrainian poetry conducted by Professor Amelia Glaser at the University of California San Diego. My collaborations with Professor Glaser have impressed on me the value of developing digital archives of Slavic poetry. In the future, scholars of language and literature, anthropology, and digital culture may turn to archives of Ukrainian and Russian poetry for artifacts of this time.

Because content on the internet can be so easily transformed and destroyed, there are no guarantees that the poems in the collection can be found again in the future in the same way and in the same places I found them. Posts get taken down, accounts are deactivated, messages are deleted, and websites eventually die if they are not maintained by the developer. The loss of internet ephemera is particularly likely during a time of war, especially when Russia is actively suppressing oppositional media. Although some of these poems will likely find their way into books and anthologies in the future, newly printed copies will be transformed by intentional revisions and the changes from digital to print media. I captured poems as I found them during my collection period from September to December 2022. One main contribution of this thesis is the creation of an archive of Russophone poetry posted to the internet from 2017 to 2022. February 2022 is among several moments from recent history where Russian discourse changed. The 2011-2012 anti-government protests in Russia, 2014 annexation of Crimea, and 2019 protests in Moscow may also serve as good turning points for the

---

analysis of linguistic change.\textsuperscript{19,20} Expanding the archive to include Russophone poetry posted before and after these pivotal events would open new possibilities for tracking the rhetorical changes occurring in response to dramatic geopolitical events and the increasing repression of the freedom of speech inside Russia.

1.4. Introduction of Approach

This thesis uses digital humanities approaches to study a poetic corpus by incorporating both quantitative and qualitative methods. Computational methods are especially helpful for tasks of trend and pattern detection, something I was especially interested in for this study of poetry's large-scale transformations. Because poetry is not studied widely through a computational lens, I experimented with a number of existing NLP methods to see how they could aid in the investigation of my research question. After trial and error with methods such as word co-occurrence networks, contextualized word embeddings, and named entity recognition, I opted to focus on word frequency for its simplicity and interpretability. For more information on the specifics of the project's quantitative methods, see the Methodology section.

I draw inspiration for a word-based study of language change from the Ukrainian academic and poet Ostap Slyvynsky, who began qualitatively recording semantic changes in words in 2022 at the Lviv railway station. In his 2022 book \textit{War Vocabulary}, he documented how the war has changed people's relationships to their language.\textsuperscript{22} For example, the word "bath" changes meaning as the bath transforms from a household comfort into a makeshift bomb shelter. Another speaker describes how war changes the significance of "beauty":

\begin{flushleft}
\textsuperscript{22} Ostap Slyvynsky, \textit{Slovnyk Viyniy} (Kharkiv, Ukraine: Vivat, 2023).
\end{flushleft}
"I read a story about WWII not so long ago. There was this girl who wore her mom's worst clothes to pass by the Nazis unnoticed, to avoid being raped. I pause near my wardrobe; is it time to wear the worst already, or can I still make it? Things change so quickly... In a time of war, beauty becomes dangerous. Beautiful things, people, relationships—nowadays they don't exist to inspire. They exist to be annihilated. Not for admiration and loving touches, but for pain."23

The passage provides a compelling example of how a concept can acquire new meanings, connotations, and associations during times of violence. Slyvynsky's texts also suggest that changes in meaning can be found within the words themselves, rather than in larger constructions of language like sentences, paragraphs, or novels. While Slyvynsky writes about how people's experiences of war have changed language and meaning within Ukraine, my study looks at changes in Russian-language poetry across countries. These language practices intersect but have distinct social and political qualities.

I incorporate quantitative methods in the study to investigate the phenomenon of word-level semantic shift using a linguistic corpus. Corpus-based text analysis raises new possibilities for identifying and quantifying patterns on a larger scale. The incorporation of methods from natural language processing (NLP)—the application of computers to human language—allows us to analyze thousands of poems at once, opening the door to the analysis of more texts than a typical study of close reading. The expansion of the scale of analysis allows for the inclusion of less popular texts from a wider diversity of authors. NLP also defamiliarizes the study of literature, forcing us to reconsider how humans understand the construction of meaning and potentially raising new opportunities for literary interpretation.

Researchers may use NLP as a tool for quantitatively studying language along a number of linguistic dimensions, such as semantic, syntactic, and morphological. I have chosen to focus on the semantics of language shift, so I

incorporate text processing and computational methods such as frequency calculations and word embedding vectors. These methods use metrics like word counts as a foundation for understanding the relative significance and relationships between words. These types of analyses, which often rely on counting, vectors, and matrices, allow us to explore the construction of linguistic meaning from a computational angle. Computational semantic analyses of a large corpus then guide the process of individual close readings, suggesting starting points for qualitative investigations. A detailed description of these methods and calculations can be found in the Methodology section.

Together, quantitative and qualitative methods yielded a variety of expected and unexpected results about the changing relationships between language and politics in Russophone poetry. Global results based on word counts and occurrences show that the meta-conversation about language itself has shifted based on geopolitical circumstances. From pre-invasion to post-invasion, Russophone poetry across authors generally shifted away from themes of nature and mundanity towards military and philosophical topics. The word-level results show a poetry grown disembodied, lacking a physical and aesthetic basis in reality. Authors express a tense and estranged relationship with the language of their art because of its proximity to the war. The database, despite being composed of many unrelated texts, itself mirrors literary and non-literary conversations about the state of the Russian language in 2022 and the diversity of voices which comprise those conversations. Poetry, as usual, pushes the limits of what can be expressed and how. More than just being "barometers" of reality, contemporary Russophone poets use language to create experimental spaces for processing the present moment.

1.5. Contributions
The main contributions of this thesis are:

1. Creating a novel database of Russian-language poetry posted to the internet from 2017-2022, which serves as a valuable resource for studying the evolution of language and themes in Russophone literature during the Russian War in Ukraine.
2. Adapting interdisciplinary methods to conduct an analysis of Russophone poetry before and after the February 24th, 2022, Russian invasion of Ukraine, to uncover the impact of war on language, poetic themes, and the poetic imagination.

3. Developing a web platform to facilitate browsing the poetry database by author and keyword, improving the accessibility and utility of the dataset for researchers and readers.

In general, this thesis also demonstrates the potential for interdisciplinary research by combining Slavic studies, linguistics, and computer science to uncover patterns and trends in poetic expression, showcasing the value of natural language methods, such as frequency calculations and word embedding vectors, to explore poetic meaning, highlighting the potential for new opportunities in literary interpretation and Slavic studies. The results of the thesis also underscore the importance of poetry as a subject of study which allows us to better understand the relationship between language and reality during times of crisis.
2. Historical Background

This thesis explores the reciprocal relationship between reality and literature, particularly as the circumstances of wartime shape writers' use of language. Because this thesis is focused on the effects of the Russo-Ukrainian War on Russian-language poetry, I provide background on the events of the war and the current situations of Russians and Russian culture. I mostly focus on Russia and Ukraine, since they are the war's major actors, but I must add that many of the authors in the archive come from other countries, including but not limited to Belarus, Georgia, Kazakhstan, Uzbekistan, Turkmenistan, and Moldova.

2.1. Maidan Revolution and Donbas War

Although the Russo-Ukrainian War emerged most clearly to Western viewers in 2022, it began eight years earlier when Russia annexed Crimea, a culturally and strategically important peninsula in the Black Sea. The annexation of Crimea in February 2014 came at the culmination of the Euromaidan: three months of protests and demonstrations by Ukrainian citizens speaking out against the duplicity of the corrupt and pro-Russian regime led by President Viktor Yanukovych. Protests began in November 2013 following Yanukovych's abrupt refusal to sign the European Union–Ukraine Association Agreement, instead favoring closer ties with Russia and their Eurasian Economic Union. From November 2013 to February 2014, Independence Square (Maidan Nezalezhnosti) in Kyiv became the focal point of protests, as protesters occupied the square (Maidan) for weeks on end in the dead of Ukrainian winter. In the end, the Euromaidan protests reached their conclusion in Ukraine's Revolution of Dignity, also known as the Maidan Revolution, where protesters ousted the pro-Russian Ukrainian president Viktor Yanukovych and won concessions in favor of the protesters' agenda. This occurred in February 2014. The people's victory was historic, but catastrophically, the Russian government took swift countermeasures in efforts to maintain its influence over Ukraine.

In April 2014, pro-Russian separatists in Ukraine, with military support from Russia, declared the establishment of the Donetsk People's Republic and the Luhansk People's Republic in the eastern provinces of Ukraine bordering
Russia. The easternmost area of Ukraine is called the Donbas, derived from a portmanteau of the phrase "Donets Coal Basin," and from 2014 to 2022, the Russo-Ukrainian War was generally known as the War in Donbas (or in Russian state discourse as the "Armed Conflict in Donbas"). Donbas is home to a significant minority of ethnic Russians, and it is a predominantly Russophone territory — according to the 2001 census, Russian was the primary language of 74.9% of residents in Donetsk Oblast and 68.8% in Luhansk Oblast.24 Despite such large numbers of ethnic Russians and Russian speakers in the area, in October 2014, 78% of survey respondents from Eastern Ukraine opposed Russia sending troops there to protect Russian-speaking citizens.25 The vast majority of citizens in the Donbas, including Russian speakers, opposed Russian intervention at that time.26

The War in Donbas was generally overlooked by Western audiences, particularly because it was viewed as an internal Ukrainian conflict rather than a conflict between nations. However, Russia was a central actor in the conflict. During the War in Donbas, the Russian military armed, trained, and led separatist forces, according to the Rand Corporation.27 Vladimir Putin repeatedly legitimized the pro-Russian separatist regions of Ukraine, and the Russian state took significant measures to deflect responsibility.28

The official Russian line about the conflict in Donbas was that ethnic Russians and Russian speakers in eastern Ukraine faced persecution from a nationalist, neo-nazi, pro-European Ukrainian government, and Russia was coming to their people's protection. In keeping with this fiction, Russia did not declare war on Ukraine or invade the Donbas all at once. Instead, Russia

26 “IRI Ukraine Pre-Election Poll Shows Strong Opposition to Russian Aggression, Support for Kyiv Government | International Republican Institute (IRI).”
employed tactics from hybrid warfare—a military strategy which combines aspects of conventional warfare with irregular warfare, political warfare, and cyberwarfare—and employed other influence operations like disinformation and electoral intervention. The United Nations Human Rights Monitoring Mission estimates the total number of casualties in Ukraine related to the war in Donbas from April 14th, 2014, to December 31st, 2021, to be 51,000–54,000 people. This includes 14,200-14,400 killed (including at least 3,404 civilians), and 37-39,000 injured (including 7,000–9,000 civilians).

2.2. Russia's 2022 Invasion and War in Ukraine

Russia's shocking full-scale invasion of Ukraine on February 24th, 2022, occurred following months of military and rhetorical escalation from Russia. During this time, Russia amassed troops on Ukraine's borders, escalated rhetoric delegitimizing Ukraine's status as a sovereign nation, pushed false claims about rampant neo-Nazism in Ukraine's government, and made security demands to the West. One of these demands was that NATO and the United States sign on to a legally-binding agreement to prevent Ukraine from ever joining NATO. The United States and NATO rejected this proposal on the grounds that every country, including Ukraine, has the right to decide its own foreign policy.

On February 21st, Vladimir Putin announced Russia's official recognition of the Donetsk and Luhansk People's Republics and ordered Russian troops to be deployed into Donbas for a "peacekeeping mission." On February 24th, Putin announced a "special military operation" in Ukraine to "demilitarize

---

and denazify” the country. In a matter of minutes, missiles began to strike Ukrainian cities, and Russian ground troops entered the country on multiple fronts. People around the world learned in real time that the invasion had begun.

The full-scale invasion commenced on February 24th, 2022. During the first year of this phase of the war, Ukraine suffered immense casualties, destruction and displacement. It is incredibly difficult to estimate how many people have been killed, particularly because of the lack of information from Russian-occupied territories of Ukraine and the political prerogative (of all actors involved) to suppress certain information to maintain public support. Estimates vary wildly between different accounts. As of December 2022, the United Nations confirmed the deaths of 6,755 Ukrainian civilians (and 10,607 wounded), although the real numbers are thought to be higher. Also as of December 2022, the Ukrainian government reported that 10,000 Ukrainian troops have been killed, while the Russian government reported that the real number of Ukrainian troops killed is more than 60,000. Tens of thousands of both civilians and troops have been wounded, and there are many reports from on the ground of heinous, debilitating injuries from weapons like mines and cluster bombs. Third parties repeatedly confirm that Russia has deliberately ordered airstrikes on civilian targets. Grotesque evidence of torture and rape continue to be uncovered in liberated territories after Russian occupation. After the liberation of Bucha, a suburb of Kyiv, 458

---

bodies were uncovered in the community, of which 419 were "shot, tortured or bludgeoned to death." Many cities were yet to be liberated as of March 2023, and cities free from Russian occupation still face mortal threats from long-range Russian missiles. On December 13th, 2022, the UN reported that 50% of Ukraine's energy infrastructure was destroyed by Russian airstrikes. Millions of Ukrainians faced a brutally cold winter with a lack of heat, running water, and food. Countless Ukrainian families became refugees in their own country, including poets in my archive like Iya Kiva, a Ukrainian poet born in Donetsk who fled to Kyiv during the Donbas War, and then was forced to flee again to Lviv in 2022.

