

## Balancing freedom – Navigating the nuances of speech regulation for equality and justice

Marina Munoz Ayuso<sup>1</sup>

<sup>1</sup>Studentin der Alice Salomon Hochschule

### Contents

- I. Unravelling new complexities of free speech
- II. Legal framework
- III. Debating freedom of speech and its constraints
- IV. Exploring the challenges of speech regulation
- V. Mapping the paths to responsible expression
- VI. Conclusion

### Abstract

In a controversial context concerning the implications of freedom of speech, this paper argues that regulations are needed in order to prevent ideas and actions that aim to suppress freedom and equality. For this purpose, a review of the new challenges that highlight this need is conducted, along with some older yet still relevant justifications. With that established, it is proposed that the main challenge within this topic does not lie in the regulation or non-regulation of freedom of speech. Indeed, the real debate is found in determining the circumstances and conditions under which limitations can be justified in this domain, seeking lawful and proportionate resolutions. The aim of this text is not to provide a final answer to the discussion, rather a wider and clearer sight on important matters that must be taken into account for limiting speech, and some input on the later proposals that must be further developed.

### Keywords

free speech, hate speech, regulations, democracy, discrimination

#### Citation:

Marina Munoz Ayuso, Balancing freedom – Navigating the nuances of speech regulation for equality and justice, in: MRM 29 (2024) 2, pp. 94–107.  
<https://doi.org/10.60935/mrm2024.29.2.10>.

Received: 2024-08-13

Accepted: 2024-09-16

Published: 2024-11-19

#### Copyright:

Unless otherwise indicated, this work is licensed under a Creative Commons License Attribution 4.0 International. This does not apply to quoted content and works based on other permissions.

## I. Unravelling new complexities of free speech

Changes in the past years like the introduction of new forms of expressions like social media have set a wider, more intense and polarised debate<sup>1</sup> on freedom of speech and its limitations. It is, for instance, nowadays evident that the internet serves as a platform for some of the most negative social behaviours,<sup>2</sup> as it provides a wide-ranging platform for individuals who seek to perpetuate oppression against minority groups. Through it, anyone can reach and influence large audiences. This means that those advocating for social progress, as much as those promoting racist violence, can expand the scope, diversity, and reach of their audience.<sup>3</sup> According to *Laub*,<sup>4</sup> an increasing frequency of assaults targeting immigrants and other minority groups has raised new concerns regarding the link between online hate speech and acts of violence. This has also led to urgent discussions about the responsibilities of corporations and governments in regulating speech.

The aforementioned has been demonstrated through several studies, such as

the one carried out by *Müller and Schwarz*<sup>5</sup> about the link between anti-refugee sentiment on Facebook and hate crimes against refugees in Germany. They point to the influence of the AfD (Alternative für Deutschland) that positioned itself as an anti-refugee and anti-immigration party, in promoting anti-refugee sentiment on social media and the correlation with the anti-refugee violent incidents. Because of this connection, the authors call attention to the recent policy debates about whether and how to “regulate” hate speech on social media.

Another instance was the ethnic cleansing targeting the Muslim community, specifically the Rohingya men, in Myanmar. In this scenario, Facebook posts and messages served as effective tools for those aiming to propagate hate. The influence of the rhetoric employed – rooted in hatred and discrimination – was amplified by the dissemination of false or incomplete information and explicit calls for nationalist actions. Meanwhile, the government’s response did not address hate speech. Consequently, the authorities in Myanmar fostered an environment where hate speech could spread easily, human rights abuses were legitimised, and incitement to discrimination and violence were facilitated.<sup>6</sup>

These cases manifest not only the need for specific regulations on this topic, but a requirement for new laws that adapt to the current context that the era of internet and social media has created. As a result, nations acknowledge hate speech as a significant issue, prompting various

<sup>1</sup> *Alexander Tsesis*, Hate in Cyberspace: Regulating Hate Speech On the Internet, in: *Loyola University Chicago Law Journal* (2001), pp. 817–874 (818); *Zachary Laub*, Hate Speech in Social Media: Global Comparisons, in: *CFR Backgrounders of 7 June 2019*, available at: <https://www.cfr.org/backgrounders/hate-speech-social-media-global-comparisons> (last visited 19 April 2024).

<sup>2</sup> *Nazmine/Khan Manan/Hannan Khan Tareen/Sidra Noreen/Muhammad Tariq*, Hate Speech and social media: A Systematic Review, in: *Turkish Online Journal of Qualitative Inquiry* 2021, pp.5285-5294.

<sup>3</sup> *Tsesis* (Fn. 1).

<sup>4</sup> *Laub* (Fn.1).

<sup>5</sup> *Karsten Müller/Carlo Schwarz*, Fanning the Flames of Hate: Social Media and Hate Crime, in: *Journal of the European Economic Association* 2020, pp. 2131–2167.

<sup>6</sup> Cf. UN Doc. A/HRC/39/64.

international and European proposals designed to define the problem and devise effective strategies to tackle it. Germany for instance has recently passed a law known as NetzDG (Netzwerkdurchsetzungsgesetz)<sup>7</sup>. Through it, social media companies are required to take down hate speech within twenty-four hours of a complaint. *Kohl*<sup>8</sup> would argue that NetzDG can be understood as a compromise between the two opposite speech standards. Yet, there is still a debate inside it regarding the transfer of public dimensions into private processes on the one hand, and the incompatibility with the First Amendment as it seeks to protect the integrity of the act of expression from government intrusion on the other.

