Hate speech – ethical framework of communication and subjective accountability towards the fighting words

KEY WORDS
ethics of words, hating, hate speech, persuasion, rhetoric of empathy, freedom of speech

ABSTRACT
The purpose of the article is to show the complexity of hate speech as a colluded notion, as regards both the philosophy of language and praxis. It is a well know phenomenon, but its new form has definitely emerged along with the existence of the Internet. Defining the framework for the communication and the ethical aspect thereof as a normative determinant in the process of indirect communication is of considerable theoretical and academic significance and it also has certain practical implications. This is so because hate speech is used to recount facts and present social reality, but also to manipulate and persuade.

When considering the ethics of the word, it should be recalled that the scientific reflection and debate on these issues are not new among philosophers of language and philologists. It is worth quoting linguistic inquiries from 1500 BC or the arguments of Socrates, Plato and Aristotle in the fifth century. Extremely intensive period of research on language was the nineteenth century, when e.g. Danish philosopher, Søren Kierkegaard, wrote that language is a set of elements that are given originally and other elements that develop in a free manner\(^1\). Another important breakthrough was the publication of Ferdinand de Saussure’s work entitled *Course in General Linguistics* in 1916. The researchers were interested in the key problems related to the nature of meaning, the use of language, its understanding, relation to reality, as well as language as a component in other fields: logic, history, politics, as well as axiology. The reflection over the ethics of the word is a recurring phenomenon, especially in the face of emerging phenomena, such as, for example, *hate speech* that is the subject of this article.

---

Both the followers of analytic philosophy and representatives of continental philosophy perceived language as a tool for communicating used between the sender and the recipient in the social process and thought about the phenomenon – how the meanings relate to the truth and to reality. In this context, it is worth recalling the theory of direct reference of John Stuart Mill\(^2\) to the truth. It was not so much a dispute about whether the sentence is actually true (it was rather interesting for linguists and philosophers of cognition), but a dispute to decide what kinds of meanings are true (e.g. how sentences express the truth about non-existent things).

For language philosophers and linguists, the important issue was to discover a relationship between language and thought. Mind philosophers asked whether language influences thinking or thinking influences language. Linguistics – Edward Sapir and Benjamin Lee Whorf\(^3\) – saw the functioning of specific topics in the area of “linguistic community”, while the primary role of thinking was emphasised by, e.g. Paul Grice\(^4\), writing about the language coded in the mind. From the perspective of subsequent decades of research, we can talk about the merging and – as some say – blurring of the boundaries between what is a linguistic description of reality and what is a “meta-language performative creation”\(^5\).

An important role in the context of reflections on hate speech is also taken by rhetoric, which examines the use of language to achieve a specific goal (evoking emotions or stimulating thinking), which may be, for example, convincing to one’s own reasons, supporting own opinions by changing the attitudes of the other person (e.g. propaganda), inspiring reflection on current opinions (didacticism) or – importantly in the context of hate speech – provoking the recipient to react to social reality. In the case of the last of these areas, rhetoric focuses on the analysis of pejorative expressions, the study of vulgar or offensive words and phrases, and the consequences of such acts and actions like changes in relationships and behaviour in people subject to this phenomenon. In the case of hate, related hate speech, understood as any action on the Internet, which is a manifestation of hatred and aggression by hurting another person mainly with a word, but also with a photo/graphic/meme or film, performative utterances (speech acts) would be particularly important, having the

---


power of “acting within a language ”The phenomenon of hate would also be a break with the conversational principles of Herbert Paul and the politeness maxims of Geoffrey N. Leech.

Going beyond certain language structures, in the discussion on hate speech basic for this text, we should also mention the disciplines that examine the interpretation of whole messages, especially the twentieth-century hermeneutics of Martin Heidegger, Hans-Georg Gadamer and Wilhelm Dilthey, and phenomenology with the rule of “experiencing the truth”. It was Gadamer who saw language as a place where “there is a concrete understanding and agreement between two people”6. By analogy, one could say that upon misunderstanding there is a kind of conflict between people.