2.3. Situation on the Russian Home Front
The official Russian state line has been to call the war a "special military operation" or just a "military operation" ("[Специальная] военная операция"). Russian territory has remained virtually unscathed, but the home front is where the Kremlin dominates the information war. During his 23-year reign as president (as of 2023), Putin has gradually suppressed more and more independent Russian media. Since the invasion, he has ramped up censorship severely. On March 1st, 2022, the Russian government blocked TV Dozhd (Russian for "rain"), Russia's last independent TV channel, declaring it a "foreign agent." Television is one of the most popular ways for Russians to get the news, and the Kremlin has expended significant effort in controlling the Russian news narrative. Social media and online poetry are outlets for protest, but they aren't without their challenges. In 2022, the Russian government blocked Facebook (an important source in this project) and


The East European historian Timothy Snyder has described Putin's rhetorical strategy as one of "implausible deniability," where the government makes false claims which lack any ground for being plausibly denied. Some of the major themes which echoed in Russian state television in 2022 include a (false) genocide against ethnic Russians and Russian speakers in Donbas, portraying the Ukrainian government as a neo-Nazi junta installed by the West, and defining the Kremlin's war in Ukraine as a war against not only Ukraine, but all of NATO.

On March 3rd, 2022, Putin signed a bill into law introducing fines and prison sentences up to 15 years for those who publish "knowingly false information" about the Russian military and its operations, leading many media outlets to stop reporting on Ukraine. Roskomnadzor (the Russian federal executive agency responsible for monitoring, controlling and censoring Russian mass media) blocked access to several foreign media outlets, including BBC News Russian, Voice of America, Meduza, Facebook, and Twitter. Some of the aforementioned companies and others like Bloomberg, Spotify, Netflix, Ikea, Adidas, Nike, American Express, Mastercard, IBM, and Sony left Russia, suspended new investments, or discontinued access to their services in Russia in protest.\footnote{The New York Times, “Companies Are Getting Out of Russia, Sometimes at a Cost,” \textit{The New York Times}, October 14, 2022, sec. Business, https://www.nytimes.com/article/russia-invasion-companies.html.} Many Russians have begun to use virtual private networks (VPNs) to access blocked content. VPN downloads in Russia increased by 50 times from January to April 2022.\footnote{Anthony Faiola, “Millions of Russians Are Tearing Holes in the Digital Iron Curtain,” \textit{The Washington Post}, 2022.}
Censorship and intimidation have made independent polling extremely difficult and unreliable in Russia.\textsuperscript{46} Leaked information from Kremlin polling has shown that as of November 2022, only 25% of Russians were in favor of continuing the war, down from 57% in July—such numbers should be viewed with a significant amount of skepticism.\textsuperscript{47} Experts attributed this decrease in Russian support for the war to Putin's partial mobilization beginning September 21st, 2022. Denis Volkov, the director of the Levada Center (Russia's only large independent sociological institute), commented on decreased Russian support for the war:

"This is sheer reluctance to take part in the war personally. [Russians] continue to support it, but they have very little desire to participate themselves. Besides, their support was, from the very start, something they declared with regard to what they perceived as having nothing to do with themselves: “Life goes on — it’s even getting better.” Now, the risks are greater, and people want to start the talks. Still, the majority of people leave this to the government: “We’d like it, but it’s up to them to decide.”\textsuperscript{48}

Sociologist Grigory Yudin also credits the draft with decreasing public support for Russia's continuation of the war and cites Russian "loss of faith in the victory" and the “absence of a convincing account of how exactly Russia might win.”\textsuperscript{49} On October 28th, the Russian government reported that they reached their goal of mobilizing 300,000 troops. The mobilization process has been characterized by its chaos—families have complained of mobilized family members receiving inadequate gear and minimal training before being sent to the front line. In at least once instance, several dozen newly mobilized


\textsuperscript{48}Pertsev.

\textsuperscript{49}Pertsev.
soldiers were reportedly killed within days of their arrival in Ukraine. In this way, popular Russian opposition to the war appears to often be driven by the war’s impact on the home front, not by outrage at the treatment of Ukrainians.

In response to the military draft, hundreds of thousands of Russians— in particular, young men avoiding the draft— left the country. Citing a Kremlin source, Forbes Russia reported that 700,000 Russians fled the country during the period from September 21st to October 4th. Russians fled to the European Union as well as to countries which Russians may enter without a visa, such as Georgia and Kazakhstan. In the week ending on September 27th, 2022, there were triple-digit increases in one-way flights from Russia to Tbilisi, Almaty, Istanbul, Tel Aviv, and Dubai.

Besides potential draftees, it's unclear how many Russians have left the country as political refugees or as refugees of conscience. Many intellectuals have emigrated from Russia since the beginning of the war, who join countless others who emigrated during the twilight years of the Soviet Union. In Russia in 2022, it is dangerous to be openly against the war, with thousands of protestors detained and arrested. Putin has made clear how he regards Russians opposed to war. In a March 16th, 2022, televised news conference, Putin called protestors "traitors and scum" and said that he is “convinced that this natural and necessary self-cleansing of society will only strengthen our country.” His dehumanizing characterization of protestors echoes chilling labels for the opposition which were hallmarks of repressive Soviet propaganda, like "enemies of the people" (Russian: враги народа, vragi

---

52 Reuters.
Putin's tactics on the home front are reminiscent of Soviet censorship and repression, and apt comparisons between Putin and Stalin paint an eerie picture of the state of Russia in 2022.  

---

3. Related Work

3.1. Violence and Poetic Language

Violence and its specter have the potential to prompt unexpected changes to language. In this section, I will discuss several ways of understanding the poetic "response" to violence and violent times.

Authors may reshape language in the face of violence with the practice of defamiliarization, or dehabitualization. One can point to the Russian literary theorist Victor Shklovsky's concept of "estrangement" (Russian: остра нен ие; ostranenie), which is when an author uses language in unexpected and unusual ways to force a reader to think more deeply about and truly "see" the object at hand. Shklovsky suggests that the constant reinvention of language can fight against the dangers of "getting used to" reality. He also wrote that "habitualization devours work, clothes, furniture, one's wife, and the fear of war."

The renewal of language, then, is a means of de-habitualizing the subject of the conversation, like war, and pointing out its true colors. Shklovsky writes, "...art exists so that one may recover the sensation of life; it exists to make one feel things, to make the stone stony." The linguistic device of "defamiliarization" is useful during wartime when it counteracts the de-sensitization to violence that may occur through other media and makes the reader more aware of the insanity of war. This presentation of violence as surreal, which seems to be more common among contemporary Russophone poets, differs from portrayals of violence in epic poems throughout the history of poetry.

In the 2020 article "Mined Words: An Unimaginable Reality and the Search for a New Language in the Poetry of Maidan," the Ukrainian scholar Olena Haleta describes how violence deconstructs linguistic potential:

"Reality goes beyond the boundaries of the possible and thus elevates even the so-called lingual metaphor, or the established idiom. No one can declare what remains imaginable in these changed circumstances, and what belongs to the sphere of the potentially possible."  

---


57 Haleta, “Mined Words.”
Using examples from the poetry written during the Maidan, Haleta illustrates how linguistic conventions "collapse" in the face of violence. She writes about how violence deprives the narrator of his or her language, and in this sense words degenerate into "a shriek, a moan or other similar speechless reactions."

Some authors in the database are more directly affected by the violence of Russia's war than others, for example, if they live in Ukrainian cities experiencing Russian bombardment, or if they know soldiers on the front line. One would expect that their poetry might experience more deviations from conventional language as Haleta describes because they are closer to the violence. This situation is relevant to a subset of authors in the dataset, who respond to violence first-hand. How does this linguistic breakdown manifest when written in the language of the aggressor?

At the same time, violence can lead to a reconstruction of language, often out of a need to create new symbols to interpret a changing reality. People at a loss for words must find new ways to express feelings and experiences which previously were not part of their reality. In other circumstances, people seeking new justifications for unprecedented acts of violence will manipulate the language to meet their needs. In a public journal entry from December 2022, the Ukrainian writer and artist Yevgenia Belorusets describes how a breakdown of law and order precipitates language construction:

"In such incidents and countless others, I see that this war is intended to teach my society that all previous rules of life have been suspended, that there is no law anymore...And in place of this great absence—for which there is no serious explanation—there is only delusion, propaganda. It works like a procession of empty spaces that are randomly and reactively filled with new concepts, words, and radical emotions."58

A quick review of Russian state media from 2022 reveals a propagandistic overhaul of the language. State media, blogs, and social media are rife with

58 Belorusets, “THE WAR DIARY OF YEVGENIA BELORUSETS.”
examples of newspeak, novoyaz (новояз) in Russian. Most notably, as of March 2023 the Russian state continues to refer to the 2022 invasion of Ukraine as a "special military operation" (or spetsoperatsiya for short in Russian). One just needs to open the webpage of Russia Today or flip on Channel One to see words and phrases such as "the Collective West," "foreign agents," "nuclear blackmail," and "denazification." The state enforces its linguistic propaganda by prosecuting, jailing, and instituting huge fines on people or institutions who violate any of the conveniently ambiguous war censorship laws which prohibit the dissemination of "unreliable information" about the Russian Armed Forces. Free speech inside of Russia is costly, and the language itself faces a number of linguistic hurdles to properly describe the horrors at hand.

However, what happens to poetry when the author, who may or not have experienced violence first-hand, shares the same language as the perpetrator(s) of violence? This question is discussed at most length in the context of studies of German poetry after World War II. German, like Russian, was co-opted by authoritarian states in order to spread propaganda and enable ethnic violence. Additionally, a few powerful speakers of the language declared sole ownership over the language in order to create in- and out-groups. Language was and is certainly a tool of the regime. Adrian del Caro describes the phenomenon further in the essay "Paul Celan and the German of the Non-German:"

"Before we humans mobilize as armies we mobilize as nations, and before we mobilize as nations, we cement our exclusivity and identity through language. Whatever is hated and despised in the Other, we have formulated in our respective languages."

In this sense, native speakers must deal with the curse of writing in the language of the aggressor. Paul Celan, a major poet from the post-World War II era, was a Jew and a native German speaker whose work was heavily influenced by the horrors he experienced and witnessed during the

---

59 Russian: Россия Сегодня, Rossiya Segodnya
60 Russian: Первый Канал, Perviy Kanal
Holocaust. His poetry was written in a "non-German German" in the sense that he was not ethnically German, but he spoke German. Del Caro's article, in summary, describes Celan's poetry as undertaking the project of "rescuing of language from the clutches of hate speech" by exposing the Christian-nationalist aspects of German through poetry. Celan was marginalized in some respects within German literature because of his ethnicity, which may have given him greater access to a German language that could be antithetical to German nationalism. He wrote in the "same" language as the perpetrators, yet it was a language that did something very different. The reconstruction (or assertion) of a certain kind of German language within Celan's texts rendered German more "humane."

It is impossible to draw direct parallels between Celan and any of the Russophone authors writing in the 21st century, but the relationship between language and state violence shares some aspects, as mentioned above. In the database, some of the authors write in "Russian" and others what one might call "non-Russian Russian" (although the line between "Russian" and "non-Russian" are essentially blurry because of the mixed forces of ancestry, emigration, and geopolitics). Russia's violence in Ukraine is ongoing as of March 2023, so it is possible that such reflections on the Russian language were not possible in 2022 and will not be possible for years until there is more distance from the violence. However, considering the outpouring of poetry from Russophone authors in 2022, why not pose the question: are any Russophone poets writing poetry which "rescues" language from the "clutches of hate speech, from oppression per se, from violent death by genocide?" If so, how? In what forms can we recognize this?

In summary, language's evolution during and after violence does not mimic or directly translate current events. Instead, the cycle of violence and repair during war implicates language, changing how people express themselves in light of a changing reality. From a reader's perspective, it is difficult to know which aspects of the text are intended or coincidental: did the poet intend for this particular line to be interpreted as a deconstruction of government rhetoric? Establishing which aspects of the language have transformed on

---

62 Del Caro and Ward.
both general and specific scales offers a window to the poetic "response" to state violence.

3.2. Digital Literary Analysis

In this thesis, I explore the intersection of quantitative cultural analysis, applied natural language processing (NLP), and digital humanities. In general, the project falls under the umbrella of the digital humanities, a field of inquiry which applies digital concepts to the humanities and humanities concepts to the digital. In this section I will introduce the field from a literary perspective and explain the rationale for incorporating digital humanities methods into this project.

Digital humanities (DH) has gained prominence since the mid-20th century, evolving from earlier disciplines such as scientific philology.

Philology—the intersection of textual criticism, literary criticism, history, and linguistics—goes back many centuries, with the more recent emergence of "scientific philology" during the 19th century. In scientific philology, practitioners brought together aspects of the scientific method (such testable hypotheses, experimental sample collection, and measurement-based statistical testing) with the study of literary and historical texts. Empirical methods found their way into myriad domains of literary studies, such as stylistic, narrative, structural, and genealogical analyses.

Using DH can look very different in different circumstances because of the breadth of digital resources and possible questions in the humanities in fields from architecture to music. Because of that, DH projects can run the gamut from digital archival to map making, computational stylistic analysis, and the creation of virtual environments for the dissemination of research.63 Just within literary studies, there are many ways of studying literature with digital tools. Notably, Professor Ted Underwood demonstrates the potential of machine learning and DH in literary analysis in his book *Distant Horizons: Digital Evidence and Literary Change*.64 The book discusses hypotheses,

theories, and real experiments, which are mostly conducted on corpora of Anglophone novels from 1800 to 2000. His choice of methods vary depending on the questions at hand. In the chapter "Metamorphoses of Gender," Underwood collaborated with NLP expert David Bamman to study the history of gender roles in Anglophone literature. At times, he considers metadata such as the gender of the author of a text, or he uses existing literary theories as a starting point for closely studying the behavior of a few words. In the chapter on methods, Underwood discusses the challenges of using statistics to work with "unstructured humanistic evidence." Based on his observations and analysis, he asserts that a "model that connects linguistic variables to social ones can develop a stronger web of hermeneutic connections" than one which analyzes linguistic variables in an isolated context. In the project, I weave together linguistic variables with the social and political reality of the Russian language from 2017 to 2022.