Despite the new concerns introduced by the development of technology, the core arguments in the debate have seen little alteration. Arguments for unrestricted speech are still rooted in the work of classical authors like *Mill*,<sup>9</sup> or *Dworkin*,<sup>10</sup> emphasising its essential role in democratic governance, autonomy and intellectual growth. Conversely, advocates for hate speech regulations highlight concerns about its potential to incite violence and perpetuate discrimination<sup>11</sup>. As the discourse unfolds, challenges emerge in de-

lineating the boundaries of permissible speech. However, two essential points will be considered in order not to mischaracterize the debate:

First, the crucial debate is not about whether we should infringe free speech in order to stamp out hateful attitudes and the various evils they engender. A revision has demonstrated that hate speech does not even constitute the sort of expression that the right to freedom of expression exists to protect. Consequently, the debate lies less in endorsing unrestricted free speech or not, but rather around the definition of hate speech and the specific areas of protection.

As for the second most relevant point, it is argued that a comprehensive resolution of this debate should strive for more than merely proposing a balancing model between the commitment to free speech and with other normative commitments, such as the social equality, dignity, or the security of historically marginalised citizens. It should at least aim to find a clear and well-supported perspective that offers clear normative guidance.<sup>12</sup>

On that basis, the discussion is divided in different sections for the purpose of a clearer analysis. Hence, a review of the rules applicable to hate speech will first be conducted, followed by the discourse surrounding freedom of speech and its potential limitations, the challenges that the establishment of limitations to free speech may face, and finally some input on the later proposals to face this problem.

<sup>7</sup> *Bundesministerium der Justiz*, Netzwerkdurchsetzungsgesetz of 1. September 2017 (Federal Law Gazette I pp. 3352), updated by Article 29 of the act of 6 May 2024 (Federal Law Gazette 2014 I No. 149).

<sup>8</sup> *Uta Kohl*, Platform regulation of hate speech – a transatlantic speech compromise?, in: *Journal of Media Law* 14 (2022), pp. 25–49.

<sup>9</sup> *John Stuart Mill*, *On Liberty*, 1978, pp. 45, 82.

<sup>10</sup> *Ronald Dworkin*, A New Map of Censorship, in: *Index on Censorship* 23 (1994), pp. 9–15.

<sup>11</sup> *Hebert Marcuse*, Repressive Tolerance, in: Robert Paul Wolff/Barrington Moore, J.R./Hebert Marcuse, *A Critique of Pure Tolerance*, 1969, pp. 95–137; *Mill* (Fn.9); *Jeffrey W. Howard*, Free

speech and hate speech, in: *Annual Review of Political Science* 22 (2019), pp. 93–109.

<sup>12</sup> *Jeremy Waldron*, *The Harm in Hate Speech*, 2012, p. 47.

## II. Legal framework

### 1. Protection against hate speech in international law

Freedom of speech is safeguarded by written laws like constitutions and bills of rights, considered a fundamental freedom that governments cannot suppress or control. This right is assured by various international agreements, including the Universal Declaration of Human Rights (UDHR) Art. 19<sup>13</sup> and important conventions like the European Convention on Human Rights and Fundamental Freedoms (ECHR) Art. 10<sup>14</sup>, and the International Covenant on civil and political rights (ICCPR) Art. 19<sup>15</sup>.

These instruments mainly ensure that individuals have the right to speak their minds without interference from the government or other authorities.<sup>16</sup> In addition, they often include freedom of speech not just as a specific right, but as well as a component of other rights protected under the same agreements, such as the freedom of assembly.

Nonetheless, these same instruments also incorporate that the exercise of these rights carries with it special duties and responsibilities, which makes them subject to certain restrictions.

In both the UDHR and the ICCPR, freedom of expression is enshrined in Article

19, laying out precise parameters for its limitations in para. 3: Restrictions must be legally prescribed and essential for safeguarding the rights or reputations of others, or for preserving national security, public order, or public health and morals. Another instance is Art. 10 of the ECHR, which stipulates the right to freedom of expression, but in para. 2 takes into account the duties and responsibilities associated with its exercise. Under this pretext, Art. 17 of ECHR has given the Court a legal frame to hold on to, so that freedom of expression may not be used to lead to the destruction of the rights and freedoms granted by the same Convention. These decisions are rooted in the concept of the paradox of tolerance: complete tolerance can result in the acceptance of ideas that advocate intolerance, finally endangering the very tolerance it seeks to uphold<sup>17</sup>.

The protection against hate speech is mainly a process that took place after the Second World War and the impact in society that Nazi propaganda had on society. The founding of the United Nations (UN) in 1945 and, in particular, the UDHR in 1948 were key events in the protection of the right to equality between persons, without discrimination of any kind, as well as the protection of human dignity. Since then, the ICCPR has been the most relevant international instrument and, unlike the previous declaration, it is legally binding on states that have ratified it. Its significance lies not only in how Article 19(3) of this international treaty sets out the circumstances in which freedom of expression may be limited, but also in the explicit prohibition of what is currently

<sup>13</sup> UN Doc. Resolution 217 A III.

