



Hate speech as a threat to human rights

April 2019

Facultad de Derecho

Centro de Estudios en Libertad
de Expresión y Acceso a la Información

UP
Universidad
de Palermo

Hate speech as a threat to human rights

Natalia Torres¹ y Víctor Taricco²

I. Introducción

In recent times there has been a disturbing increase in the circulation of so-called hate speech in the public sphere. The discussion on the legalization of abortion, the debates on immigration policy and on the actions of security forces in Argentina have caused the circulation of violent, aggressive or discriminatory expressions in public conversation. The inflamed political confrontation in the last presidential elections in Brazil; the escalation of harassment and threats that led to the resignation of the Brazilian congressman Jean Wyllys; the situations of racial violence in the United States; hate speech related to migrants in Hungary; the attacks against the Muslim community in Myanmar; the Brexit campaign in Great Britain and more recently the attacks suffered by Venezuelan migrants in Ecuador, have shed a light on the relevance of the issue and the need to advance in its conceptualization and the systematic collection of data.

States, international multilateral governance bodies and civil society organizations have made efforts to understand and deal with this phenomenon. Those analysis and attempts at answers have relied mostly on tools provided by law to analyze a topic that presents a special complexity, as hate speech presents a tension between the right to express all kinds of ideas in the public debate and the need to guarantee to all citizens the possibility of exercising their rights on an equal footing. The way in which hate speech has been defined in various current legal frameworks explains the way in which the conflict between freedom of expression and the promotion of equality has been resolved.

Faced with this question, we assume that between freedom of expression and the right to equality there is an affirmative, positive and complementary relationship, as the Camden principles state: "The realization of the right to freedom of expression facilitates a vibrant and multifaceted public interest debate that gives voice to different perspectives and points of view. Inequality results in the exclusion of certain voices, undermining this debate. People's rights to be heard, to speak and to participate in political, artistic and social life are, in turn, integral to the attainment and enjoyment of equality".³

The need to conduct theoretical research on hate speech and to collect data on the issue has been highlighted — as this document will show — by different multilateral organizations and leaders in the promotion of freedom of expression and equality. This work aims to contribute to that discussion. Our objective is to consider the wide

¹ Natalia Torres is a researcher at CELE, a public policy consultant specializing in the right to information and gender. She is a political scientist graduated from the University of Buenos Aires and holds a master's degree in public policy from University College London.

² Victor Taricco holds a degree in Communication Sciences from the University of Buenos Aires and a Master's Degree in MIES-CA-UBA. As an independent consultant he participated in different political campaigns.

³ Article 19, "Los principios de Camden sobre libertad de expresión e igualdad", [The Camden Principles on Freedom of Expression and Equality], 2009, retrieved from: <https://www.article19.org/data/files/pdfs/standards/los-principios-de-camden-on-the-freedom-of-expression-and-the-equality.pdf>, last access: April 8, 2019.

range of phenomena contained under the concept of “hate speech” that threaten the effective exercise of human rights to move forward in designing policies that are not limited to prohibition and which propose an active role for the states in the deconstruction of stereotypes, in actions aimed at countering discriminatory discourses, in the promotion of pluralism and in guaranteeing everybody an unrestricted access to the public debate.

In the first section, we present an approach to hate speech as social discourse, as discursive articulations that are based on a conception of the world that seeks to exclude and segregate diversities, differences and dissent. In this section, we will also present a possible distinction between different types of hate speech.

In the second section, we will analyze how hate speech has been conceptualized in the Inter-American, European and international legal frameworks, considering how it is approached from its defense and the promotion of freedom of expression as well as from the fight against discrimination and racism.

In the third section, we will review current policies to counter hate speech, expanding on the concept of social hate speech and the differentiation of phenomena encompassed in them. In the final section, we will present conclusions and a series of recommendations on the matter.

II. Hate speech as social discourse

A first and general approach to the issue of hate speech will allow us to point out that these expressions, in their multiple levels, are used to harass, persecute, segregate, justify violence or deprivation of rights, generating an environment of prejudices and intolerance that encourages discrimination, hostility or violent attacks against certain people or groups of people; for reasons of race, color, sex, language, religion, political opinions or of any other nature, national or social origin, economic position or any other social condition.⁴ This ability of hate speech to generate an environment of intolerance and encourage discrimination and violence can be further understood if analyzed as social discourses.

According to Marc Angenot,⁵ social discourses can be defined as everything that is said and written at a given historical moment in a given society, everything that is “narrated and argued” at a certain moment through the media, public conversations or social networks. Another, more specific, way of describing social discourses would be to assume that these are constituted by certain rules of succession that organize what is available, what can be said at a given historical moment. These discursive series prescribe legitimate ways of saying that, by making themselves a place in the uniformity of social murmur, have “social efficacy and captive audiences”.

According to the Belgian-Canadian historian, social discourses form discursive memories loaded with ways of outlining the functioning of the world, they bear the marks of the ways of knowing and representing what is known, they manifest social interests and the appropriate norms of conduct for that community, generating a discursive memory, made up of forms and contents, that globally overdetermine what can legitimately be said and what cannot be said.

This overdetermination is slowly forming a *doxa*, “what is obvious”, what is necessary to be able to think what is thought and say what is said. The discursive memories that constitute this *doxa* are loaded with a series of ideas

⁴ Gagliardone, Ignio, Gal, Danit, Alves, Thiago and Martínez, Gabriela, “Countering Online Hate Speech”, UNESCO, 2015, retrieved from: <https://unesdoc.unesco.org/ark:/48223/pf0000233231>, last access: April 8, 2019.

⁵ Angenot, Marc, *El discurso social: los límites históricos de lo pensable y lo decible* [Social discourse: the historical limits of what is thinkable and what can be said], Buenos Aires, Siglo XXI, 2012.

and preconceptions about the characteristics and intentions of the “other”. This other, as different studies have shown, is always less known for its own attributes than for the fantasies and fears it embodies. As Slavoj Zizek states, echoing Jaques Lacan, the other can only be known through a presumption, as “an object of belief”.⁶

These presuppositions about the other, which exist in the form of fears from time immemorial, in periods of political and social uncertainty, articulate a series of evident explanations about the difficulties that are being faced, through some outlines about the functioning of the world. As the philosophers Patricia Berrotarán and Alejandro Kaufman put forward, these explanations constitute a “doxical precondition” about the reasons that account for the events that surround us.⁷

The doxic prerequisite then functions as the necessary symbolic field for acts of responsibility, defamation, harassment, discrimination, denial of rights or violence. Every new act of speech takes part in the formation of the ideological and discursive framework that shapes the social world. As Angenot points out, social discourses create “a gnoseology about the world, a series of outlines that constitute the precondition of judgments”⁸

In the specific case of hate speech, as social discourses, we understand that there is a strong unity forged within this hate speech between a certain conception of the world and rules of conduct that adjust to this conception. These are forms that reject diversity, difference or dissidence or, as Ezequiel Ipar asserts, “Freedom as exercised by an ‘other’”.⁹ Hate speech, in any of its levels, consists of discursive articulations that try to prevent the other to exercise the right to freedom and equality.

Hate discourses, as social discourses, seek to impose a unique way of interpreting events, as well as a correct, usually traditional, way of working in the community. Due to these characteristics, hate speech is usually directed against dissident, vulnerable, migrant groups or any person or group of people that is seen as threatening (or responsible for the loss of) a political and social order that must be restored.

Hate speech articulates historically constituted fears, socially held prejudices and organizes a political will to unify over the elimination of everything that does not correspond to a certain way of understanding and inhabiting the world. According to Antonio Gramsci,¹⁰ who inspired much of Angenot’s reflections, hate speech is organized as an “iron doctrine” where all diversities, differences and dissidence must be silenced, segregated or eliminated to reestablish that ideal (and idealized) social order that has been lost by the action of others.

Hate speech, as it does not seek to incorporate all that is different under a logic of hegemonic domination, tries to bring back a “unifying function” similar to that which religious discourses exercised in pre-modern social formations. A non-hegemonic will of domination over the other.

One of the most common forms of exercising that will of non-hegemonic domination is defamation. According

⁶ Zizek, Slavoj, *Sobre la violencia: seis reflexiones marginales* [Violence: Six Sideways Reflections], Buenos Aires, Paidós, 2009.