The focus of this thesis was "stylometry," the term coined by the scholar Wincenty Lutoslawski (1863–1954) to describe stylistics, the study of linguistic style, in the context of statistics. Lutoslawski brought basic statistical ideas of word usage to the problem of dating the dialogues of Plato. The inventions of the computer, digitized text, and natural language processing sparked exponential growth in stylometrics. Resources such as Lib.Rus.Ec, Project Gutenberg, and Google Books are large repositories of digital texts often used for quantitative literary studies. Many stylometric studies involve the process of counting and the representation of text as feature vectors. Depending on the hypotheses of the stylometric study, one might count different parts of speech, content-free words (prepositions,

---

65 Underwood.
67 https://www.gutenberg.org/
68 https://books.google.com/
conjunctions, etc.), verb tense, sentence or word length, lexical allusions, collocations, words from specialized categories, and more.  

I studied how the language of poetry changes when it is also the language of the aggressor state. How does my investigation benefit from the application of quantitative methods?

Leveraging computational tools in this project facilitates the identification of significant poetic passages for close reading, addressing the challenges posed by the constantly evolving nature of contemporary Russian poetry. Because I worked with thousands of individual poems in the project, computation enabled me to select a few of them for close reading. I conducted high-level analyses on language data to locate which poetic passages and texts had the most potential for close reading. This process of "location" was especially important in this case because the recency of contemporary Russian poetry and the speed of poetry posted to the internet means that these works have not yet entered the literary canon. The literary canon—that list of "good" and "classic" literature which is passed from one generation to the next—is not without its flaws, but it often serves as a foundation for academic discussions about literature. In the case of this project, the guiding or foundational canon does not exist.

Such little time has passed since the invasion of Ukraine, and so much has changed in East European society and culture that a retrospective canon was nonexistent or unreliable as a navigator of this research. Certain poets in the collection are more well-known in 2022 than others, which provided some indication of their value to culture and society, but of course, some "classic" authors only become widely read long after their deaths. For example, Mikhail Bulgakov's magnum opus The Master and Margarita was published

---


26 years posthumously. As so, there was no consensus and only a few theories about which contemporary Russophone poetry captured the zeitgeist or was changing the state of poetry. In order to address the uncertainty that stems from the contemporaneity of this research, I used DH as a source of navigation through many possible noteworthy texts.

In addition to helping me locate which texts adhere to or deviate from global patterns of word usage, DH also allowed me to conduct a more inclusive study of literary transformations. Quantitative methods excel at scaling up, as noted by Underwood and other DH scholars. Using the computer to "read" text expedites the process of drawing connections between aspects of an extremely large number of data points. The computational power and speed of computers enable the analysis of a vast number of data points (words, lines, poems), far more than a human could typically manage. This allows for a broader range of perspectives in literary discussions and ideally helps to capture a more comprehensive view of language shifts that transcend individual authors.

Adopting a balanced level of skepticism is crucial when utilizing digital humanities (DH). While computers offer increased speed and memory, they also have inherent limitations. Computers can only represent specific features of literary texts and communicate results in particular ways. Despite their advanced capabilities, computers still operate on binary code and rely on human-designed algorithms. Converting text into data involves transforming textual symbols into complex numerical sequences, a process driven by human innovation and effort. The encoding of text into bytes has advanced to the point where we can capture various features of language such as morphology, semantics, syntax, and phonetics. However, a computer's quantitative understanding of a text differs significantly from our qualitative comprehension.

Both the computer's ability to understand language and our capacity to grasp the implications of its quantitative analysis have limitations. As a result, it is essential to maintain a healthy skepticism when translating between languages and numbers. Numerical results should be considered within the context of the dataset's inherent randomness and imperfections. Ultimately,
the meaning of numbers, like words in literature, depends on human interpretation.
4. Methodology

4.1. Data Collection

Mоление о весне превращается в стихотворение, распадающееся на единицы информации, летящие по интернету к тебе и ко всем другим, способным читать, любить и ненавидеть.\textsuperscript{72}

A prayer for spring turns into a poem broken down into units of information, flying over the internet to you and to all others able to read, love and hate.

—Maksim Semyonov, 2022

The internet has become a rich and diverse source for contemporary Russian-language poetry, offering unique insights into the evolving landscape of language and culture during times of political upheaval. By examining poetry from various online platforms, including Facebook, Telegram, and blogs, this study explored the role of digital spaces in shaping and disseminating poetic works. The immediacy and accessibility of these platforms not only foster a more dynamic and interactive relationship between poets and their audience, but also reveal a complex interplay between literature, politics, and social media in the 21st century.

I scraped the text and metadata of the Telegram channels and No War Poetry with the help of a Python scraper and collected poem data from ROAR and Facebook by hand. Because I was interested in the transformations of literary language over a specific period, I only collected poetry posted to the internet from 2017 to 2022.

For Facebook, I collected text and metadata from the public Facebook profiles of a few prolific poets who publish work in Russian, including Aleksandr Skidan (from St. Petersburg), Boris Khersonsky and Ludmila Khersonskaya (from Odessa), Dmitry Strotsev (from Minsk), Elena Fanailova (from Voronezh), Iya Kiva (from Donetsk), and others.\textsuperscript{73} I collected most of the Russian-language poetry on their profile I could find from the period from 2017 to 2022. Their accounts are particularly interesting because of the high


\textsuperscript{73} Several of the authors have emigrated or fled the locations mentioned.
degree of interaction they enjoy with other Facebook users. Although the Facebook poetry ecosystem was not the main focus of this project, I was especially interested in looking at poetry which authors may intend to amplify using social media, as this poetry may speak more directly to the popular conversation.

Similarly, I scraped data from Telegram, a social media platform based around group chats. Public group chats on Telegram, also called channels, allow users to send text, images, and links to a wide audience. I drew from channels which were explicitly designated for sharing poetry. On advice from Stephanie Sandler and Anna Ivanov, scholars studying Russian internet poetry at the Harvard Slavic Department, I sourced Telegram poetry from channels called Metajournal (created by the Russian poet Yevgeniy Nikitin), Kabanov_slovo (Aleksandr Kabanov’s poetry channel), and essentialpoetry. Kabanov_slovo and essentialpoetry have 1.5K and 1.4K subscribers respectively, and Metajournal has 2.4K. On the Telegram channels, poems are posted by the moderator of the channel, but the author of the poem is not necessarily the moderator. Metajournal goes back to 2019, but since 2021, the content on Metajournal has been curated by a board of around twenty rotating editors from different Russian-language poetry circles. These editors include Aleksei Masalov, Maksim Dremov, Yevgeniya Ulyanikina, Zoya Falkova, Anton Platonov, Ivan Poltoratskiy, Pavel Bannikov, and others. Metajournal is an explicitly anti-war publication and has held anti-war literary events online since February 24th, 2022.

The blogs—No War Poetry and Russian Oppositional Arts Review (ROAR)—went live online after February 24th, 2022. The first issue of ROAR was posted on April 24th, 2022, and the second and third volumes were

---

76 Alexandr Kabanov Kabanov, “Poetry Telegram Channel of Alexandr Kabanov,” Telegram, accessed December 7, 2022, https://t.me/s/kabanov_slovo?q=%D1%82%D0%B0%D0%B2%D0%B0%D1%80%D1%8B%D1%88%D1%8B%E2%80%93%D1%82%D0%B0%D0%B2%D0%B0%D1%80%D1%8B%D1%88%D1%8B%20&before=248.
78 https://www.youtube.com/watch?v=IsAskTtn18E
subsequently released at two month intervals. The purpose of ROAR, according to its site,

"is to introduce its readers to the artifacts of the contemporary Russian-language culture...opposing the loyalist and servile official culture, which in extremes merges with the blatant propaganda serving the current criminal political regime in Russia."

The *No War Poetry* publishing project is run by the KriK Publishing House, a Russian language publishing house founded by Gennady Katsov and Rika Katsova. Like ROAR, *No War Poetry* has political aims. A note from KriK on the *No War Poetry* website says,

"We...cannot remain indifferent to the terrible events of today. We have launched a new poetry project NO WAR - POETS AGAINST WAR. It is a playground for writers of poetry united by a common hatred of war. In our previous projects, when selecting publications, we focused solely on aesthetic criteria. Unfortunately, today we are all faced with a choice, and this cannot but affect the opinion of the editors. In the horrific Russian-Ukrainian conflict, 'We are Ukraine!'"

Both websites are online publication projects which signal a poetic response to the Russo-Ukrainian War. The statements of purpose for both websites clarify the publishers' oppositional stance and their views on the importance of politics to poetry and vice versa.

The freedom and multidimensionality of online poetry make it an excellent source for studying changes in language during times of political turmoil. Based on studies of Russian internet poetry, the researcher Julia Bartosh has written that "network poetry is characterized by the following features: the absence of editors and, as a result, greater freedom in choosing a topic and

---

form, the ability to correct the text after it has been published, hypertextuality, multimedia, polycode\(^{82}\), and the use of special kinds of linguistic devices."\(^{83,84}\) The absence of the editor is useful for a study of language change in a sensitive political environment, since authors may (more) freely express their thoughts and emotions regarding the ongoing political situation without any censorship or restrictions. This freedom in choosing topics and forms results in a more accurate reflection of the evolving language and sentiments during times of turmoil.

Bartosh notes that the most important feature of online poetry is "the exaggerated role of the reader, in fact, a fusion of the poet, the reader with his response-reviews, and the poetic text itself." The internet shortens the distance between writer and reader, ideation and publication. Scholars of the internet such as Whitney Phillips, Ryan M. Miller, and danah boyd have written extensively on the blurred digital boundaries which characterize networked interactions, including the boundaries of then and now, formal and folk, commercial and populist, and public and private.\(^{85}\) This interconnectedness allows for a more comprehensive view of language change as it reflects the complexity of the socio-political environment.

### 4.2. Frequency Calculations

In order to analyze the changing presence of words and concepts throughout wartime, I looked at words which became more or less frequently used in poems posted online after February 24th, 2022. To do this, I calculated and visualized word frequencies.

Here is the general algorithm for a simple version of word frequency for some word \(w\):

\[^{82}\text{Polycode texts are defined as a combination of the verbal code with the codes of some other semiotic systems (images, music).}\]

\[^{83}\text{Russian: "Кроме того, для сетевой поэзии характерны следующие черты: отсутствие редакторов и, как следствие, — большая свобода в выборе темы и формы, возможность корректировать текст уже после его опубликования, гипертекстуальность, мультимедийность, поликодowość, использование особого рода языковых средств."}\]

\[^{84}\text{Julia Bartosh, “Специфика языка русской сетевой поэзии,” Русистика 17, no. 2 (2019): 229–42.}\]

\[^{85}\text{Zizi Papacharissi, }\text{A Networked Self: Identity, Community and Culture on Social Network Sites} (\text{New York: Routledge, 2011}).\]
Naive frequency algorithm:

- Given the total number of words in the dataset: $T$
- Given the total occurrences of the word $w$: $o_w$
- In the simple version, frequency of word $w$ is calculated by dividing the total occurrences of the word $o_w$ by the total number of words in the dataset $T$: $\frac{o_w}{T}$

Unfortunately, the simple version of word frequency is not suitable for a poetry dataset because of the presence of stylistic repetition within poems. Because poets use repetition as a poetic device, poems with strong repetition will "artificially" inflate the overall word frequency for those words. A poem by Igor Satanovskiy exemplifies the case of poems with strong repetitive tendencies:

Пушкин и война

Война идёт, а Пушкин пишет
Война идёт, а Пушкин пишет
Война идёт, а Пушкин пишет
Война идёт идёт идёт

А Пушкин пишет пишет пишет
А Пушкин пишет пишет пишет
А Пушкин пишет пишет пишет
Война идёт идёт идёт

Pushkin and War
The war goes on, and Pushkin writes
The war goes on, and Pushkin writes
The war goes on, and Pushkin writes
The war goes goes goes

And Pushkin writes writes writes
And Pushkin writes writes writes
And Pushkin writes writes writes
The war goes goes goes\textsuperscript{86}
—Igor Satanovsky, 2022

Although war and Pushkin should be recognized as significant themes in Satanovsky's poem, counting every instance of these repeated words towards the total word count in the post-invasion corpus would skew word frequencies for war, Pushkin, writes, and go on the basis of only one poem. In order to account for the biasing effects of poetic repetition, my word frequencies calculations differed from the simple version. I calculated word frequencies for a word $w$ on a poem by poem basis. The modified algorithm is given below:

**Modified frequency algorithm:**

- Total number of poems in the dataset: $P$
- Total number of poems containing the word at least once: $p_w$
- Frequency of the word in question is calculated by dividing the total number of poems $p_w$ by the total number of poems $P$. This is the percentage of total poems which contain the word $w$: \( \frac{p_w}{P} \)

So in the case of Satanovsky's poem above, $p_{\text{\{war\}}}$ and $p_{\text{\{pushkin\}}}$ would both increase by one, instead of by 5 and 6, respectively.