<sup>14</sup> Convention for the Protection of Human Rights and Fundamental Freedoms of 4 November 1950, ETS No. 005.

<sup>15</sup> International Covenant on Civil and Political Rights of 16 December 1966, UNTS vol. 999, p. 171.

<sup>16</sup> *Martín Risso Ferrand*, Freedom speech and combat to hate speech, in: *Estudios Constitucionales* 2020, pp. 51–89 (56).

<sup>17</sup> *Dominika Bychawska-Siniarska*, Protecting the right to Freedom of Expression under the European Convention on Human Rights: A handbook for legal practitioners, 2017, p. 12.

known as hate speech in Article 20(2)<sup>18</sup>. In the American context, the American Convention on Human Rights (ACHR) of 1969 is equally important,<sup>19</sup> and so are the ECHR in the European context and the African Charter on Human and People's Rights<sup>20</sup> for African countries. Additionally, other treaties have been important for Freedom of expression in the international sphere such as: the "Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW)"<sup>21</sup>, the "Declaration on the Elimination of All Forms of Intolerance and Discrimination Based on Religion or Belief"<sup>22</sup>, and especially the International Convention on the Elimination of All Forms of Racial Discrimination (ICERD)<sup>23</sup>.

These examples are proof that since many years, the international community has already given relevance to the regulation of free speech as a tool not only to prevent atrocities witnessed in the last century but also to ensure and effectively protect the dignity of historically marginalised groups.

<sup>18</sup> *Juan P. Cajigal Germain*, *Los discursos de odio como límite a la libertad de expresión*, 2018, pp.45–48.

<sup>19</sup> Organization of American States (OAS), *American Convention on Human Rights (ACHR)*, treaty No. 1144 of 22 November 1969.

<sup>20</sup> *African Charter on Human and People's Rights*, treaty CAB/LEG/67/3 rev. 5, 21 I.L.M. 58 (1982) of 21 October 1986.

<sup>21</sup> UN Committee for the Elimination of All Forms of Discrimination against Women, *CEDAW General Recommendation No. 19: Violence against women*, in UN Doc. HRI/GEN/1/Rev.1.

<sup>22</sup> UN Doc. A/36/PV.73.

<sup>23</sup> *International Convention on the Elimination of All Forms of Racial Discrimination* of 7 March 1966, UNTS vol. 660, p.195.

## 2. Divergent paths: a comparative analysis of hate speech regulation in Germany and the United States (US)

As reviewed, since the Universal Declaration of Human Rights in 1948, international human rights instruments have included clauses condemning discriminatory conduct. However, it is the duty of States to amend their national legislation to bring it into line with this principle in order to effectively guarantee the right to equality and non-discrimination. In this way, States parties have adopted various standards against arbitrary discrimination, but without international consensus.

There are two main models of protection against hate speech: On the one hand, Western European states, despite their declared commitment to fundamental rights, lean more towards a restrictive state policy that emphasises the main constitutional value of human dignity. In Denmark, for instance, hate speech is linked to Section 266b of the Danish Penal Code<sup>24</sup>, which makes it a criminal offense to express statements that "publicly or with intent to disseminate to a 'wider circle, threaten, insult or degrade a group of persons on the basis of race, skin colour, nationality, ethnicity, faith or sexual orientation'. Italy, on its side, has the Law 205/1993 which punishes hate speech and tackles discrimination on the grounds of race, religion and nationality."<sup>25</sup>

The German case stands out due to Germany's strict attitude towards any form of racist behaviour, alongside a steadfast commitment to the principles of a free

<sup>24</sup> *Criminal Code of Denmark*, LBK No. 1007 of 24 October 2012, para. 266b.

<sup>25</sup> *Italy Law No. 205 of 25 June 1993* referred to as the "Mancino Law".

and democratic society<sup>26</sup>. The German Basic Law's Art. 5<sup>27</sup> establishes the foundation for freedom of expression in the country, which also specifically outlines limitations on free speech in section 2. Besides, some other restrictions, implicit and explicit, are included in the Basic Law through several articles (Art. 1; Art. 18; Art. 21, section 2). Sections 130 and 131 of the German Criminal Code (Strafgesetzbuch)<sup>28</sup> also criminalise hate speech, specifically incitement to hatred based on race, ethnicity, or religion, which stems from the country's historical context and lessons learned after World War II. Thus, Germany upholds criminal responsibility for those who employ hate speech.