One of the variables that can trigger such dispute is the use of a double-edged or multi-edged (e.g. symbolic) language, as mentioned by Paul Ricoeur7 and Jacques Derrida8. According to the well-known formula of semiotics saying that everything can bear meaning (any context or nuance), reception of signs and language symbols in the process of building meaning – which was described, among others, by Charles Sanders Peirce, Roland Barthes, Roman Jakobson and Umberto Eco – is a compiled activity. In the case of hate speech, which uses subtexts, hidden meanings, understatements etc., also multi-layer activity. It is impossible not to quote the well-known thesis of colloquial language philosophers with John Langshaw Austin, the creator of the theory of speech acts9 from the middle of the last century, saying that the utterances have secondary meaning, and primary meaning is why they are used in the message (e.g. to hurt or fight), as pragmatics deals with it. The language has enormous causative power and has a social impact, as Jerzy Bartmiński reminds us: “We can – referring to the assumptions of anthropological and cultural linguistics – reconstruct the linguistic image of speaking and the cultural pattern of communication. The methodology of researching the linguistic image of the world prefers in this respect vocabulary and phraseology, which are the most-distinguished classifiers of social experiences. We find in them information about socially accepted scenarios of communication behaviours, about language patterns”10.

---

9 J.L. Austin, How to do things with words: the William James lectures delivered at Harvard University in 1955, Oxford 1962.
Title of the book by Dwight Bolinger. *Language – the loaded weapon*\(^{11}\) is confirmed by the language with which we deal in the media discourse. “Language of the media on one hand reflects the primitiveness and brutality of public-political discourse, and on the other hand itself contributes to the tabloidisation of social communication. One of the elements of this phenomenon is the so-called hate speech, used to label selected statements, activities, attitudes and actions of public life, which is more and more often used at the metalinguistic level, assessing the language of the media”\(^{12}\).

From the perspective of axiology of communication and the media, it can be said that hate speech becomes an autotelic phenomenon. Such specific labelling in the media is used both to describe and purify the public sphere, and – unfortunately – to create and enlarge areas of hostility in the media, and thus to promote hatred in the cultural sphere and social communication (e.g. as a tool to increase hatred).

**From the ethics of the word (in language) to the ethics of discourse (in communication)**

An extremely interesting discussion about the ethics of words in Polish linguistic thought was triggered in the 1980s by Jadwiga Puzynina and Michał Głowiński, writing about newspeak\(^{13}\), and then again in 1995 the aforementioned researcher, together with Anna Pajdzińska, provoked discussion by presenting “ethical principles of words based on values such as kindness, need of contact with other people, honesty, truth, integrity, respect for otherness, willingness to act for the common good”\(^{14}\). Authors of the so-called eight rules wrote: “In order for human communication to be healthy, we must respect human right – both of the sender and the recipient – to the absence of fear (against revealing own opinions, against violation of intimacy of acts of communication, against violation of personal dignity, lying, distortion, manipulation, etc.) and the right to choose attitudes”\(^{15}\). Over time, there appeared


\(^{12}\) M. Drożdż, *Hate speech…*, op. cit., p. 21.


\(^{15}\) Ibid, p. 39.
numerous committees of ethics and a living word\textsuperscript{16}, and their aftermath were publications on ethical topics. Scientists, as Anna Cegiela wrote in the introduction to the book \textit{Słowa i ludzie. Wprowadzenie do etyki słowa}\textsuperscript{17} [People and words. Introduction to the ethics of word] had doubts about research on ethics using the methodology of strictly linguistic sense. First of all, pragmatic issues and communication dimension connected with ethical issues. Linguistics got the support of philosophers, sociologists, political scientists, historians and, in time, specialists in the field of media sciences, e.g. in the study of political communication or manipulation, media lies, concealment and avoiding socially relevant issues and the discussed subject of ethics in the media. It is worth mentioning the concern for ethics as an important part of the discussion about the media, included in the text by Michał Drożdż, \textit{Etyczne aspekty mediów integralną częścią nauk o mediach}\textsuperscript{18} [Ethical aspects of the media are an integral part of the media sciences]. The researcher postulates the ethical purpose of dialogue, which is aimed at agreement, and not – as in the case of the discussed phenomenon – that conflict and confrontation.

