

FREEDOM OF EXPRESSION AND HATE SPEECH: WHEN VALUES COLLIDE IN DIVIDED SOCIETIES

Bertus de Villiers*

Law School, University of Johannesburg
bertusdv@uj.ac.za

Received: 27 September 2022 | Last Revised: 24 November 2022 | Accepted: 9 December 2022

Abstract

One of the thorniest issues in law, especially concerning the boundaries of what is reasonable and proportionate, is the distinction between freedom of expression and hate speech. Striking a balance between freedom of expression and hate speech is, however, not a mere exercise in theory; it goes to the core of respect of individual rights and freedoms. To one person, uttering speech pursuant to the right to free expression is essential for a free and open democratic society; whereas another person, offended by what they perceive as hatred, can experience such speech as an attack on their identity and self-worth, causing harm, fear and anxiety that deny their individual rights to equality, identity and dignity. This paper gives a brief overview of jurisprudential developments in international law concerning speech that may fall within the category of hate speech, whereafter two prominent South African judgments by the Equality Court are discussed. Those two judgments highlight the complexities in determining when speech can be regarded as hate speech; what test is applied to ascertain whether speech constitutes hate speech; what evidence is required for a finding to be made; and the effect of a declaratory order. The two judgments discussed, the Nelson Mandela Trust and *Ors v. AfriForum and Ors (Old Flag case 2019)*

* Professor of the Law School of the University of Johannesburg (South Africa); Member of the State Administrative Tribunal of Western Australia (Australia); and Fellow of the Alexander von Humboldt Stiftung (Germany).

and the AfriForum and Economic Freedom Fighters and Ors (Kill the Boer Case 2022), attempted to determine the line that separates freedom of expression from hate speech. The judgments, perhaps not unexpectedly, have given rise to more questions than answers. The inconsistency in comparative jurisprudence reaffirms that the labelling of speech as hate speech should be reserved for the most extreme forms of speech; it should be proportionate to the speech, including who expressed it, where and when; and any declaration should only be directed at the specific incident and not restrict speech in general.

Keywords: Freedom of Expression, Hate Speech, Kill-the-Farmer, Old Flag, Proportionality, Signage.

I. INTRODUCTION

In few areas of law are the boundaries of what is reasonable and proportionate so difficult to determine as in the clash between freedom of expression and hate speech. If ever there was a case that reasonableness (as in beauty) is in the eye of the beholder, it applies to the undefinable line that tells us where freedom of expression ends and hate speech begins. The reality is that hate speech is circumstantial. There is no universal definition for it. And to the extent that there are attempts to settle on a universal statutory definition of hate speech, such efforts at a normative level differ in their practical application. The Supreme Court of India has aptly observed the difficulty of “confining the prohibition [against hate speech] to some manageable standard.”¹

Hate speech is determined by factors such as place, time, history, perception, population composition and circumstance. It is particularly in deeply divided societies with high ethno-cultural plurality and a poorly developed common identity, that one person’s freedom of expression may be perceived by another as hate speech. Hate speech is, however, not limited to inter-ethnic utterances. Hate speech can also be directed to persons in their social circumstances, for example, regarding their gender.²² The increased use of social media is giving rise

¹ Amish Devgan v. Union of India, SCC OnLine SC 994 (2020).

² Qwelane v. South African Human Rights Commission and Another, ZACC 22 (2021).

to an exponential growth in hate speech toward individuals and communities.³ The Supreme Court of Canada has described the effect of hate speech in the following way:

“Using expression that exposes the group to hatred, hate speech seeks to delegitimize group members in the eyes of the majority, reducing their social standing and acceptance within society. Hate speech, therefore, rises beyond causing distress to individual group members. It can have a societal impact. Hate speech lays the groundwork for later, broad attacks on vulnerable groups that can range from discrimination, to ostracism, segregation, deportation, violence and, in the most extreme cases, to genocide. Hate speech also impacts on a protected group’s ability to respond to the substantive ideas under debate, thereby placing a serious barrier to their full participation in our democracy.”⁴

Striking a balance between freedom of expression and hate speech is not a mere exercise in theory, it goes to the core of respect of individual rights and freedoms. To the one person, uttering speech pursuant to the right to free expression is essential for a free and open democratic society; whereas a person offended by what they perceive as hatred can feel such speech tearing into their identity and self-worth, causing harm, fear and anxiety that deny their individual rights to equality, identity and dignity. Sanjeev Khanna of the Supreme Court of India has described the denial of dignity as follows: “Loss of dignity and self-worth of the targeted group members contributes to disharmony amongst groups, erodes tolerance and open-mindedness which are a must for multi-cultural society committed to the idea of equality. It affects an individual as a member of a group.”⁵ Critics caution, however, that relying on the feelings of injury to the dignity of a person may open the door to restrictions on free speech and that “it will not be long before this definition will be used against

³ A. Tontodimamma et al., “Thirty Years of Research into Hate Speech: Topics of Interest and Their Evolution,” *Scientometrics* 126 (2021): 157–79. In Norway (HR-2018-871-A), it has been held that a statement made via social media, even in the context of a closed group, met the requirements of the statement being ‘public’ pursuant to the relevant legislation I.N. Duy, “The Limits to Free Speech on Social Media: On Two Recent Decisions of the Supreme Court of Norway,” *Nordic Journal of Human Rights* 38 (2020): 240. The court did, however, note that a distinction must be drawn between speech directed at a specific person or community, and speech that contributes to the public debate and discourse (HR-2020-2133-A, § 58 ff).