⁷ Kaufman, Alejandro and Berrotarán, Patricia, “La construcción de la tiranía: el Libro Negro” [The construction of tyranny: the Black Book], in: Bonet, María Teresita and Ciappina, Carlos, *Representaciones, discurso y comunicación: el peronismo entre 1945-1970* [Representations, speech and communication: Peronism between 1945-1970], La Plata, Editorial de la Universidad de La Plata, 2014, pp. 23-42.

⁸ Angenot, *op. cit.*

⁹ Ipar, Ezequiel, “Instituciones desfondadas generan un enorme espacio para el autoritarismo social” [Degraded institutions generate a giant space for social authoritarianism], interview by Nicolás Tereschuk, in: *Arte y política*, 2018, retrieved from: <http://artepolitica.com/articulos/ezequiel-ipar-instituciones-desfondadas-generan-un-enorme-espacio-para-el-autoritarismo-social>, last access: April 8, 2019.

¹⁰ Gramsci, Antonio, *El materialismo histórico y la filosofía de Benedetto Croce* [The historical materialism and philosophy of Benedetto Croce] Buenos Aires, Nueva Visión, 2008.

to Jeremy Waldron,¹¹ what constitutes hate speech is the regular practice of defamation of certain persons or groups of people, which affects “the normative basis of equality” for the exercise of rights. Defamation of a person or group of people directly affects their reputation and the possibilities of exercising their rights to education, work, health and everything they are entitled to as human beings. These attacks on reputation constitute an “assault on dignity”, understanding by dignity the basic social position that allows the recognition of the other as an equal, as a subject of rights, as a citizen.

We understand that it is important to point out that hate speech as social discourse must reach legitimate ways of expression, with social efficacy and a receptive audience. To achieve this status, hate speech must leave the private sphere and gain legitimacy in the public arena. The beginning of this process, the construction of the doxical precondition, requires an authorization, permission, (pre) recognition of social legitimacy for those statements.

Waldron concludes that the search for legitimacy would be enunciated through a speech stating:

We know that some of you agree that these people are not wanted here. We know that some of you feel that they are dirty (or dangerously criminal or terrorist). Know that you are not alone. Notwithstanding what the government says, there are enough of us to make sure that these people are not welcome. There are enough of us to draw attention to how these people really are. Talk to your neighbors, talk to your customers. And above all, do not let more of them in.¹²

1. Towards a typology of hate speech

Identifying and differentiating hate speech has been the object of multiple studies and theories. The different classifications of hate speech combine the regularity or systematicity of the statements, their specific contents, the conditions of enunciation, the contexts in which they circulate and their capacity to harm.

One of the first contributions in this regard was provided by a group of civil society organizations based on an initiative of the organization Article 19. The Camden principles on freedom of expression and equality recommend that:

*The national legal systems should make clear, either explicitly or through authoritative interpretation, that: (i) The terms “hatred” and “hostility” refer to intense and irrational emotions of shaming, showing enmity and aversion towards the target group. (ii) The term “promotion” shall be understood as requiring the intention to publicly promote hatred against the target group. (iii) The term “incitement” refers to statements about national, racial or religious groups that may create an imminent risk of discrimination, hostility or violence against persons belonging to such groups.*¹³

Returning to some issues analyzed by the Camden principles, the UN Rabat Plan of Action (2013) establishes a series of criteria for identifying hate speech and its eventual criminalization: (i) the prevalent social and political context at the time the speech was articulated and disseminated; (ii) the position or social status of the speaker, including the position of the individual or the organization in the context of the audience to which the speech is addressed; (iii) the intention of the speaker; (iv) the content or form of the speech, which may include the evaluation

¹¹ Waldron, Jeremy, *The Harm in Hate Speech*, Cambridge, Harvard University Press, 2012.

¹² Waldron, *op. cit.*, p. 13. .

¹³ Article 19, *op. cit.*

of the extent to which the speech was provocative and direct, as well as a focus on the form, style and nature of the arguments expressed in the speech in question or in the balance reached between the arguments expressed; (v) the attributes of the discourse, including elements such as the scope of the discourse, its public nature, the relevance and size of the audience; and (vi) the possibility, including the imminence, that there is a reasonable probability that the discourse will succeed in inciting real action against the target group, recognizing that this causal relationship must be rather direct.¹⁴

This interpretive framework — similar to the one generated that same year by the UN Committee on the Elimination of Racial Discrimination (CERD) —¹⁵ is what allows the Plan of Action to distinguish between three types of speeches: (i) expressions that constitute a crime, (ii) expressions that are not punishable criminally but that could justify a civil process or administrative sanctions, and (iii) expressions that are not legally sanctionable “but that still generate concern in terms of tolerance, civility and respect for the rights of others”. That is to say, a differentiation of hate speech according to the damage produced allows us to distinguish the range of possible actions and not limit them to the criminalization, prohibition and restriction of discourse.

After the violent elections of 2007 in Kenya (which left more than 1,000 dead and 600,000 people were displaced), Project UMATI was created in 2013 in order to analyze the circulation of hate speech on the Internet. The project developed a specific methodology to identify, collect and classify this type of discourse based on the definitions provided by Susan Benesch¹⁶ about dangerous discourse, as a discourse with the potential to catalyze collective violence. The key variables of Benesch’s perspective indicate that one should take into account: (i) the influence of the speaker, (ii) the audience’s receptivity, (iii) understanding the content of the speech as a call to action, (iv) the social and historical context in which the discourse unfolds and (v) the means of dissemination used to issue the information.

Using these variables, the UMATI project defined three categories: (i) offensive discourse, (ii) moderately dangerous discourse, or (iii) extremely dangerous discourse, especially depending on the speaker’s level of influence and what is perceived as a call to action.¹⁷ In offensive speech, according to UMATI, the main intention is to insult a member of a certain group due to their belonging to it, or directly insult the entire group. In this classification, the speaker has little influence on the audience, what the speaker says generates few reactions and the statements do not encourage to commit harmful actions against the offended target group. Therefore, the statements in this category are described as “offensive” because they aim to discriminate verbally, but have a low potential to provoke violence. It was also observed that, if these words or statements were repeated by more influential speakers and more receptive audiences, they could increase their level of danger until violence was ignited.

In moderately dangerous speeches, speakers generate interventions with little or moderate influence on their audience. The content has a mixed effect: some could see it as inflammatory, while others find it simply offensive. For UMATI, the influence that a speaker has on a particular audience has greater relevance than the content of the statement.

¹⁴ United Nations High Commissioner for Human Rights, “Rabat Plan of Action”, 2013, retrieved from: https://www.ohchr.org/Documents/Issues/Opinion/SeminarRabat/Rabat_draft_outcome.pdf, last access: April 8, 2019.

¹⁵ Committee on the Elimination of Racial Discrimination (CERD), “Recomendación general N° 35. La lucha contra el discurso de odio racista” [General Recommendation No. 35. Combatting racist hate speech], United Nations Human Rights, Office of the High Commissioner, 2013, retrieved from: <https://www.ohchr.org/sp/hrbodies/cerd/pages/cerdindex.aspx>, last access: April 8, 2019.

¹⁶ Benesch, Susan, “Vile Crime or Inalienable Right: Defining Incitement to Genocide”, in: *Virginia Journal of International Law*, Vol. 48, No. 3, April, 2008, retrieved from: https://www.researchgate.net/publication/228149554_Vile_Crime_or_Inalienable_Right_Defining_Incitemen, last access: April 8, 2019.

¹⁷ iHub Research, *Umati Final Report*, 2013, retrieved from: <https://dangerouspeech.org/umati-final-report>, last access: April 8, 2019.

Statements that fall into the category of dangerous speeches are made by speakers with moderate or high influence over their audiences. Their statements had a strong inflammatory character and pose a great risk of triggering violence. UMATI describes as dangerous speeches those expressions that are an explicit or implicit call to overcome, forcibly evict or kill a person or a group of people. Comments that fall into the extremely dangerous speech category have the greatest potential to generate situations of violence by providing a clear plan of action that can be well understood by the target audience. As we can see, the different classifications on available hate speech combine the regularity or systematicity of the statements, their specific contents, as well as the conditions of enunciation and the contexts in which they are circulated, which helps to clearly determine their capacity to hurt.