Today, we have a scientific background – a multitude of publications on the subject of the ethics of words, scientific initiatives (including a series of “Media ethics” conferences held regularly for over a dozen years), and institutional (e.g. the Association of Ethics of the Word). As Bartmiński wrote, appreciating both the conference “Rhetoric and ethics” and the \textit{Association of Ethics of the Word} established in 2008: “The subject »ethics of the word«, included in the program of our Polish Language Congress, is part of a series of similar topics, increasingly undertaken in the environment of people of science, media, politicians, not to mention traditional philosophers and theologians. This interest results from the growing role of the word in social activities and from the emerging abuses”\textsuperscript{19}.

It is also worth mentioning the definition quoted by Professor Puzynina, a promoter of axiological topics in Polish linguistics, in the article “\textit{Hate speech – and the ethics of the word ZES}”, who writes: “The expression \textit{hate speech} is now an international term defining negatively evaluated and tackled aggressive statements addressed to individuals or collectives,

\textsuperscript{16} \textit{Rhetoric and ethics} – conference organised on 16-18 April 2007 in Poznań by the Committee for the Culture of the Living Word of the Polish Language Council. A book with such title was published in Poznań in 2009, edited by Barbara Sobczak and Halina Zgółkowa.
\textsuperscript{17} A. Cegiela, \textit{Słowa i ludzie. Wprowadzenie do etyki słowa} [People and words. Introduction to the ethics of word], Warszawa 2009.
\textsuperscript{18} M. Drożdż, \textit{Etyczne aspekty mediów integralną częścią nauk o mediach} [Ethical aspects of media as integral part of media science], “\textit{Studia Medioznawcze}” 2013, no. 4 (55), pp. 11–25.
but actually addressed always to collectives, a group of those who are basically excluded. These are primarily natural groups, participation in which is determined by biological factors (skin colour, sex, sexual preference, ethnicity and disability) or social factors (language, citizenship, religion, place of residence), factors which are often difficult to change, or profession, in many cases "inherited" within family. Some authors of articles and books dealing with hate speech see it as stigmatizing for people and groups with different (chosen by them) views, especially political orientations. Without the last sentence, the final part of the definition would sound rather anachronistic, especially in the context of the last of the arguments mentioned above, concerning inheritance of professions and the emerging thinking about hate speech in terms of the different political views that currently constitute a strong thematic dominant of this phenomenon. This author’s contribution to the criticism of the different political orientations concerns a very important article by Michał Głowiński entitled Retoryka nienawiści [Rhetoric of Hatred], to which the researcher refers.

The author, which is very valuable, defines the concept of hate speech as the opposite of rhetoric of empathy and excludes those utterances in case of which – in her opinion – it is an exaggeration to speak of hate. Hate is the manifestation of hostility towards a person and wishing him/her bad luck, and what appears in the media space are usually contemptuous, humiliating, offensive, ridiculing or simply unfriendly statements. Confirmation of such a way of defining the hate speech by the expert in Polish ethics of the word is her statement quoted in “Rzeczpospolita” in 2017, when in an interview with Magdalena Bajer, she pointed out: “Well, we put emphasis on attitudes expressed in the language used, not the language itself. If we hear the term hate speech, it is a signal that there is hatred that must be tackled. I believe that this is an exaggerated concept, but it has already become an international term. Many words of insulting and mocking nature are not hate speech, but from an ethical perspective they are reprehensible and should be avoided.”

In addition, this category of hate speech features, emphasised by linguists, must be completed by issues inextricably connected with this phenomenon, and considered mainly in the philosophical field, i.e. stating false (lack of knowledge about spoken issues - e.g. misleading) and truth understood as the opposite of lying (statements, which the sender does

---

20 J. Puzynina, Mowa nienawiści – a etyka słowa ZES [Hate speech - and the ethics of the word ZES], http://www.etykaslowa.edu.pl/wp-content/uploads/2016/05/Mowa-nienawi%C5%9Bci-a-etyka-s%C5%82owa.pdf [access: 11.10.2017].