To prepare the dataset for calculation, I lemmatized each poem using the Spacy package in Python, which converts each word form to its lemma (the word's singular, present tense, or "dictionary" form).\textsuperscript{87} I calculated word frequencies for every lemma in the corpus. I ran a two-sample two-sided T-test to determine whether a lemma's frequency changed significantly from the pre-invasion period to the post-invasion period.\textsuperscript{88}

4.3. Word Embedding Vectors

I generated word embedding vectors for the analysis and visualized them using Principal Component Analysis (PCA). In this section, I will give an overview of how these vectors and vector dimensionality reduction algorithms work in order to explain the quantitative underpinnings of the visual analysis. Those with familiarity with word embeddings and vector projections may want to skip this section.

Figure 2: Examples of word embedding vectors

In natural language processing, a word embedding is a vector representation of a word which encodes information about the word's meaning in such a way that similar words have similar vectors, and relationships between the vectors correspond to semantic relationships. Word embedding vectors may have an arbitrary number of dimensions (typically ranging from the hundreds to thousands), but they are often projected down to fewer dimensions (often two or three) for visual inspection by humans, as shown in Figure 2. The different dimensions of vectors correspond roughly to different "dimensions of meaning," but since the vectors are automatically generated by computers, humans may have difficulty interpreting specific dimensions. Ongoing research in NLP investigates the interpretability and meaning.

behind the dimensions in word embedding vectors and other products of neural networks.90

Word embeddings are generally calculated by running an algorithm over a large corpus of text; this is the training corpus. There are different algorithms for calculating word embeddings for words. Several popular word embedding tools and algorithms include Tomas Mikolov’s Word2vec, Stanford University’s GloVe, AllenNLP’s ELMo, fastText, and Gensim.9192939495 Word2vec (2013), for example, calculates word embeddings by training a neural network to predict word context.96 The training process involves a skip-gram model, which predicts a word’s surrounding context for a given word, or Continuous Bag of Words (CBoW), which predicts a word given its context. GloVe, on the other hand, develops word embeddings on calculations over word co-occurrence probabilities (the probability of two words occurring together in the same window of \( n \) words in the training text).

For my analyses, I primarily used Navec, a set of pre-trained Russian word embedding vectors developed by Alexander Kukushkin. Navec is part of Natasha, an extensive Russian NLP project on Github.97 To generate word embedding vectors for 500,000 Russian words, Kukushkin trained the GloVe algorithm on a 150 GB sample of the lib.rus.ec book collection, a digital library.98 The sample from lib.rus.ec consisted of 12 billion tokens (words).

---

98 “Librusec.”
With such a large training corpus, Kukushkin generated 300-dimension word embeddings for Russian words which are useful for different semantic tasks, like finding synonyms and similar words. The Navec embeddings were evaluated for their accuracy on sample tasks and performed the same or better than Rusvectores, a popular Russian word embedding model. I visualize the Navec vectors by projecting them from 300 dimensions onto two dimensions using PCA.

4.4. Principal Component Analysis

Principal Component Analysis (PCA) is a widely-used method in data analysis and machine learning that simplifies intricate data by cutting down its dimensions. Imagine having a massive dataset with numerous variables (like 300-dimensional word embeddings) that you want to simplify while keeping the most crucial information intact. PCA first identifies principal components, which are essentially the directions of data which capture the most information. Then, it uses these principal components to project the data onto a lower dimension, such as a 2-dimensional vector space, preserving the data’s information while making it more manageable to visualize and analyze.

Figure 3: Transformation of vector space with PCA
In the case of word embeddings, PCA is a way of grasping the core of words' meanings while maintaining their relationships in a more straightforward, lower-dimensional space. As mentioned previously, word embeddings are numerical representations of words that convey information about their meanings. Reducing the dimensions helps us see how these meanings relate to each other. For my study, PCA transformed the Navec vectors from 300 dimensions to two. Figure 3 depicts the transformation in vector space. During this transformation, PCA works to retain the semantic relationships between words as much as possible. Words with similar meanings will still be close to each other in the new 2-dimensional space. This lets us visualize and comprehend the semantic relationships between words more easily, even though some data may be lost due to the reduction in dimensions.

In short, PCA is an effective mathematical technique that simplifies complex data by reducing its dimensions while preserving the most significant information. When used with word embeddings, PCA helps us better understand the relationships between words and their meanings by projecting them onto a lower-dimensional space which is easier to visualize and analyze.
5. Results

5.1. Overview of the Metadata

In short, I assembled a database of 3,222 poems found on the internet. The poetry is from the period from 2017 to 2022, generally anti-war and anti-authoritarian, and largely written in Russian. Because the database contains poetry from hundreds of authors, many poems deviate from those descriptions. I quantified the most important metadata about the poetry in the database to better characterize the subset of literature in the study. An explanation of the breadth of online sources and regional diversity of authors contextualizes the poetic findings, as the experiences of Russian speakers vary widely from region to region, platform to platform, and author to author.

The breakdown of poems per source is shown in Table 1. The essentialpoetry Telegram channel contributed the most poetry to the study, followed by the No War Poetry blog and the metajournal Telegram channel.

<table>
<thead>
<tr>
<th>Source</th>
<th>Poem Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>essentialpoetry Telegram channel</td>
<td>945</td>
</tr>
<tr>
<td>No War Poetry blog</td>
<td>724</td>
</tr>
<tr>
<td>metajournal Telegram channel</td>
<td>662</td>
</tr>
<tr>
<td>ROAR V3 blog</td>
<td>350</td>
</tr>
<tr>
<td>Public Facebook profiles</td>
<td>253</td>
</tr>
<tr>
<td>ROAR V2 blog</td>
<td>234</td>
</tr>
<tr>
<td>Public Telegram profile</td>
<td>54</td>
</tr>
</tbody>
</table>

In general, the poetry I collected was posted to the internet at some point over a five-year span from 2017 to 2022, since I am interested in the differences in literary language before and after the February 24th invasion of Ukraine. As much as possible, I removed poetry explicitly written prior to 2017 in order to reduce additional temporal variability in the dataset.
However, the contemporaneity of internet ephemera is sometimes difficult to determine. Since I collected from internet blogs, there was not always a "published on" notice to know when a poem was posted online. Additionally, the day a poem gets posted online may be quite different from the date it was originally written; this can be the same for print publication. I generally use the date of web publication as an indication of the relevant time period of a poem, since that date is the one on which it was deemed relevant for public dissemination by the author or editor. The online publication date, then, is a serendipitous measure of the relevance of language (both for the writer and audience) over time.

There are slightly more poems from the pre-invasion period versus the post-invasion period due to variations among the sources. Ideally, one would compare a set of poems from a single one-year span to another set of poems from a one-year span. However, because of the different qualities of my data sources (blogs posted on a single date vs. Telegram channels which were active over time), I had to include a longer time period in the past direction to collect the same number of poems during both time periods. A line graph of poems over time is shown in Figure 4, and the number of poems per time period is shown in Table 2. The line graph displays circles for No War Poetry and ROAR because the poems from those sources were essentially published on one date, not continuously.102

Figure 4: Publication date of poems in database (circles for one-time publications)

---

102 The poems from Alexander Kabanov's personal Telegram channel are not included in the graph because they are not numerous enough.
Table 2: Time period breakdown of the database

<table>
<thead>
<tr>
<th>Time of Online Publication</th>
<th>Poem Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before February 24th, 2022</td>
<td>1661</td>
</tr>
<tr>
<td>On or after February 24th, 2022</td>
<td>1561</td>
</tr>
</tbody>
</table>

The poems come in a variety of styles and lengths. The histogram in Figure 5 shows the distribution of poems by the total number of lines in the text. The average number of lines per poem is 24.92 lines, and the median is 20 lines. The length of poems ranges from one line to 262 lines.\(^{103}\)

Figure 5: Histogram of poems by number of lines

---

\(^{103}\) The 262 line poem (which could potentially be broken down into smaller poems) is here: https://eelegiap.github.io/thesis-code/search/?q=160. This is an example of a one-line poem: https://eelegiap.github.io/thesis-code/search/?q=3105.
It is occasionally difficult to determine what constitutes just one poem as opposed to several, and whether a piece of text on social media is a poem or rather some other kind of writing. Outliers are generally easy to find using text parsing tools. There are some pieces of prose poetry without line breaks which resemble paragraphs, like this text by Elena Fanailova, which requires a closer reading to determine that its genre is poetic:

"rain, wind and snow outside the window, clear air and bare branches of the garden that my father once planted. apple trees, cherries, one huge apricot, they once cooked two buckets of jam from it. three years ago I filmed through the window, how my brother and I were burning my dead parents' unnecessary things, the film remained in the old gadget. warm house. I think about its fate and history."

To emphasize the diversity of forms present in the database, at least one poem, written by Dmitriy Gerchikov, is presented as a questionnaire:

1. Do you feel used to war? Why? Yes / No. Please circle one.

2. Вы замечаете, что делаете меньше записей в социальных сетях?  
2. Do you find yourself posting less on social media?

---

104 Elena Fanailova, Facebook. https://eelegiap.github.io/thesis-code/search/?q=621
3. Do you realize that you have begun to work better, wake up earlier, distinguish colors? [...]

In addition to a medley of poetic styles, the database also includes a few poems that contain words and phrases in languages other than Russian. A 2017 poem by Andrei Sen-Senkov includes lines written in Russian transliterated to the Latin alphabet: "...neizvestniy mne chelovek..." (English: "a person unknown to me"). A poem by Irina Valerina from Babruysk, Belarus ends with a series of renditions of the phrase "How to live well/How good it is to live" in various Slavic languages (Bulgarian, Belarusian, Ukrainian, Serbian, Polish, Slovak, Slovenian):

...Как естe добро живот...
...Як добра жыць...
...Як добре жити...
...Как да живем добре...
...Jak dobrze żyć...
...Како добро животи...
...Ako žit’ dobre...
...Kako dobro živeti.

Valerina's text, for example, comments on the connections between different Slavic cultures by using the language of the text itself as a carrier of meaning. The use of words in languages other than Russian is more difficult to incorporate into the word frequency analyses because of how the practice breaks linguistic conventions, since it interrupts certain steps in the text parsing process. However, it is possible to detect the presence of anomalous text using Spacy's language detection models, and for this reason could be the focus of another research project about language switching within poetry.

---

The poetry is written in Russian, but it is not necessarily Russian in the sense of nationally Russian or originating from Russia, since many poets were born in countries other than Russia around Eastern Europe, the Caucasus, and Central Asia. Many emigrated outside of these regions over time and now live in locations such as the United States, Israel, Germany, and Canada. Therefore, the dataset showcases a distinct community of Russian-language poets who are united by their online literary connections rather than being defined by geographical boundaries. The geographical diversity of authors means that (self-)censorship may play less of a role in their work.

My poetry dataset contains at least 738 authors from at least 17 countries. The countries with the most poets are Russia, Ukraine, Kazakhstan, and Belarus. The exact counts can be found below in Table 3. Several poems were published anonymously or under pseudonyms, and occasionally, it was impossible to trace the origins of certain authors. The authors in the database span several generations, and their birth dates span from the 1930s to the 2000s (to my knowledge, the youngest poet in the dataset was born in 2003). Based on Wikipedia articles and biographies on poetry blogs, I identified the birthplaces of 677 out of 738 authors. The birth city and country provide some indication of an author's cultural background, although many moved homes at a young age and thus grew up in a different environment than indicated by the map. Figure 1a shows the distribution of birthplaces of authors in the database.109

Figure 1a: Authors by birthplace (excluding the Russian Far East)

---

108 A full breakdown can be found in the appendix.
109 For information on the sources for birthplaces and other author information, see the spreadsheet:
https://docs.google.com/spreadsheets/d/1XWr2cgFiRLZ-mO5wyCaK23wlqSBYGeFgbsZ7OODnsAc/edit?usp=sharing
110 An interactive version of the map can be found at
The information shown on the map is also reflected in Table 3, which shows the exact counts of authors born in each country. There were 61 authors for which I could not find biographical information.

Table 3: Authors by birthplace

<table>
<thead>
<tr>
<th>Country of Birth</th>
<th>Author Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Russia</td>
<td>370</td>
</tr>
<tr>
<td>Ukraine</td>
<td>86</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>26</td>
</tr>
<tr>
<td>Belarus</td>
<td>16</td>
</tr>
<tr>
<td>Uzbekistan</td>
<td>6</td>
</tr>
<tr>
<td>Moldova</td>
<td>5</td>
</tr>
<tr>
<td>Latvia</td>
<td>5</td>
</tr>
<tr>
<td>Estonia</td>
<td>4</td>
</tr>
<tr>
<td>Kyrgyzstan</td>
<td>2</td>
</tr>
<tr>
<td>Azerbaijan</td>
<td>2</td>
</tr>
<tr>
<td>Lithuania</td>
<td>1</td>
</tr>
<tr>
<td>Iran</td>
<td>1</td>
</tr>
<tr>
<td>Georgia</td>
<td>1</td>
</tr>
<tr>
<td>Tajikistan</td>
<td>1</td>
</tr>
<tr>
<td>Country</td>
<td>Count</td>
</tr>
<tr>
<td>---------------</td>
<td>-------</td>
</tr>
<tr>
<td>Turkmenistan</td>
<td>1</td>
</tr>
<tr>
<td>USA</td>
<td>1</td>
</tr>
<tr>
<td>Armenia</td>
<td>1</td>
</tr>
</tbody>
</table>

My collection process of internet poetry was guided by the credibility and popularity of the online sources involved. However, the sources turned out to encompass more generational and geographic diversity than one might expect from Russian-language publications. Such diversity of authors and publishing sources in the database help capture a more comprehensive view of Russian language use. Although falling outside the scope of this study, demographic data may be used in future research to study regional differences in the themes and styles of Russian-language poetry.

