Psycholinguistic Approach to the Analysis of Manipulative and Indirect Hate Speech in Media

Yuliya Krylova-Grek *

University of Hradec Králové, the Czech Republic

Abstract. he present study takes a psycholinguistic approach to the analysis of Russian media texts published between December 1, 2020 to May 31, 2021. I aimed to provide a scientific basis for the existence of manipulative and indirect hate speech using an interdisciplinary methodology comprising linguistic, psycholinguistic, and other analytical methods such as fact-checking and logical analysis. This facilitated the identification of techniques employed by the authors of the respective texts. In the article, I describe how I use the methodology to analyse media texts. I discovered that three basic types of hate speech were used to influence the audience's consciousness: (1) direct hate speech; (2) indirect (hidden) hate speech; and (3) manipulative hate speech. The first and second types were the most common. This may be explained by the fact that direct hate speech is condemned by international organisations and its use may be a reason for lawsuits against media outlets and their further penalisation. Texts with evidence of the second and third types of hate speech aimed to create a negative attitude toward a particular nationality, race, citizen, and so on. I consider such behaviour to be an early manifestation of widespread discrimination and other forms of intolerance, including possible violence and genocide. The present study was carried out in collaboration with a Crimean human rights group. The author was invited to participate as an expert in the field of psycholinguistic textual analysis. The research was prepared and completed at the beginning of February 2022, on the eve of the invasion of Ukraine. We have gathered evidence of indirect and manipulative hate speech that dehumanised, demonised, and marginalised Ukrainian citizens. This has led to violence against the civilian population and high numbers of casualties. The aforementioned methodology will continue to be used in the analysis of current media content.

Keywords: media text, psycholinguistic analysis, Ukraine, war, hate speech.

Крилова-Грек Юлія. Психолінгвістичний підхід до аналізу маніпулятивної та прихованої мови ворожнечі в медіа.

Анотація. У роботі описано психолінгвістичний підхід до аналізу медіаконтенту, зокрема представлено результати дослідження текстів ЗМІ, що акредитовані Роскомнадзором за період з 1 грудня 2020 до 31 травня 2021 року. Мета дослідження полягала в науковому обґрунтуванні наявності маніпулятивної та непрямої мови ворожнечі серед поданих на експертизу медіатекстів. Для проведення дослідження використовувалась авторська методика психолінгвістичного аналізу тексту, яка базується на міждисциплінарному підході та включає психологічні, лінгвістичні та психолінгвістичні методи аналізу текстів, а також інші аналітичні методи, наприклад, факт-чекінг, аналіз логіки побудови тексту, відповідність стандартам журналістики. Метод психолінгвістичного

* Yuliya Krylova-Grek, ⓘ https://orcid.org/0000-0002-2377-3781, E-mail: yulgrek@gmail.com

аналізу тексту дає змогу виявити методи та прийоми, якими послуговується журналіст для розповсюдження прихованої та маніпулятивної мови ворожнечі. У результаті роботи авторка дійшла висновку, що сучасні медіа послуговуються трьома типами мови ворожнечі, розподіленими нами залежно від мовленнєвих та немовленнєвих засобів, що використовуються у тексті для здійснення впливу на свідомість читача: 1 тип – пряма мова ворожнечі; 2 тип – непряма мова ворожнечі; 3 тип – маніпулятивна мова ворожнечі. Виявлено, що журналісти здебільшого використовували другий та третій тип мови ворожнечі, оскільки використання прямої мови ворожнечі засуджується на законодавчому рівні, та може призвести до судових позовів та стягування штрафів з медіавидання. Тексти з ознаками мови ворожнечі є інструментом маніпуляції та впливу на свідомість читацької аудиторії з метою створення негативного ставлення до національної, расової приналежності, громадянства тощо. Таку поведінку медіа розглядаємо як початковий етап роздмухування дискримінації, насилля та інших проявів нетерпимості. Робота виконувалася у співпраці з Кримською правозахисною групою в межах міжнародного проєкту, до якого авторка була залучена як експерт-психолінгвіст. Дослідження було завершено на початку лютого 2022 напередодні військового вторгнення в Україну. У межах цього дослідження з'ясовано, що систематичне розповсюдження мови ворожнечі вплинуло на сприйняття російським суспільством українців та України, сприяло дегуманізації, демонізації та маргіналізації, що в підсумку вилилося в насильницькі дії проти цивільного населення України. Наразі методологія продовжує використовуватися для аналізу медіаконтенту в теперішній ситуації.

Ключові слова: *медіатекст, психолінгвістичний аналіз, Україна, війна, мова ворожнечі.*

Introduction

Under modern conditions, the media play an increasingly important role in the formation of public opinion during conflicts and crises, both internally and internationally. The situation in Ukraine has become an example of the foundation on which it is possible to observe the consequences of the informational war and the methods journalists use in their work.

The media's responsibility for inciting hate speech, which can lead to discrimination, violence, and genocide, is stated in Recommendation No. (97) 20 of the Committee of Ministers of the Council of Europe. The Recommendation states that member states, together with local authorities and other officials, are responsible for statements in the media which can contribute to the dissemination of racial hatred, xenophobia, and other forms of discrimination. The recommendations emphasise that while evaluating the work of a media professional, a clear distinction should be driven between the responsibility of a person expressing information that incites hatred and the responsibility of media professionals involved in disseminating such ideas (Recommendation No. R (97) 20).