On the other hand, proponents of free speech as a key constitutional value in the US, often draw on *Justice Holmes*<sup>29</sup> and *Mill's* marketplace of ideas theory<sup>30</sup> and the self-government argument of *Meiklejohn*.<sup>31</sup> The US example is relevant to this article as it is the major country following an approach that protects hate speech for the sake of freedom of expression<sup>32</sup>. Its legal principles governing freedom of expression originated from the inter-

pretation of the First Amendment to the US Constitution, which states: "Congress shall make no law ... . abridging the freedom of speech."<sup>33</sup> For that reason, the US Supreme Court generally classifies hate speech as falling within the realm of protected speech, "regardless of the effect it has on the listener and society."<sup>34</sup>

Yet, cases of lawsuits where courts have awarded damages for racial insult happened,<sup>35</sup> but these were only contemplated under Tort law. In order for this to happen, the speech has to be addressed by the Brandenburg rule, which requires it to present a "clear and present danger".<sup>36</sup> For instance, in *Alcorn v. Anbro Engineering*,<sup>37</sup> the California Supreme Court ruled that a plaintiff could recover damages for emotional distress caused by severe racial insults in the workplace. The plaintiff, Alcorn, was subjected to racial slurs and fired in a humiliating manner by his employer. The court found that such extreme and outrageous conduct could constitute intentional infliction of emotional distress, even without physical harm. This decision marked an important recognition of the emotional impact of racial discrimination and allowed for compensation under California law.

<sup>26</sup> Yulia A. Timofeeva, *Hate Speech Online: Restricted or protected? Comparison of Regulations in the United States and Germany*, in: J. Transnational Law & Policy 12 (2003), pp. 253–286 (270).

<sup>27</sup> Art. 1; Art. 5; Art. 18; Art. 21, section 2 of the Basic Law for the Federal Republic of Germany of 23 May 1949, updated by Article 1 of the Act of 19 December 2022 (Federal Law Gazette I, p. 2478).

<sup>28</sup> Criminal Code of 13 November 1998 (Federal Law Gazette I, p. 3322), as last amended by Article 2 of the Act of 22 November 2021 (Federal Law Gazette I, p. 4906).

<sup>29</sup> *Oliver W. Holmes, Abrams v. United States*, decision of 10 November 1919, para. 58.

<sup>30</sup> *Mill* (Fn. 9).

<sup>31</sup> *Alexander Meiklejohn, Free speech and its relation to self-government*, first edition, 1948.

<sup>32</sup> *Timofeeva* (Fn. 26).

<sup>33</sup> U.S. CONST. amend. I.

<sup>34</sup> *Rachel Weintraub-Reiter, Hate Speech over the Internet: A Traditional Constitutional Analysis or a New Cyber Constitution?*, in: Boston University Public Interest Law Journal 1998, pp.145-174 (161), citing *Samuel Walker, Hate Speech: The History of an American controversy*, 1994.

<sup>35</sup> See, e.g., *Wiggs v. Courshon*, 485 F.2d 1281 (5th Cir. 1973); *Alcorn v. Anbro Eng'g*, 2 Cal. 3d 49 3 (Cal. 1970); *Agarwal v. Johnson*, 25 Cal. 3d 932 (Cal. 1979).

<sup>36</sup> See U.S. Supreme Court, *Schenck v. United States*, decision of 3 March 1919, para. 5.

<sup>37</sup> Supreme Court of California, *Alcorn v. Anbro Eng'g*, 2 Cal. 3d 49 3 of 24 April 1970.

However, in more recent decisions, the Supreme Court has displayed a reduced willingness to acknowledge this legal claim<sup>38</sup>. In this regard, the *R.A.V. v. St. Paul*<sup>39</sup> ruling was very significant as it reinforced the broad protections of the First Amendment, limiting the government's ability to regulate speech based on its content. In this case, the Supreme Court struck down a St. Paul, Minnesota ordinance that prohibited the display of symbols like burning crosses or swastikas that incite racial or gender-based hatred. The decision emphasized that, even though hate speech is offensive, laws regulating speech must be content-neutral and cannot target specific viewpoints.

Thus, the opposition of these two models, well represented by Germany with a more restrictive approach, and the US with a more liberal approach, is clear. The liberal approach prioritises safeguarding safety and public order for the prohibition of hate speech, whereas the restrictive approach emphasises safeguarding dignity along with fostering tolerance and non-discrimination, crucial elements in a democratic and pluralistic society<sup>40</sup>.

### III. Debating freedom of speech and its constraints

Freedom of expression is a fundamental right deeply rooted in Western cultures. Its origins can be traced back to the Declaration of the Rights and Duties of Man and

the Citizen<sup>41</sup> following the French Revolution of 1789, where representatives of the French people emphasised the importance of various rights and responsibilities associated with liberty and equality.

Through history, and intensified in the 19th and 20th centuries, some classical authors have contributed to the definition and defence of free speech and have helped shape modern conceptions of its importance in democratic societies:

*Mill*<sup>42</sup> with his work 'On Liberty' and *Justice Holmes* in *Abrams v. United States*<sup>43</sup>, both argued the importance of protecting deviating opinions and ideas because of the role of such speech in the pursuit of truth. In addition, they followed the promotion of a 'marketplace of ideas', as a metaphor for a situation in which people speak and exchange ideas freely, still predominating in modern First Amendment jurisprudence in the US. Aligned with this conception, *Meiklejohn*<sup>44</sup> with his theory of self-government, based his defence of freedom of speech on democratic theory. For him, freedom of speech – understood as the right to know – is the fundamental requirement for democratic governance. On his side, *Milton*<sup>45</sup> advocates in his work 'Areopagitica' for freedom of speech as a condition for the development of morality.