5.2. Significant Word Frequency Changes

From the word frequency calculations and significance tests, I identified all of the words which authors used much more frequently or much less frequently in works from 2022. In this section, I present these words in tables and figures to give the reader an overview of the word frequency results. Later, in the Discussion section, I will discuss the implications of these results in social and historical context with examples from the text.

The words presented in this section are those which had a significant change in frequency from pre- to post-invasion. In this context, "significance" refers to a statistical measure of whether or not a variable, such as frequency, has fluctuated due to randomness in the dataset. Thus, among the words which changed in frequency from before to after the invasion, these are the words which had the largest changes in frequency. Frequency for a word like "war" was calculated for each time period by counting the total number of poems containing the word "war" in any of its forms ("war" or "wars") and dividing that count by the total number of poems from that time period. More information on the calculation process can be found in the Methodology section.

Based on the frequency calculations and significance tests, 179 words experienced a significant increase in frequency from pre- to post-invasion, and 242 words experienced a significant decrease in frequency during the
same period. The tables below show the top words from the "significant increase" and "significant decrease" categories. All of the words shown had significant ($p < .05$) changes in frequency between the two time periods in question.

The tables (Table 4 and Table 5) offer a numerical snapshot of changes in individual word use. For both tables, each row corresponds to a single word. The keyword is the Russian word (in dictionary form/lemma) whose frequency is in question. The translated keyword is my translation of the keyword. The ratio (after/before) is the number of poems containing the keyword in the post-invasion corpus divided by the number of poems containing the keyword in the pre-invasion corpus. For example, there were 20.48 times more poems containing the word "rocket" in the post-corpus than in the pre-corpus. The frequency before and frequency after refer to the percentage of poems during that time period (before or after the invasion) which contained the keyword at least once.

Table 4 (partially shown below) shows words which increased significantly in frequency in the post-invasion corpus, while Table 5 (also partially shown below) shows words which decreased significantly in frequency during the same period. These are partial tables—more complete tables can be found in the Appendix or at the link in the footnotes.  

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Translation</th>
<th>Ratio</th>
<th>Before %</th>
<th>After %</th>
</tr>
</thead>
<tbody>
<tr>
<td>мариуполь</td>
<td>Mariupol</td>
<td>28.20</td>
<td>0.12%</td>
<td>3.40%</td>
</tr>
<tr>
<td>бомбить</td>
<td>to bomb</td>
<td>27.67</td>
<td>0.06%</td>
<td>1.67%</td>
</tr>
<tr>
<td>ракета</td>
<td>rocket</td>
<td>20.48</td>
<td>0.24%</td>
<td>4.93%</td>
</tr>
<tr>
<td>обстрел</td>
<td>artillery fire</td>
<td>N/A</td>
<td>0%</td>
<td>2.05%</td>
</tr>
<tr>
<td>буча</td>
<td>Bucha</td>
<td>N/A</td>
<td>0%</td>
<td>2.24%</td>
</tr>
</tbody>
</table>

Table 4 - Partial: Words significantly ($p < .05$) more frequent in post-invasion Russophone poetry

111 Link to significant frequency change tables: https://eelegiap.github.io/thesis-code/sigFreqTables/.
Table 5 - Partial: Words significantly \((p < .05)\) less frequent in post-invasion Russophone poetry

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Translation</th>
<th>Ratio</th>
<th>Before %</th>
<th>After %</th>
</tr>
</thead>
<tbody>
<tr>
<td>бабочка</td>
<td>butterfly</td>
<td>0.09</td>
<td>2.05%</td>
<td>0.19%</td>
</tr>
<tr>
<td>осенний</td>
<td>fall (autumnal)</td>
<td>0.12</td>
<td>1.57%</td>
<td>0.19%</td>
</tr>
<tr>
<td>сплошной</td>
<td>complete/continuous/solid</td>
<td>0.18</td>
<td>1.75%</td>
<td>0.32%</td>
</tr>
<tr>
<td>предмет</td>
<td>subject</td>
<td>0.20</td>
<td>1.57%</td>
<td>0.32%</td>
</tr>
<tr>
<td>лапа</td>
<td>paw</td>
<td>0.22</td>
<td>1.75%</td>
<td>0.38%</td>
</tr>
</tbody>
</table>

As seen in Table 4, the word "Mariupol" (Russian: Мариуполь; Mariupol)—the Ukrainian city under siege in 2022 and still under Russian occupation in February 2023—had the largest increase in frequency, with 28.2 times more mentions in the post- vs. pre-invasion period. A poem from Pavel Avtomenko shows one example of its use: "all the roads to Russian culture / clogged with Mariupol's dust."\(^{112}\) Conversely, the word with the most drastic decrease in frequency was butterfly (Russian: бабочка; babochka): "A butterfly beats on the cocoon of the day."\(^{113}\)

At face value, these two words, Mariupol and butterfly, fit the expected view of how the composition of poems shifts from pre- to post-invasion. During wartime, more literature is written about war, violence, and suffering, concepts which were embodied by the occupied city of Mariupol in 2022. Conversely, one expects a decrease in "lighthearted" subjects such as butterflies and flowers as poets reckon with dark subjects. While the Mariupol-butterfly dichotomy confirms our expectations about war's effects on poetry at a high level, the dataset has a lot more nuance to offer about the poetry of war.


In order to better characterize the trends present in the word frequency results, I used the mathematical representations of these words to make visualizations. More specifically, I visualized words using their pre-trained word embeddings vectors from Navec (Section 4.3). The Navec vectors are encoded with the "standard" usage of words as they appear in a large corpus of literary texts (not my own corpus). The plot of word embeddings is essentially a semantic spread of the words and their similarities to each other. The distances between words should roughly correspond to the "distance" in meaning between them.

It's important to note that a 2-dimensional word embedding plot must be interpreted much differently than a regular 2-dimensional scatterplot with X- and Y-axes. Rather than drawing a trend line and trying to find correlation among the points as one typically does with an X-Y plot, the word embedding plot is interpreted by examining the relative positions of points to each other. Are there clusters of similar words? Gradients of meaning? Poles of extreme opposites? The ordering and mapping of words into a vector space allows one to understand the semantic landscape of the selected words while taking into account features such as parts of speech, synonyms, and specialized vocabularies which map onto the graph. The word embedding plot allows us to confirm or question our hypotheses in the linguistic space. What is expected, and what is unexpected?

5.2.1. Words with Significantly Increased Post-Invasion Frequency

First, I analyze Figure 6, a visualization of all the words which experienced a significant increase in frequency during the post-invasion period.

Figure 6: PCA plot of words (by Navec word embedding vector) which experienced a significant \( p < .05 \) post-invasion increase in frequency
For a zoomable version of the graph with both English and Russian, see the web version of the graph in the footnotes.\textsuperscript{114} A bird's-eye view of the word embedding plot reveals semantic clusters and axes within the dataset. Based on my own interpretation of the plot, I delineated the space into quadrants from a completely qualitative standpoint. The goal of the delineation is to characterize the types of words which experienced significant increases in frequency. These quadrants include:

I. War/Military  
II. Pain/Negativity  
III. Life/Death  
IV. History/News

I give examples of words from each quadrant in Table 6.

\textsuperscript{114} Web version of Figure 6: https://eelegiap.github.io/thesis-code/sigPCA/
Table 6: Examples of significantly \((p < .05)\) more frequent words by quadrant

<table>
<thead>
<tr>
<th>I. War/Military</th>
<th>II. Pain/Negativity</th>
<th>III. Life/Death</th>
<th>IV: History/News</th>
</tr>
</thead>
<tbody>
<tr>
<td>battle (бой)</td>
<td>pain (боль)</td>
<td>you (вы)</td>
<td>Russia (Россия)</td>
</tr>
<tr>
<td>explosion (взрыв)</td>
<td>scream (крик)</td>
<td>conscience (совесть)</td>
<td>west (запад)</td>
</tr>
<tr>
<td>column (колонна)</td>
<td>corpse (труп)</td>
<td>faith (вера)</td>
<td>Kyiv (Киев)</td>
</tr>
<tr>
<td>ditch (ров)</td>
<td>hell (ад)</td>
<td>god (бог)</td>
<td>Mariupol (Мариуполь)</td>
</tr>
<tr>
<td>nuclear (ядерный)</td>
<td>shame (стыд)</td>
<td>house (дом)</td>
<td>victory (победа)</td>
</tr>
<tr>
<td>to bomb (бомбить)</td>
<td>to kill (убивать)</td>
<td>survive (выжить)</td>
<td>Putin (Путин)</td>
</tr>
</tbody>
</table>

5.2.2. Words with Significantly Decreased Post-Invasion Frequency

Second, I categorize words which experienced a significant decrease in frequency in the post-invasion period. We potentially expect to see a decrease in lighthearted and jovial words (in contrast with the gravity and pain of war). The actual spread of words is unexpected; there is a heavy emphasis on nature as well as sensory words and descriptors. In other words, the graph tells us that these categories have disappeared in the subcorpus of wartime poetry.

Figure 7: PCA plot of words (by Navec word embedding vector) which experienced a significant \((p < .05)\) post-invasion decrease in frequency
Following the same method as Figure 6, I broke up Figure 7 into four quadrants in order to characterize the dataset shown, which are of words which experienced significant decreases in frequency in 2022. I noticed that items towards the left are related to the human experience, and items to the right have more to do with nature and the environment. We can roughly separate the graph into quadrants:

I. Descriptions of the environment.
II. Human artifacts and experiences
III. Mundanity
IV. Nature
Like before, for a zoomable version of the graph with both English and Russian, see the web version of the graph in the footnotes.\textsuperscript{115} Tables 7 and 8 provide some examples of words which adhere to these categories.

Table 7: Examples of significantly ($p < .05$) less frequent words by quadrant — Descriptions of Environment and Nature

<table>
<thead>
<tr>
<th>I. Descriptions of Environment</th>
<th>IV. Nature</th>
</tr>
</thead>
<tbody>
<tr>
<td>dark (тёмный)</td>
<td>shore (берег)</td>
</tr>
<tr>
<td>warm (тёплый)</td>
<td>dawn (заря)</td>
</tr>
<tr>
<td>bridge (мост)</td>
<td>snow (снег)</td>
</tr>
<tr>
<td>wild (дикий)</td>
<td>foliage (листва)</td>
</tr>
<tr>
<td>round (круглой)</td>
<td>grass (трава)</td>
</tr>
</tbody>
</table>

Table 8: Examples of significantly ($p < .05$) less frequent words by quadrant — Human Artifacts/Experiences and Mundanity

<table>
<thead>
<tr>
<th>II. Human Artifacts/Experiences</th>
<th>III: Mundanity</th>
</tr>
</thead>
<tbody>
<tr>
<td>eyesight (зрение)</td>
<td>room (комната)</td>
</tr>
<tr>
<td>face [anatomy] (лицо)</td>
<td>cup (стакан)</td>
</tr>
<tr>
<td>letter [mail] (письмо)</td>
<td>to exist (существовать)</td>
</tr>
<tr>
<td>art (искусство)</td>
<td>to become (становиться)</td>
</tr>
<tr>
<td>experience (опыт)</td>
<td>to swim (плыть)</td>
</tr>
</tbody>
</table>

As a note, the word frequency results are generally consistent when comparing word frequencies between authors born in Russia versus authors not born in Russia. Out of all the words with significant frequency changes detected in the corpus as a whole, only one of them had its frequency move in different directions between the two demographic groups: this word is "узhas" (Russian: ужас; English: horror/terror). Among the Russian-born

\textsuperscript{115} Web version of Figure 7: https://eelegiap.github.io/thesis-code/sigPCA/
group, there was no significant difference between the before and after frequencies (2% to 1.94%). However, among the non-Russian-born group, there was a 66% increase in the use of *uzhas*, from 1.8% of poems to 3% of poems containing the word. Such frequencies are already small, so it is difficult to speak with certainty about whether *uzhas* is a salient topic in the rhetoric around Russianness. I do not examine the regional linguistic differences in depth in this thesis, but this particular finding suggests that the question of birthplace and demographics require more nuanced study.

In the Discussion section, I will respond to my research questions with a discussion of the word frequency results in this section in social and historical context.

5.3. **Web Interface**

To make the database accessible to a wider audience, I developed a web platform for browsing the archive. The web platform is particularly useful for searching and close reading. Figure 8 shows the layout of the page when searching by keyword, and Figure 9 shows the layout of the page when searching by author. On the right, there is a search bar and a results section underneath. The left side of the page is the space for displaying the selected poem.

The web interface can be used to browse all 3,222 poems in the digital archive. The user may search by author or by keyword according to the specifications of the tool:

- Author: search by first and last name or just last name
- Keyword: search in the singular/infinitive/masculine form (lemma)

Searches by lemma return results which contain the lemma in any of its word forms (e.g., a search on the verb *to speak* should return all of the poems which contain the word *speak, speaks, spoke, spoken*, or *speaking*, etc.). I pre-processed the poem texts for their lemmas using the Spacy Python package.