⁴ Saskatchewan (Human Rights Commission) v. Whatcott, 2013 SCC 11 (CanLII), 1 SCR 467 (2013).

⁵ Amish Devgan v. Union of India, SCC OnLine SC 994 (2020).

individuals speaking against the government”.⁶ The Constitutional Court of South Africa has also declared that the use of the word ‘hurt’ as a requirement for hate speech is too vague and lacks the clarity required for a finding of hate speech.⁷

Hate speech is not necessarily limited to actual words. It can also find expression via a sign, a symbol, a song, an insignia, a publication or something in writing. ‘Speech’ is given a wide meaning that relates to observable senses rather than just the spoken word. This is consistent with the ‘purposive interpretation’ of legislation.⁸ But this wide approach also presents challenges since caution must be shown before the mere display of a symbol is classified as hate speech. As is discussed below, in South Africa the waving of the previous, pre-democracy, national flag without any words being uttered, has been regarded as factual evidence of hate speech without any other evidence being required, even though the flag had not been banned or criminalized by Parliament.⁹ On the other hand, in the United States of America (US), the burning of the Christian cross, which is typically associated with the racism of the Ku Klux Klan, has been treated as falling within freedom of expression.¹⁰

Since there is no requirement in law for words to be uttered to meet the threshold of hate speech, the challenge that courts face is to ascertain what constitutes hate speech in the absence of an agreed definition, either in international law or domestic constitutional law. The term hate speech is easily used by those who are offended, while those who speak are equally hasty to deny that their speech is hate speech. One principle that is clear, is that hate speech is the most extreme and objectionable form of speech that dehumanizes a person or a community. The Canadian Supreme Court has said three principles should be applied in interpreting prohibitions on hate speech: first, the prohibition must be applied objectively; second, hatred must be interpreted as being restricted to

⁶ P.V. Rao, “After Supreme Court Judgment, We Must Combat Hate Speech at Social and Political Levels,” The Leaflet, published December 8, 2020.

⁷ Qwelane v. South African Human Rights Commission and Another, ZACC 22 (2021), para. 155.

⁸ Qwelane v. South African Human Rights Commission and Another, para. 115.

⁹ Nelson Mandela Trust and Ors v. AfriForum and Ors, EQ 02/18 (2019).

¹⁰ Virginia v. Black, 538 U.S. 343 (2003).

extreme manifestations of harmful speech; and third, courts must focus on the effect of the expression at issue, rather than the nature of the ideas expressed.¹¹

This paper provides a brief overview of jurisprudential developments in international law about speech that may fall within the category of hate speech, whereafter two prominent South African judgments by the Equality Court are discussed. The methodology employed is based on comparative literature and jurisprudential analysis, whereby applicable judgments from various parts of the world are used to highlight the complexity and inconsistency between jurisdictions to identify and curb hate speech. The two South African judgments are used as points of reference to highlight the complexity in determining when speech can be regarded as hate speech; what test is applied to ascertain whether speech constitutes hate speech; and what evidence is required for a finding to be made. The two judgments under discussion, *Nelson Mandela Trust and Ors v AfriForum and Ors*,¹² and *AfriForum and Economic Freedom Fighters and Ors*¹³ attempted to determine the line that separates freedom of expression from hate speech. The judgments, perhaps not unexpectedly, have given rise to more questions than answers.

II. HATE SPEECH IN INTERNATIONAL LAW

The menace that hate speech poses to stability, human rights and democratic peace and order has recently been described by the Secretary General of the United Nations, Antonio Guterres, as follows:

“Public discourse is being weaponized for political gain with incendiary rhetoric that stigmatizes and dehumanizes minorities, migrants, refugees, women and any so called “other’...Hate speech is a menace to democratic values, social stability and peace. As a matter of principles, the United Nations must confront hate speech at every turn. Silence can signal indifference to bigotry and intolerance, even as a situation escalates and the vulnerable become victims... Addressing hate speech does not mean limiting or prohibiting freedom of speech... By enhancing global resilience against

¹¹ *Saskatchewan (Human Rights Commission) v. Whatcott*, 2013 SCC 11 (CanLII), 1 SCR 467 (2013).

¹² *Nelson Mandela Trust and Ors v. AfriForum and Ors*, EQ 02/18 (2019).

¹³ *AfriForum v. Economic Freedom Fighters and Others* (EQ 04/2020), ZAGPJHC 599 (2022). ‘Boer’ in the Afrikaans language has principally two meanings: a farmer or a white, Afrikaans-speaking person.

this global phenomenon [of hate speech], we can strengthen the bonds of society and build a better world for all.”¹⁴

There is no agreed universal definition of hate speech. Since hate speech is not a term of art, the meanings attached to it can range from the over-sensitive to the highly offensive. Hate speech is not to be equated to speech that we hate. As observed by the US Supreme Court: “Speech that demeans on the basis of race, ethnicity, gender, religion, age, disability, or any other similar ground is hateful; but the proudest boast of our free speech jurisprudence is that we protect the freedom to express “the thought that we hate.”¹⁵ Whilst it is acknowledged in jurisprudence that hate speech is the most extreme form of speech and should not be confused with other forms of objectional speech, the dilemma is that what constitutes hate speech is often dependent upon the circumstances, context and history of the speech in question and those who are exposed to it.