In this research, I examined the role of the media in shaping negative and hostile attitudes toward certain groups. Specifically, it relates to how the media create the preconditions for future discrimination and possible genocide that generates considerable interest. In the current research, I considered the example of neighbouring communities and government agencies that coexisted peacefully before the informational aggression.

There are numerous examples of informational aggression throughout history when the dissemination of hate speech would lead to murders on the basis of racial, national, ideological, and other differences. For example, the massive Tutsi genocide in Rwanda in 1994 had been provoked by large-scale anti-Tutsi rhetoric on local radio over almost a year (Yanagizawa-Drott, 2014; Melvern, 2004), the Srebrenica massacre (13-22 July 1995) which was preceded by an anti-Muslim information campaign in the Serbian media, which was controlled by Milosevic's government (DellaVigna et al., 2014), and the state terror of Muslims in Myanmar (Selth, 2004).

In the current study of hate speech, certain difficulties in defining the concept of "hate speech" are noted. The main reasons for the difficulties in interpreting the concept lie in the legal and ethical planes: 1) nowadays, there is no unification of the "hate speech" concept in the legislation of different countries; 2) differences exist in the interpretation of the concept in academic science and legal practice (Howard, 2019; Tontodimamma et al., 2021; Strossen et al., 2016; Waldron, 2012, etc.); 3) while determining whether the language used contains hostility, the problem lies in how to delineate the boundary between where freedom of speech ends and the language of hostility begins (Sellars, 2016); and 4) the same statements are perceived differently in various cultures and so may be considered as both offensive and an expression of freedom of speech from different viewpoints. (Salminen et al. 2018).

Douglas (2012) and Ben-David et al. (2016) have considered hate speech in social media. In particular, Douglas (2012) stresses that despite a number of studies investigating hate speech, there is a lack of research exploring its consequences, so this problem should be considered more thoroughly. Ben-David et al. (2016) studied overt and covert hate speech on Facebook in Spanish society and stressed that despite the website's algorithm settings blocking overt hate, covert hate speech is widespread on social media via the "comment," "like," "share," or "report" buttons. In this regard, he stresses that it is not sufficient to use as the only method of content analysis for the identification of covert hate speech; instead, non-linguistic tools, technological affordances, and so on should be taken into consideration. In my research, I consider a set of extra factors such as language means, non-linguistic tools, text building, and social content.

One of the ways to analyse hate speech in the text is linguistic analysis, which often involves identifying the lexical meanings of words in context. At the same time, it is often insufficient, as it is important to consider other factors (current situation, state-society relations, channels of spreading, affected emotions etc.) which influence the creation, perception, and interpretation of texts (Leets, 2002; Fairclough, 2003; Machikova, 2015; Matsuda et al., 1993; Whillock & Slayden, 1995; Paz et al, 2020).

Thus, substantiation of the presence of hate speech in the text requires an interdisciplinary scientific approach that will be applicable in legal practice.

In the current research, journalists' work was investigated from the psycholinguistic point of view; specifically, the work considered the linguistic and non-linguistic tools used to influence the audience.

I considered Russian-language media texts covering the period 1 December 2020 to 31 May 2021. These media outlets were officially registered in Russia and transmitted information on the territory of the Russian Federation and Crimea. The texts were selected using a content analysis programme.

By analysing the selected texts, I took notice of the fact that journalists often use hate speech bypassing formal legal prohibitions (direct insults and calls for action). I deem this manner of hate speech dissemination to be indirect or manipulative.

The aim of the research is 1) to substantiate the existence of hate speech in officially registered Russian online media outlets; 2) to identify the types of hate speech used by modern journalists; 3) to show the methods and techniques used by journalists to spread hidden and manipulative hate speech; and 4) to identify whether hate speech in the analysed publications is systematic and purposeful in order to discriminate, dehumanise, and marginalise the community by language and nationality.

The hypothesis is that the psycholinguistic analysis of media outlet content will allow the study to reveal the hidden and manipulative language of hostility and to determine whether it is systemic, which allows the prediction of possible manifestations of discrimination, mass killing, or genocide.

Method

To detect hate speech, I used the author's methodology of psycholinguistic text analysis (patent registration #30563 / 3M / 22). The methodology helps to provide a scientific basis for the existence of manipulative and indirect hate speech with the help of an interdisciplinary methodology. The methodology comprises linguistic, psycholinguistic, method of content analysis and analytical methods such as fact-checking and logical and structural analysis.

The methodology involves analytical linguistic means and non-linguistic tools that are used by journalists, because non-linguistic tools are an important instrument of visual impact on readers' consciousness (photographs, pictures, font and colour, and so on).

In addition, I consider adherence to journalism standards and ethical norms (standards). The methodology requires the consideration of both words and sentences both in the context of the whole text and in the context of the current situation.

For text selection, I used the method of *content analysis* that was carried out with a computer programme technically developed by a specialist belonging to a Crimea Human Rights group. The programme selected texts according to key units that include words and word combinations from ten Russian online media outlets published in Russia and in Ukrainian territories occupied by the Russian Federation. These sites have more than 1 million visitors per month: "Forpost Sevastopol", "RIA Krym", "KP Krym", "Novosti Kryma", "Krym Realii", "Russkaia vesna", "Politnavigator", "Novorossinform", "Krym24", and "Vesti Krym" (see Fig. 1).

