10.17951/j.2020.33.4.239-248

ANNALES
UNIVERSITATIS MARIAE CURIE-SKŁODOWSKA
LUBLIN – POLONIA

VOL. XXXIII, 4 SECTIO J 2020

The Maria Grzegorzewska University. Institute of Education

ANNA ANDRZEJEWSKA

ORCID 0000-0002-0743-3517,
an.andrzejewska@o2.pl

Challenges for Educational, Rehabilitation and Penitentiary Measures Against Hate Speech Use in Social Media

Wyzwania dla oddziaływań wychowawczo-resocjalizacyjnych i prawnych przeciwko mowie nienawiści w mediach społecznościowych

How to quote this paper: Andrzejewska, A. (2020). Challenges for Educational, Rehabilitation and Penitentiary Measures Against Hate Speech Use in Social Media. *Annales Universitatis Mariae Curie-Skłodowska. Sectio J, Paedagogia-Psychologia, 33*(4), 239–248. DOI: https://doi.org/10.17951/j.2020.33.4.239-248.

ABSTRACT

Hate speech should be ranked among new phenomena occurring in digital space, especially in social media. This issue is of interdisciplinary nature. Its theoretical and empirical analysis was presented within the context of especially media pedagogy; however, references to the achievements of legal, sociological and psychological sciences were made in the paper to a certain extent. This new and dynamically developing phenomenon was characterised in the present article. The recommendations of the Commissioner of Civil Rights Protection of the Republic of Poland were shown as challenges for educational, rehabilitation and legal measures within this area. The analyses included aim at persuading the reader to consider the harmfulness of hateful language in individual and social aspects.

Keywords: hate speech; youth; social media; rehabilitation; penitentiary science

INTRODUCTION

Social networking sites have revolutionised individual, social and cultural life in global terms. They help us at work, provide entertainment, facilitate social relations as well as broaden our horizons due to the possibility of sharing knowledge, work and ideas with other people. Reality takes place on two planes – online and offline. When analysing the issue of social media it is worth to focus on the

view suggested by Maciej Tanaś, who states that "the development of information and communication technology has broadened the scope of potential contacts and has provided us with attractive forms of communication, enabling forming and keeping such contacts irrespective of territorial and temporal boundaries. It has also become possible to cross cultural boundaries by the people who know either pidgins or the mother tongue of the interlocutor" (Tanaś, 2006, p. 5).

"Social networking services are used for building social relations and networks often based on mutual opinions. The way we achieve self-presentation, the kind of content we share, and, especially, how we are evaluated by other users, who do not necessarily have to be our real-life acquaintances – it all has impact on fundamental part of our identity. Facebook and Instagram help millions of people consciously create a neurotic identity captured in the confinement of their own image" (Andrzejewski, 2018, p.22).

It is worth noting that one is more and more frequently exposed to a growing scale of hate speech in social networking sites. It can be encountered in web forums, below the photos published online, in the form of memes, movies and songs available in popular networking sites. The use of hate speech by the internauts is happening more and more often, even on a mass scale, despite the rise of organisations trying to fight it. Adolescents and children are most susceptible to hate speech in social media and they are more and more willing to apply it in cyberspace.

The aim of the article is a theoretical and empirical analysis of the issue of hate speech in the following areas:

1. The characteristics of hate speech phenomenon.
2. The recommendations of the Commissioner of Civil Rights Protection.
3. Rehabilitation and penitentiary measures in the context of hate speech.

1. THE CHARACTERISTICS OF HATE SPEECH PHENOMENON

Despite the fact that the term hate speech has been present in the public discourse for a number of years, it has not been defined in an unequivocal way. Within the provisions of law in Poland there is no exact definition of "hate speech", thus a lot of criminal acts caused by it cannot be prosecuted efficiently. To a limited extent some acts can be prosecuted – they take into account the characteristics of description of a hate-motivated criminal act. There are also those which are difficult to be held liable due to the mode of prosecution (for example, a private prosecution). A crucial aspect is a perpetrator's motivation for a crime based on hate.

The term "hate speech" was formed following the term "hate crime", which constitutes a penal act. Among such acts are included, for instance, a homicide, health impairment, property damage, threats. These offences need to be preceded by adequate motivation of the perpetrator. Hate speech is not always classified as

a crime, but it is worth noting that certain types of speech, such as a threat, can be recognised as one.

The term "hate crime" was first used in legal publications in the 1980s in the United States (Duda, 2016, p. 480). The reason for the creation of that term was the increasing sensitivity of the public towards prejudice and violation of law in the racial, religious, ethnic and sexual orientation contexts. Nevertheless, in Poland it has not been defined by legislation despite in-depth studies in the literature on the subject started in 21st century (Duda, 2016, p. 48).