In short, what constitutes hate speech is influenced by various factors, for example: the history of a country or a community; the timing when the speech is uttered since a speech may lose or gain its hate-status with the passing of time; the circumstances when and where the speech is uttered, for example, at a political event, on social media, in the privacy of home, or at a scientific or educational forum; and the perception of those to whom the speech is directed, because not all persons from the same community may display the same sensitivity when they observe the speech. Added to the definitional complexity, hate speech is often directed at ethno-cultural minorities, but it is not necessarily solely directed at minorities. Hate speech can also be directed at a majority, a dominant community, or an individual. Minorities may, however, be more vulnerable because they do not have the institutional or political means to combat hate speech or to declare it a criminal act through legislation. The majority, on the other hand, can criminalize hate speech directed at them, or they may ban certain symbols or insignia that offend them, while the principal recourse of minorities is to rely on courts to make a finding under common law or statutory law that speech

¹⁴ UN Strategy Hate Speech, “UN Strategy and Plan of Action on Hate Speech” (Report, Geneva, 2019), 2.

¹⁵ *United States v. Schwimmer*, 279 U. S. 644, 655 (US Supreme Court, 1929).

constitutes hate speech. The reality is, however, that the justices are appointed by the majority. This often leaves the minority not only exposed to hate speech, but potentially without effective remedy to challenge it.

There is a risk that the categorization of speech as hate speech can also become a mechanism for censorship. A dominant political party or an approach adopted by the courts may unduly limit freedom of expression and in effect use curbs against hate speech to suppress criticism or controversial views. Free and open debate can then become stifled or even criminalized behind the veil of combatting hate speech. In considering the complexity of hate speech in countries with deep-seated historical conflicts, it is salient to note the US Supreme Court's observation that in deciding whether freedom of expression should be limited, the court must not adopt a position of favoritism, whereby the sensitivities of one group are given preference over those of another group.¹⁶ If the court were to adopt a favoritism approach, where the history or sensitivities of one community are treated more favorably than the history or sensitivities of another community, judgments would merely become an extension of the dominant view in society, rather than protecting the plurality of opinions.¹⁷

In order to give greater guidance to what is meant by hate speech, the United Nations through the office of the Secretary General, has suggested the following working definition of hate speech:

“The term hate speech is understood as any kind of communication in speech, writing or behavior, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, color, descent, gender or other identify factor.”¹⁸

This proposal clearly leaves a lot of room for courts to maneuver. Striking a balance between the limits of freedom of expression and the utterance of hate speech, is challenging. Freedom of expression is at the core of any free

¹⁶ R.A.V. v. City of St. Paul, 505 U.S. 377 (1992).

¹⁷ B. De Villiers, “Flying a Flag for Freedom of Expression: When Does a Historic Symbol of a Minority Turn into Hate Speech? The Case of the Old Flag of South Africa,” in *Navigating the Unknown – Essays on Selected Case Studies about the Rights of Minorities* (Leiden: Brill, 2022), 246.

¹⁸ UN Strategy Hate Speech, “UN Strategy and Plan of Action on Hate Speech” (Report, Geneva, 2019), 2.

and democratic society. This is because, in the words of the Canadian Supreme Court, the connection between freedom of expression and the political process is “perhaps the linchpin” of a free and democratic society.¹⁹ The freedom of expression includes the right to say and publish things that may cause offense, that are objectionable, that may hurt, that may lack sensitivity, or that may be derogatory.²⁰ In this regard, the caution expressed by the European Court of Human Rights must be noted, namely that a hurtful. At the same time, however, if a hateful comment is made and the person making the statement fails to remove it from a social media site, the imposition of a penalty is justified in a free and democratic society.²¹ At the same time, however, if a hateful comment is made and the person making the statement fails to remove it from a social media site, the imposition of a penalty is justified in a free and democratic society.²²

Freedom of expression is an ambit right of which the boundaries are wide and should not readily be restricted. As has been acknowledged by the United Nations Special Rapporteur on Minorities, Fernand de Varennes, “the freedoms of opinion and expression should be viewed as a default starting point with only very strictly constructed restrictions”.²³ In an earlier publication, I have described the balance between free expression and hate speech as “a mirage that remains elusive in international and constitutional law”, but “one aspect that is however shared by the respective case studies is that of all the different types of unacceptable speech, hate speech is the highest threshold to meet” and “if freedom of expression is to be curtailed, it should be narrow rather than wide”.²⁴

Several instruments of international law – both hard and soft law – refer to the primacy of freedom of expression in a free and democratic society, for example: Article 19 of the Universal Declaration of Human Rights; Article 19 of the International Covenant on Civil and Political Rights; Article 5 of the

¹⁹ R. v. Keegstra, 3 SCR 697 (1990).

²⁰ M. Herz and P. Molnar, *The Content and Context of Hate Speech: Rethinking Regulation and Responses* (Cambridge: Cambridge University Press, 2012).

²¹ Refah Partisi (the Welfare Party) and Others v. Turkey App, No. 41340/98 ECTHR 13 (2003), para. 57.

²² Sanchez v. France, 45581/15 (2021).