not consider to be true – for example in the case of duplication of stereotypes). Along with the development of the ethics of the word, a set of principles has been developed that relate to the positively understood scope of it. These are real statements free of aggression and manipulation. Creating a list of such rules on the basis of codes, one should mention first of all: using the language with keeping the personal dimension of a human being, and thus expressing about the other person with inextricable dignity. Another important principle is the anti-reductionist attitude – treating a person in his/her multidimensionality as a complex being the labelling of whom is most often an attempt to discredit. And the third of the most important rules of the ethics of word is the ability to distinguish the necessary substantial criticism from an unjustified personal attack. Ethicists, when considering a teleological issue, speak here of the purpose of a critical statement, which can be human action or its effect, but not a person. In the case of hate speech, verbal aggression focuses on the generalisation and emotionally valued human being, not on his/her actions or their effects. “Hatred is not only a matter of emotional antipathy, but it is an expression of the voluntary rejection of another person in attitude, language and actions. The nature of hatred is determined by the purpose of rejection of the other person: either because of the danger to me (odium abominationis), or I feel revulsion because of the other person, wishing him/her bad luck (odium inimicitiae). In both cases, I deny the value and dignity of the other person”23.

Even if ethicists have developed a set of rules, linguists and media experts who study the practical use of language ask about a “socially accepted pattern of linguistic communication. I assume that it exists, even if it is not always respected in practice, and sometimes even ostentatiously broken”24 – claims Barmiński. The researcher reminds that in the public discourse already in the conditions of a democratic state in indirect media communication there was something that Marek Czyżewski, Sergiusz Kowalski and Andrzej Piotrowski called “ritual chaos”25, and then came a time of extraordinary brutalization of language behaviour. The effect of this in the 21st century is – as Bartmiński points out – „the substantive sterilization of public (political) discourse and the loss of its credibility”26.

23 M. Drożdż, Hate speech..., op. cit., p. 22.
24 J. Bartmiński, Etyka słowa a potoczny wzorzec komunikacji [Word ethics and common communication pattern], op. cit.
26 J. Bartmiński, Etyka słowa a potoczny wzorzec komunikacji [Word ethics and common communication pattern], op. cit.
Hate speech on the Internet

As Bożena Witosz wrote: “Hating and hate speech was not created by today’s media fashion for aggression and insults, but it certainly contributed to their dissemination”\(^{27}\). Media in which hate speech appears, not only fulfil the rudimentary epistemological function, which is describing reality, but they create their own image of the world. Media contents that appear in all forms – from words in the printed press, to memes in virtual space – are created not only by journalists, but most often with the participation of “users ”themselves. There has been a strong fusion of expression from the private sphere to public or mediated sphere\(^{28}\). It is a space for research within not only social psychology, but even a phenomenon in the field of clinical psychology, as Zbigniew Bauer wrote in the article Polish Internet forums. Example of “hate speech” and “speech of aggression”\(^{29}\).

“Today, the Internet space allows us to present the most diverse senses – including those that once functioned in the margins and which we currently evaluate negatively. As users, we pay the price for the network and opportunities it gives – unfortunately, it is the right of free speech for aggressive or simply unwise people”\(^{30}\). Bogusław Skowronek adds to this statement a whole range of rudimentary features of the Internet: dialogue, contact, universality, accessibility, democracy and anonymity as well as the possibility of covering all taboos that are conducive to this phenomenon.

Hate speech built on the expression of negative emotions (mainly cognitive structures) is a complex phenomenon, because – as psychologists say – it is not ultimately hatred directed at a specific person, but impersonal violation of norms through the medium. It is most often self-serving and conscious attack disguised as emotionally expressed statements that are supposed to humiliate someone. An extreme example is the functioning of “professional” haters within political parties or competitive companies. The aforementioned hate – cyberbullying\(^{31}\), although it often appears in colloquial, journalistic and scientific opinions interchangeably with the name “hate speech”, “speech of aggression” or “language of

---


\(^{30}\) B. Skowronek, Hejt jako zjawisko lingwistyczno-medialne. Zarys problemu [Hate as a linguistic and media phenomenon. Outline of the problem] [in:] Genres in media…, op. cit., p. 184.

hostility”[^32] – is more widely understood than hate speech, because the latter has a specific thematic set. Hating refers not only to “public figures propagating fascist or other totalitarian state system or calling for hatred on the basis of national, ethnic, racial or religious differences”[^33], but in the selection of topics it is almost unlimited. However, the distinguishing feature is the unit-directed attack (insult and destruction of the opponent as a performative goal in the perlocutional aspect), while hate speech is most often collective, even institutional. However, similarly to the hate speech, it is strongly axiologically marked (hate uses negative axiology), because it tends to polarize people and deprecate another person. The vision of the world that is being portrayed is extreme (good-evil, hater’s domination through “obliteration” of the opponent), and the ultimate goal is to attack and to hurt and to attack the dignity of a person. In this context, the boundaries between the concepts of hate speech and hatred can also be determined. As Agnieszka Więckiewicz wrote: “While dangerous speech, involving encouragement to take physical violence against the offended people, can be described as a more radical version of hate speech, hatred could be preliminarily described as a lighter form that may or may not transform into it, as well as a kind of broadly hurtful statement, which does not use preconceived ideology that is a factor constituting the hate speech”[^34].