---

On the webpage, poems are associated with their author, source of publication, birthplace of the author, (approximate) date of publication, and the poem's unique index in my database.
Future versions of the tool should include functionalities like searching by more than one keyword, filtering by time period, searching for themes, or recommending similar poems. As of March 2023, the tool is best suited for users with knowledge of both English and Russian. Expanding the tool to include translations into English and other languages would be an exciting opportunity to bring contemporary Russophone poetry to new audiences.
6. Discussion

6.1. Questions on Morality and Mortality

«Это когда правда — не следствие факта, но предмет веры.»

"This is when truth is not a consequence of a fact, but an object of faith."
—from "Zazuan about War" by Sandzhar Yanyshev, 2022¹¹⁷

Understanding the types of words which have "entered" or "departed" the poetic vocabulary allows us to assess the flow of ideas in poetic speech at the turning point of the February 2022 invasion. In this section, I will interpret the statistically significant word frequency results initially presented in Section 5.2 (Results). First, I will discuss the increased frequency of words in the post-invasion dataset which pertain to mortality, morality, and religion. Next, I will turn to the decreased frequency of words describing nature and the environment. These high-level trends in the dataset, supported by close readings, indicate a shift in focus in contemporary Russophone poetry towards themes of mortality, morality, and religion, while attention to nature and the environment seems to have diminished.

To begin, Figure 6 presents all of the words which experienced a significant increase in poem frequency in the post-invasion period: explosion, wound, prayer, news, Kyiv, etc. Graphing the words using PCA enabled me to use the semantic space to divide the words into different regions of meaning and association. I was able to visually identify four categories which roughly coincided with the quadrants of the graph. These categories—"War/Military," "Pain/Negativity," "Life/Death," and "History/News"—describe the general direction of semantic movement in Russophone poetry since February 24, 2022.

Figure 6a: PCA plot of words (by Navec word embedding vector) which experienced a significant post-invasion increase in frequency [zoomed into III. Life/Death]

Figure 6a shows Quadrant III, which I titled "Life/Death" based on the types of words present. Although the content-free words (such as conjunctions, prepositions, pronouns, etc.) in this quadrant provide potentially interesting information about changes which occurred in the modes and functions of speech, they are not as informative about thematic change. The words with explicit military connotations (soldier, enemy, target) belong more to Quadrant I. "War/Military." I argue that the remaining words in this quadrant, such as conscience, faith, survive, and god, are related to big philosophical and spiritual questions of life and death. Within the "Life/Death" quadrant, we can roughly divide words into the categories of "Mortality," "Morality," and "Religion," although there is much overlap among the concepts involved. Below we see some of these words (which, to reiterate, all experienced a significant increase in frequency in the post-invasion period):

Significantly (p < .05) increased words related to mortality, morality, and religion
Mortality: to kill, to die/death, to survive, to destroy

Morality: good, evil, lie, conscience, guilt, honor, fear, hatred/to hate, to forgive

Religion: god, faith, cross, hope/to hope, holy, eternal, prayer/to pray, glory

The connection between these terms and the war is more abstract than, say, a missile launcher. The words in this quadrant are related to more lofty questions raised by one's proximity to the horrors of large-scale, state-sanctioned violence. How does one respond poetically to topics such as lies, murder, and evil? What is the poet's relationship to God and faith when the world exhibits a capacity for such horror? Concepts of death, guilt, fate, conscience, and suffering all feature in a poem by Ksenia Kirillova:

Застыли часы, метроном отбивает такт
За всех, кто погиб под каскадами «братских» бомб.
И в общем, неважно, ты дворник иль депутат,
Ты совесть страны навсегда воплотил собой.

И каждый в итоге решает свою судьбу
Как минимум в том, для чего и зачем страдать.
Минута молчанья и слово длиной в пять букв,
И чистая совесть, и пытка длиной в года. 118

The clock froze, the metronome beats time
For all those who died under the cascades of "brotherly" bombs.
And in general, it doesn’t matter if you are a janitor or a deputy,
You have embodied the conscience of the country forever.

And in the end everyone decides their own fate
At least in what and why to suffer.
A moment of silence and a five-letter word
And a clear conscience, and years of torture.

Poets have never shied away from philosophizing, but after February 24th, 2022, Russophone poets are discussing philosophy and spirituality more

frequently than before the invasion. In a poem published in *No War Poetry*, the poet Oleg Tupitskiy from Belarus writes,

Если мир за тысячи лет
не устал воевать —
значит, бога в помине нет
или ему плевать.

If the world has not tired of fighting
for thousands of years
it means there is no god
or he doesn't care.119

As seen in Tupitskiy's poem, highly evocative words such as *god*, *guilt*, *hope*, *evil*, and *conscience* can be used in a critical context, an important observation for interpreting these word frequency results. For example, the increased frequency of the word "hope" does not necessarily translate to a significant increase in hopefulness in the poetry. Instead, the increased frequency of some words points more specifically to an increase in discussion *about* those particular words or topics. This increase in discussion could critique or refute, rather than support, the topics in question. The absence of human qualities like faith, hope, and conscience is central to wartime poetry. We see this phenomenon across wartime poems, like "There is no place for hope. There is no peaceful life." ("Нет места надежде. Нет мирного быта.")120 and "You can love by killing, kill by loving, / without regard to conscience or reason." ("Можно любить, убивая, убить любя, / не обращаясь ни к совести, ни к уму.").121

The words which appear more frequently in the post-invasion corpus speak volumes about the questions poets were grappling with in their work. Given how death and suffering have penetrated the reality of people close to the war, metaphysical concerns of fate and faith understandably became more salient than they would be in "normal times." For similar reasons, discussions of "War/Military" (Quadrant I) also increased in frequency.

However, the proliferation of moral and military themes becomes more interesting, and in some ways, deeply troubling, when viewed in contrast with the themes which disappeared in the post-invasion period. As I will discuss in the following section, the issue is rooted in the finding that increased discussion of philosophy and war are met by decreased discussion of the everyday and the natural world. The tradeoff in themes points to a tradeoff of poetic priorities.

6.2. Nature's Vanishing Voice

Ирпень, Гостомель, Буча, Бородянка, ни в вашем воздухе, ни в водах ваших рек, ни в чёрной пахоте, ни в мареве морей отрава наших тел не растворима.

Irpin, Hostomel, Bucha, Borodyanka, neither in your air, nor in the waters of your rivers, nor in the black plowing, nor in the haze of the seas the poison of our bodies is insoluble.

—Anna Germanova, 2022

Figure 7 showed words which decreased significantly in frequency in the post-invasion period. Similar to the previous section, these words can be roughly split into the categories of "Descriptions of Environment," "Human Artifacts/Experiences", "Mundanity," and "Nature." The two quadrants discussed in this section ("Descriptions of Environment" and "Nature") are shown in more detail in Figure 7a. As shown in the figure, words related to one's environment (both natural and manmade) disappeared in poetry post-invasion.

Figure 7a: PCA plot of words (by Navec word embedding vector) which experienced a significant post-invasion decrease in frequency (I. Descriptions of Environment, IV: Nature)

---

In contrast, nature and the environment are prominent topics in pre-invasion Russophone poetry: 7.04% of pre-invasion poems contained the word *tree*, compared to the 7.17% share of the post-invasion poems which contained the word *Ukraine*.

Animals, plants, geography, and the climate all experienced significant decreases, like *landscape, bird, air, pine, river, trail, grass, and snow*. Even the word *nature* (Russian: природа; *priroda*) itself experienced a significant decrease in frequency. Most of the adjectives which experienced frequency decreases are used to describe physical properties such as color, texture, size, shape, and temperature. The disappearance of so many adjectives describing sensation in 2022 is striking.

**Adjectives with Significant Post-Invasion Decreases**

**Color:** light blue, steel blue, white, green, transparent, dark, light

**Texture:** soft, wet

**Size:** small, long

---

123 The percentage of poems containing *tree* decreased to 2.95% in the post-invasion period.
**Shape:** round, narrow

**Temperature:** hot, warm, icy

The loss of so many words referring to nature and the environment points to a deprioritization of the natural world with respect to war. This finding is surprising when viewed in light of the fact that nature and war are actually close companions. Nature plays a strategic role in warfare. During and prior to the invasion, military experts talked at length about the role Ukraine's geography would play in a potential Russian invasion, such as the frozen ground on the Pripyat/Pinsk Marshes. \(^{124}\) Words like *marsh* and *freeze* did not experience any significant frequency changes, and the word for *icy* actually decreased significantly in frequency. The natural and urban environments are also receivers of violence on a mass scale. Russian bombardments tear open apartment buildings across Ukraine and burn acres of unharvested wheat. \(^{126}\) The UN Environment Programme (UNEP) anticipates that air, water, and land pollution from the war will linger for generations to come. \(^{128}\) Words like *pollution* and *to pollute* are essentially nonexistent in the database. The word for *earth* does not experience a significant change in frequency from pre- to post-invasion, and words like *water*, *forest*, *air*, and *shore* all experienced significant decreases in frequency. Considering the huge


impact of war on the environment, why are many words related to nature being used less often?

The overall changes in word frequency are not attributable to any particular author or publishing house, as the results reflect general thematic changes which are only perceptible from a distance. The war transformed poetic rhetoric into one which must attend to questions of human fate at the expense of nature (and everyday human experiences). The poetry in the database is often explicitly and implicitly oppositional, and anti-war themes naturally seem to spur discussions of pacifism, spirituality, and philosophy. In other words, the appearance of these themes isn't all too surprising. Yet, something is lost when the poetic discussion leaves out substantial references to nature, a theme which is not only relevant to unfolding geopolitical events but also central to poetry as an empathetic act.

In these ways, the collective poetic conversation, as observed in the database of contemporary Russophone poetry, did not confront the reality of war from all of its angles. Certainly, human life and consciousness deserve a central role in poetry written about the war. However, if poetry truly attempts to take account of the war's perpetrators and victims, the works of this time are incomplete if they do not account for the wholesale destruction of nature and the lasting effects of a toxic environment on future generations. The war's effects on the psyche may distract even poets from "a profound empathy for the suffering of all living beings."129 Or, rather than being a "distraction" from the goal of poetic empathy, the war may have presented so much human suffering that poets simply do not have the attention or time to devote to nature, which has been pushed out of the conversation. This missing piece of the conversation may mirror larger gaps in the news and popular narratives about war's deleterious effects on the environment.

6.3. War's Degradation of Russian

<table>
<thead>
<tr>
<th>Что Вы сказали, принц? Я говорил слова!</th>
<th>Word-verb, word-conjunction,</th>
</tr>
</thead>
</table>

Another important result from this study was evidence of a changing relationship between Russophone authors and the language of their work. Similar to the results in previous sections, warfare seems to be depriving the Russian language of some aspect of its empathetic or communicative potential. Russophone authors are struggling with the awareness of their language's associations with violence. In the social and political environment of 2022, the Russian language is stunted in its ability to speak about Russia's violence.

Taking a look at poems which mention yazyk (the Russian word for language), the reader finds a stark picture of a degraded language:

[...], myself
Born in a killer country
A native speaker of the language of the rapist,
When you mumble even this requiem
In orcish language.

The above excerpt from a poem by Alya Khaytlina highlights the estrangement she feels from her language. Instead of writing that she speaks or talks in Russian, she writes that she "mumbles" in "orcish" language. She cannot bring herself to use her whole voice to express her thoughts in Russian, which she emphasizes is no longer her own; it is the language "of

---

132 Referring to the orcs, a brutish, violent, and ugly race of creatures created by the dark lord Sauron in J.R.R. Tolkien's Lord of the Rings. Since 2022, the term "orcs" has been especially in vogue among Ukrainians and their allies for describing Russian soldiers.
the rapist." Similar ideas about estrangement from the language arise in other poems: Vera Pavlova writes, "In the language of the aggressor / I cry for its victims," and Sandzhar Yanyshev refers to his "traitorous native language." 133 134 Marina Gurman writes that the Russian language now "writhe[s] in the lips of non-humans / Smell[ing] of burning forever:"

Искать ответа
В великом русском языке.
Язык мой, пушкинский, мой русский,
Большая, страшная беда.
Ты корчишься в устах нелюдских,
Пропахший гарью навсегда.135

Looking for an answer
In the great Russian language.
My language, Pushkin's
language, my Russian,
Big, terrible tragedy.
You writhe in the lips of
non-humans,
Smelling of burning forever.

The discomfort present in the poem speaks to the internal conflict of writing poetry related to the injustices of the war in the language of its perpetrators. The war has made the Russian language property of the regime, rendering it incompatible with the humanitarian sentiments authors hope to express.

Word frequency results provide other angles to view how war changes the author's relationship with his or her language. When looking at voice-related words which increased and decreased in frequency, we find a stark line separating speech from non-speech.

**Words with a Significant Decrease ↓ in Frequency:**
- speak (говорить)
- say (сказать)

**Words with a Significant Increase ↑ in Frequency:**
- yell (орать)
- howl (выть/вой)
- sob (рыдать)
- scream (кричать/крик)

---

The results here show that the verbs *to speak* and *to say* appear less frequently post-invasion than pre-invasion, and the verbs and nouns *yell*, *howl*, *sob*, *scream*, and *moan* all appear more frequently post-invasion than before. We see some examples:

Твой крик, твой шёпот — это глина.

(...) Вдали, где нет ни слёз, ни смеха, где отголоском — боль и стыд, твой деликатный доктор Чехов с тобой беззвучно говорит.  

—Pavel Gudanets, 2022

Сводный хор скорбящих людей и зверей живых
На разорванный воздух криком наложит швы,
Но замолкнет вскоре, запутавшись в огласовках.  