²³ F. De Varennes, “Minority Issues: Report of the Special Rapporteur on Minority Issues, Fernand de Varennes” (Report, United Nations, 2021), 5.

²⁴ Villiers, “Flying a Flag,” 248-49.

International Convention on the Elimination of all Forms of Racial Discrimination; Article 11 of the European Charter of Fundamental Rights; Article 13 of the American Convention on Human Rights; Article 9 of the African Charter of Human and Peoples' Rights; Article 32 of the Arab Charter on Human Rights; and Article 10 of the European Convention for the Protection of Human Rights and Fundamental Freedoms. Added to these international instruments, an explicit or implied protection of the right to freedom of expression is associated with modern constitutions and constitutionalism. The high regard in which freedom of expression is held, means that whenever an allegation is made of hate speech, the starting point is protection of freedom of expression, with the burden resting on the aggrieved person to prove that the objectionable speech was disproportionate to freedom of expression; that it constituted hate speech; and that a proportionate restriction or remedy should be granted. This is because it is universally accepted that the exercise of the right to free speech "without fear or unlawful interference – is central to living in an open and fair society; one in which people can access justice and enjoy their human rights".²⁵

Although hate speech often has a suggestion of violence, or that it may seek to normalize or encourage violence, evidence of actual violence or a threat of violence is not a requirement for hate speech to be proven.²⁶ Speech itself is adequate to constitute hate speech without any actions or threats arising from it. Experiences show, however, that hate speech often has the effect, if not the intent, to dehumanize persons or a community, to encourage or justify violence against them, to rationalize discrimination against them.²⁷ States are therefore encouraged by international instruments to enact legislation to combat hate speech. There are two main categories of permissible limitations on the scope of freedom of expression to protect persons against hate speech, namely: criminalization of speech that constitutes incitement to genocide; and prohibition

²⁵ Amnesty International, "Freedom of Expression," Amnesty International.

²⁶ *Qwelane v. South African Human Rights Commission and Another*, ZACC 22, para. 181.

²⁷ F. De Varennes, "Minority Issues: Report of the Special Rapporteur on Minority Issues, Fernand de Varennes" (Report, United Nations, 2021), 8.

of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility, or violence.²⁸

In summary: the principles arising from international law to protect individuals against hate speech are as follows: first, the right to freedom of expression is a fundamental right in a free and democratic society and any restriction placed on the right must be proportionate and limited to extreme forms of speech; second, speech refers not only to spoken words, but includes symbols, insignia and publications; third, there is no universal definition of hate speech; and fourth, while states are encouraged to enact legislation to prohibit hate speech, a finding of hate speech is based on an objective assessment of the speech in question and is dependent on the specific circumstances of the grievous speech.

III. JURISPRUDENTIAL STRUGGLE TO DEAL WITH ALLEGATIONS OF HATE SPEECH

International experiences in dealing with freedom of expression vis-à-vis hate speech make for fascinating, albeit confusing, analysis. An overview of judgments shows a lack of consistency not only between nations, but also within nations about determining when speech constitutes hate speech.²⁹ There is neither an agreed international law definition of hate speech, nor a closed list of words or symbols of what constitutes hate speech.³⁰ International law and statutes enacted by states deal with hate speech by way of generality, with courts having to ascertain if, in a particular context, freedom of expression should be curtailed. Although the meaning of the term hate speech is to be ascertained objectively in a factual situation, international jurisprudence suggests that the term ‘hate’ to categorize speech in a legal sense is only to be found in extreme circumstances.³¹ One can agree with Brown when he observes as follows:

²⁸ Hogan Lovells, “The Global Regulation of Online Hate: A Survey of Applicable Laws” (Report, Hogan Lovells, 2020), 42.

²⁹ K. Topidi, “Words That Hurt (1): Normative and Institutional Considerations in the Regulation of Hate Speech in Europe” (Paper, SSRN, 2019), 6.

³⁰ Topidi, “Words That Hurt,” 32.

³¹ J. Schweppe and D. Walsh. *Combating Racism and Xenophobia Through the Criminal Law* (Limerick: Centre of Criminal Justice, 2008), 72.

“Given that hate speech laws provoke such strong moral reactions, on the part of defenders and critics alike, and given that legal meanings will themselves draw on a range of deeper values and principles about which people reasonably disagree, it is no surprise that there remains such divergence over how to define the very term that stands at the epicenter of the disagreement.”³²

The difficulty to define hate speech and the relatively low number of successful prosecutions of hate speech have caused Brown to observe that “there are many instances of hate speech where this putative connection is not merely weak but non-existent”.³³ An important principle identified in the *Faurisson judgment* of the European Committee on Human Rights is that any power to place restrictions on freedom of expression “must not be interpreted as license to prohibit unpopular speech, or speech which some sections of the population find offensive”.³⁴ Sensitivity by a person or community to speech should therefore not be used as an excuse to censor speech. A democracy is built on the pillar of free and robust speech that may cause offense.