Figure 1
The Audience of Selected Sites
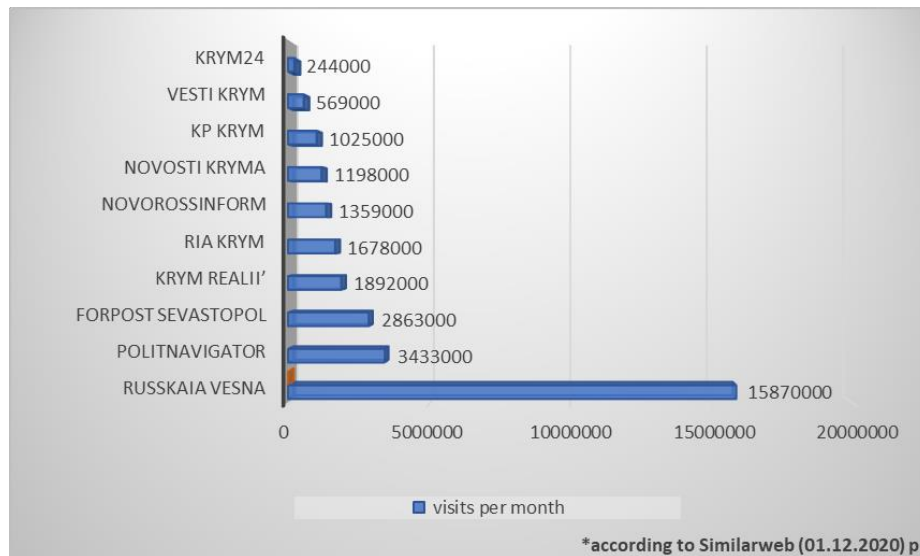
The key units (words and word combinations) were gathered from 2014-2017 and express hate, humiliation, negative sarcasm, offence, and other manifestations of hate speech to ethnic groups, nationalities, languages, or gender. During the research the list of words has been constantly supplemented by new words and word combinations.

All selected key words and combinations were gathered in the hate speech dictionary that comprises more than 400 words and word combinations (Sedova & Pechonchik, 2018). Moreover, because the specifics of the newly-created words are unclear for people who are not Russian or Ukrainian native speakers, I constructed a special dictionary where the meaning of these words and expressions are explained (Sedova & Krylova-Grek, 2021).

Among linguistic methods used were lexical and semantic analysis. I use lexical and semantic analysis to identify the meaning of the words and their combinations in context. To carry out such an analysis, researchers usually use dictionaries that offer different explanations and propose several meanings of the word. The lexical approach to text analysis can also be used to consider ambiguous texts to clarify the meaning of certain words or expressions in the context (Kukushkina, 2016). At the same time, while analysing the text we should take into consideration word sense disambiguation, the relationship between them in the text, and the situation in society and culture, because in one situation text that sounds like a joke can be highly painful or offensive in another.

Psycholinguistic methods. I used the word association test (WAT) and adapted it to the study. In the text, the associations were connected with the archetypes of the Second World War that are used for the description of the Ukrainians and Ukraine government, for example, such words as "fascist", "fascism", "nazi".

Moreover, based on the phonetic similarity of the words "Nazi" and "nationalism" in Ukrainian and Russian, journalists use the word [natsist] instead of [natsionalist] (Nazi, nationalist).

Considering text as a tool of psychological influence, we also pay attention to media headlines. These attract auditory attention and can form the audience's opinion and influence its point of view before they read the larger text (Ungerer, 2000). Schneider (2000) considers the headlines and tags of publications as markers of meanings that influence the perception of information before the text is read. Modern information technologies also influence the process of perceiving information: readers are inclined to perceive information quickly and look at headlines, abstracts, and pictures (Outing, 2004). Therefore, if we are talking about the psychological influence surrounding perceiving information, headlines and tags can be considered as tools to influence readers' consciousness.

The analysis of the text logic allows us to identify whether there are logical errors, as well as the faults that affect the meaning of the text and manipulate the facts. Stenberg & Karasik (1993) considered subverted logic within text to be language manipulation, because it leads to fact distortion and misrepresentation; for example, the fallacies related to argumentation, informal fallacies, causation fallacy, and so on.

The analysis of non-linguistic means, such as photographs, drawings, and iconography, show all of these serve to reinforce the textual material. These tools need additional verification for the conformity of the information they illustrate and the conformity of the photo to its original appearance (whether there are added details or whether certain elements of the photograph were erased; whether the iconographic information corresponds to facts, etc.).

Standards of journalism. The detection of violations of journalistic standards is an extra sign of a journalist's involvement, a lack of objective coverage of events, and support of negative rhetoric towards certain groups or individuals. While evaluating the text, I used the Principles on the Conduct of Journalists adopted by the Bordeaux Declaration in 1954, amended in 1986 (IFJ Declaration of Principles on the Conduct of Journalists, 1954 [1986]).

Our methodology was tested and implemented in a project aimed at studying the language of hostility in the editions published in the occupied territory of Crimea. In the present study, I applied the methodology to identify and prove the presence of hate speech in media outlets.

The information aggression against Ukraine and Ukrainians has been in force since 2014, and on February 24, 2022, it transformed into military aggression. On the example of Ukraine, we saw the Russian media create a dehumanised image of Ukrainians. Consequently, we observed how the Russian military's perception of Ukrainians as "under-people" led to the mass killing based on the people's citizenship, nationality, and language.