**Style of the public discourse: hate speech and rhetoric of empathy**

Trying to determine the mechanisms that govern hate speech, we can repeat after Głowiński: “Means of rhetoric of hate are used by people of various places in the public space, in certain situations it is identical with the language of power and its acolytes; it is used by political leaders on the most important positions in the state, activists of more or less serious organisations and occupations, and by propagandists and journalists dealing with commenting on current events. A subject of hatred can be anyone who is considered an enemy that must be destroyed (the sooner, the better!). In this view, being an enemy is a predetermined role, completely non-subjective, that is, independent of the person to whom such a title would be assigned (things are similar in the case of institutions and organisations)”[^35]. Interestingly, as

[^32]: B. Witosz, Czy ,,hejt ”to problem genologiczny [Is hate a genological problem?] op. cit., p. 194.
[^33]: A. Naruszewicz-Duchlińska, Nienawiść w czasach Internetu [Hatred in the era of Internet], Gdynia 2015, p. 16.
Głowiński adds in the cited text from 2007: “In the current situation in Poland, there is basically only one man who is entirely safe – John Paul II”\(^\text{36}\).

As current examples show (for example the play “Curse” by Oliver Frljic in Teatr Powszechny in Warsaw and subsequent debate, among others in the press and on the internet forums of opinion-forming magazines), already a decade later – both the current head of the Catholic Church and the aforementioned John Paul II – they no longer enjoy such safety. Contemporary language of public debate – as never before – is brutalized, which according to Jerzy Bralczyk is caused by the clash of two strong ideas. He claims that language exacerbates when the public (and political) scene becomes a battlefield (arena) of two clearly outlined sides. “Two metaphors usually try to describe our public discourse, especially political one: of war and theatre. They refer to not only communication offenses: aggression and lie”\(^\text{37}\).

Analysing numerous examples of hate speech on the Internet, it can be said that the aforementioned metaphor (referring to Lakoff’s and Johnson’s breakthrough research) sets the participants in a fight position. However, whether it is an enemy (whom we hate), or the opponent (with whom we polemize) it is just – as Głowiński believes – only the difference of degree. “Hating is a form of deviant behaviour during public internet discussions. It involves abusive language, contemptuous assessment of various phenomena, insulting both interlocutors and various other entities and expressing aggression and hatred against them”\(^\text{38}\) – writes Marta Juza. The goal of the fight is to destroy the other one, the goal of polemic with the opponent may even be coming to an agreement. It seems that in the case of hate speech there is no such plane of communication with the adversary. Sides, influenced by the language of hostility, are never seeking agreement, but become even more polarized.

The second conclusion, which results from the conducted analyses, is that over time users of online forums are no longer fighting a specific person (e.g. protagonist of an article), but above all each other within a given medium, using mainly emotionally charged and semantically imprecise terms.

We can list the following indicators of hate speech\(^\text{39}\) after Głowiński:

- rhetoric of absolute opinions (the world in the extreme form: black and white)

\(^{36}\) Ibid.
\(^{38}\) M. Juza, *Hejterstwo w komunikacji internetowej: charakterystyka zjawiska, przyczyny i sposoby przeciwdziałania* [Hating in Internet communication: characteristics of the phenomenon, causes and methods of counteracting], "Profilaktyka Społeczna i Resocjalizacja" 2005, no. 25, p. 29.
• no persuasion, because we do not want to change and convince anybody, we just want to “morally ridicule ”and destroy their image
• ruthlessness of opinions - dichotomous divisions (e.g. universals: us-them, connecting directly with axiology)
• conspiracy view of the world (from the material sphere to the moral sphere)
• uniform and unambiguous evaluation and labelling
• the authoritarian nature of speaking subject and at the same time the lack of individual references (depersonalized entity does not legitimize itself, it uses an online nickname, hiding a person under its veil).