Their thoughts were relevant not only for understanding the broad scope that the right to free expression includes, but also for distinguishing its different dimensions. When the discussion about restrictions to

<sup>38</sup> *Timofeeva* (Fn. 26).

<sup>39</sup> U.S. Supreme Court, *R.A.V. v. City of St. Paul*, 505 U.S. 377, of 22 June, 1992.

<sup>40</sup> *Cajigal* (Fn. 18).

<sup>41</sup> Declaration of Man and the Citizen of 26 August 1789.

<sup>42</sup> *Mill* (Fn. 9).

<sup>43</sup> *Holmes* (Fn. 29).

<sup>44</sup> *Meiklejohn* (Fn. 31).

<sup>45</sup> *John Milton*, *Areopagitica*, 1964, p. 275.

this right, like rules against hate speech, arises, those with a more liberal stance typically advocate for freedom of expression, emphasising its role in fostering autonomy, democracy, and the exchange of ideas in the marketplace of thoughts<sup>46</sup>.

### 1. The defence of unrestricted speech

Following *Howard's*<sup>47</sup> analysis, there are currently four main arguments supporting the idea of freedom of speech against the regulation and criminalization of hate speech:

a) First, and defended by *Dworkin*,<sup>48</sup> freedom of speech from a listener's autonomy standpoint values individuals' ability to think independently. This means people should have the freedom to decide their beliefs and access diverse opinions without government interference. According to *Nagel*<sup>49</sup> restrictions on freedom of expression, like those targeting hate speech, derive from a form of authoritarianism that does not regard citizens as free and equal. Such limitations impose criteria on individuals and prevent them from their decision-making ability. Advocates argue that even potentially harmful or offensive views should be heard by autonomous individuals. *Feinberg*<sup>50</sup> adds to this by suggesting that the harm principle alone might not comprehensively

cover all aspects necessary for a rule concerning free speech.

- b) The second argument regards the speaker's autonomy: In line with *Barendt*,<sup>51</sup> limitations on what one can say or read hinder personal growth. Without the freedom to express political views and engage in public debate, individuals cannot fully develop intellectually and spiritually. Conversely, free speech enables individuals to express themselves and shape their lives. It is about personal freedom and the ability to share beliefs. Protecting free speech allows people to discover themselves. *Baker*<sup>52</sup> argues that even hate speech can be a form of self-expression, so suppressing it might hinder people from showing who they truly are.
- c) There is a third argument from a Democracy perspective: The literature review reveals a common concern among the authors (*Meiklejohn*<sup>53</sup>; *Weinstein*<sup>54</sup>) that is related to the protection of free speech, which is democracy. Freedom of expression is always presented as crucial for democracy because it enables an exchange of ideas, promotes the discovery of truth through open debate, and empowers individuals to participate fully in the democratic process. Therefore, when we consider the debate around hate speech, we face a series of problems inherent to debate in the framework of democracy.

<sup>46</sup> *Cajigal* (Fn. 18).

<sup>47</sup> *Howard* (Fn.11).

<sup>48</sup> *Dworkin* (Fn. 10).

<sup>49</sup> *Thomas Nagel*, Personal Rights and Public Space, in: *Philosophy & Public Affairs* 24 (1995), pp. 96-97.

<sup>50</sup> *Joel Feinberg*, Harm to Others: The Moral Limits of the Criminal Law, 1984.

<sup>51</sup> *Eric Barendt*, Freedom of Speech, 1985, p. 14.

<sup>52</sup> *C. Edwin Baker*, Human Liberty and Freedom of Speech, 1992, p. 69.

<sup>53</sup> *Meiklejohn* (Fn. 31).

<sup>54</sup> *James Weinstein*, Extreme Speech, Public Order, and Democracy: Lessons from the Masses, in: *Extreme Speech and Democracy* 2009, pp. 25-30, 47, 48, 61.



Authors like *Heinze*<sup>55</sup> clearly state that society's democratic character is reduced when the law prevents citizens from advancing hateful policy proposals. In recent discussions, some influential thinkers<sup>56</sup> argue that freedom of speech is crucial for our role as citizens in a democratic society. They believe that without free speech, we cannot properly engage in making the laws that affect us. This idea emphasises that public discussions about political matters are what freedom of speech is mainly meant to protect. According to these thinkers, democracy allows individuals to reconcile different perspectives through reasoning, promoting self-determination.

- d) Lastly, there is a position related to the marketplace of ideas that *Cueva Fernández*<sup>57</sup> named the 'epistemic argument'. In this case, relevant support for free speech emphasises its role in accurately conveying our thoughts.<sup>58</sup> This perspective stresses that speech is crucial for understanding others' genuine beliefs and to seek for the truth. Without it, grasping the world and discussing moral issues becomes challenging. It asserts that limiting free communication and suppressing genuine expressions is unjust. According to *Mill*, such liberty of expression is necessary for the dignity of persons. If liberty of expression is stifled,

the price paid is "a sort of intellectual pacification" that sacrifices "the entire moral courage of the human mind".<sup>59</sup>

## 2. Advocating for hate speech regulations

However, even very liberal authors and great defenders of free speech consider that speech will have to be somehow limited for the sake of order. *Mill* is one of them, with his statement "the only purpose for which power can be rightfully exercised over any member of a civilised community, against his will, is to prevent harm to others"<sup>60</sup>.