In this paper, the results of the study of the online media content produced in occupied Crimea are presented. Eleven popular Russian-language online publications, with an average readership of more than one million per month, of which the share of Ukrainian readers is at least 25%, were selected for the study. We used the AI (artificial intelligence) platform of media monitoring called "Semantrum" to identify quantitative indicators of attendance (https://promo.semantrum.net/en/main/). The complexity of the work was twofold: defining the concept of hate speech and proving the presence of manipulative and hidden hate speech, as from the legal point of view hate speech includes direct insults and calls for violent or discriminatory actions against a group of people based on their race, ethnicity, gender, etc.

Taking into consideration the definition of hate speech in such reputable sources as the Cambridge Dictionary and the United Nations Strategy and Plan of Action on Hate Speech, I define hate speech as a concept with several common traits.

According to the Cambridge Dictionary, hate speech includes public statements that spread hatred or incitement to violence against a person, group of people on the basis of race, religion, gender and other characteristics (Cambridge Dictionary). Wardon (2012) states that hate speech refers to any public communication that attacks or uses derogatory or discriminatory statements about a person or a group of persons on the basis of who they are. The United Nations considers hate rhetoric to be any communication, be it oral, written, or behavioural, that attacks or uses derogatory or discriminatory words and expressions against a person or a group of persons on the basis of who they are; in other words, based on their religion, ethnicity, race, nationality, social origin, gender, and other factors of identity pain. Such expressions and words create intolerance, hatred, and can be humiliating and cause pain (UN Strategy and Plan of Action on Hate Speech, 2019).

As a result, we define hate speech in the media as public statements that humiliate, marginalise, dehumanise, and demonise groups of people based on their race, religion, gender, ethnicity, nationality, language, and other factors. In addition, statements containing hate speech may include calls for violence and discrimination. Publications containing hate speech create intolerance, hatred, and can lead to discrimination and violence.

The difficulty of detecting and proving the presence of hate speech is related to the methods and techniques used by journalists to circumvent formal legal prohibitions and avoid the use of profanity or outright calls for violence. We define this way of spreading hate speech indirect or manipulative.

The hate speech used by the Russian media in the publications selected for our study was void of indirect and manipulative forms of expression, which formally relieved news agencies of legal responsibility for disseminating hate speech.

Depending on the methods and techniques used by journalists in their publications, three types of hate speech used in the Russian media were identified:

Type#1 direct hate speech;

Type#2 indirect hate speech;

Type#3 manipulative hate speech.

The evidence from our study highlighted the techniques that have been increasingly used by online news outlets to discriminate against certain social groups based on people's nationality and religion.

#1 direct hate speech: This is defined as incitement to hatred through the use of obscenities, direct insults, dehumanisation (e.g., comparison of humans with animals), calls for action on discrimination, and violence.

It is vital to emphasise that even if a text does not contain direct insults, discriminatory statements, or direct calls for violence, its general content does contain signs of humiliation and marginalisation by a person's nationality, language, and citizenship. Such texts are categorised as those possessing hidden and manipulative language of hostility. With the help of the author's methodology of psycholinguistic text analysis, the instances of both direct and hidden hate speech usage by journalists were singled out and explained.

#2 indirect or hidden hate speech: This includes dehumanisation and marginalisation of members of ethnic groups, demonstrations of contempt for these groups or their culture or religion, the distortion of historical facts, sarcasm and humiliation, offensive ethnonyms, the separation of the "in-group" from the "out-group", the generalisation of negative stereotypes as a typical trait of the whole group, and creating new words with negative connotations for members of ethnic groups. To spread the second type of hate speech, modern mass media exploits contempt and ridicule, deliberate exaggeration, or bracketing, which gives the word a figurative meaning or so-called affirmative forms to elucidate historical events.

For example, the idea of separation of the "in-group" and the "out-group" can be illustrated with the following citation "…the country has long been firmly divided into two parts. And none of our people will again stand shoulder to shoulder bearing pitchforks with the Maidan's trash"[1].

Another example of hate speech shows how different regions of Ukraine (Galicia and other regions) oppose to each other: "I hate these people, who brought the Galician ideology, occupied the positions in Kiev, and put an information hood on your head"[2].

The example of offensive ethnonym is "rahuli" [3] («рагули»), the word marginalized and humiliate Ukrainian speaking population. In the same article there is also the example of artificially created negatively connoted concepts that humiliate supporters of a national identity: 'Ukroparanoiki', 'Ukromarazm' (can be translated as "Ukronoids", "Ukromarasmus". The words consist of two parts: the word

---

[1] Belaya, T. (09.01.2021). Nado privykat k zhizni na Ukraine»? Ili bunt, ili ne vyzhit! [To Get Accustomed to Life in Ukraine? Riot or Death!], *Politnavigator*. https://cutt.ly/aNwaL5J

[2] Gladkov, V. (14.12.2020). Peregovorshchik LNR o skandale iz-za poiavleniia v ukrainskom efire: «Zelenskii derzhit narod za bydlo» [The negotiator from LPR about the scandal due to his appearance in the Ukrainian broadcast: «Zelensky considers people as lemming»]. *Politnavigator*. https://cutt.ly/RNwa16I

[3] Slovnik UA. Rahul(or Rohul) Invective: 1) an ill-mannered, primitive, stupid, uneducated person, a country bumpkin; 2) an inhabitant of any settlement or city who arrived there mostly from the countryside, sloppily dressed, with primitive habits, Web Portal of Ukrainian Language and Culture. Retrieved from https://clck.ru/Z4faz (Accessed 20.10.2022)

'Ukrainian' (shortened to ukro-) and words that name mental disorders (degeneracy, paranoia) and create a concept intended to show that Ukrainians who support the national identity suffer mental disorders[1].