—Mikhael' Scherb, 2022

These two results together represent how during wartime, regular human speech is superseded by non-linguistic vocal expression. This swap speaks directly to Haleta's argument that linguistic conventions "collapse" in the face of violence. In "Mined Words," Haleta writes,

"Violence changes the narrator's situation, since he or she is literally lacking words, and is to have his or her voice changed to a shriek, a moan or other similar speechless reactions, including muteness, which,

---

136 Full data on these words can be found at https://eelegiap.github.io/thesis-code/sigFreqTables/


contrary to silence, has an anthropic character...The human is placed into an inhuman situation impossible either to be explained, or to be settled in a rational way. The text responds to the course of the events, but it is also a desperate attempt to hold them on the brink of ultimately breaking off into overall chaos.”

Haleta is referring to poetic responses to violence which occurred on the Ukrainian Maidan. The Ukrainian poets she studied were generally closer (physically and psychologically) to violence than most of the authors in my database, who mostly live in Russia, at a distance from the violence in Ukraine in 2022. The poetic responses to the violence on the Maidan vs. the violence of the Russian invasion of Ukraine are different in that in the latter case, Russophone poets feel their language is somehow complicit with the perpetrators of violence. It appears that poets struggle to speak straightforwardly in a language entangled with violence.

The close readings and word frequency results reveal a number of findings about the state of the Russian language. As discussed by Haleta, expressions of physical and emotional suffering are not well-suited for conventional language, and we see unconventional, non-linguistic expressions of voice take precedent in the post-invasion corpus. In the complicated case of Russophone poets writing during the Russian war in Ukraine, conventional language often appears not to be sufficient for addressing the ongoing events. The contamination of the language by its proximity to the perpetrator also causes poets to struggle to express humanitarian sentiments. How can one make amends in a certain language if the language is complicit? (And what does it mean for a language to be complicit?)

In general, authors have not been able to reconcile their individual poetic language as unique from the state's language of fear and violence. Many poets reject the language, expressing that the language is no longer their own, because of its ties to the regime's violence. It is possible that poetic discussions of estrangement from the language are a method of denouncing the regime and its official language of violence. However, authors' attempts to distance themselves from the language can feel self-defeating. There is no

139 Haleta, “Mined Words.”
clear route to establishing one's own language. Perhaps such associations are impossible to escape: in one poem, Ramil' Niyazov-Adyldzhyan circularly writes in Russian, "There will not be a single Russian word in this language."\textsuperscript{140}

6.4. Turning Away From Memory

Мы любили повторять в 80-е годы девиз и молитву: «Надо освобождаться». От всего.
Но тогда может вообще ничего не остаться, кроме движения к освобождению. Ничего и не осталось, кроме чистого движения и время от времени — сожаления о прошлом, как бывает ложная память.

In the 80s we loved to repeat the motto and prayer: "We must be freed." From everything. But then there might be nothing at all other than moving towards liberation. Nothing remains but pure movement and, from time to time, regrets about the past, as there is a false memory.\textsuperscript{141}

—Aleksandr Barash, 2022

Earlier, we learned that morality and mortality became important new themes in the post-invasion time period. In this section, I will discuss how Russophone poetry in 2022 appears to de-emphasize explicit references to memory in poetry, and why this is a crucial development in wartime Russophone poetry. During the post-invasion period, poets began to use fewer words related to memory and experience, such as memory, experience (noun), recollection, and to remember, and more words related to the present and

\begin{footnotesize}
\footnotespace
\begin{enumerate}
\item\textsuperscript{140} Ramil' Niyazov-Adyldzhyan, "CoRusPo," ROAR V3, Aug 24, 2022, https://eelegiap.github.io/thesis-code/search/?q=523
\item\textsuperscript{141} Aleksandr Barash, "CoRusPo," ROAR V3, May 31, 2022, https://eelegiap.github.io/thesis-code/search/?q=235
\end{enumerate}
\end{footnotesize}
future. These words and their statistics are shown in Table 9. Additional word frequency results about the tenses of the verb to be signify a shift in poetry towards the future and away from the past, suggesting a temporal re-orientation of Russophone poets writing in 2022.

The most relevant time period in 2022 Russophone poetry is almost certainly the present. The unfolding events in Ukraine are at the forefront of poetic discussion, as seen with a 28 times increase in mentions of "Mariupol." Other cities in Ukraine, such as Kyiv, Kharkiv, Bucha, and Irpen, feature widely in post-invasion poetry. Moreover, newspapers and the news have become increasingly prominent in the poetry—the words for "the news" (Russian: новости; novosti) and newsfeed (Russian: лента; lenta) more than doubled in frequency in the post-invasion period (2.38x and 2.04x respectively). In 2022, the news, authors write, is awful, inescapable, and vitally necessary. Olga Andreeva writes, "The news comes straight to my blood / bypassing websites, blogs and Facebook." Many poems mention cities, people, and organizations from the time (Kyiv, Putin, NATO etc.), giving the poetry a documentary quality. Most of the poetry since the invasion is about the war in some way; it’s impossible to not be aware of it. The present moment is central in 2022 Russophone poetry. In this section, I will discuss close readings and word frequency results which talk about other temporal aspects of the poetry: are poets more oriented towards discussions of the past or future?

The events of 2022 seem to have drawn poetic attention away from memory; it is simply being discussed less in the poetic discourse. Words such as memory, experience (noun), recollection, and to remember are all being used significantly less frequently in the post-invasion period. For example, the post-invasion frequency for recollection is only 1/4th of its pre-invasion frequency. Such linguistic losses are likely not due to random fluctuations in the vocabulary because they pass the threshold of statistical significance.

---

142 Russian: память, опыт, воспоминание, помнить
143 The word for newsfeed, lenta, also means ribbon, and its increase in frequency is partially attributed to references to ribbons in the military context.
Table 9: Significant ($p < .05$) decreased frequency of words related to memory and experience

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Translation</th>
<th>Ratio</th>
<th>Before %</th>
<th>After %</th>
</tr>
</thead>
<tbody>
<tr>
<td>воспоминание</td>
<td>recollection</td>
<td>0.24</td>
<td>1.32%</td>
<td>0.32%</td>
</tr>
<tr>
<td>опыт</td>
<td>experience</td>
<td>0.46</td>
<td>1.69%</td>
<td>0.77%</td>
</tr>
<tr>
<td>память</td>
<td>memory</td>
<td>0.66</td>
<td>6.08%</td>
<td>4.04%</td>
</tr>
<tr>
<td>помнить</td>
<td>remember</td>
<td>0.72</td>
<td>7.28%</td>
<td>5.25%</td>
</tr>
</tbody>
</table>

However, the decrease in references to memory does not necessarily mean that memory is a taboo topic among Russophone poets. Many do discuss individual and collective memories and the role memory may play in the future.

Мне не уйти от этого огня.
Христа помянешь — помяни Иуду.
Ирпень и Буча, вы теперь повсюду.
Помянут вас — помянут и меня.145

I can't get away from this fire.
If you remember Christ, remember Judas.
Irpin and Bucha, you are everywhere now.
If they remember you, they will remember me.

However, as seen, when memory is discussed, it is often used to express hopelessness and difficulty. Yekaterina Zadirko writes, "open your notebook and start writing / memory is work / memory is revenge."146 Pavel Grushko, from Odessa, writes in a hopeless tone,

Ой-ё-ёй, как много снега... Wo-wo-wow, so much snow...

---

This is not to say that the concept of memory has completely transformed in Russophone poetry in 2022. Mentions of memory in the pre-invasion period are also not uncontroversial or uncomplicated—memory is often obscured or false. In June 2018, Sergei Gandlevskiy writes, "Tell lies, memory, lie serenely: There are no eyewitnesses, I am the last one." However, there are significantly more references to memory among authors in the pre-invasion period than in the post-invasion period, and these references often incorporate nature. In 2021, Anna Glazova writes, "let the memory / when rolled back / returns the tides to the foliage, / a little noise can be heard / under the trees..." There are a number of other connections between nature and memory in the pre-invasion period. Other authors write, "the forest remained askance like a lasting memory" and "I guess it's something to do with atmospheric memory. / Maybe she has some kind of memory, the atmosphere..." As shown, when memory was mentioned in 2022, it was being used in different contexts than before. However, we must keep in mind that the use of the word memory also decreased in frequency in the post-invasion period, pointing to a potentially diminished role of memory in the discourse during this time.

Other quantitative results point to other possible evidence that post-invasion poetry is looking away from memory. The changing frequencies of the verb *to be* (Russian: быть; byt) offer some insight on this question. Perhaps not coincidentally, the past tense of *to be* decreases in frequency while occurrences of the future tense increase during the post-invasion period. In

---

particular, the past tense form "it/there was" decreased significantly in frequency, while the future tense forms "we will," "they will," and "he/she/it will" increased significantly, as shown in Table 10. That is, the word frequencies reflect an increase in future tense and a decrease in past tense of the verb to be, an incredibly important verb for description and expression.

**Table 10: Forms of "to be" which experienced significant (p < .05) changes in frequency**

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Translation</th>
<th>Ratio</th>
<th>Before %</th>
<th>After %</th>
</tr>
</thead>
<tbody>
<tr>
<td>было</td>
<td>it/there was</td>
<td>0.75</td>
<td>10.78%</td>
<td>8.14%</td>
</tr>
<tr>
<td>будем</td>
<td>we will</td>
<td>2.13</td>
<td>1.32%</td>
<td>2.82%</td>
</tr>
<tr>
<td>будут</td>
<td>they will</td>
<td>1.67</td>
<td>1.81%</td>
<td>3.01%</td>
</tr>
<tr>
<td>будет</td>
<td>he/she/it will</td>
<td>1.46</td>
<td>9.63%</td>
<td>14.03%</td>
</tr>
</tbody>
</table>

Close readings of the text offer some explanations of how the future fits into post-invasion Russophone poetry. Some authors write somewhat optimistically about how the war will end and what the future will bring. Some imagine victory parades in Kyiv and rebuilding Ukraine's cities:

Когда закончится война,  
Украина будет спасена  
От нас — и, гарь отхаркав,  
Мы будем строить Харьков.

When the war ends  
Ukraine will be saved  
From us — and, after coughing up the fumes,  
We will build Kharkiv.

Когда растает страшный сон,  
В цветах поднимется Херсон,  
Изюм и Мариуполь.  

When the terrible dream melts away,  
Kherson, Izyum and Mariupol  
Will rise in flowers.  
—Tatyana Voltskaya, 2022

Authors are concerned with what will happen to both Ukraine and Russia. The poet Grigoriy Blanshteyn writes, "Who started this mess / In a foreign

---

152 Tatyana Voltskaya, "CoRusPo," No War Poetry, Apr 21, 2022,  
https://eelegiap.github.io/thesis-code/search/?q=1680
land? Who will receive retribution / On what scale? It appears to be easier for poets to write about the future than the past, which is in keeping with results which show decreased references to memory. The future, although uncertain, is still unwritten. In contrast with difficult pasts and a difficult present, the future is at least a place where one can imagine something better. This type of ideation appears at times as hopefulness and at others as a coping mechanism.

Discussions of individual and shared memories are crucial for maintaining a society's freedom. Russia, of course, is far from being a free society. The country's volatile history throughout the 20th and 21st centuries problematizes memory for those who lived through it or live with its history. It is difficult to form a clear and consistent understanding of the past when regimes and borders change. In 2022, as well as at different points in the past, the Russian government aimed to erase certain memories which did not support its agenda and used propaganda to mold a reality which adheres to fixed narratives. For example, we see this in the Russian government's persecution and dismemberment of the Russian human rights organization "Memorial," which was working to document the crimes against humanity committed in the Soviet Union under Joseph Stalin.

As demonstrated in this section through quantitative and qualitative means, post-invasion Russophone poetry tends to focus heavily on the present and future, with a decreased emphasis on memories. This sets off alarm bells for those concerned with poetry's role as a keeper of memory. Why is memory not mentioned more often in post-invasion Russophone poetry? The poetry from 2022 responds quickly to the war as it unfolds, but it struggles to process and acknowledge memory as it relates to the war. Of course, one might talk about memory without referencing memory itself, but it would be hard to say that poetry is becoming more subtle at the moment. Poetry written in reaction to the 2022 invasion of Ukraine appears to face the future more confidently than the past; in other words, some poets do not appear able to "face" memory. The way that memory and time are addressed in the poetry may


unintentionally play, in small part, into the regime's agenda of continued memorylessness.
7. Conclusion

To analyze changes in language through poetry, this thesis compares Russophone poetry before and after the February 24th, 2022, Russian invasion of Ukraine, using a collection of Russian-language poetry posted to the internet from 2017-2022 as a novel database. The study aims to understand what happens to a subset of literature when the common language is under duress, and how poetry can serve as a worthwhile source for studying language change as it reflects societal conditions. This work incorporated both quantitative and qualitative methods, including computational semantic analyses of a large corpus and individual close readings of selected poems, to explore the changing relationships between language and politics in Russophone poetry.

7.1. Overview of Methodology and Results

To summarize the methodology, I used quantitative analyses corroborated by close readings to shed light on war-induced shifts in word choice in Russophone poetry. Quantifying and comparing the linguistic changes which occurred on the level of word choice give us access to trends which stretch across a corpus of text. From this, we learn about the significance and associations of words, ideas, and topics which transcend individual poems. I visualized quantitative representations of these words to better understand the semantic space they occupied. Lastly, I developed a web platform to more easily browse the poetry database by author and keyword and close read.