Hate crimes are considered to be "every offence of a criminal nature, including offences aimed at people and their property, as a result of which the victim, venue or any other offence target are chosen with regard to a real or alleged connection, relation, affiliation, membership or support towards a group distinguished by characteristics common to all its members, such as real or alleged race, nationality or ethnicity, language, skin colour, religion, sex, age, mental or physical disability, sexual orientation or other similar features" (Pływaczewski, 2012, p. 46).

One can thus agree with Mateusz Woiński on the fact that acts possessing the character of hate speech, simultaneously meeting the legal criteria of a crime, can be described as "hate speech crimes" (Woiński, 2014, p. 32). According to Maciej Duda "hate speech crimes are offences under articles 119, 256 and 257 of the Penal Code committed as verbal acts (speech acts)" (Duda, 2016, p. 53).

Taking the above considerations into account, one should refer to the definition of "hate speech" as proposed by the European Council in 1997. The definition comprises "all forms of speech that spread, promote or justify racial hatred, xenophobia, anti-Semitism as well as other forms of hatred based on intolerance, i.a. intolerance expressed through aggressive nationalism and ethnocentrism, discrimination and hostility towards minorities, immigrants and people of immigrant origin" (The Council of Europe, 1997). This definition is most often referred to in their documents by European Institutions, i.a. the European Court of Human Rights, various non-government organisations and human rights defenders.

One should pay attention to the fact that the phenomenon does not only concern lack of tolerance of racist or xenophobic nature, but also manifests itself in the escalation of hatred in the form of violent acts and exhortations to criminal offences. It is not only realised through the exhortations to hate others on the basis of sex, race, nationality, sexual orientation, disability etc., but also through violent acts directed at those groups and popularisation of offensive content, humiliating images and materials about a given person (Uniwersytet Warszawski, 2020).

In the Polish society, the brutalisation of language has been taking place to a larger and larger extent. Instead of using dialogue, opinions characterised by negative emotions are expressed more and more often. Such opinions evoke hostility and hatred among recipients towards people with different worldviews. The freedom of speech is not considered a boundary and human dignity is not taken

into consideration. It is especially evident in social media, where the quality of communication is reduced to a minimum. The language used is overflowing with obscenities and abuse. The commonplace character of hate speech is confused with broadly understood freedom of speech and the right to express one's own opinions.

As it was mentioned before, setting a boundary between hate speech, criticism and common freedom of speech is not easy, but it is crucial from the point of view of internauts. In cyberspace, especially in social networking sites, the problem is even greater, since there is a lot of the content that is extremely hard to classify in one way or the other. Criticism is aimed at showing mistakes or even irregularities in a message or behaviour, without using stigmatising language; hate speech, however, refers to humiliation, abuse and inciting fear in another person.

It is possible to notice that in internet circles hate speech is widely used, i.a. in discussion forums, below celebrities' and private individuals' photos describing their creative activity, etc. A person applying hate speech feels allowed to formulate thoughts in this way. Very few react to it and it is rarely reported, thus the consent is given.

It should be stressed that this phenomenon is widespread among young internauts – users of social networking services. The research conducted by the Bathory Foundation in 2014 point to this fact. During the research a group of 653 Polish youth aged 16–18 was randomly examined. The results show that, among other things, 2% of the young people permit to say such anti-Semitic claims as "Jews need to understand that their treachery and crimes caused the hatred of Polish people. Today, Jews are trying to hide their guilt and put responsibility on us". 14% of young people encounter such statements frequently, 59% of which mostly on the internet. 19% of respondents allow for hateful language in relation to the Ukrainians, which is manifested, for instance, through the acceptance of the web comment "the Ukrainians can murder helpless women and children at best". 46% of youth have contact with this kind of language on the internet. It is similar with the hateful language used against Romani people, which is encountered online by 60% of the youth. 16% of young Poles consider a statement "A Gypsy is a thief, and it will always be like that" justifiable. Hate speech directed at people of different skin colour is widespread on the internet and it is expressed by statements like "A Negro is not a human but a monkey". 70% of young respondents actually witnessed it online. It is also essential to address the hateful language popularised on the internet with relation to non-heterosexual people. 77% of adolescents declare having had contact with it on the internet. A statement "I loathe faggots, they are degeneration of humanity and should undergo treatment" was accepted by 20%. There is also wide consent for hate speech against the Muslim community. 19% of young respondents deem statements such as "Muslims are filthy cowards, they murder only women, children and the innocent" as acceptable and, in 55% of

cases, point to the internet as the main source of contact with such utterances (see also: Bilewicz, Marchlewska, Soral, Winiewski, 2014, pp. 4–7).