Following these premises and precedents, it can be said that hate speeches, in its plural form, constitute a kind of generic discourse, in the style of political, legal or academic discourse, composed of other specific discourses that can be classified based on the type of damage that they generate. We will say “hate speech” (in the singular form) when the discourse is articulated in an expository manner of incitement to commit acts of violence that threaten the life and safety of a person or group of people. “Discriminatory discourses”, which claim that a person, or group of people, be excluded, segregated or unable to exercise their rights are another specific type of this genre of discourse. This type of hate speech will not threaten the life or physical integrity of people, but will threaten their “civic dignity”, that is, the right to freely exercise their rights, their citizenship. As Waldron points out,¹⁸ hate speech transmits defamatory messages that say: “Do not be fooled into thinking that you are welcome here. You are not wanted, and you and your families will be rejected, excluded, beaten and expelled, as long as we can get away with it.” A third type of hate speech would be the “harassing speech” as a systematic discursive practice carried out with the intention of preventing or limiting the use of expression in a public space. This type of discourse seeks to damage the right to freedom of expression through harassment or intimidation. This sort of practice is very common in social networks and has been analyzed in Argentina and other parts of the world, by civil society organizations and academic institutions.¹⁹ According to the report “*Contrarrestar el discurso de odio en línea*” [Counteracting hate speech online], the rise and dissemination of hate speech on digital platforms “is an evolving phenomenon and collective efforts are needed to understand its importance and consequences, as well as to develop effective responses.”²⁰

¹⁸ Waldron, *op. cit.*

¹⁹ Based on a study carried out by the Argentine office of Amnesty International, this type of phenomenon was analyzed and it was concluded that “critical information, as well as the defense of human rights, have been recently exposed to attacks on digital social networks that in many cases are coordinated and seek to inhibit the expression of pluralistic perspectives and limit the circulation of diverse opinions on fundamental issues of public space”. The Amnesty International study “*El debate público limitado. Trolling y agresiones a la libre expresión de periodistas y defensores de DD.HH. en Twitter Argentina*” [Limited public debate. Trolling and attacks on the free expression of journalists and human rights defenders on Twitter Argentina], 2018, retrieved from: <https://amnistia.org.ar/wp-content/uploads/delightful-downloads/2018/03/online-pre1.pdf>, last access: April 10, 2019. The study “Trolls Just Want to Have Fun” carried out by Erin Buckels, Paul Trapnell and Delroy Paulhus presents a complete bibliographical review on the subject and is available at: http://www2.psych.ubc.ca/~dpaulhus/research/DARK_TRIAD/ARTICLES/PAID.2014.with.Buckels-Trapnell.pdf, last access: April 10, 2019.

²⁰ Gagliardone *et al.*, *op. cit.*

III. Hate speech in the international legal framework

The way in which hate speech has been discussed in different legal frameworks in force shows how the tension between freedom of expression and the promotion of equality has been resolved with different perspectives, from the right to express all kinds of ideas in the public debate to the need to guarantee all citizens the possibility of exercising their rights on an equal footing. To analyze the state of the issue of hate speech in existing legal frameworks, be it American, European or international, requires analyzing both the normative body linked to freedom of expression and those instruments that address issues related to the fight against racism and discrimination, which are generally in tension with one another so they must be harmonized in an integral and systemic way to guarantee the full validity of international human rights treaties.

1. Hate speech in the freedom of expression legal framework

The International Covenant on Civil and Political Rights (ICCPR) offers a specific response to the tension between freedom of expression and equality. The ICCPR gives freedom of expression a place of relevance and establishes restrictions in specific situations. Article 19 states that the exercise of freedom of expression “entails special duties and responsibilities. Therefore, it may be subject to certain restrictions, which must, however, be expressly established by law and which are necessary: a) to contribute to securing respect for the rights or reputations of others; b) to ensure the protection of national security, public order or public health or morals”.²¹ On the other hand Article 20 provides that “any defense or praise of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law.” As can be seen, in the international arena, the ICCPR does not establish incitement to violence as an exclusive requirement to exclude a particular discourse from the protection established in Article 19. For the ICCPR, social hate discourses that incite discrimination or hostility should also be prohibited by law.

The European Convention on Human Rights (ECHR) presents a broad recognition of freedom of expression in its Article 10: “Everyone has the right to freedom of expression. This right includes the freedom of opinion and the freedom to receive or communicate information or ideas without interference from public authorities and regardless of frontiers”. Similar to Article 19 of the ICCPR, the ECHR establishes that:

The exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic society, in the interests of national security, territorial integrity or public safety, for the prevention of disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, for preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary.²²

²¹ United Nations High Commissioner for Human Rights, “Pacto Internacional de Derechos Civiles y Políticos” [International Covenant on Civil and Political Rights], 1966, retrieved from: <https://www.ohchr.org/sp/professionalinterest/pages/ccpr.aspx>, last access: April 8, 2019.

²² European Commission of Human Rights, Convenio Europeo de Derechos Humanos [European Convention on Human Rights], 1950, retrieved from: https://www.echr.coe.int/Documents/Convention_SPA.pdf, last access: April 8, 2019.

The ECHR did not include any reference to the treatment of hate speech, which has been compensated by the adoption of various documents²³ and by the relevance given to the cases analyzed by the European Court of Human Rights:

The European Court of Human Rights' jurisprudence has analyzed extensively the subject of hate speech based on the intersection of Article 10 of the European Convention and domestic laws banning these forms of incitement.

In these decisions, the Court has utilized the standards of Article 10(2) to determine when restrictions on freedom of expression are justified: an interference with freedom of expression violates Article 10 unless it is "prescribed by law," is designed to carry out at least one of the aims laid out in Article 10(2) and is "necessary in a democratic society."²⁴

In turn, the Inter-American System has been described as one of the most protective legal frameworks for freedom of expression. This broad recognition is reflected in the wording of Article 13 of the American Convention on Human Rights (ACHR), which guarantees everyone the right to "seek, receive and impart information and ideas of all kinds,"²⁵ which will have substantial practical consequences on how issues related to any potential restriction on the dissemination of speech will be analyzed.

Article 13.2 of the ACHR prohibits prior censorship and, on the other hand, authorizes a system of subsequent liability in specifically delimited situations. The application of this system of subsequent liability must always comply with three conditions established by the ACHR:

(a) the limitations must be established by clearly and precisely written laws; (b) the limitations must be aimed *at achieving the imperative objectives authorized by the American Convention; and (c) the limitations must be necessary in a democratic society for the achievement of the objective they pursue, strictly proportional to the purpose they seek, and ideal to achieve said objective.*

²³ In addition to reading the provisions of the ECHR there is a series of documents that contribute to the development of a relevant legal framework to understand the way in which cases linked to hate speech are dealt with at a European level. Perhaps the absence of an explicit mention on the subject is what has prompted the adoption of different documents, among which are the Recommendations and Resolutions of the Parliamentary Assembly of the Council of Europe; the Council Framework decision 2008/913 / JHA of 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law; Directive 2010/13 / EU of the European Parliament and of the Council of 2010 on the coordination of certain provisions foreseen in legislation, regulation or administrative actions in the member states on the provision of audiovisual communication services; the Additional Protocol to the Convention on Cybercrime, which requires the member states to adopt measures to criminalize the dissemination of racist and xenophobic material through computerized means and the use of computer systems to threaten or insult for racist or xenophobic reasons and the denial, manifest minimization, approval or justification of any genocide or crimes against humanity; the European Convention on Transfrontier Television that requires that programs cannot incite racial hatred and the Council of Europe Convention on Preventing and Combating Violence against Women and Domestic Violence that refers to forms of violence against women that can also be manifestations of sexist hate speech both online and offline (European Commission against Racism and Intolerance, "General Policy Recommendation No. 15 on combating hate speech and explanatory memorandum", 2016, retrieved from: http://www.mitramiss.gob.es/oberaxe/files/documents/2016_12_21-Recomendation_ECRI_NO_15_Discurso_odi-ES.pdf, last access: April 8, 2019).

²⁴ RFOE-OAS, "Informe Anual de la Relatoría Especial para la Libertad de Expresión" [Annual Report of the Office of the Special Rapporteur for Freedom of Expression], 2004, retrieved from: <http://www.oas.org/es/cidh/expresion/docs/informes/anuales/Informe%20Anual%202004.pdf>, last access: April 8, 2019.

²⁵ Organization of American States (OAS), Convención Americana sobre Derechos Humanos [American Convention on Human Rights], 1969, retrieved from: https://www.oas.org/dil/esp/tratados_b-32_convencion_americana_sobre_derechos_humanos.htm, last access: April 8, 2019.