In many articles journalists used WWII archetypes and such words as Nazis, Fascist, for example, "Ukrainian Nazism" or *"Galician Nationalism"*[2].

#3 manipulative hate speech: This category employs means of influencing the emotional state of the individual, whipping up negativity, in particular:

- distortion and subjective interpretation of historical facts;
- citing biased "experts".

For example, we can see an example of distortion of historical facts by an invited expert: "The rollout of Ukrainians and the Ukrainianness has not been launched today, this dates back to Austria-Hungary, non-existing now... First, the Ukrainian language was invented, since there was no Ukrainian language"; "...the Ukraine Western project was concocted by Austria-Hungary…". Andrey Konovalov represented as "a Donetsk's philosopher and expert of Izborsky club". On the site of "Izborsky club" such a person is not on the list of experts[3]: Andrey Konovalov[4] works at Donetsk university as a Lecturer of the Department of Philosophy. The university fanctions on Russia controlled territory (DNR); he supports anti-Ukrainian and pro-Rusian narratives in his comments found on the internet. Moreover, his comments were unilateral and repeated the thesis of Putin's speech[5]. It is also worth saying that he has no academic rank and does not engage in scientific activities (he has been working as a university teacher for 35 years and has only one publication in a local journal without any impact factor, two tutorials for local students and one conference thesis) and this several works has nothing common with the topic he commented. The abovementioned facts indicated the person represented as an expert is biased and doesn't have enough qualification to be considered as an expert in the field of history he commented.

- using the technique of substituting the meanings of concepts by other concept that create negative associations and negative images;

The technique of substituting concept "nationalists" can be seen in many media texts: concept "nationalism" is deliberately replaced by the word "Nazis" due to the similarity of the phonetic system; or instead of "peaceful protest" a journalist used "coup d'etat" a concept with a negative connotation, which means a violent and unconstitutional change of power in the state as a result of a secret conspiracy.[6]

---

[1] Toporov, A. (6.01.21). Nashi na Ukraine: geroi-odinochki, zapugannye oppozitsionery i dremliushchie obyvateli [Ours in Ukraine: Lone Rangers, Intimidated Oppositionists and Ordinary People]. *Novorossinform,* Retrieved from https://bit.ly/3GGDnKl

[2]. Gladkov, V. (23.05.2021). Ostalsia odin sposob ostanovit ukrainskiĭ natsizm [Single Way Left to Stop Ukrainian Nazism], *Politnavigator.* Retrieved from https://www.politnavigator.net/ostalsya-odin-sposob-ostanovit-ukrainskijj-nacizm.html

[3] Izborsky club. List of Experts. Retrieved from https://izborsk-club.ru/experts (Accessed 20.10.2022)

[4] Konovalov Andrey Grigorievich. Personal page at Donetsk University site. Retrieved from https://donnu.ru/ud/ph/konovalov-andrey-grigoryevich (Accessed 20.10.2022)

[5] The Russian President's official website. (2023). Vladimir Putin's annual news conference, 23 December 2021. Retrieved from http://en.kremlin.ru/events/president/news/67438/videos

Putin, V. (12.07.2021). On the Historical Unity of Russians and Ukrainians. Retrieved from http://en.kremlin.ru/events/president/news/66181

[6] Zelen'ko H. (2011). Coup d'état. *Political Encyclopedia*, 549

- the use of an artificially-created negative statement (fake) or a statement given by a third party as a fact on the basis of which the message creating negative associations and negative images is built. For example, the dissemination of information that Ukraine prepared terrorist attacks in Donbas[1].

- justification of aggression or violence against a certain group of persons on the basis of speculative conclusions about them presenting danger and the justification of aggression as preventative actions;

For example, by calling Ukraine fascist state a journalist justified war aggression: The fascist Ukraine has no right to exist. Therefore, this is our existential threat, we cannot allow the revival of fascism in the centre of Europe, and if our partners in the anti-Hitler coalition keep silence, we will not be silent, we will act."[2]

- strengthening information by non-linguistic means, which are used to enhance the impact of textual material on the emotional and sensory sphere. For example, photo manipulation and using non-event photos, deleting or adding photos, highlighting certain information, and using photo caricatures; using photograph with negative associations that has no relation to the event;

- use of manipulative names, which make a reader form a certain idea before reading the article. A title is considered manipulative if it does not match or distorts the information presented in the text of the article. According to the Poynter Institute, only two thirds of users read the text to the end. The perception of information by scanning is common: the user pays attention to pictures, photographs, headlines, and the synopsis of the article (Outing, 2004). Therefore, if a title contains an inference or statement with a negative meaning, some readers will form their opinion solely on the basis of the title and a cursory glance at the publication (title, photograph, font, etc.).

All three types are characterised by the use of tools that are designed to influence the emotional and sensory fields of the recipient, to evoke persistent stereotypes, and form a polarised worldview. There is often a violation of causation and the logic of presenting events that distort information in a way that contributes to the formation of hostile, superficial, and contemptuous attitudes based on a person's nationality, language, citizenship, or region of residence.