As it is observable in the research presented, hateful language has entered the internet discourse of adolescents. That is why it is extremely important, for pedagogical reasons, to take widespread educational action within this social group so as to raise awareness in terms of widely understood harmfulness of its use.

2. THE RECOMMENDATIONS OF THE COMMISSIONER OF CIVIL RIGHTS PROTECTION

With reference to the problematics of hate speech, in 2019 the Commissioner of Civil Rights Protection submitted recommendations to the Prime Minister of Poland concerning effective struggle with the issue along with its accompanying crimes, after previously making them public on Facebook. The Commissioner of Civil Rights Protection also treated them as a prelude to a debate on the public authorities' involvement in the fight against hate speech.

They concern (*Jak walczyć z mową…*, 2019):

I. Changes in law.
   1. Prosecution *ex officio* of the crimes based on prejudice against disability, age, sexual orientation and gender.
   2. Penalisation of sole membership in organisations promoting or inciting to racial hate.
   3. The introduction of the legal definition of hate speech.

II. The fight against hate speech on the internet.
   1. The creation of codes of good practice by the internet providers and non-governmental organisations.
   2. The appointment of independent contact points and advisory bodies monitoring the enforcement of responsibilities by internet providers with respect to counteracting hate speech.
   3. Making the administrators of commercial internet information portals aware of their responsibility to moderate, filter and delete comments that might be hate speech.
   4. The introduction of the internet providers' duty to inform the law enforcement authorities about hate speech cases.
   5. The internet providers' duty to implement an online form allowing to report online illegal content and clarification concerning the dates of their deletion. The introduction of a so-called blind lawsuit.
   6. Development and maintenance of a publicly available list of trustworthy reporting entities.

III. The changes concerning social and legal awareness.
   1. Running social campaigns aimed at school-age youth.

2. Carrying out trainings, workshops and instructions in schools concerning the harmfulness of hate speech.
3. Organising methodological trainings for teachers; police officers, prosecutors and judges should be provided with trainings on legal aspects of hate speech, similarly to lawyers and legal advisers.

IV. The changes related to the activity of public authorities.

1. The publication of recommendations on hate speech by the National Broadcasting Council.
2. A complex, independent analysis of the Parliamentary Ethics Committee as regards its reaction to hate speech.
3. A complex analysis of the way the public prosecutor's office deals with hate speech, i.a. with regard to the implementation of the Public Prosecutor General's guidelines of 26th February 2014 on said matter.
4. The review of all the proceedings discontinued by the public prosecutor's office of the last 3 years, in which it was possible to press charges of promoting hate speech.
5. The encouragement of political parties to sign the Charter of European Political Parties for a non-racist society.
6. The involvement of city authorities and local government units for the fight against hate speech.

It is worth stressing, however, that the issue encompasses the internet users regardless of age. It occurs most often among young people, though, which is shown by numerous research results as well as those presented in the present paper.

The actions proposed are extremely valuable, especially in the context of children and youth, who have not been equipped with the critical faculty towards publishing unwanted content full of hatred and contempt as well as recognition and proper reaction to such content. It can be achieved, as pointed out by the Commissioner of the Civil Rights Protection, through the introduction of changes concerning raising social and legal awareness (i.a. thanks to running social, educational and training campaigns for various milieus on the harmfulness of hate speech) and through changes related to the activity of public authorities (*Jak walczyć z mową…*, 2019).

3. REHABILITATION AND PENITENTIARY MEASURES IN THE CONTEXT OF HATE SPEECH

The main normative act that regulates the forms of responsibility for using hate speech is the penal code. These are crimes under the Articles 119, 256 and 257 (*Ustawa z dnia 6 czerwca 1997 r. Kodeks karny*), belonging to the category of conventional crimes, which result from international conventions ratified by

Poland. Additionally, the signs of online hate are prosecuted under Article 212 for libel and under Article 216 for insult (*Ustawa z dnia 6 czerwca 1997 r. Kodeks karny*). The abovementioned crimes are subject to fine, restriction of personal liberty or imprisonment. Jakub Czarkowski proves that "apart from questions of general nature there is a vital social necessity to reflect upon the issue of executing a penalty in individual as well as systemic terms. The issue of penalty execution is crucial for rational and effective functioning of judiciary system. It is also essential for the sense of justice and safety of citizens, not only in a personal and individual sense, but also when it comes to the perception of society as a sovereign who is entitled to demand from the justice system both efficiency in action and acting in their name with dignity and the respect for rules and values important to the citizens" (Czarkowski, 2018, pp. 123–140).