In response to the four arguments earlier presented, it is discussed that some hate speech does not just try to persuade people to hate others; it includes direct attacks, threats, and harassment against vulnerable groups. Even if people want to make their own choices, they also want protection from discrimination and violence that hate speech might encourage. 'One potential rationale of hate speech restrictions is to stop the spread of the kind of xenophobic, nativist, hate-filled discourse that leads to flagrantly unjust policies'<sup>61</sup>. So, it might make sense for citizens in a democratic society to ask the government to stop hate speech to protect them from harm, even if it limits their ability to hear different opinions.

Also, not all speech is about presenting reasons for people to think or act in certain ways. For example, some argue against violent pornography because it might affect people's desires or beliefs without

<sup>55</sup> *Eric Heinze*, *Hate Speech and Democratic Citizenship*, 2016, p. 7.

<sup>56</sup> *Corey Brettschneider*, *When the State Speaks, What Should It Say? How Democracies Can Protect Expression and Promote Equality*, 2012; *John Rawls*, *Political Liberalism*, 2005.

<sup>57</sup> *Ricardo Cueva Fernández*, *El discurso de odio y su prohibición*, in: *Cuadernos de Filosofía del Derecho* 35 (2012), pp. 437–455 (442).

<sup>58</sup> *Seana Valentine Shiffrin*, *Speech Matters: On Lying, Morality and the Law*, 2014, pp. 1–4.

<sup>59</sup> *Mill* (Fn. 9), p. 31.

<sup>60</sup> *Mill* (Fn. 9), p. 9.

<sup>61</sup> *Howard* (Fn. 11), p. 99.

them consciously thinking about it. If hate speech operates in a way that influences people's minds without engaging their ability to think for themselves, then stopping it might still respect people's ability to think independently.<sup>62</sup>

On the other hand, the democracy argument is referred in every liberal text and might be considered a strong one, but *Heinze's*<sup>63</sup> statement on how the society's democratic character is reduced when the law prevents citizens from advancing hateful policy proposals raises the question of 'why we should want to live in a democracy, or at least a democracy of this demanding kind?'<sup>64</sup> While political participation is crucial for collective self-governance, other rights also contribute to sustaining individual autonomy, as reflected in constitutions.<sup>65</sup> Moreover, not all individuals possessing moral autonomy have access to the same rights, which calls into question the traditional democratic argument.

Finally, the epistemic argument could only be considered in a perfect world in which a perfect marketplace of ideas exists without the predominance of certain dominant groups that start from an advantageous position to impose their ideas. In any case, the most important point here is that the argument lacks validity as long as the subjects do not have access to perfect information and the possibility to act freely in providing and choosing the options available.

Overall, the prevalent argument suggests that while personal freedom is typically significant, its value diminishes when it is employed for harmful or mistaken intentions.

Therefore, as the philosopher *Marcuse*<sup>66</sup> expressed already half a century ago in his contribution to the book '*Repressive Tolerance*': while tolerance is essential in many situations like harmless debates or academic discussions, it cannot extend to protecting false words or harmful actions that oppose freedom and liberation. In matters where freedom and happiness are threatened, society cannot afford to be indiscriminate, as it would perpetuate oppression rather than fostering genuine freedom.

*Marcuse*<sup>67</sup> posits that in contemporary Western societies, tolerance has been distorted to tolerate the intolerable, allowing oppressive ideologies and practices to persist under the guise of free speech and diversity of opinion. In any case, true tolerance should not extend to ideas and actions that aim to suppress freedom and equality. Instead, the author advocates for a form of 'liberating tolerance', which involves actively opposing and restricting intolerant ideologies in order to create a more just and equitable society. In order to achieve genuine freedom, society must be willing to be intolerant towards oppressive forces, even if it means restricting their expression.

<sup>62</sup> *Howard* (Fn. 11), pp. 93–109.

<sup>63</sup> *Heinze* (fn. 55), p. 7.

<sup>64</sup> *Amanda R. Greene/Robert Mark Simpson*, *Tolerating hate in the name of democracy*, in: *Mod. Law Rev.* 80 (2017), pp. 746–765 (757).

<sup>65</sup> *Cueva Fernandez* (fn. 57), p. 444.

<sup>66</sup> *Marcuse* (Fn. 11).

<sup>67</sup> *Ibid* (Fn.66).

## IV. Exploring the challenges of speech regulation

The discussion is therefore not so much about whether or not to support the idea of completely unrestricted freedom of speech. Rather, the focus should be on exploring and making tough decisions about what should or shouldn't be said, and determining the specific areas or instances where limitations on speech are necessary. Indeed, the biggest challenge here is around the definition of hate speech.