The study consisted of the following stages:

I. Monitoring of information publications in the media space through the use of a content analysis programme. Machine monitoring can select texts without hate speech. Subsequently, in the second stage all the selected texts are double-checked manually to identify the errors.

II. General evaluation of information in the text and exclusion of texts that do not contain hate speech.

III. Psycholinguistic analysis of the text, identification of signs of hate speech using the three aforementioned types.

---

[1] Goncharov, T. (24.03.2021). Dalshe tianut nekogda: vremia deistvii dlia Rossii v Donbasse prishlo [There is no possibility to proceed: it's time for Russia to act in Donbas], https://cutt.ly/cNwadsw

[2] Moskalenko, V. (29.04.2021). V Moskve prishli k vyvodu: «Fashistskaia Ukraina ne imeet prava na sushchestvovanie» [It Was Concluded in Moscow: "Fascist Ukraine Has No Right to Exist"], Politnavigator, https://cutt.ly/yNwavrN

IV. Detection of violations of journalistic standards (if they were violated).

V. Conclusion on each text.

VI. General conclusions for the whole monitoring period.

Results and Discussion

The results of our study reveal that the media often uses the second and third categories of hate speech in order to complicate the exposure of hate speech and to evade legislation. The legal systems of many European states prohibit hate speech; however, these laws can still be evaded through the use of sophisticated tools that disseminate hidden hate speech, which can be better detected as a result of combining psycholinguistic science with media analyses.

Stage I. According to the given keywords, the content analysis programme selected 1,284 publications which could contain hate speech.

Stage II. Upon reviewing the selected materials, 724 texts that did not contain hate speech were rejected. Thus, 560 texts remained from the total sample. The reasons for the error (724) are related to the algorithm for configuring the content analysis programme which scans the page together with comments and other information. Therefore, the reasons for the errors are justified by the following factors: 1) comments under the text included hate vocabulary. As the research aimed to analyse the products of the media specialists' activities, the comments were not taken into account and such texts were attributed to error. Among other things, comments can be a product distributed by bots or specifically hired people, which requires additional technical methods for their analysis; 2) texts in which keywords have a direct meaning; for example, the word "fascists" used in the text give a factual retrospective to the military events of the Second World War. At the same time, there were only a few such texts (2%) that were removed from the list for further analysis.

Thus, from the initial sample of 1,284, 560 texts were selected and subjected to psycholinguistic analysis.

Stage III. At this stage, the psycholinguistic analysis of the text as a product of information and communication activities of the media specialist is carried out in order to identify and explain the presence of hate speech in the text, and indicate its type. We analysed the linguistic and non-linguistic means used by journalists in texts in order to dehumanise, marginalise, and demonise people on the basis of their nationality, language, and citizenship.

Stage IV. Detection of journalism standard violations and one-sided coverage of events, including the substitution of facts by judgments, the distortion and falsification of facts, and the use of fake news. Manipulations with hashtags, for example, when the words "#Nazism", "#Punishers", "#Punitive operation", or "#Fascism" are added to the article about Ukraine. Therefore, in terms of our research

the violation of journalistic standards can be seen as deliberately disseminating inaccurate information and inciting hostility.

Stage V. Conclusion on each text: determining the type of hate speech.

Stage VI. General conclusions.

Subsequently, the 560 publications that contain manifestations of hate speech were categorised according to type. The language of hostility comprising the first type was present in 16 publications, whilst the second type of hate speech was found in 341 publications and the third type was present in 203 publications (Fig. 2)

Figure 2
Distribution of Hate Speech in Media Texts for the Period of 12/1/2020-05/30/2021
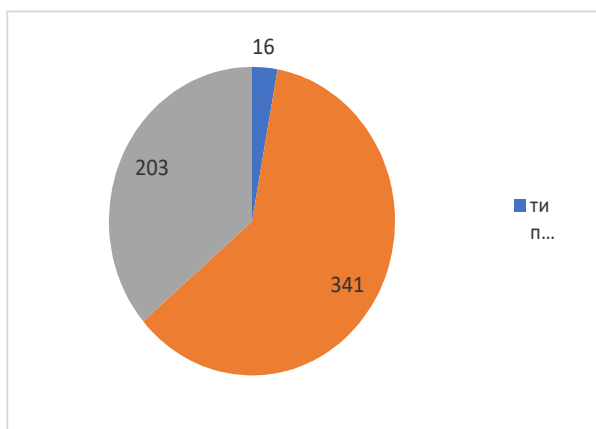
The results of the study show that currently, media professionals, in most cases, do not use direct hate speech with direct insults, calls, or incitement to actions against a person or a group of people. On the other hand, in most cases, media professionals utilise hidden or manipulative language of hostility which does not contain direct images or manifestations of intolerance on national, racial, gender, or religious grounds. Simultaneously, however, their content presents readers with a negative attitude towards certain groups and individuals, which can be used by stakeholders to incite violent actions.

Conclusions

In the analysed content, hate speech includes direct attacks on the target (type 1), indirect attacks by means of ridicule, sarcasm, marginalisation, negative associations based on Second World War archetypes (type 2), and manipulation and suggestion (type 3).

Among the main manifestations of hate speech are messages related to anti-Ukrainian rhetoric. The main negative in the analysed media is directed towards such objects as Ukraine as a state entity, Ukrainians, patriotic citizens, the Ukrainian-speaking population, and participants in the 2014 revolution of dignity.