The following examples are used to highlight for the purposes of this paper how challenging it is to lay down general norms to determine when freedom of expression should be restricted because of its potential hate element:

3.1. Can a Symbol Constitute Hate Speech?

The display of a symbol or the way a symbol is dealt with, for example, the waving of a flag or the burning of a national flag, is treated as a form of speech. No words are required for hate speech to be established. The question is when, or if, the display of a symbol does in fact constitute hate speech. Gelber, for example, reflects on efforts in the US, Australia and New Zealand to restrict the burning of the national flag. He says those efforts to restrict freedom of expression or to criminalize the burning of a flag, expose the “fragility of freedom of speech” and that freedom of speech is “culturally at risk”.³⁵ In the *Texas judgment*, the US Supreme Court acknowledged that the burning of the flag is, as a “political

³² A. Brown, “What Is Hate Speech? Part 1: The Myth of Hate’,” *Law and Philosophy* 36 (2017): 422.

³³ Brown, “What Is Hate Speech? Part 1,” 466.

³⁴ *Faurisson v. France* Communication, No 550/93 (1996).

³⁵ K. Gelber, “Political Culture, Flag Use and Freedom of Speech,” *Political Studies* 60 (2012): 176.

statement”, a form of speech that is protected by the Constitution regardless of the hurt, fear, intimidation or hate that onlookers may experience.³⁶ On the other hand, the Equality Court of South Africa has declared that the waving of the pre-democracy national flag at a public rally against high levels of crime, was not only hate speech on the specific day, but that any display of the country’s old flag other than for scientific purposes also constitutes hate speech into the future.³⁷ The Equality Court in effect banned the old flag – something that the Parliament of South Africa had not thought fit to do. The different standards adopted by countries to regulate hate speech are also reflected in the manner that legislation deals with the most controversial of symbols, namely Nazi symbols. The display of the swastika flag and other Nazi insignia in some European states is highly regulated and even prohibited, particularly so in Germany, whilst in other countries such as the US and Australia there is not a general prohibition against those symbols since their display is protected by the right to freedom of speech.

3.2. Must Hate Speech Include Violence or Threat of Violence?

Hate speech does not necessarily have to be accompanied by a form of violence or threat of violence. The International Criminal Tribunal for Rwanda declared that hate speech does not necessarily require an incitement to violence since the nature of the speech in itself can meet the threshold of ‘hate’.³⁸ On the other hand, in the *Kill the Boer* case, as discussed below, the Equality Court of South Africa relied on the lack of adequate evidence between the song that encourages the killing of Boers (white farmers) and actual violence on farms to conclude that the song did not meet the threshold of hate speech. The Court found that there was inadequate evidence to show that the controversial song contributed to actual violence or murders against farmers.³⁹ If, however, one adopts the reasoning of the UN Special Rapporteur for Minorities, who describes hate

³⁶ *Texas v. Johnson*, 491 US 397 (1989).

³⁷ *Nelson Mandela Trust and Ors v. AfriForum and Ors*, EQ 02/18 (2019); Villiers, "Flying a Flag," 215-50.

³⁸ *The Prosecutor v. Ferdinand Nahimana Jean-Bosco Barayagwiza Hassan Ngeze*, ICTR 99-52-T (2003).

³⁹ *AfriForum v. Economic Freedom Fighters and Others* (EQ 04/2020), ZAGPJHC 599 (2022), para. 101.

speech as dehumanizing language that normalizes violence against a minority,⁴⁰ it seems as if a compelling case can be made that a song calling upon people to kill (white) farmers, meets the threshold of hate speech, even if that song may have been appropriate in the pre-democracy liberation struggle. In the *Qwelane* case, on the other hand, the Constitutional Court of South Africa took into account complaints lodged against a newspaper article that criticized homosexuality and evidence given in court on the effects of the speech.⁴¹ The Constitutional Court accepted the presence of hate speech in the article even in the absence of evidence that violence or threats of violence arose from the article.⁴²

Hate speech is impacted by time and circumstance, which implies that speech may lose its hate speech status as time effluxes, or speech may become hate speech considering new circumstances. In Australia, for example, it has been found that the word ‘nigger’ does not always constitute racism since in the context of the historic naming of a sports stadium, the word ‘Nigger’ was used as a nickname of a celebrated rugby player.⁴³ In another Australian judgment, in the *Catch the Fire* case, Justice Robert French explained that what constitutes offensive speech in one circumstance may not constitute offense speech in another, even if the same symbol is shown or the same words are uttered. It must also be borne in mind that as circumstances change, the offense that is experienced by the use of certain words or expressions may change.⁴⁴

3.3. Is the Test for Hate Speech Subjective or Objective?

The test whether speech constitutes hate speech is an objective test by considering what a reasonable person observing the speech would experience, not the subjective intentions of the speaker or the sensitivity of a single onlooker. The perception of a fictional, reasonable onlooker who is adequately informed about the issue, is ascertained by the court. The application of this test may be more complex than it may seem at first glance since some members of a

⁴⁰ F. De Varennes, “Minority Issues: Report of the Special Rapporteur on Minority Issues, Fernand de Varennes” (Report, United Nations, 2021), 8.

⁴¹ *Qwelane v. South African Human Rights Commission and Another*, ZACC 22 (2021), para 6.

⁴² *Qwelane v. South African Human Rights Commission and Another*, para. 187.

⁴³ *Hagan v. Trustees of the Toowoomba Sports Ground Trust*, FCA 1615 (2000).

