The representation of Concept of Hate speech in mass media

Krylova-Grek Yuliya[0000-0002-2377-3781]

Kostiuk Institute of Psychology
of the National Academy of Educational Sciences of Ukraine. Kyiv
yulgrek@gmail.com

Introduction. In terms of the exponential growth of mass access to sources of information, hate speech has become a discussible concept nowadays.

In my research, I give attention to the concept of hate speech and its complexity in terms of its influence on people's consciousness and behaviour as a result of the media's impact. Despite hate speech being a discussible concept at an international level, a definition of it has not yet been standardised. According to the Cambridge Dictionary, it is public speech that expresses hate or encourages violence towards a person or group based on something such as race, religion, sex, or sexual orientation ("hate speech". dictionary.cambridge.org.). At the same time, in terms of mass media, it uses many ways of influencing and changing the worldview of a person and can be seen in a wide variety of manifestations, for example, fake news, propaganda, spread of violence, xenophobia, hatred against certain groups or individuals and so on. Moreover, hate speech can be used in different forms: direct, indirect and manipulative. It is important to consider hate speech as a concept and its manifestation in the media in order to monitor the situation and take preventive measures.

I consider hate speech a concept because it contains characteristics of a concept and includes a wide range of notions (ideas) that are used in media and influence people's values and worldviews [9; 10].

Hate speech as a phenomenon has been mentioned in terms of laws and freedom of the press. In particular, Howard is concerned with where the border between freedom of speech and hate speech begins [1]. So, to avoid the abovementioned ambiguity, it is important to consider hate speech as a concept from different sides.

Hate speech as an object of scientific research is considered in political and social sciences (Howard, 2019; Paasch-Colberg et al., 2021 and others), legal sciences (Waldron, 2012; Ghanea,2012; Howard, 2019 and others) and media communication Communication (Bahador and Kerchner, 2019; Bahador, 2020 and others). The manifestation of hate speech in comments and posts on social networks is studied by Chen, 2011; Paasch-Colbergetal, 2021; Schmidt and Wiegand, 2017 and others.

Thus, hate speech has been a subject of scientific study but has not been considered as a concept in the cross-disciplinary research of the humanities and computer science. In this paper, we consider hate speech as an object of psycholinguistics that we have detected in online media texts with the help of a computer bot created for this aim.

In my research, I examine the concept of hate speech in mass media as a tool of manipulation and influence on a person's consciousness from a psycholinguistic

perspective. Of particular interest are linguistic means, methods and techniques that aim to change the reader's worldview and build mental representations. The core focus of our current research is hate speech in print media.

Aim. To consider hate speech as a concept within a cross-disciplinary setting, identify its type and methods of influencing an individual's consciousness by the author's method of psycholinguistic analysis and the AI-based bot.

The work was performed within an international project "Free Voices: Promoting Independent Media in the Target Region" featuring the Crimean Human Rights Group. This paper presents an intermediate study of hate speech in the texts of online publications dated November–December 2020.

Methods. Thus, it was decided that the best procedure for this investigation was to join the author's psycholinguistic method of textual analysis and ICT, namely an artificial intelligence-based bot for selecting texts, following the given parameters.

The bot performs lexical analysis of content. Keyword selection strategy was based on the semantic meaning of the words in phrases and the entire publication with a negative tone.

The methodology of psycholinguistic analysis of texts ("PLAT") is the author's methodology developed by Yu.M. Krylova-Grek. It is presently being registered at the patent office (application № m 2, UI Property Institute). The methodology was applied as an initial approbation during the pilot course "Psycholinguistic basics of media literacy" in 2018 [2].

In the intermediate study of the research, I demonstrated the main approach to studying hate speech as a concept in the cross-disciplinary field and a possible way for cooperation between humanities and computer science for studying hate speech as a concept.

Our research consists of four stages. During the first stage, a team of journalists and volunteers with the help of a software bot carried out search query mining to find negative information. Later, through the application of the author's method of psycholinguistic analysis of the text, we scrutinised the material obtained to identify the type of hate speech, its methods and techniques of influencing consciousness. We selected 405 texts with hate speech.