Psycholinguistic manipulations are actively used in the media for indirect influence, the essence of which is to use both verbal and non-verbal means to influence the emotional and sensory sphere of the content's readers. Repetitions of the same narratives and their frequency indicate the purposeful nature of the impact. The repetition method is used to convince the audience and fix negative information in the minds of the content's users.

Among the actively repeated narratives in the analysed Russian-language media are: calls for the overthrow of the government; non-recognition of Ukraine as an independent state and Ukrainian as a national identity; dehumanisation, marginalisation, and demonisation of activists and Ukrainian-speaking citizens; marginalisation of Ukraine as a sovereign state, the Ukrainian language, and the western regions of Ukraine; and artificial division into good (Russian-speaking) and bad (Ukrainian-speaking).

Thus, hate speech in the Russian media is used to form a critical attitude towards the Ukrainian nationality, language, and citizenship.

The negative rhetoric disseminated by journalists in the analysed Russian publications contained numerous examples of manipulations and indirect hate speech: 1) strengthening the existing negative prejudices and stereotypes with negative rhetoric; 2) creating negative associations on the basis of negative archetypes of the past; 3) creating new stereotypes, prejudices, and fake news and 4) a systemic and frequent repetition of negative information.

The analysis of the hate speech dissemination in the Russian media demonstrates that media texts have become an instrument of influence on the public consciousness. Moreover, they have formed aggressive attitudes and created the preconditions for mass killing, discrimination, and hatred on the basis of national and linguistic grounds, in addition to civic identity.

References

Кукушкина О. В. (2016). Методы анализа, применяемые в судебной лингвистической экспертизе. *Теория и практика судебной экспертизы, 1*(41), 118-126. https://doi.org/10.30764/64/1819-2785-2016-1-118-126

Мачикова М. (2015). Медіа в контексті психолінгвістики та медіапсихології. *East European Journal of Psycholinguistics, 2*(2), 118–126. URL: https://eejpl.vnu.edu.ua/index.php/eejpl/article/view/154

Седова, И. & Печончик, Т. (2018). Язык вражды в информационном пространстве Крыма: Информационно-аналитический доклад о распространении языка вражды на территории Крымского полуострова (Март 2014-Июль 2017). Киев.

Сєдова І, Крилова-Грек Ю. (2022). Мова ворожнечі в онлайн-медіа, які висвітлюють події у Криму: Інформаційно-аналітична доповідь про поширення мови ворожнечі в російськомовних онлайн медіа, які регулярно висвітлюють збройний конфлікт України та РФ і пов'язані з ним події у Криму (грудень 2020–травень 2021). Київ.

Сентенберг И. В., Карасик, В. И. (1993). Псевдоаргументация: некоторые виды речевых манипуляций. *Речевое общение и аргуменация, 1,* 30-38.

DellaVigna, S., Enikolopov, R., Mironova, V., Petrova, M., & Zhuravskaya, E. (2014). Cross-Border Media and Nationalism: Evidence from Serbian Radio in Croatia. *American Economic Journal: Applied Economics, 6*(3), 103–132. https://doi.org/10.1257/app.6.3.103

Fairclough, N. (2003). Analysing Discourse: Textual Analysis for Social Research. Routledge

Leets, L. (2002). Experiencing hate speech: Perceptions and responses to anti-Semitism and antigay speech. *Journal of Social Issues, 58*, 341–361.

Matsuda, M. J., Lawrence, C. R., Delgado, R., Williams Crenshaw, K. (1993). Words That Wound. Westview.

Outing, S. (2004). Eyetrack III: What News Websites Look Like Through Readers' Eyes. The Poynter Institute. https://www.poynter.org/archive/2004/eyetrack-iii-what-news-websites-look-like-through-readers-eyes

Paz, M. A., Montero-Díaz, J., & Moreno-Delgado, A. (2020). Hate Speech: A Systematized Review. SAGE, https://doi.org/10.1177/2158244020973022

Salminen, J., Veronesi, F., Almerekhi, H., Jun, S., & Jansen, BJ. (2018). Online Hate Interpretation Varies by Country, But More by Individual: A Statistical Analysis Using Crowdsourced Ratings. Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS), 88–94, https://doi.org/10.1109/SNAMS.2018.8554954

Schneider, K. (2000). The emergence and development of headlines in English newspapers. In F. Ungerer, (Ed.). English Media Texts – Past and Present: Language and textual structure. (pp. 45-66). John Benjamins. https://doi.org/10.1075/pbns.80

Sellars, A. F. (2016). Defining Hate Speech. Berkman Klein Center Research Publication, (pp. 16-48). Boston University School of Law, Public Law Research, Boston University School of Law. https://doi.org/10.2139/ssrn.2882244

Selth, A. (2004) Burma's Muslims and the War on Terror. *Studies in Conflict & Terrorism, 27*(2), 107-126, https://doi.org/10.1080/10576100490275094

Tontodimamma, A., Nissi, E., Sarra, A., & Fontanella, L. (2021). Thirty years of research into hate speech: Topics of interest and their evolution. *Scientometrics, 126*(1), 157–179. https://doi.org/10.1007/s11192-020-03737-6

Ungerer, F. (2000). English Media Texts – Past and Present: Language and Textual Structure. John Benjamins. https://doi.org/10.1075/pbns.80

Waldron, J. (2012). The Harm in Hate Speech. Cambridge, MA: Harvard Univ. Press

Whillock, R. K., Slayden, D. (Eds.). (1995). Hate Speech. Sage.