While 13.2 establishes a system of subsequent liabilities to guarantee respect for the rights or reputation of others, or the protection of national security, public order or public health or morals, Article 13.5 establishes that:

Any propaganda in favor of war and any defense of national, racial or religious hatred that constitutes incitement to violence or any other similar illegal action against any person or group of persons, for any reason, including those of race, color, religion, language or national origin is prohibited.²⁶

We understand that Article 13.5 indicates that the ACHR contemplates the offense known as *apología* [to defend or praise criminal acts] of hate as a phenomenon comprising different situations or levels of discourse among which those expressions that constitute incitement to violence or any other illegal action must be excluded from the protection. In this matter, it should be noted that an expression that fits within what is called hate speech but which does not incite violence can be protected under the terms of Article 13 but subject to subsequent liability: “Article 13(2) of the American Convention considers other intolerant expressions or comments that do not strictly amount to “incitement to violence” could be subject to the imposition of subsequent liability to ensure the rights to dignity and non-discrimination of a particular group in society.”²⁷

Returning to Article 13.5 and the ACHR’s instruction of prohibiting hate speech that constitutes incitement to violence, and given the differences between the ACHR’s English and Spanish versions, a controversy has arisen in the framework of the Inter-American System over whether the expressions of hatred should be considered criminally punishable offenses or if they should be prohibited by law, which means allowing censorship measures. Although there is no specific answer on the subject, some specialists maintain that neither the advisory opinions nor the judgments of the Inter-American Court of Human Rights (the I/A Court H.R.) have admitted prior restrictions to any type of speech “even when the text reads ‘will be prohibited’”.²⁸

Thus, the ACHR — interpreted exclusively from the perspective of the protection of freedom of expression — classifies and organizes the protection and promotion of human rights in the Inter-American system and provides an answer to the tension between freedom of expression and the fight against discrimination and racism, insofar as, in principle, the system only accepts the restriction of the circulation of hate speech in specific cases where the incitement to violence can be proven irrefutably.

It is important to mention that the expression “hate speech” used explicitly to refer to expressions that incite hatred or violence against people or groups of people is not present in the ACHR — nor in the International Covenant on Civil and Political Rights (ICCPR), as we saw before. Thus, while the ACHR does not provide a precise definition for the term, and while the I/A Court H.R. has not yet issued an opinion on the matter, the reports of the Special Rapporteurs on Freedom of Expression of the years 2004 and 2015 have taken on that subject, providing valuable approximations for the understanding of the term.²⁹

²⁶ OAS, 1969, *op. cit.*

²⁷ RFOE-OAS, “Informe Anual de la Relatoría Especial para la Libertad de Expresión” [Annual Report of the Office of the Special Rapporteur for Freedom of Expression], 2015, retrieved from: <http://www.oas.org/es/cidh/expresion/docs/informes/anuales/InformeAnual2015RELE.pdf>, last access: April 8, 2019.

²⁸ Loreti, Damián, *Tensiones entre libertad de expresión y protección contra la discriminación: la incidencia de las regulaciones sobre censura previa y el debate sobre el rol del Estado* [Tensions between freedom of expression and protection against discrimination: the impact of regulations on prior censorship and the debate on the role of the State], Red Universitaria sobre Derechos Humanos y Democratización para América Latina, Buenos Aires, July, 2012, retrieved from: http://www.unsam.edu.ar/ciep/wp-content/uploads/pdf/damian_loreti.pdf, last access: April 8, 2019.

²⁹ RFOE-OAS, 2004, *op. cit.*; RFOE-OAS, 2015, *op. cit.*

Therefore, the Inter-American System is much more protective of freedom of expression than the other systems, although they all share the importance of complying with three-part procedures in cases where it is appropriate to apply regimes of subsequent liability. Along these lines, a recent example would be the Joint Declaration on Freedom of Expression and “Fake News”, Disinformation and Propaganda which proclaims that “States may only impose restrictions on the right to freedom of expression in accordance with the test for such restrictions under international law, namely that they be provided for by law, serve one of the legitimate interests recognized under international law, and be necessary and proportionate to protect that interest.” The same document stipulates that “Restrictions on freedom of expression may also be imposed, as long as they are consistent with the requirements noted in paragraph 1(a), to prohibit advocacy of hatred on protected grounds that constitutes incitement to violence, discrimination or hostility (in accordance with Article 20(2) of the International Covenant on Civil and Political Rights)”.³⁰

2. Hate speech in the legal framework of combating racism, discrimination and intolerance

However, social hate speech has not been covered only from the perspective of the defense and promotion of freedom of expression. In turn, the fight against discrimination and racism has built a body of standards that seeks to respond to the phenomenon of hate speech as a way to guarantee the equal exercise of rights.

On December 21, 1965, the General Assembly of the United Nations adopted the International Convention on the Elimination of All Forms of Racial Discrimination (ICERD). The Convention is based on the 1963 Declaration on the Elimination of All Forms of Racial Discrimination and presents a strong stance against any doctrine of racial differentiation or superiority, which it classifies as “scientifically false, morally condemnable, socially unjust and dangerous and it has no theoretical or practical justification.”

Understanding that racial discrimination harms not only those who are its victims, but also those who practice it, Article 4 of the Convention calls on states to condemn all propaganda and all organizations based on ideas of racial superiority, to act in the elimination of all incitement to discrimination, to prohibit the dissemination of ideas based on racial superiority and acts of violence or incitement to violence against any race. Thus, with the objective of preventing racial hatred, the Convention establishes a greater margin for restrictions on freedom of expression.

Hence, article 4 provides that:

States Parties condemn all propaganda and all organizations which are based on ideas or theories of superiority of one race or group of persons of one colour or ethnic origin, or which attempt to justify or promote racial hatred and discrimination in any form, and undertake to adopt immediate and positive measures designed to eradicate all incitement to, or acts of, such discrimination and, to this end, with due regard to the principles embodied in the Universal Declaration of Human Rights and the rights expressly set forth in article 5 of this Convention, inter alia: (a) Shall declare an offence punishable by law all dissemination of ideas based on racial superiority or hatred, incitement to racial discrimination, as well as all acts of violence or incitement to such acts against any race or group of persons of another colour or ethnic origin, and also the provision of any assistance to racist activities, including the

³⁰ RFOE-UN, RFOE-OSCE, RFOE-OAS and ACHPR, Declaración Conjunta sobre Libertad de Expresión y “Noticias Falsas” (“Fake News”), Desinformación y Propaganda [Joint Declaration on Freedom of Expression and “Fake News”, Disinformation and Propaganda], 2017, retrieved from: <http://www.oas.org/es/cidh/expresion/showarticle.asp?artID=1056&ID=2>, last access: April 8, 2019.

financing thereof; (b) Shall declare illegal and prohibit organizations, and also organized and all other propaganda activities, which promote and incite racial discrimination, and shall recognize participation in such organizations or activities as an offence punishable by law; (c) Shall not permit public authorities or public institutions, national or local, to promote or incite racial discrimination.³¹

Another outstanding event in relation to the fight against racism is the World Conference against Racism held in Durban, South Africa, in 2001. This conference produced the Durban Declaration and its Programme of Action, documents that address a wide range of problems and propose concrete policies and global efforts to fight against racism, racial discrimination and xenophobia. The Durban Review Conference, held in 2009 in Geneva, assessed the progress in implementing the measures of the Durban Declaration and Programme of Action. During that conference, the terms analyzed by the ICCPR were reinforced and it was resolved to “fully and effectively prohibit any advocacy of national, racial, or religious hatred that constitutes incitement to discrimination, hostility or violence and implement it through all necessary legislative, policy and judicial measures.”³² These examples show that restrictions on hate speech in a broad sense are ratified at an international level, although the role of freedom of expression is highlighted in the fight against racism, racial discrimination, xenophobia and related forms of intolerance.³³

Another international document relevant to the conceptualization of hate speech from the perspective of discrimination is the Convention on the Elimination of all Forms of Discrimination against Women adopted by the General Assembly of the United Nations on December 18, 1979, coming into force on September 3, 1981. The Convention sets obligations on states to condemn and prohibit any act of discrimination against women. While the Convention does not make specific reference to the treatment to be given to hate speech in consideration to freedom of expression, it does reference this right regarding those expressions which are discriminatory towards women.