The examples of hate speech were split into three types based on the use of specific linguistic and graphic tools:

1) direct hate speech

2) indirect (hidden) hate speech: opponents' dehumanisation and marginalisation, without direct insults or call to action, but with using sarcasm, demonstrating arrogance and mentioning negative generalisations (e.g. all of them are lazy, greedy or cunning).

3) manipulative hate speech

Results & Discussion. The evidence from this study intimates that the media tends to use examples of hate speech of the second and third types and avoids direct discrimination of the first type.

According to the data obtained, slightly more than half of the sample was presented by type #3 (55.3%), approximately a fifth was type #2 (20.1%) and less than 1% was type #1 (0.2%)

In addition, there were mixed types. Type #2+3, which combined the second and third types, included indirect statements that provoked negative emotions towards a certain group of people (19.4 %). Less common (5 %) was mixed Type#1+3, the combination of the first with the third or second type. It involved a negative word characteristic used as an axiom to affect the content of the entire text (fig.1).

We assumed the third and mixed types of hate speech to be the methods of psycholinguistic manipulation. So, the definition of the concept of hate speech in online media should include the abovementioned ways of its manifestation. Moreover, while defining the concept of hate speech, it is important to mention factors that serve as a prerequisite for building the negative mental representation of the group of people or a person.

Conclusion. Taken together, the findings of the study indicate that in the media, hate speech is currently presented by indirect influence and manipulation (79.7%). In this regard, we emphasise that this concept is not a stable mental representation, and it is under the great influence of external factors, such as mass media. Nevertheless, it is possible to draw conclusions about the common and general features, identify intention and offer a definition. As we can see, the concept strongly depends on the kind of mental representation suggested by mass media. Therefore, hate speech is a complicated concept that needs further study.

Keywords: Hate speech, Media manipulation, Text, Psycholinguistics.

References

1. Howard, Jeffrey W.: "Free Speech and Hate Speech". In: Annual Review of Political Science. 22, pp. 93–94. doi:10.1146/annurev-polisci-051517-012343 (2019).
2. Krylova-Grek Yu.: Peculiarities of information science based on the pilot course "Psycholinguistic basics of media literacy". In: The proceedings of13th conference reader: DisCo 2018: Overcoming the Challenges and the Barriers in Open Education. June 25-26, Prague: Centre for Higher Educa-tion Studies, pp. 199-206 (2018).
3. Bahador, B., Kerchner, D.: Monitoring hate speech in the US media (Working Paper). Washington, DC: The George Washington University (2019).
4. Chen.Y.: Detecting Offensive Language in Social Medias for Protection of Adolescent Online Safety. Ph.D. Dissertation. The Pennsylvania State University (2011).
5. Ghanea, N.: The Concept of Racist Hate Speech and its Evolution over time. In: UN CERD. Thematic Discussion on Racist hate speech, 81st session, 28 August 2012, Geneva. https://www.ohchr.org/documents/hrbodies/cerd/discussions/racisthatespeech/ nazilaghanea.pdf), last accessed 2021/04/07
6. Paasch-Colberg, S., Strippel, Chr., Trebbe, J., Emmer, M.: From Insult to Hate Speech: Mapping Offensive Language in German User Comments on Immigration. In: Media and Communication. 9. pp.171-180. doi 10.17645/mac.v9i1.3399 (2021).
7. Schmidt, A. and Wiegand, M.: A survey on hate speech detection using natural language processing. In: Proceedings of the Workshop on Natural Language Processing for Social Media (SocialNLP'17). 1 (2017).
8. Waldron J.: The Harm in Hate Speech. Cambridge, MA: Harvard Univ. Press (2012).

9. Seel, N.M. Concept Formation: Characteristics and Functions. Encyclopedia of the Sciences of Learning. Springer, Boston, MA. doi.org/10.1007/978-1-4419-1428-6_1866 (2012).
10. Stepanov, Y.S. Constants. Dictionary of Russian culture. Research experience. M., pp. 40-43 (1997).