References (translated and transliterated)

Kukushkina, O.V. (2016). Metody analiza, primenyaemye v sudebnoj lingvisticheskoj ekspertize [The methods of analysis used in forensic linguistic expertise]. *Teoriya i Praktika Sudebnoy Ekspertizy, 1*(41), 118-126. https://doi.org/10.30764/64/1819-2785-2016-1-118-126

Sedova, I. & Pechonchik (2018). Yazyk vrazhdy v informacionnom prostranstve Kryma: Informacionnoanaliticheskij doklad o rasprostranenii yazyka vrazhdy na territorii Krymskogo poluostrova (Mart-Iyun, 2017) [Hate speech in the information space of Crimea: an information report about spreading of hate speech across the territory of Crimea (March 2014 – July 2017)]. Kyiv.

Sedova, I & Krylova-Grek, Yu. (2022). Mova vorozhnechi v onlain-media, yaki vysvitliuiut podii u Krymu: Informatsiino-analitychna dopovid pro poshyrennia movy vorozhnechi v rosiiskomovnykh onlain media, yaki rehuliarno vysvitliuiut zbroinyi konflikt Ukrainy ta RF i poviazani z nym podii u Krymu (Hruden' 2020–Traven' 2021) [Hate speech in online media covering events in Crimea: Information and analytical report on the spread of hate

speech in Russian-language online media, which regularly cover the armed conflict between Ukraine and Russia and related events in Crimea (December 2020-May 2021)]. Kyiv.

Stenberg, I. V. & Karasik, V. I. (1993). Psevdoargumentaciya: nekotorye vidy rechevyh manipulyacij. [Pseudo-argumentation: some types of speech manipulations]. *Rechevoe obshchenie i Argumenaciya, 1,* 30-38.

DellaVigna, S., Enikolopov, R., Mironova, V., Petrova, M., & Zhuravskaya, E. (2014). Cross-Border Media and Nationalism: Evidence from Serbian Radio in Croatia. *American Economic Journal: Applied Economics, 6*(3), 103–132. https://doi.org/10.1257/app.6.3.103

Fairclough, N. (2003). Analysing Discourse: Textual Analysis for Social Research. Routledge

Leets, L. (2002). Experiencing hate speech: Perceptions and responses to anti-Semitism and antigay speech. *Journal of Social Issues, 58,* 341–361.

Machikova M. (2015). Media v konteksti psykholinguistyky ta media psykholohii [Media Within Psycholinguistic and Media Psychological Context]. *East European Journal of Psycholinguistics, 2*(2), 118–126. Retrieved from https://eejpl.vnu.edu.ua/index.php/eejpl/article/view/154

Matsuda, M. J., Lawrence, C. R., Delgado, R., Williams Crenshaw, K. (1993). Words That Wound. Westview.

Outing, S. (2004). Eyetrack III: What News Websites Look Like Through Readers' Eyes. The Poynter Institute. https://www.poynter.org/archive/2004/eyetrack-iii-what-news-websites-look-like-through-readers-eyes

Paz, M. A., Montero-Díaz, J., & Moreno-Delgado, A. (2020). Hate Speech: A Systematized Review. SAGE, https://doi.org/10.1177/2158244020973022

Salminen, J., Veronesi, F., Almerekhi, H., Jun, S., & Jansen, BJ. (2018). Online Hate Interpretation Varies by Country, But More by Individual: A Statistical Analysis Using Crowdsourced Ratings. Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS), 88–94, https://doi.org/10.1109/SNAMS.2018.8554954

Schneider, K. (2000). The emergence and development of headlines in English newspapers. In F. Ungerer, (Ed.). English Media Texts – Past and Present: Language and textual structure. (pp. 45-66). John Benjamins. https://doi.org/10.1075/pbns.80

Sellars, AF, (2016). Defining Hate Speech. Berkman Klein Center Research Publication, 16-48, Boston University School of Law, Public Law Research, Boston University School of Law. https://doi.org/10.2139/ssrn.2882244

Selth, A. (2004) Burma's Muslims and the War on Terror. Studies in Conflict & Terrorism, 27(2), 107-126. https://doi.org/10.1080/10576100490275094

Tontodimamma, A., Nissi, E., Sarra, A., & Fontanella, L. (2021). Thirty years of research into hate speech: Topics of interest and their evolution. *Scientometrics, 126*(1), 157–179. https://doi.org/10.1007/s11192-020-03737-6

Ungerer, F. (2000). English Media Texts – Past and Present: Language and Textual Structure. John Benjamins. https://doi.org/10.1075/pbns.80

Waldron, J. (2012). The Harm in Hate Speech. Cambridge, MA: Harvard Univ. Press

Whillock, R. K., Slayden, D. (Eds.). (1995). Hate Speech. Sage.

Sources

Cambridge Dictionary. Retrieved from https://dictionary.cambridge.org

Recommendation No. R (97) 20 of the Committee of Ministers to Member States on "Hate Speech" (30 October 1997). https://rm.coe.int/1680505d5b

UN Strategy and Plan of Action on Hate Speech. (2019). Retrieved from
        https://www.un.org/en/genocideprevention/hate-speech-strategy.shtml