⁴⁴ *Catch the Fire Ministries Inc v. Islamic Council of Victoria Inc*, VSCA 284; 15 VR 207; 235 ALR 750 (2006).

community may objectively view a speech differently from other members of society. For example, the opinion of a reasonable person who is heterosexual may differ to the opinion of a reasonable person who is homosexual when exposed to degrading speech. It is not a given that a judge can really feel or comprehend the hurt felt by a minority community in response to certain speech. Whilst some suggest the reasonable person test is akin to the test of reasonable person in tort or delict, this is not the case since in tort or delict the community standard of conduct is indeed appropriate, while with hate speech the hurt experienced by an individual or a minority community is at issue. This raises the question of whether the opinion of an objective person must arise from that reasonable person being part of the community at whom the speech is directed.

The Supreme Court of India has emphasized that the courts must apply the hate speech prohibition objectively. The question courts must ask is whether a reasonable person, aware of the context and circumstances, would view the expression as exposing the protected group to hatred.⁴⁵ In Norway, the Supreme Court has formulated the objective test as being the average reader or listener's "natural perception of the statement made from the context" (HR-2020-184-A) The average person's perception of the grievous speech is therefore essential.⁴⁶ But courts have also accepted that the context within which speech takes place is a relevant consideration.⁴⁷ The Constitutional Court of South Africa emphasized that the hate speech must be "reasonably construed" and cannot be based on an inference or assumption made by the targeted individual or group.⁴⁸ The US Supreme Court has noted the court must examine the "content, form, and context" of the speech and in doing so the court must evaluate all the circumstances of the speech, "including what was said, where it was said, and how it was said".⁴⁹ In Canada, the Supreme Court described the test to ascertain hate speech as follows:

⁴⁵ *Pravasi Bhalai Sangathan v. Union of India and Ors.*, No. 157 (Supreme Court, 2014).

⁴⁶ I.N. Duy, "The Limits to Free Speech on Social Media: On Two Recent Decisions of the Supreme Court of Norway," *Nordic Journal of Human Rights* 38 (2020): 242.

⁴⁷ *Qwelane v. South African Human Rights Commission and Another*, ZACC 22 (2021), para 20.

⁴⁸ *Qwelane v. South African Human Rights Commission and Another*, para. 96.

⁴⁹ *Snyder v. Phelps*, 131 S. Ct. 1207 (US Supreme Court 2011).

“The repugnancy of the ideas being expressed is not sufficient to justify restricting the expression, and whether or not the author of the expression intended to incite hatred or discriminatory treatment is irrelevant. The key is to determine the likely effect of the expression on its audience.”⁵⁰

The Supreme Court of India in the *Amish Devgan* case observed that both context and content of a speech are important to determine whether it amounts to hate speech. ‘Content’ has more to do with the expression, language and message, which should be to vilify, demean and incite psychosocial hatred or physical violence against the targeted group, the judgment observed. The ‘context’ has a certain key variable, namely, ‘who’ and ‘what’ is involved and ‘where’ and the ‘occasion, time and under what circumstances’ the case arises.⁵¹ The Court went on to say consideration must also be had for the position and status of the person making the statement. For example, the statement of an ordinary individual at a private event may not be met with the same scrutiny as that of a public figure at a public rally.⁵² Rather than to clarify the legal position about hate speech, these judgments add new layers of complexity.

The test of objectivity therefore implies objectivity within the nation or circumstances where the speech is expressed, not objectivity from the perspective of universal norms. This is a reasonable proposition, since speech taken out of context is impossible to categorize as breaching the limits of freedom of expression. For example, it has been held in France by the European Human Rights Committee that in the context of the “conditions of present-day France, Holocaust denial may constitute a form of incitement to anti-semitism”, albeit that in another circumstance or time or place the same denial may be judged otherwise.⁵³ In the *Faurisson judgment* concerning a person’s denial of the Holocaust, the Committee found that the intention of the person who expresses the denial is a relevant consideration, but that in the particular case the denial did constitute hate speech.⁵⁴ The Committee considered the context in which

⁵⁰ Saskatchewan (Human Rights Commission) v. Whatcott, 2013 SCC 11 (CanLII), 1 SCR 467 (2013).

⁵¹ *Amish Devgan v. Union of India*, SCC OnLine SC 994 (2020).

⁵² *Amish Devgan v. Union of India*, SCC OnLine SC 994 (2020), para. 52.

⁵³ *Faurisson v. France Communication*, No 550/93 (1996).

⁵⁴ *Faurisson v. France Communication*, No 550/93.

the Holocaust denial occurred as well as the intention of the person making the statement, for example, whether the statement was made in the context of a scientific discourse. The Constitutional Court of Germany has also noted that the question whether speech ought to be prohibited must be assessed in light of the circumstances of each case.⁵⁵ This approach is also reflected in the *Perincek* case in Switzerland, in which the court took detailed account of the context in which potentially hateful speech occurred.⁵⁶

The foregoing cases highlight that hate speech is not only be determined by the jurisdiction (country) where the speech is expressed, but also the context and the time in history when it is used. In India, for example, it has been stated that the right to free expression contributes to society being able to reflect on its past and present social reality. In the *FA Picture* case, the court found that although a section of society may perceive a film or image as hate speech,⁵⁷ the public debate caused by such a display may contribute to discourse about the wrongs committed under the symbol and those discussions could in turn deepen and strengthen democracy. This approach is consistent with that of Rosenfeld, who says that in the US, the display of Nazi or Ku Klux Klan symbols has become isolated due to popular irrelevancy and rejection of those symbols as a result of informed debate based on freedom of speech, rather than by legislative or judicial intervention.⁵⁸ Whilst the test for hate speech is objective, it remains fictional in nature because what is reasonable and what is proportionate are highly discretionary, and added thereto is that not all persons may display the same sensitivity to public speech.