The Special Rapporteur on Contemporary Forms of Racism, Racial Discrimination, Xenophobia and Related Intolerance, whose report³⁴ exhorted states to take advantage of all opportunities, including those on the Internet, to combat racism and promote the values of equality, non-discrimination, diversity and democracy, while respecting the obligations laid down by the ICCPR. The Rapporteur urged states to strengthen freedom of expression, “which contributes in an essential way to promoting democracy and combating racist and xenophobic ideologies based on racial superiority”. The report prepared by the Special Rapporteur on questions of minorities is of great relevance, it analyzes the impact of the circulation of hate speech against minorities in the media and its author has been a permanent reference in this text.³⁵

At this point, we can see that at an international level, the tension between the protection of freedom of expression and the fight against racism and discrimination finds its reference points in the ICCPR, on the one hand,

³¹ UN, Convención Internacional sobre la Eliminación de Todas las Formas de Discriminación Racial [International Convention on the Elimination of All Forms of Racial Discrimination], 1965, retrieved from: <https://www.ohchr.org/SP/ProfessionalInterest/Pages/CERD.aspx>, last access: April 10, 2019.

³² UN, “Informe de la Conferencia de Examen de Durban” [Outcome document of the Durban Review Conference], 2009, retrieved from: <http://www.un.org/es/durbanreview2009/ddpa.shtml>, last access: April 8, 2019.

³³ *Ibid.*

³⁴ UN, “Formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia” [Contemporary Forms of Racism, Racial Discrimination, Xenophobia and Related Intolerance], 2016, retrieved from: <https://www.acnur.org/fileadmin/Documentos/BDL/2016/10808.pdf>, last access: April 10, 2019.

³⁵ UN, “Informe de la Relatora Especial sobre cuestiones de las minorías, Rita Izsák” [Report of the Special Rapporteur on minority issues, Rita Izsák], 2015, retrieved from: <https://www.acnur.org/fileadmin/Documentos/BDL/2015/10012.pdf>, last access: April 10, 2019.

and the conventions against racial and ethnic discrimination and discrimination against women, on the other.

The UN's Committee on the Elimination of Racial Discrimination (CERD) presented in 2013 Recommendation No. 35 on combating racist hate speech, acknowledging the need to coordinate efforts and with the intention of providing a comprehensive and systemic perspective in defense of freedom of expression and the fight against racism. In that document, the Committee states that:

In gauging the scope of freedom of expression, it should be recalled that the right is integrated into the Convention and is not simply articulated outside it: the principles of the Convention contribute to a fuller understanding of the parameters of the right in contemporary international human rights law. The Committee has integrated this right to freedom of expression into its work on combating hate speech, commenting where appropriate on its lack of effective implementation and, where necessary, drawing upon its elaboration in other entities created by human rights treaties. (...) The Committee recommends that the criminalization of forms of racist expression should be reserved for serious cases, to be proven beyond reasonable doubt, while less serious cases should be addressed by means other than criminal law, taking into account, *inter alia*, the nature and extent of the impact on targeted persons and groups.³⁶

Recommendation No. 35 is particularly relevant for coordinating these issues when placing freedom of expression as a fundamental tool for the defense of human dignity:

In addition to underpinning and safeguarding the exercise of other rights and freedoms, freedom of opinion and expression has particular salience in the context of the Convention. The protection of persons from racist hate speech is not simply one of opposition between the right to freedom of expression and its restriction for the benefit of protected groups; the persons and groups entitled to the protection of the Convention also enjoy the right to freedom of expression and freedom from racial discrimination in the exercise of that right. Racist hate speech potentially silences the free speech of its victims. (...) Freedom of expression, indispensable for the articulation of human rights and the dissemination of knowledge regarding the state of enjoyment of civil, political, economic, social and cultural rights, assists vulnerable groups in redressing the balance of power among the components of society, promotes intercultural understanding and tolerance, assists in the deconstruction of racial stereotypes, facilitates the free exchange of ideas, and offers alternative views and counterpoints. States parties should adopt policies empowering all groups within the purview of the Convention to exercise their right to freedom of expression.³⁷

In Europe, the European Commission against Racism and Intolerance (ECRI) issued in 2016 General Policy Recommendation No. 15 on combating hate speech and its explanatory memorandum. In this document, it is proposed that:

The starting point for the Recommendation is the recognition of the fundamental importance of freedom of expression, tolerance and respect for equal dignity, all of which are guaranteed under nu-

³⁶ CERD, 2013, *op. cit.*

³⁷ CERD, 2013, *op. cit.*

merous international instruments accepted by member States of the Council of Europe. ECRI is aware, in particular, that any efforts to tackle hate speech should never exceed the limitations to which freedom of expression, as a qualified right, can legitimately be subjected. It is also aware that in some cases hate speech can be effectively responded to without restricting freedom of expression. For this reason, the Recommendation has a graduated approach to the measures that need to be undertaken. In particular, the view that the use of criminal sanctions should not be the primary focus of action against the use of hate speech reflects not only the importance of respecting the rights to freedom of expression and association but also an appreciation that addressing the conditions conducive to the use of hate speech and vigorously countering such use are much more likely to prove effective in ultimately eradicating it.³⁸

In the Americas, the treatment of issues related to combating discrimination and intolerance appears in its founding texts: both the Charter of the Organization of American States and the American Declaration of The Rights and Duties of Man and the American Convention on Human Rights highlight the importance of the prohibition of discrimination on the basis of race, color, sex, language, religion, political opinions or of any other nature, national or social origin, economic position, birth or any other social condition. In addition to these founding texts, there have been resolutions issued by the General Assembly since 1994 to promote the adoption of effective measures to promote tolerance and the eradication of racist and discriminatory practices in the Inter-American System and to urge states to draft and implement policies to prevent and avoid all forms of racism, discrimination, xenophobia and intolerance.

On June 5, 2013, the General Assembly of the Organization of American States at its forty-third regular session of the General Assembly adopted the Inter-American Convention against Racism, Racial Discrimination and Related forms of Intolerance — which has not yet entered into force, and the Inter-American Convention against All Forms of Discrimination and Intolerance, in force since 2017.³⁹

The Inter-American Convention against All Forms of Discrimination and Intolerance states in Article 4: “The states undertake to prevent, eliminate, prohibit, and punish, in accordance with their constitutional norms and the provisions of this Convention, all acts and manifestations of discrimination and intolerance.” These acts include “Publication, circulation or dissemination, by any form and/or means of communication, including the Internet, of any materials that: a. advocate, promote, or incite hatred, discrimination, and intolerance; b. condone, justify, or defend acts that constitute or have constituted genocide or crimes against humanity as defined in international law, or promote or incite the commission of such acts.” In similar terms, the Inter-American Convention against Racism, Racial Discrimination and Related forms of Intolerance establishes similar restrictions in Article 4 to prevent, eliminate, prohibit and punish all acts and manifestations of racism, racial discrimination and related intolerance.

Another source in the fight against discrimination is the Inter-American Convention on the Prevention, Punishment

³⁸ European Commission against Racism and Intolerance, op. cit.

³⁹ At the time of writing this document, the Inter-American Convention against All Forms of Discrimination and Intolerance had been signed by Argentina, Barbados, Bolivia, Brazil, Chile, Colombia, Ecuador, Haiti, Mexico, Panama, Peru and Uruguay and only Uruguay had ratified it, so the Convention has not yet entered into force. In turn, the Inter-American Convention against Racism, Racial Discrimination and Related forms of Intolerance had been signed by Antigua and Barbuda, Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Ecuador, Haiti, Panama, Peru and Uruguay and Uruguay, Costa Rica and Antigua and Barbuda had ratified it, for which the Convention entered into force on November 11, 2017.

and Eradication of Violence against Women, adopted on September 6, 1994 by the General Assembly of the OAS and in force since 1995. The Convention of Belem Do Para is a comprehensive document for the fight against violence against women, which it defines as “any act or conduct, based on gender, which causes death or physical, sexual or psychological harm or suffering to women, whether in the public or the private sphere”. In similar terms to those previously analyzed the Convention requests states to adopt policies aimed at preventing, punishing and eradicating such violence.

As can be seen in the Inter-American system, the protection of freedom of expression established in the American Convention on Human Rights seems to contradict the provisions of the conventions analyzed. In the first place, the conventions establish a treatment of hate speech that covers a range of actions: prevention, eradication, elimination, prohibition and punishment. Second, by not distinguishing between the types of situations on which each of these measures would apply, these could be in tension with the provisions of Article 13.2 of the ACHR. Due to how new these instruments are, it is understandable that there still has not been an action to coordinate them that would allow us to advance in the construction of instruments and resources that facilitate hermeneutical actions in the treatment of hate speech, but it is clearly a task that becomes necessary for the future.