This methodology allows us to examine high-level trends with a concrete basis in authors’ choices of words. Beginning with the question of how the language of poetry changes when it is the language of the aggressor, we learn how authors feel estranged from the Russian language in reaction to the violence and absurdity permeating official Russian discourse in 2022. Authors are painfully aware of the limitations of poetry: as Dmitriy Rastaev writes, "you cannot wash away blood with words."\(^{155}\) Returning to discussions of Paul Celan's "non-German German," one finds many authors in the database who pine for a Russian which is not Putin's Russian. Poets seek a

---

language that is *russkiy* (русский) rather than *rossiyskiy* (российский), but they don't manage to find it.\(^{156}\) Close reading of the texts provided us with examples of authors' worsening relationships towards the Russian language, and my analysis of trends in word frequency illustrated language's decreased expressive potential during times of violence: language less often "speaks."

While close readings were generally more helpful for analyzing the discussion of linguistic estrangement, computational methods were more useful for uncovering thematic shifts. In the period following the invasion, there was a marked increase in themes of war, morality, and mortality in the poetry. Simultaneously, themes of nature, mundanity, and memory became less frequent, reflecting the impact of war on authors' lives and priorities. Animals, plants, features of the landscape and the words describing them disappeared in frequency despite the violence inflicted on nature during war (and the environment's strategic role in warfare).

Thematic results such as these reflect an accumulation of individual poetic choices which become cohesive only when viewing the corpus of poetry as a whole. By using quantitative methods, we begin to understand the common threads which pass through many different poems. The poems in the database, written by a group of interlinked authors during a particular historical moment, are part of a large, connected opus. This contributes to the appeal of studying internet poetry, especially on social media sites, since online poetry is part of a public conversation. The conversational aspect of this corpus underscores its value in relation to society at large, as poets and publishers aim to share poetry with wider audiences.

Overall, contemporary wartime Russophone poetry seems to serve a number of different purposes, such as emotional catharsis, social critique, and historical documentation. We see acute emotional responses to the highly troubling events of 2022. These deeply painful poems are directed at both internal and external suffering. This poetry also levies political critiques of Russian government and society. For authors who are Russian citizens still living in Russia, the action of publishing to a blacklisted site such as *ROAR*

\(^{156}\) In essence, *russkiy* (русский) is used to describe ethnic and cultural aspects of Russia, while *rossiyskiy* (российский) refers to the political and institutional aspects of the country.
could potentially lead to social or legal consequences. And most centrally, Russian-language poetry in 2022 documents the present moment. It forms a record of the horrific tragedies perpetrated by Russia in Ukraine and the harrowing political reality unfolding in Russia. Unlike a news archive, poetry documents experiences which must be expressed creatively.

7.2. Limitations and Future Work

This study showcases the potential for interdisciplinary research by combining Slavic studies, linguistics, and computer science to uncover patterns and trends in poetic expression. The experimental combination of quantitative and qualitative methods showcased the potential for natural language processing to detect subtle, large-scale thematic shifts in text, which can be valuable for various literary and non-literary applications, such as sentiment analysis and trend detection in news and social media.

The interdisciplinary approach also emphasizes the importance of poetry in relation to society. Language and reality are mutually constitutive: reality shapes our language and language shapes our reality. This project is a case study of how poetry can be used as a worthwhile source for understanding the relationship between language and reality during times of crisis, such as war. As seen in the archive, contemporary poetry is a literary genre imbued with highly concentrated emotion and meaning. Because of these qualities, contemporary poetry can speak to the limits of what can be expressed and how we can talk about subjects which appear unspeakable. Poetic texts are often a place for authors to process and reflect on reality, and, for those of us who engage with poetry, poems have a way of altering our perceptions of the world.

There are many future avenues for research. Naturally, more work can be done on the archive itself to expand the time period, data sources, and authors included. With these changes, rich work may be conducted during other periods of history. Collecting all of the necessary metadata would allow researchers to study particular groups of poets, like only those born in Belarus or only those born after 1991. Although 3,222 (the number of poems in my analysis) is a large number of poems, my quantitative analyses are limited in comparison to other studies which use larger digital archives, like
news websites, established digital libraries, and Wikipedia data, which provide hundreds of thousands or millions of examples. Larger datasets combined with more sophisticated methods such as transformer-based machine learning could reveal more subtle trends and test specific hypotheses with greater confidence. For example, based on a poet's citizenship or residence, one might train machine learning models to detect instances of self-censorship or coded language in poems written under strict censorship, examining how poets adapted their language and themes to navigate restrictions on creative expression.

In conclusion, this thesis presents a broad analysis of Russophone poetry before and after the February 24th, 2022 Russian invasion of Ukraine, highlighting the profound impact of war on language, poetic themes, and the poetic imagination. Ultimately, the study of contemporary wartime Russophone poetry serves as a testament to the importance of documenting the human experience, even during times of extreme turmoil and suffering.

Мысленно бродя по кладбищу
друзей и знакомых
думаешь, кому повезло больше:
тем, кто умер до февраля 20-го
и не узнал новый всемирный смысл
слова «корона»,
или тем, кто не дожил до февраля 22-го
и не узнал новый позорный смысл
слова «война».157

Mentally wandering through a cemetery of friends and acquaintances
Who do you think is luckier?
those who died before February 2020
and did not learn a new universal meaning
of the word "corona,"
or those who did not live to see February 2022
and did not learn a new shameful meaning
of the word "war."

—Iosif Galperin, 2022

### 8. Appendix

Table 4: Top 25 words significantly \((p < .05)\) more frequent in poems post-invasion\(^{158}\)

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Translation</th>
<th>Ratio</th>
<th>Before %</th>
<th>After %</th>
</tr>
</thead>
<tbody>
<tr>
<td>маринуполь</td>
<td>Mariupol</td>
<td>28.2</td>
<td>0.12%</td>
<td>3.40%</td>
</tr>
<tr>
<td>бомбить</td>
<td>to bomb</td>
<td>27.67</td>
<td>0.06%</td>
<td>1.67%</td>
</tr>
<tr>
<td>ракета</td>
<td>rocket</td>
<td>20.48</td>
<td>0.24%</td>
<td>4.93%</td>
</tr>
<tr>
<td>обстрел</td>
<td>artillery fire</td>
<td>N/A</td>
<td>0%</td>
<td>2.05%</td>
</tr>
<tr>
<td>буча</td>
<td>bucha</td>
<td>N/A</td>
<td>0%</td>
<td>2.24%</td>
</tr>
<tr>
<td>харьков</td>
<td>Kharkiv</td>
<td>19.86</td>
<td>0.18%</td>
<td>3.59%</td>
</tr>
<tr>
<td>киев</td>
<td>Kyiv</td>
<td>18.36</td>
<td>0.24%</td>
<td>4.42%</td>
</tr>
<tr>
<td>сирена</td>
<td>siren</td>
<td>15.96</td>
<td>0.18%</td>
<td>2.88%</td>
</tr>
<tr>
<td>украина</td>
<td>Ukraine</td>
<td>13.24</td>
<td>0.54%</td>
<td>7.17%</td>
</tr>
<tr>
<td>российский</td>
<td>Russian</td>
<td>8.94</td>
<td>0.30%</td>
<td>2.69%</td>
</tr>
<tr>
<td>танк</td>
<td>tank</td>
<td>8.83</td>
<td>0.60%</td>
<td>5.32%</td>
</tr>
<tr>
<td>ядерный</td>
<td>nuclear</td>
<td>8.51</td>
<td>0.18%</td>
<td>1.54%</td>
</tr>
<tr>
<td>бомба</td>
<td>bomb</td>
<td>7.63</td>
<td>0.36%</td>
<td>2.75%</td>
</tr>
<tr>
<td>град</td>
<td>Grad (weapon)</td>
<td>6.38</td>
<td>0.36%</td>
<td>2.31%</td>
</tr>
<tr>
<td>кровавый</td>
<td>bloody</td>
<td>6.08</td>
<td>0.84%</td>
<td>5.12%</td>
</tr>
<tr>
<td>пуля</td>
<td>bullet</td>
<td>6.07</td>
<td>0.60%</td>
<td>3.65%</td>
</tr>
<tr>
<td>совесть</td>
<td>conscience</td>
<td>5.5</td>
<td>0.36%</td>
<td>1.99%</td>
</tr>
<tr>
<td>ложь</td>
<td>lie</td>
<td>5.43</td>
<td>0.60%</td>
<td>3.27%</td>
</tr>
</tbody>
</table>

\(^{158}\) Full table (with all 179 words with significant increases) available at https://eelegiap.github.io/thesis-code/sigFreqTables/
### Table 5: Top 25 words significantly \( (p < .05) \) less frequent in poems post-invasion

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Translation</th>
<th>Ratio</th>
<th>Before %</th>
<th>After %</th>
</tr>
</thead>
<tbody>
<tr>
<td>бабочка</td>
<td>butterfly</td>
<td>0.09</td>
<td>2.05%</td>
<td>0.19%</td>
</tr>
<tr>
<td>осенний</td>
<td>fall (autumnal)</td>
<td>0.12</td>
<td>1.57%</td>
<td>0.19%</td>
</tr>
<tr>
<td>сплошной</td>
<td>continuous/universal</td>
<td>0.18</td>
<td>1.75%</td>
<td>0.32%</td>
</tr>
<tr>
<td>предмет</td>
<td>subject</td>
<td>0.2</td>
<td>1.57%</td>
<td>0.32%</td>
</tr>
<tr>
<td>лапа</td>
<td>paw</td>
<td>0.22</td>
<td>1.75%</td>
<td>0.38%</td>
</tr>
<tr>
<td>небольшой</td>
<td>small (not big)</td>
<td>0.22</td>
<td>1.44%</td>
<td>0.32%</td>
</tr>
<tr>
<td>сосна</td>
<td>pine</td>
<td>0.23</td>
<td>1.93%</td>
<td>0.45%</td>
</tr>
<tr>
<td>ягода</td>
<td>berry</td>
<td>0.23</td>
<td>1.38%</td>
<td>0.32%</td>
</tr>
<tr>
<td>воспоминани</td>
<td>memory</td>
<td>0.24</td>
<td>1.32%</td>
<td>0.32%</td>
</tr>
</tbody>
</table>

\(^{159}\) Full table (with all 242 words with significant decreases) available at [https://eelegiap.github.io/thesis-code/sigFreqTables/](https://eelegiap.github.io/thesis-code/sigFreqTables/)
<table>
<thead>
<tr>
<th>существо</th>
<th>essence/being</th>
<th>0.24</th>
<th>1.32%</th>
<th>0.32%</th>
</tr>
</thead>
<tbody>
<tr>
<td>кое</td>
<td>something</td>
<td>0.25</td>
<td>1.26%</td>
<td>0.32%</td>
</tr>
<tr>
<td>круглый</td>
<td>round/total</td>
<td>0.28</td>
<td>1.81%</td>
<td>0.51%</td>
</tr>
<tr>
<td>стакан</td>
<td>cup</td>
<td>0.28</td>
<td>1.63%</td>
<td>0.45%</td>
</tr>
<tr>
<td>выдох</td>
<td>exhalation</td>
<td>0.28</td>
<td>1.38%</td>
<td>0.38%</td>
</tr>
<tr>
<td>блестеть</td>
<td>to shine</td>
<td>0.28</td>
<td>1.63%</td>
<td>0.45%</td>
</tr>
<tr>
<td>жар</td>
<td>heat</td>
<td>0.32</td>
<td>1.38%</td>
<td>0.45%</td>
</tr>
<tr>
<td>дикий</td>
<td>wild</td>
<td>0.32</td>
<td>1.38%</td>
<td>0.45%</td>
</tr>
<tr>
<td>тропа</td>
<td>pathway</td>
<td>0.32</td>
<td>1.38%</td>
<td>0.45%</td>
</tr>
<tr>
<td>стальной</td>
<td>steel</td>
<td>0.32</td>
<td>1.20%</td>
<td>0.38%</td>
</tr>
<tr>
<td>трамвай</td>
<td>tram</td>
<td>0.32</td>
<td>1.38%</td>
<td>0.45%</td>
</tr>
<tr>
<td>высоко</td>
<td>high</td>
<td>0.33</td>
<td>2.11%</td>
<td>0.70%</td>
</tr>
<tr>
<td>форма</td>
<td>form</td>
<td>0.33</td>
<td>2.71%</td>
<td>0.90%</td>
</tr>
<tr>
<td>листва</td>
<td>foliage</td>
<td>0.34</td>
<td>2.29%</td>
<td>0.77%</td>
</tr>
<tr>
<td>состоять</td>
<td>to comprise</td>
<td>0.34</td>
<td>1.14%</td>
<td>0.38%</td>
</tr>
</tbody>
</table>
Figure 6: PCA plot of words (by Navec word embedding vector) which experienced a significant ($p < .05$) post-invasion increase in frequency.
Figure 7: PCA plot of words (by Navec word embedding vector) which experienced a significant ($p < .05$) post-invasion decrease in frequency.
Bibliography

https://jalammar.github.io/illustrated-word2vec/.
https://literature.ucsd.edu/people/faculty/aglaser.html.
October 12, 2012. 
https://doi.org/10.1162/tacl_a_00051.


Kabanov, Alexandr Kabanov. “Poetry Telegram Channel of Alexandr Kabanov.” Telegram. Accessed December 7, 2022. https://t.me/s/kabanov_slovo?q=%D1%82%D0%B0%D1%80%D1%8B%D1%88%D1%8B%20&before=248.