Conviction by virtue of law and the placement of online hate perpetrator in a penal institution initiates a whole range of rehabilitation measures while being sentenced to imprisonment, which may be diverse in terms of pedagogical efficiency for the imprisoned. Every man in such conditions reacts differently to various penal treatments. Institutional conditions of imprisonment efficiency, determinants of isolation of such criminals, therapy efficiency, effectiveness of correctional methods, prisonisation process, prison staff competence, integration of penitentiary and post-penitentiary measures – these are some of numerous intervening variables pertaining to the issue of reasonableness of isolation sanctions for hate speech use.

One should notice that every instigator or exponent of online hate speech who enters the penitentiary system generates additional load for the prison system apparatus. It is not always the case that rehabilitation proves effective. A perpetrator after release from prison may not become a better person, understand their mistake, draw conclusions, make amends, adjust to society and "recover" from the need to spread online hate. One should take into consideration the risk of creating a vicious circle. A hater sentenced to imprisonment for an insulting online comment, coming against a toxic environment and harsh conditions of a penal institution, may become even more depraved as a result of environmental impact. Having left the confines of the penal institution, in time they might give vent to their anger towards the world in a worse way than just hate speech, thus turning to recidivism.

In order to find a solution to the abovementioned dilemmas, one should consider a hypothetical systemic modification in the form of changing the classification of a deed from a crime to a petty offence. Taking the current legislation into account, it would require the amendment of legal and systemic regulations. Such a change would enable the replacement of criminal sanctions in the form of imprisonment with punitive measures prescribed by the procedures in the Code of Petty Offences. A solution proposed is to impose financial sanctions on the peo-

ple who commit an act of hate speech online. In practice, it would require additional involvement of the officers from the Cyber Crime Bureau, running as a part of the National Police Headquarters, and the officers from the provincial police headquarters, where there are units specialising in this regard. So as to identify the occurrence and the authors of offensive and vulgar content, a greater activity of law enforcement bodies on online forums and social media would be essential. Based on joint cooperation with social networking sites administrators and internet providers, the officers would be able to determine the IP address of the computer, which was used to send a post and, consequently, personal information of the sender, issuing a penalty notice.

It is worth to analyse such an operating model for a several reasons. A penalty notice is a form a simplified procedure in a petty offence cases and also it involves imposing a fine for the minor offence (Journal of Laws of 2018, item 475, as amended). It is a type of pre-trial procedure which allows for minimising the social costs of proceedings. Moreover, it also fulfils preventive and educational goals in relation to a perpetrator. (Marek, 2008, pp. 214–215). This approach would permit to resolve the case almost immediately. According to the reinforcement theory, devised by Burrhus F. Skinner, a negative consequence (positive punishment) should take place immediately after an undesired behaviour. From the point of view of behaviourism, the time criterion and fast reaction towards the behaviour is essential to attain corrective effects of the punishment (Zimbardo, Ruch, 1997). Currently the proceedings take a long time, and a good deal of cases are remitted (*Wyciąg ze sprawozdania*…, 2017).

It can be, thus, assumed that such a fast punishment may be proper from the preventive and educational perspective, especially in case of perpetrators whose manifestations and consequences of hate speech can be placed as minor social danger of the deed. Another advantage is unburdening of the penitentiary apparatus, which has been struggling with overcrowding for a while. Attention should be paid to the fact that the number of prisoners in penal institutions is constantly rising, which makes the Polish penitentiary system lead in statistics compared to other European countries (Adamczyk, 2015, p. 10).

The resignation to transfer the perpetrator into prison system may happen to be a valuable solution not only for the economy of the system itself, but also in pedagogical and social reflection. With reference to the above considerations, what is debatable is the efficiency of penitentiary measures, which can do more harm than good in case of the individuals convicted of spreading hate. The prospective advantage of the solution is the improvement of the issues related to legal regulations and the reinforcement of the cooperation between the law enforcement authorities and the entities creating information ecosystems in the cyberspace. In the light of civilisational, cultural, social and technological changes, the tightening of those relations for the good of internet users and the general public seems to be inevitable.

SUMMARY

Hate speech can be placed among new phenomena occurring in cyberspace, especially in the domain of social media. Very young people are more and more often its perpetrators, because it is in cyberspace that they display the greatest activity. There have been more and more actions taken by numerous academic bodies as well as by practitioners analysing the requirements, course and consequences of internet hate in individual and social aspects. Nevertheless, these attempts are insufficient in comparison with the scale and range of its use. It is so because cyberspace can be characterised by, for instance, its globality and, what is more important, anonymity which fosters disinhibition, giving rise to a greater array of hate speech uses.