IV. Policies regarding hate speech

Prohibition should not be the first or the only measure to combat the use of hate speech. Punitive measures on the circulation of discourses affect freedom of expression and association and can be ineffective in identifying and understanding the conditions that favor them, a fundamental aspect to vigorously combat their use.⁴⁰ We consider it essential to address the wide range of phenomena that are included under the categorization of hate speech, in order to advance in the design of actions without reducing the role of politics to prohibition. If we understand that hate speech is a social construction that threatens the lives of people, undermines public dignity or seeks self-censorship and silence, and it marginalizes public debate to specific sectors, the mere prohibition of circulation will probably not reach the objective to counteract or eradicate it, but rather the opposite would happen.

A significant contribution along these lines is the one provided by the principles of Camden, where it is stated that:

States should engage in comprehensive efforts to combat negative stereotypes of, and discrimination against, individuals and groups and to promote intercultural understanding and evaluation, including providing teacher training on human rights values and principles and introducing or strengthening intercultural understanding as a part of the school curriculum for pupils of all ages.⁴¹

In this regard, the Commission and the Office of the Special Rapporteur for Freedom of Expression of the OAS, in their 2015 report, insist “to effectively combat hate speech, a comprehensive and sustained approach that goes beyond legal measures and includes preventive and educational mechanisms should be adopted”. These types of measures point to the cultural roots of hate speech and, as such, can be valuable instruments to identify and refute hate speech, encouraging the development of a society based on the principles of diversity, pluralism and tolerance. It would be advantageous for the state to carry out positive policies that allow the deconstruction of stereotypes,

⁴⁰ European Commission against Racism and Intolerance, *op. cit.*

⁴¹ Article 19, *op. cit.*

counteracting discriminatory discourses, promoting pluralism and guaranteeing everyone access to public debate.

In this context, preventive mechanisms could include: education to promote understanding and to combat negative stereotypes and discrimination, including programs aimed at children of school age, as well as information campaigns; training law enforcement agents and justice operators regarding the prohibition of incitement to violence; etc. To develop these types of policies it is necessary to improve the knowledge on hate speech as social discourses, as well as the collection of data on this type of statements. This position coincides with the one adopted by the Special Rapporteur of the United Nations who, in their special report on the subject in 2012, proposes the creation of a series of areas where progress must be made to prevent and counteract the circulation of hate speech: education and awareness, social debate and dialogue, conducting research to add information to the public debate and the role of the media.⁴²

The UN Special Rapporteur on Minority Issues agrees with this need by pointing out that the same attention and interest should be given to social or non-legal responses as to those strictly from a legal perspective. His 2015 report reviews some community experiences that have creatively sought to counteract hate speech.⁴³

The UNESCO report of 2015 compiles a wide range of actions to counter hate speech. The UN agency for education and culture reports on the efforts undertaken in the design of early warning systems and methods to distinguish between different types of discourse,⁴⁴ reviews the actions implemented by civil society to face the problem,⁴⁵ presents actions that seek responses to online discourse and describes media literacy initiatives and awareness for citizen empowerment in the interpretation and reaction to hate speech.⁴⁶

In relation to judicial remedies and civil sanctions — monetary and non-monetary damages, rectification and reply — and administrative measures that could be implemented to combat hate speech, the following considerations should be taken into account:

Expressions that openly denigrate, stigmatize or discriminate against persons or groups of persons based on their sexual orientation or identity of actual or perceived gender, not reaching the threshold of hate speech that constitutes incitement to illegal violence in accordance with Article 13.5 of the American Convention (hate speech) may be subject to the imposition of subsequent sanctions of a civil or administrative nature, or to remedies such as the right to rectification and reply. Nonetheless, the IACtHR and the Office of the Special Rapporteur for Freedom of Expression emphasize that sanctions cannot be aimed at inhibiting or restricting the dissemination of ideas or information on matters of public interest.⁴⁷

⁴² Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression of the UN, "Informe Anual" [Annual Report], 2012, retrieved from: <https://www.palermo.edu/cele/pdf/SRs-Informe-Espa.pdf>, last access: April 8, 2019.

⁴³ UN, 2015, *op. cit.* The report highlights the Studio Ijambo initiative to combat inter-ethnic violence in Burundi. Norikoe Net is a group settled in Japan that is trying to end hate speech and racism, often against Koreans who are in the country, or the actions taken by the International Network against CyberHate (INACH), among others.

⁴⁴ The research initiatives highlighted by the report include the UMATI Project and the project "Online Comment Moderation: Emerging Best Practices", produced by the World Association of Newspaper.

⁴⁵ The research activities highlighted by the report include Panzagar, a campaign developed by activist Nay Phone Latt that openly opposes hate speech, actions undertaken by the Save Darfur Coalition, Invisible Children with the Kony 2012, the Anti-Defamation League (ADL), Women, Action and the Media (WAM!), The Australian-based Online and Hate Prevention Institute, the Canada-based Sentinel Groups for Genocide Prevention, the British-based Tell Mama-Measuring Anti-Muslim Attacks, HateBase, Tell Mama's Islamophobic and Online Hate Prevention Institute's Fight Against Hate.

⁴⁶ Gagliardone et al., *op. cit.* The report highlights the actions undertaken by Global Citizenship Education (GCED), one of the strategic work areas of UNESCO's education program.

⁴⁷ RFOE-OAS, 2015, *op. cit.*

The Special Rapporteur for Freedom of Expression of the OAS highlighted the non-punitive resources implemented and adopted by the Defensoría del Pùblico de Servicios de Comunicación Audiovisual de Argentina [Office of the Public Defender of Audiovisual Communication Services of Argentina].⁴⁸ Similarly, Damián Loreti⁴⁹ emphasized the creation and operation of the Observatorio de la Discriminación en Radio y Televisión en Argentina [Observatory of Discrimination in Radio and Television in Argentina] in 2006:

The entity analyzes the objected content and defines whether or not the material reviewed contains a discriminatory message. Its intervention is not intended to impose sanctions — it does not even have the power to do so — but is limited to communicating the result of the evaluation to the parties involved in the production and distribution of the objected content, and then publicize the analysis.

We consider it vitally important to follow the recommendations of various international entities that call for the implementation of policies that allow us to repair the general ignorance over the magnitude of the circulation of hate speech, the conditions caused by its emergence or its effects. The Reports of the European Commission against Racism and Intolerance (2016),⁵⁰ the Office of the Special Rapporteur for Freedom of Expression of the OAS (2015),⁵¹ the Special Rapporteur on Freedom of Expression of the United Nations (2012)⁵² and the UN Special Rapporteur on Minority Issues (2015)⁵³ have expressed coincidentally the need to develop studies to advance the design of preventive policies that collect and analyze data on these phenomena and thus strengthen the processes of decision making, the design, development and implementation of policies to better protect at-risk population groups.

V. Conclusions and recommendations

Conceptualizing hate speech, identifying its effects and the most appropriate actions to counter it requires dealing with a type of expression that generally confronts fundamental rights such as freedom of expression and the right to equality. We strove to provide an understanding of instances of hate speech as social discourses, that is to say, as systematic, structured and structuring social practices of subjectivity and not as isolated or fleeting manifestations. According to the points argued in this document, we consider that the systematic social practices of circulation, enabling and legitimization construct the necessary doxical precondition to carry out acts of harassment, discrimination, denial of rights and, in its most extreme cases, violence.

Hate speech contains a doctrine that articulates communal memories with a conception of the world and a series of rules of conduct. This doctrinal unity rejects someone else's possibilities of exercising their freedom, any diversity of conceptions about the world, and difference of opinions or dissidence in matters of sexual behavior.

⁴⁸ *Ibid.*

⁴⁹ Loreti, *op. cit.*

⁵⁰ European Commission against Racism and Intolerance, *op. cit.*

⁵¹ RFOE-OAS, 2015, *op. cit.*

⁵² Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression of the UN, "Informe Anual" [Annual Report], 2012, *op. cit.*

⁵³ Special Rapporteur on Minority Issues, "Informe Anual" [Annual Report], 2015, retrieved from: <https://www.acnur.org/fileadmin/Documentos/BDL/2015/10012.pdf>, last access: April 8, 2019.