*Waldron*,<sup>68</sup> for instance, states that the harm in hate speech results primarily from speech that is written rather than spoken. This assertion requires updating due to the emergence of new communication avenues like online platforms and social media. Hate speech online presents distinct features in terms of how interactions happen and how particular content is used and disseminated. Hateful content online can persist longer, experience fluctuating popularity, establish connections with new networks, resurface, and be anonymous. As a result, there is ongoing debate about the responsibility of online space moderators and the criteria for removing content.<sup>69</sup> Indeed, the previously mentioned law NetzDG clearly represents this debate as it mandates that bigger social media platforms moderate illegal content, including hate speech. The main issues for criticism were, firstly, the risk that this law implies for turning platforms into "privatized censors" by forcing them to make decisions that should be handled by courts. On the other hand, it could also encourage excessive content removal to avoid fines or due to the time limits that

have been set (illegal content must be removed within 24 hours).

In addition, social media companies have shifted to automated systems as a technique for monitoring hate speech at large scale. This also presents challenges, since errors in detection can lead to the removal of non-hateful content that could stifle open dialogue and threaten free speech. Besides, these systems are often limited to English, and research on multilingual methods is lacking. That, and the fact that most studies and monitoring efforts focus on the U.S. and Europe, results in a lack of tools, data, and a clear understanding of how hate speech spreads in other regions. Bridging this gap is especially important due to the context-specific nature of hate speech.<sup>70</sup>

Moreover, there is a subjectivity and context dependence of hate speech that makes it difficult to establish universal definitions. Different cultural norms and societal sensitivities influence perceptions of hate speech. *Waldron*, for example, defends that while feelings of offence, even if strongly felt, should not be regulated by law, the main focus of hate speech laws should be to safeguard dignity – as a person's fundamental right to be seen as a valued member of society, regardless of their belonging to a minority group, and to ensure that they are not excluded from regular social interactions<sup>71</sup>. On the other hand, *Feinberg* advocates that "the prevention of offensive conduct is properly the state's business"<sup>72</sup>, and that many factors need to be taken into account, such as the extent, duration and social value of

<sup>68</sup> *Waldron* (Fn.12) p. 45.

<sup>69</sup> UN Doc. CI/FEJ/2021/DP/01.

<sup>70</sup> UN Doc. (Fn. 69) p. 6

<sup>71</sup> *Waldron* (Fn. 12) p. 15.

<sup>72</sup> *Joel Feinberg*, *Offense to Others: The Moral Limits of the Criminal Law*, 1985, p. 1.

the speech, the ease with which it can be avoided, the motives of the speaker, the number of people offended, the intensity of the offence, and the general interest of the community.

It is true that some legislation has proven to adjust to the standards of clarity, just like Article 20 of the ICCPR<sup>73</sup>, which links “advocacy” of national, racial or religious hatred with “incitement” to discrimination, hostility or violence. However, the translation of this same Covenant from Spanish into English already gives space for different interpretations. *Rollnert*<sup>74</sup> identifies for instance how the word ‘Advocacy’ in the English version doesn’t mean the same as the word ‘Apología’ in the Spanish version, which also includes justification or praise. Moreover, linking advocacy to incitement (to discrimination, hostility and violence), requires that there is a risk where the courts must assess whether there was a reasonable likelihood that the incitement would succeed. This ambiguity in the legislation is repeatedly found over several examples already reviewed, such as the Tort laws in the US, where no elements are listed for what constitutes “inciting, aiding and abetting”, nor is there a reference to intent.<sup>75</sup>

After reviewing the main challenges to the definition of hate speech, it is clear that

vagueness and ambiguity is a base problem related to the other challenges. The language in hate speech laws or definitions can be unclear or uncertain, creating difficulties in establishing precise criteria for identifying hate speech. This lack of clarity may result in inconsistent application and understanding of these laws. For this reason, the inaccuracies in the international standards must be analysed with care.

However, it should not be forgotten that the discussion revolves around establishing subsequent responsibilities for the exercise of one of the basic rights of a democratic society, namely freedom of expression.

## V. Mapping the paths to responsible expression

If one thing is clear, it is the fact that regulations require legal backing to ensure clarity, precision, and accessibility so that individuals can comprehend the consequences of their speech. Additionally, they must align with the essential requirements of a democratic society and withstand scrutiny through a proportionality assessment, evaluating suitability, indispensability, and careful balance.<sup>76</sup>

In order to achieve this, an approach reviewed by *Ferrand* is being considered. This consists of the “hate triangle”<sup>77</sup>: involving the inciter, the victim group and the incited third parties. This means that it is not solely about the content of the speech; all three elements must be present.

<sup>73</sup> Art. 20 ICCPR.

<sup>74</sup> *Göran Rollnert Liern*, El discurso del odio: una lectura crítica de la regulación internacional, in: *Revista Española de Derecho Constitucional* 115 (2019), pp. 81–109 (87).

<sup>75</sup> Report to the Ad Hoc Committee on the Elaboration of Complementary Standards to the International Convention on the Elimination of All Forms of Racial Discrimination on the intersessional legal expert consultation considering the elements of a draft additional protocol to the Convention prepared by the Ad Hoc Committee at its tenth session of 21 – 22 October 2020, p. 8.

<sup>76</sup> *Ferrand* (Fn. 16), p. 84.

<sup>77</sup> *Ferrand* (Fn.16), p. 80.