3.4. What is the Threshold of Hate Speech?

As stated above, hate speech is distinguished from other forms of inappropriate, dislikeable, offensive, hurtful and objectionable speech. Hate speech is the most extreme form of speech. Any categorization of speech as hate

⁵⁵ The Federal Constitutional Court, Federal Constitutional Court (First Senate) 15 January 1958 BVerfGE 7, 198 (1958).

⁵⁶ *Perincek v. Switzerland* App, No. 27510/08 (ECHR, 2015).

⁵⁷ *FA Picture International v. Central Board of Film Certification*, AIR 2005 Bom 145 (2005), para. 13.

⁵⁸ M. Rosenfeld, "Hate Speech in Constitutional Jurisprudence: A Comparative Analysis," *Cardozo Law Review* 24 (2003): 1538.

speech should be reserved to “cover only the most intense form of dislike”.⁵⁹ As the Supreme Court of India recently stated, “when statements are made with the objective and intent to cause public disorder or to demean dignity of the targeted groups” it should be regarded as hate speech.⁶⁰ It is important, however, that the declaration of speech as hate speech must be directed at a specific incident rather than a general declaration of hate speech into the future. It is ultimately for parliaments to ban or regulate certain speech or symbols if it is to be done prospectively, whilst a court declaration is done retrospectively with reference to a specific factual situation where it is alleged that hate speech occurred. Many nations have opted to enact legislation to deal specifically with hate speech, as well as in some instances the banning or regulation of certain insignia and symbols. This may explain why in some nations the display of Nazi insignia is banned by criminal code rather than pursuant to a hate speech litigation. Critics in Germany of legislative intervention to ban the display of national-socialist symbols have, however, cautioned that such regulation “creates a useful template for autocratic countries to use to censor political opponents and other members of marginalized groups”.⁶¹

3.5. Is Freedom of Expression a Defense against a Criminal Code?

Although a criminal code may prohibit or regulate the display of certain symbols, a constitutional defense of freedom of expression may still be raised to defend against any prosecution. For example, even if a certain form of speech is prohibited by a criminal code, as in Germany, a defense of freedom of expression may still be used by the person uttering the speech. The Constitutional Court of Germany has, for example, cautioned that although the glorification of Nazi ideology may be declared a criminal act, it does not justify a general ban on the dissemination of right-wing radical or indeed National Socialist ideas.⁶² The Court emphasized that there must be proportionality between the fundamental

⁵⁹ R. v. Keegstra, 3 SCR 697 (1990), para. 122.

⁶⁰ Amish Devgan v. Union of India, SCC OnLine SC 994 (2020).

⁶¹ J. Delcker, “Germany’s Balancing Act: Fighting Online Hate While Protecting Free Speech,” Politico, published October 01, 2020.

⁶² The Federal Constitutional Court, Order of the First Senate of 4 November 2009, 1 BvR 2150/08 (2009).

right and the restriction that is being placed on it. The Court formulated the importance of proportionality as follows:

“This setting of boundaries must also be followed by a review of proportionality. The more concretely and directly a legal interest is placed in danger by an expression of opinion, the less stringent are the requirements when it comes to an encroachment; the more indirect and distant the threatening violations of legal interests remain, the greater are the requirements to be made. Accordingly, encroachments on freedom of opinion are more to be accepted when they are restricted to the forms and circumstances of an expression of opinion in the outside world. The more, by contrast, they ultimately result in a content-related suppression of the opinion itself, the higher are the requirements as to the concrete threat of a danger to legal interests.”⁶³

In summary, the above overview of selected international jurisprudence highlights the complexity to find consistency or universality in the definition and application of hate speech. The term hate speech has a certain qualitative value, which represents an extreme form of expression, albeit situation-bound. Each country offers a unique set of experiences and even within some countries there is sometimes a lack of consistency when it comes to classifying words and acts into an appropriate category.

Several principles are, however, shared by the respective case studies, namely: (a) of all the different types of unacceptable speech, hate speech is the highest threshold to meet; (b) even if hate speech is defined by legislation, the practical circumstances of the speech will determine if hate speech did actually occur; (c) the test to be applied is objective from the perspective of the reasonably informed onlooker and not subjective from the mind of the speaker, albeit that the circumstance of the speech is a relevant factor; and (d) if freedom of expression is to be curtailed, it should be proportionate and narrow rather than wide.⁶⁴

IV. OLD FLAG CASE AND KILL THE BOER CASE

In this part, consideration is given to two recent judgments of the Equality Court of South Africa to highlight the challenges that remain in ascertaining if speech meets the threshold of hate speech. The two judgments, those of the

⁶³ The Federal Constitutional Court, Order of the First Senate of 4 November 2009, 1 BvR 2150/08 (2009), para. 52.

⁶⁴ Villiers, "Flying a Flag," 249.

Old Flag case and the *Kill the Boer* case, raise questions on the following four issues: (a) when speech can be regarded as hate speech; (b) what test – be it objective or subjective – is applied to ascertain whether speech constitutes hate speech; (c) what evidence is required for a finding to be made; and (d) whether a finding of hate speech is retrospective or prospective.