Hate speeches are attacks directed at people or groups of people whose way of understanding and inhabiting the world is seen as threatening an idealized (pre) existing social order. The concept of hate speeches (in its plural form) refers to a discourse genre composed of different types of violent speeches or aggressive toward social groups, in an attempt to segregate them. This discursive genre includes: a) “hate speech” (in its singular form), which threatens the lives of a person or group of people, and coincides with messages of incitement to hatred or violence; b) “discriminatory speeches” that threaten the dignity of a person, or group of people, and seek to segregate, discriminate or prevent the exercise of rights under equal conditions; c) “harassing speech” that seeks to limit freedom of expression of a person or group of persons, hampering their participation in public life through harassment or intimidation.

The distinction between different types of hate speech allows us to reevaluate the role of the state and expand the range of policies available to counteract hate speech. A comprehensive policy against the unifying pretense of hate speech should include, first of all, a strategy of research and data collection that increases knowledge and aids in understanding the phenomenon and contributing to the development of public policies based on the evidence; and, secondly, it must aim at the design and implementation of positive policies that allow the deconstruction of stereotypes, counteracting discriminatory discourses, promoting pluralism and guaranteeing everyone access to the public debate.

Regarding the first point, we have expressed, together with the recommendations of rapporteurs and other multilateral bodies, the need to advance in the collection and analysis of data on the circulation of hate social discourses. This document aims to support this task. We believe that increasing knowledge on the subject is particularly relevant in the region, because there have not been any experiences of the kind. The studies that took place in Italy, the United States and the UMATI Project have been successful in providing data on the specific reality of their societies. Latin America needs a similar experience to account for its peculiarities and discriminatory practices.

Regarding the second point, we consider that the conception and implementation of public policies to counteract hate speech require the constant and permanent consideration that between freedom of expression and equality there is an affirmative and complementary relationship that must be developed and that the policy cannot be reduced to punitive measures. On the contrary, this document argues in favor of the need to increase the understanding of the phenomenon as systematic practices that require a political response, a positioning that creates responses based on the distinction between types of social hate discourses and depending on the evidence. Hence, the work done by the Committee on the Elimination of Racial Discrimination of the United Nations, the contribution of the Rabat Plan of Action, and the influence of the European Commission against Racism and Intolerance can be of interest to enrich the pending discussions in our continent, where the extensive protection of freedom of expression established in The American Convention on Human Rights seems to contradict the provisions of the instruments against discrimination and racism.

We are convinced that the solutions must always be aligned with the problems to be addressed. Hate speech threatens the effective exercise of human rights, it undermines the dignity of all those who are part of a community. If hate speech is basically a type of authoritarian discourse that seeks to impose a unique way to interpret reality, if we understand it as practices articulated from historically constituted fears and socially sustained prejudices, the response should consist of a comprehensive policy that creates interventions based on the understanding of the problem and applies measures to ensure pluralism, protecting vulnerable groups and nurturing public debate with multiple voices.

Bibliography

Office of the United Nations High Commissioner for Human Rights, "Pacto Internacional de Derechos Civiles y Políticos" [International Covenant on Civil and Political Rights], 1966, retrieved from: <https://bit.ly/2uy0V0m>, last access: April 8, 2019.

United Nations High Commissioner for Human Rights, "Convención para la eliminación de la discriminación contra la mujer" [Convention on the Elimination of All Forms of Discrimination against Women], 1979, retrieved from: <https://bit.ly/2uysyqx>, last access: April 8, 2019.

Office of the United Nations High Commissioner for Human Rights, "Plan de acción de Rabat" [Rabat Plan of Action], 2013a, retrieved from: <https://bit.ly/2K2JKgq>, last access: April 8, 2019.

Office of the United Nations High Commissioner for Human Rights, "Informe Anual" [Annual Report], 2013, retrieved from: <https://bit.ly/2K2JKgq>, last access: April 8, 2019.

Amnesty International, "El debate público limitado. Trolling y agresiones a la libre expresión de periodistas y defensores de DD.HH. en Twitter Argentina" [Limited public debate. Trolling and attacks on the free expression of journalists and human rights defenders on Twitter Argentina], March, 2018, retrieved from: <https://bit.ly/2u17GKR>, last access: April 10, 2019.

Angenot, Marc, *El discurso social: los límites históricos de lo pensable y lo decible* [Social discourse: the historical limits of what is thinkable and what can be said], Buenos Aires, Siglo XXI, 2012.

Arendt, Hannah, *Los orígenes del totalitarismo* [The Origins of Totalitarianism], Madrid, Alianza, 2006.

Article 19, "Los principios de Camden sobre libertad de expresión e igualdad" [The Camden Principles on Freedom of Expression and Equality], 2009, retrieved from: <https://bit.ly/2EBMrCb>, last access: April 8, 2019.

Association of Southeast Asian Nations, ASEAN Human Rights Declaration, 2012, retrieved from: <https://bit.ly/2I3vH7J>, last access: April 8, 2019.

Benesch, Susan, "Vile Crime or Inalienable Right: Defining Incitement to Genocide", in: *Virginia Journal of International Law*, Vol. 48, No. 3, April, 2008, retrieved from: <https://bit.ly/2VSExPa>, last access: April 8, 2019.

Benesch, Susan, Buerguer, Cathy, Glavinic, Tonei and Manion, Sean, "Dangerous Speech: A Practical Guide", Dangerous Speech Project, December, 2018, retrieved from: <https://bit.ly/2l1buIL>, last access: April 8, 2019.

Bertoni, Eduardo, "Estudio sobre la prohibición de la incitación al odio en las Américas" [Study on the prohibition of hate speech in the Americas], Report prepared for the workshops organized by the United Nations High Commissioner for Human Rights, 2010, retrieved from: <https://bit.ly/2woLhpa>, last access: April 8, 2019.

Bleich, Erik, "Freedom of Expression versus Racist Hate Speech: Explaining Differences Between High Court Regulations in the USA and Europe", in: *Journal of Ethnic and Migration Studies*, Vol. 40, No. 2, 2014, retrieved from: <http://dx.doi.org/10.1080/1369183X.2013.851476>, last access: April 8, 2019.

Buckels, Erin, Trapnell, Paul and Paulhus, Delroy, "Trolls Just Want to Have Fun", in: *Personality and Individual Differences*, No. 67, 2014, retrieved from: <https://bit.ly/2AKG6Bk>, last access: April 10, 2019.

Buyse, Antoine, "Words of Violence: 'Fear Speech,' or How Violent Conflict Escalation Relates to the Freedom of Expression", in: *Human Rights Quarterly*, Vol. 36, No. 4, November, 2014, retrieved from: <https://bit.ly/2HYKk->

sN, last access: April 8, 2019.

African Commission on Human and Peoples' Rights, African Charter on Human and Peoples' Rights, 2005, retrieved from: <https://bit.ly/1NC6jis>, last access: April 8, 2019.

European Commission against Racism and Intolerance "Recomendación de política general Nº 15 relativa a la lucha contra el discurso de odio y memorándum explicativo" [General Policy Recommendation No. 15 on combating hate speech and its explanatory memorandum], 2016, retrieved from: <https://bit.ly/2M9VRLh>, last access: April 8, 2019.

European Commission of Human Rights, Convención Europea de Derechos Humanos [European Convention on Human Rights], 1950, retrieved from: <https://bit.ly/2EGCKzx>, last access: April 8, 2019.

OAS Inter-American Commission of Women, Lineamientos Interamericanos por la Igualdad de Género como Bien de la Humanidad (Proyecto) [Inter-American Guidelines on Gender Equality for the Good of Humanity (Bill)], 2017, retrieved from: <https://bit.ly/2X7IA77>, last access: April 8, 2019.

UN Committee on the Elimination of Racial Discrimination (CERD), "Recomendación general Nº 35. La lucha contra el discurso de odio racista" [General Recommendation No. 35. Combatting racist hate speech], United Nations Human Rights, Office of the High Commissioner, 2013, retrieved from: <https://bit.ly/2JJ2XEm>, last access: April 8, 2019.

Council of Europe, Convenio Europeo sobre la Televisión Trasfronteriza [European Convention on Transfrontier Television], 1989, retrieved from: <https://bit.ly/2wkQLBc>, last access: April 8, 2019.

Council of Europe, "Protocolo adicional al convenio sobre ciberdelincuencia relativo a la penalización de actos de índole racista y xenófoba cometidos por medio de sistemas informáticos" [Additional Protocol to the Convention on Cybercrime, concerning the criminalization of acts of a racist and xenophobic nature committed through computer systems], 2003, retrieved from: <https://bit.ly/2EA6HDW>, last access: April 8, 2019.