All expressions of hate speech present in contemporary media, with particular emphasis on the network, have both immediate and long-term effects. “Strong intergroup emotions lead to prejudices and discrimination, which is manifested by the greater presence of hate speech, especially on the Internet – the space of social disinhibition. (...) It cannot be ruled out that exposure to hate speech means that after some time it ceases to be considered offensive, shocking and violating social standards. Such conclusions can be drawn from studies on desensitization to images of violence. Frequent watching of violent scenes results in perceiving them as less harmful over time. Contact with hate speech may have similar effect to the “crashed glass” effect described by criminologists: by watching the violation of social norms, members of the society themselves become less inclined to observe such norms”

The long-term effect of the media hate speech is – especially on the Internet – creating separate virtual hate spaces that begin to affect non-media reality.

Trying to determine the ethical frames of hate speech, it should be emphasised that media constructs are increasingly contributing to its rooting in reality and may constitute the basis (and tool) of manipulation using negative, socially harmful values – from vulgarity, through displaying and trivializing evil, to the destruction of values for the sake of media attractiveness.

Conclusions
The communication society is also a duty of shared responsibility. The language of hatred and dispute, proposed by some of the media – both those with right-wing and left-wing editorial lines – preferred on many portals on the Internet space, tries to deprive the recipient of right to

---

personal evaluation and free thinking. In view of the multiplicity of axiological orientations in currently diversified media, this leads to a state in which media reality often refers to itself (the media are an object and context of discourse). In such a space, their role as a reference disappears – the media, using hate speech, lose the value of tools for transferring objective meanings to create intermedia communication discourse. This building of space based on hate speech, however, has consequences that ultimately directly affect non-media reality, because they are transferred to the interpersonal relations of the protagonists of the media show.

The media are created by people. And it is them who are responsible for the quality of language. This boundary between freedom (here: especially the freedom of speech) and responsibility is played out in a very sensitive area of conscience, and as ethicists emphasise: “ethical evaluation, which takes place in the conscience of man, always has a subjective dimension”\(^{41}\).

It is also worth recalling the postulate of Głowiński, who stated that “the rhetoric of empathy can be treated as opposition to rhetoric of hatred. And rhetoric of empathy, even if created for the purposes of polemics with people and institutions, while highly critical, demonstrates certain understanding for the other side, recognises that the words of the representatives of that other side do not result simply out of bad will, are not full of bandit intentions, prone to vile and ill games. [...] In the cases that we find interesting, the empathy of rhetoric consists above all in the fact that the speaker, columnist and commentator is able to at least minimally understand the other person with whom he polemizes or whom he accuses. In other words, hatred and contempt do not become the absolute dominant factors in the utterance”\(^{42}\). Otherwise, the language based on contempt, or even hate speech, will expand its circles towards whole stories, which can be described as a narrative of hatred.

Media researchers must discover and promote dialogue in language-based communication with each other; culture in personal and public life revealed in the media space, maturity in adopting different views, seeking a consensus in negotiations instead of fighting. And finally, a reminder of the axiological dimension of ethical discourse, in which – as Bartmiński stressed in *Kulturowy kanon wartości* [Cultural canon of values] – dignity is the highest value, and freedom of speech only comes with it.

---

\(^{41}\) M. Drożdż, *Hate speech*..., op. cit., p. 30.

References


Drożdż M., *Etyczne aspekty mediów integralną częścią nauk o mediach* [Ethical aspects of media aa integral part of media science], “Studia Medioznawcze” 2013, no. 4 (55).

Drożdż M., *Język nienawiści w dyskursie medialnym* [Hate speech in media discourse], “Acta Universitatis Lodziensis, Folia Litteraria Polonica ”2016, no. 1 (31).


Juza M., *Hejterstwo w komunikacji internetowej: charakterystyka zjawiska, przyczyny i sposoby przeciwdziałania* [Hating in Internet communication: characteristics of the phenomenon, causes and methods of counteracting], “Profilaktyka Społeczna i Resocjalizacja ”2005, no. 25.


Naruszewicz-Duchlińska A., *Nienawiść w czasach Internetu* [Hatred in the ear of Internet], Gdynia 2015.