Under this framework for criminalising hate speech, the sender must have a deliberate intention to incite, which means that those acts lacking this intention would not fall within the prohibited category. The motivation behind the speech matters in determining whether it qualifies as prohibited hate speech.<sup>78</sup> The difficulty here involves the complexity of determining intent. Malicious intent can be challenging to prove and may lead to legal uncertainty.

Besides, hate speech at its essence involves targeting individuals or a collective on the basis of their identity or belonging to a group. Therefore, for a law to be effective, it must specify whether all groups are covered by the law and, if not, which specific groups are included within its scope. Originally, only national, religious or racial groups were considered, as collected in Art. 20 from the ICCPR<sup>79</sup>, but nowadays, there is a need to consider broadening it to also include homophobic speech, anti-women speech, etc. But these groups must also suffer serious harm. There must be an intention to harm, to damage, to injure. This last point is not clear for example from the same Covenant, for example, and must be carefully analysed from a criminal law perspective.

The third vertex of the triangle, the incited group, needs to be clearly defined in terms of time and place. This helps us to see if there was a real risk of stirring up violence, discrimination, or hostility. For it to be considered as incitement, it has to actually cause the intended effects or pose a serious and certain risk, not just speculation.

---

<sup>78</sup> Ibid (Fn. 76).

<sup>79</sup> Art. 20 ICCPR.

Lastly, context must not be forgotten. Context is essential to accurately interpret the meaning and impact of speech, while ensuring that laws are adaptable to different situations and social changes. To be clear, it is not the same to use the ‘N-word’ in a neutral context like an academic debate about the history of the Civil Rights movement, as it is to use it in a context filled with hatred and discrimination, such as a discourse that promotes white supremacy and denigrates people of African descent.

A different proposal is to shift the focus of the debate from the protection of rights to a contemplation of duties. To support a criminal ban, it has to be proven that people have a moral duty to avoid the behaviour in question.<sup>80</sup> According to *Howard*,<sup>81</sup> five duties must be enforced to justify coercion: the duty not to threaten, the duty not to harass, the duty not to offend, the duty not to defame, and the duty not to incite wrongdoing. While citizens may lack a moral right to engage in hate speech, they may have a moral duty to refrain from it, which could be enforceable by law. However, enforcing these duties raises concerns about the reliability of the state or the hostile reactions of those inclined towards hatred that may be provoked<sup>82</sup>.

Therefore, these approaches should not be regarded as definitive guidelines, but maybe as starting points to create a clearer path for establishing restrictions on free speech.

---

<sup>80</sup> *Douglas Husak*, *Overcriminalization: The Limits of the Criminal Law*, 2008, p. 66.

<sup>81</sup> *Howard* (Fn. 11).

<sup>82</sup> *Geoffrey R. Stone*, *Perilous Times: Free Speech in Wartime from the Sedition Act of 1798 to the War on Terrorism*, 2004.

## VI. Conclusion

The issue of hate speech poses a big challenge for democracy because it seems to clash with common values of free speech and equality. Nonetheless, a simple principle emerges: it is fundamentally about how human beings should treat each other. Hate speech is fundamentally wrong because it denies the humanity of others. Consequently, when a liberal state limits this kind of speech, it does not infringe on individuals' rights; rather, it reaffirms the inherent value of all its members.<sup>83</sup>

International standards and the criminal laws of many nations currently reflect a widespread consensus that inciting hatred is detrimental and undermines dialogue. Across the globe, democratic societies consistently emphasise the importance of safeguarding human rights above allowing individuals to propagate harmful messages according to their own whims. It has therefore been recognised that while free speech is crucial for collective decision-making, hate speech that induces fear among individuals or promotes discriminatory behaviour undermines collective autonomy. Accordingly, many countries have determined that it is feasible to safeguard individuals' self-determined right to expression without embracing absolute free speech principles.

At this point, it appears that complete opposition to the criminalization of hate speech might not be feasible, and therefore hate speech must be contemplated under the right to freedom of expression. The issue is not a binary choice between

“regulating hate speech” and “unrestricted freedom of expression.” Instead, it involves beginning with the principle of freedom of expression and determining the circumstances and conditions under which limitations can be justified in this domain, seeking lawful and proportionate resolutions.

However, there is little consensus on which should be the essential elements to justify the limitations to free speech, which hinders the discussion. The only evidence is that if hate speech is to be criminalised, the laws or regulations must be exceptionally clear, specific, and narrowly defined. “Restrictions should align with the essential requirements of a democratic society and undergo a proportionality assessment: appropriateness, necessity and balancing in the strict sense of the word.”<sup>84</sup>

Anyhow, it should always be borne in mind that any limitation on freedom of expression should be considered with great caution, as it is a matter of significant gravity. Consequently, any such limitation must be subjected to a thorough and rigorous examination.

---

<sup>83</sup> Steven J. Heyman, *Hate-Speech Bans Are Consonant with Liberal Principles*, in: Eric Heinze, Natalie Alkiviadou, Tom Herrenberg, Sejal Parmar and Ioanna Tourkochoriti (ed.), *The Oxford Handbook of Hate Speech*, 2023, p. 18.

---

<sup>84</sup> *Ibid* (Fn. 76).