Council of Europe, Convenio del Consejo de Europa sobre Prevención y Lucha contra la Violencia contra las Mujeres y contra la Violencia Doméstica [Convention on preventing and combating violence against women and domestic violence], 2019, retrieved from: <https://bit.ly/2MbGxhm>, last access: April 8, 2019.

Facebook, Normas comunitarias [Community Standards], 2019, retrieved from: <https://bit.ly/2sMEnlk>, last access: April 8, 2019.

Gramsci, Antonio, *El materialismo histórico y la filosofía de Benedetto Croce* [The historical materialism and philosophy of Benedetto Croce] Buenos Aires, Nueva Visión, 2008.

Hare, Ivan and Weinstein, James (eds.), *Extreme Speech and Democracy*, Oxford, Oxford University Press, 2011.

Herz, Michael and Molnar, Peter (eds.), *The Content and Context of Hate Speech: Rethinking Regulation and Responses*, Cambridge, Cambridge University Press, 2012.

iHub Research, *Umati Final Report*, 2013, retrieved from: <https://bit.ly/2VUiF1c>, last access: April 8, 2019.

Ipar, Ezequiel, "Instituciones desfondadas generan un enorme espacio para el autoritarismo social" [Degraded institutions generate a giant space for social authoritarianism], interview by Nicolás Tereschuk, in: *Arte y política*, 2018, retrieved from: <https://bit.ly/2VUiNxI>, last access: April 8, 2019.

Kaufman, Gustavo, *Odium dicta. Libertad de expresión y protección de grupos discriminados en internet* [Freedom of Expression and Protection of Discriminated Groups on the Internet], Mexico D.F., Consejo Nacional para

Prevenir la Discriminación, 2015.

Kaufman, Alejandro and Berrotarán, Patricia, "La construcción de la tiranía: el Libro Negro" [The construction of tyranny: the Black Book], in: Bonet, María Teresita and Ciappina, Carlos, *Representaciones, discurso y comunicación: el peronismo entre 1945-1970* [Representations, speech and communication: Peronism between 1945-1970], La Plata, Editorial de la Universidad de La Plata, 2014, pp. 23-42.

League of Arab States, Arab Charter on Human Rights, 1994, retrieved from: <https://bit.ly/2hfkjYk>, last access: April 8, 2019.

Loreti, Damián, *Tensiones entre libertad de expresión y protección contra la discriminación: la incidencia de las regulaciones sobre censura previa y el debate sobre el rol del Estado* [Tensions between freedom of expression and the protection against discrimination: the impact of regulations on prior censorship and the debate on the role of the State], Red Universitaria sobre Derechos Humanos y Democratización para América Latina, Buenos Aires, July, 2012, retrieved from: <https://bit.ly/2woAEm8>, last access: April 8, 2019.

Organization of American States (OAS), Convención Americana sobre Derechos Humanos [American Convention on Human Rights], 1969, retrieved from: <https://bit.ly/2nQn5q5>, last access: April 8, 2019.

Organization of American States (OAS), Convención Interamericana contra la Violencia hacia la Mujer [Inter-American Convention on the Prevention, Punishment and Eradication of Violence against Women], 1994, retrieved from: <http://www.oas.org/juridico/spanish/tratados/a-61.html>, last access: April 8, 2019.

Organization of American States (OAS), Convención Interamericana contra el Racismo, la Discriminación Racial y Formas Conexas de Intolerancia [Inter-American Convention Against Racism, Racial Discrimination and Related Forms of Intolerance], 2013a, retrieved from: <https://bit.ly/1HO0iRn>, last access: April 8, 2019.

Organization of American States (OAS), Convención Interamericana contra Toda Forma de Discriminación e Intolerancia [Inter-American Convention against All Forms of Discrimination and Intolerance], 2013b, retrieved from: <https://bit.ly/1HO2jNB>, last access: April 8, 2019.

United Nations (UN), Convención Internacional sobre la Eliminación de Todas las Formas de Discriminación Racial [International Convention on the Elimination of All Forms of Racial Discrimination], 1965, retrieved from: <https://bit.ly/2ULZrKK>, last access: April 10, 2019.

United Nations (UN), Declaración y Programa de Acción de Durban [The Durban Declaration and Programme of Action], 2001, retrieved from: <https://bit.ly/2YRVdUg>, last access: April 8, 2019.

United Nations (UN), "Informe de la Conferencia de Examen de Durban" [Durban Review Conference Report], 2009, retrieved from: <https://bit.ly/2W9uplm>, last access: April 8, 2019.

United Nations (UN), "Informe de la Relatora Especial sobre cuestiones de las minorías, Rita Izsák" [Report of the Special Rapporteur on minority issues, Rita Izsák], 2015, retrieved from: <https://bit.ly/2HIVbZf>, last access: April 10, 2019.

United Nations (UN), "Formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia" [Contemporary Forms of Racism, Racial Discrimination, Xenophobia and Related Intolerance], 2016, retrieved from: <https://bit.ly/2WdIP4c>, last access: April 10, 2019.

Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression of the UN, "Informe Anual" [Annual Report], 2012, retrieved from: <https://bit.ly/2WyXruk>, last access: April 8, 2019.

UN Special Rapporteur on Minority Issues, “Informe Anual” [Annual Report], 2015, retrieved from: <https://bit.ly/2HIVbZf>, last access: April 8, 2019.

Organization of the American States Office of the Special Rapporteur for Freedom of Expression (RFOE-OAS), “Informe Anual de la Relatoría Especial para la Libertad de Expresión” [Annual Report of the Office of the Special Rapporteur for Freedom of Expression], 2004, retrieved from: <https://bit.ly/1Q9XfDc>, last access: April 8, 2019.

Organization of the American States Office of the Special Rapporteur for Freedom of Expression (RFOE-OAS), “Informe Anual de la Relatoría Especial para la Libertad de Expresión” [Annual Report of the Office of the Special Rapporteur for Freedom of Expression], 2015, retrieved from: <https://bit.ly/1UV3h1B>, last access: April 8, 2019.

Organization of the American States Office of the Special Rapporteur for Freedom of Expression (RFOE-OAS), *Marco jurídico interamericano sobre el derecho a la libertad de expresión* [The Inter-American Legal Framework Regarding the Right to Freedom of Expression], 2009, retrieved from: <https://bit.ly/1on89fG>, last access: April 8, 2019.

RFOE-UN, RFOE-OSCE, RFOE-OAS and ACHPR, Declaración Conjunta sobre Libertad de Expresión y “Noticias Falsas” (“Fake News”), Desinformación y Propaganda [Joint Declaration on Freedom of Expression and “Fake News”, Disinformation and Propaganda], 2017, retrieved from: <https://bit.ly/2IP3ncc>, last access: April 8, 2019.

Twitter, Política relativa a las conductas de incitación al odio [Hateful Conduct Policy], 2019, retrieved from: <https://bit.ly/2JI7abE>, last access: April 8, 2019.

Gagliardone, Ignio, Gal, Danit, Alves, Thiago and Martínez, Gabriela, “Countering Online Hate Speech”, UNESCO, 2015, retrieved from: <https://bit.ly/2HXSnpo>, last access: April 8, 2019.

European Union, Framework Decision 2008/913/JHA on combating certain forms of expressions of racism and xenophobia by means of criminal law from November 28, 2008, retrieved from: <https://bit.ly/2K7epZP>, last access: April 8, 2019.

European Union, Directive 2010/13/EU of the European Parliament and of the Council of 10 March 2010 on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services, 2010, retrieved from: <https://bit.ly/2MbK7fX>, last access: April 8, 2019.

University of Minnesota Human Rights Library, Declaration on Human Rights in Islam, 1993, retrieved from: <https://bit.ly/2EyaQIE>, last access: April 8, 2019.

University of Minnesota Human Rights Library, African Commission on Human and People’s Rights, Declaration of Principles on Freedom of Expression in Africa, 2002, retrieved from: <https://bit.ly/30MGjAj>, last access: April 8, 2019.

Waldron, Jeremy, *The Harm in Hate Speech*, Cambridge, Harvard University Press, 2012.

Zizek, Slavoj, *Sobre la violencia: seis reflexiones marginales* [Violence: Six Sideways Reflections], Buenos Aires, Paidós, 2009.