Preventive and educational measures that have been taken so far in educational institutions have not yet yielded expected results. Social awareness is still at a low level and the competences of not only children and youth but also parents, custodians and tutors still need to be improved in this regard. Formal and legal regulations have also become a significant aspect and they require a greater consideration and analysis in the light of the occurrences escalating on the internet as well as in the real public space.

REFERENCES

Adamczyk, M. (2015). System penitencjarny w Polsce, współczesne problemy oraz możliwości reformacyjne. *Horyzonty Bezpieczeństwa, 1*(1), 7–23.

Andrzejewski, A. (2018). „Lajki" i „followersi" na portalach społecznościowych sposobem na poszukiwanie własnej tożsamości. *Problemy Opiekuńczo-Wychowawcze, 1*, 21–29.

Bilewicz, M., Marchlewska, M, Soral, W., Winiewski, M. (2014). *Mowa nienawiści. Raport z badań sondażowych*. Warszawa: Fundacja im. Stefana Batorego.

Czarkowski, J. (2018). Ontologiczne i epistemologiczne uwagi o penitencjarystyce i jej autonomizacji. *International Journal of Legal Studies, 1*(3), 123–140. DOI: 10.5604/01.3001.0012.2163.

Duda, M. (2016). *Przestępstwa z nienawiści: studium prawnokarne i kryminologiczne.* Olsztyn: Uniwersytet Warmińsko-Mazurski.

Dz.U. 1997 nr 88 poz. 553)

Jak walczyć z mową nienawiści. 20 rekomendacji RPO dla premiera. (2019). Retrieved February 25, 2020, from https://rpo.gov.pl/pl/content/jak-walczyc-z-mowa-nienawisci-20-rekomendacji-rpo-dla-premiera

Marek, A. (2008). *Prawo wykroczeń (materialne i procesowe).* Warszawa: Wydawnictwo C. H. Beck.

Pływaczewski, W. (2012). Mowa nienawiści jako komponent tak zwanej przestępczości stadionowej. In: W. Pływaczewski, B. Wiśniewski (eds.), *Przestępczość stadionowa. Diagnoza i przeciwdziałanie zjawisku* (pp. 38–47). Szczytno: Wydawnictwo Wyższej Szkoły Policji.

Recommendation No. R (97) 20 of the Committee of Ministers to Member States on „hate speech", The Council of Europe, 1997.

Tanaś, M. (2016). Diagnoza funkcjonowania nastolatków w sieci – aspekty społeczne, edukacyjne i etyczne. In: M. Tanaś (ed.), *Nastolatki wobec internetu* (pp. 5–10). Warszawa: NASK.

Uniwersytet Warszawski (2020). *Mowa nienawiści*. Retrieved February 20, 2020, from http://rownowazni.uw.edu.pl/mowa-nienawisci/

Ustawa z dnia 24 sierpnia 2001 r. Kodeks postępowania w sprawach o wykroczenia. (Dz. U. z 2018 r. poz. 475, z późn. zm.)

Ustawa z dnia 6 czerwca 1997 r. Kodeks karny.

Woiński, M. (2014). *Prawnokarne aspekty zwalczania mowy.* Warszawa: LexisNexis.

Wyciąg ze sprawozdania dotyczącego spraw o przestępstwa popełnione z pobudek rasistowskich, antysemickich lub ksenofobicznych prowadzonych w I półroczu 2017 roku w jednostkach organizacyjnych prokuratury. (2017). Retrieved January 22, 2020, from https://pk.gov.pl/wp-content/uploads/2017/12/8ffee219b2f4dd56387508008eee13a5.pdf

Zimbardo P., Ruch F. L. (1997). *Psychologia i życie*. Warszawa: Wydawnictwo Naukowe PWN.

STRESZCZENIE

Do nowych zjawisk, które mają miejsce w przestrzeni cyfrowej, a zwłaszcza w mediach społecznościowych, należy zaliczyć mowę nienawiści. W artykule dokonano charakterystyki tego nowego i dynamicznie rozwijającego się zjawiska oraz przedstawiono rekomendacje Rzecznika Praw Obywatelskich zalecające wdrożenie oddziaływań wychowawczo-resocjalizacyjnych i prawnych w tym obszarze. Zamieszczone w tekście analizy mają na celu skłonienie do przemyśleń na temat szkodliwości języka pogardy w aspekcie jednostkowym i społecznym.

Słowa kluczowe: mowa nienawiści; młodzież; media społecznościowe; resocjalizacja