The factual background to the two judgments can be summarized as follows:

4.1. Old Flag Case

In the *Old Flag-Mandela Trust* judgment, the Equality Court of South Africa declared that the “gratuitous”⁶⁵ display of the old flag of South Africa on 30 October 2017 by a person participating in a public demonstration against high levels of crime in the country, constituted hate speech against black people and is therefore not protected by the constitutional right to freedom of expression.⁶⁶ The so-called old flag refers to the official South African flag that was in place during the apartheid era from 1927 to 1993. With the dawn of democracy, South Africa in 1994 adopted a new flag to represent the new democratic constitution. The old flag has not been banned by Parliament nor had the display thereof been the subject of any restrictive regulation.⁶⁷ The Equality Court declared the display of the old flag to be hate speech, not only its display on the day of the protest but also in any future display unless for journalistic, educational or scientific purposes. The case stemmed from a public protest against high levels of crime in South Africa. It is not disputed that South Africa has been experiencing an increase in crime and lawlessness. For example, South Africa has been ranked 19th in the world for organized crime.⁶⁸ Against this backdrop, a series of public marches took place on 30 October 2017. Coordinated by an Afrikaans-speaking, non-governmental association called AfriForum, the marches were attended

⁶⁵ The Court did not define ‘gratuitous’ but according to Macquarie Concise Dictionary (2019), it means ‘without good reason, cause or justification’.

⁶⁶ Nelson Mandela Trust and Ors v. AfriForum and Ors, EQ 02/18 (2019).

⁶⁷ B. De Villiers, “Flying a Flag for Freedom of Expression: When Does a Historic Symbol of a Minority Turn into Hate Speech? The Case of the Old Flag of South Africa,” in *Navigating the Unknown – Essays on Selected Case Studies about the Rights of Minorities* (Leiden: Brill, 2022).

⁶⁸ D. Delpont, “SA 19de Op Lys van Lande Met Ergste Georganiseerde Misdaad,” *Netwerk24*, published September 24, 2022.

by persons from all racial groups. During one of the protests, the old flag was displayed by some of the participants.⁶⁹ There is no suggestion that the display of the old flag during the march was orchestrated by the event organizers; or that the protest took place under the banner of the old flag; or that the protestors, including those who displayed the old flag, rejected the existing constitutional order or the new flag. No other racially offensive slogans or banners of a racial hatred nature nor a rejection of the current constitutional order accompanied the display of the old flag. The march was, according to all evidence, a peaceful protest against crime.

The Chief Executive of the Nelson Mandela Trust, Sello Hatang, filed an affidavit with the Equality Court in which he explained how he, as a black person, experienced the march and the display of the Old Flag. He had not seen the display of the old flag. He said,

“the Old Flag represents nothing other than the inhumane system of racial segregation and subjugation that governed South Africa before 27 April 1994.... To *hear* that the Old Flag has been displayed gratuitously in 2017, more than a generation after apartheid had been abolished, reminded me that some South Africans still see me and other blacks as ‘other’, and would deny us the opportunity to be human.”⁷⁰ [italics added]

The Nelson Mandela Foundation sought a declaration from the Equality Court that any display of the old flag that does not serve a genuine journalistic, academic or artistic purpose in the public interest, *ipso facto* (without any further evidence about circumstance) constitutes hate speech against black people pursuant to Section 10 of the Promotion of Equality and Prevention of Unfair Discrimination Act 4 of 2000.⁷¹ The Court found in favor of the applicant and declared the gratuitous display of the old flag on the day of the march and into the future, as hate speech – thereby effectively banning the display of the old flag at any cultural, private, sporting or political event.

⁶⁹ Tom Portner, “South Africans Hold Black Monday Protests Over Farm Murders,” *Newsweek*, published October 30, 2017.

⁷⁰ *Nelson Mandela Trust and Ors v. AfriForum and Ors*, EQ 02/18 (2019).

⁷¹ “Equality Act. Promotion of Equality and Prevention of Unfair Discrimination Act 4 of 2000,” opened for signature February 2, 2000.

4.2. Kill the Boer Case

The complainant, AfriForum, sought a declaration from the Equality Court declaring that a song sung by the respondent, the Economic Freedom Fighters (EFF), constitutes hate speech pursuant to Section 10 of the Promotion of Equality and Prevention of Unfair Discrimination Act 4 of 2000.⁷² The EFF did not deny that its members often sing the song *Kill/Kiss the Boer – Kill/Kiss the Farmer* in public, and that they sometimes display a gesture of a gun while singing it. AfriForum contended the song constituted hate speech since it “advocates hatred on the grounds of race and ethnicity, and constitute an incitement to cause harm”.⁷³ AfriForum also relied on a previous judgment of the Equality Court that declared the song constituted hate speech.⁷⁴ The precedent value of that previous case was, however, diluted by an agreement the parties had entered into on appeal at the time that a mediated outcome would substitute the earlier orders of the Equality Court. The Court in this case therefore found that the earlier judgment did not bear any relevance nor did it constitute a precedent in the current proceeding.⁷⁵

AfriForum contended that the song incites violence, that it is a form of hate speech, and that it can be linked to an increase in violence against farmers and their workers. According to AfriForum, in 2019 there were 552 farm attacks and 57 murders.⁷⁶ AfriForum said the song contributed to a political climate in which white persons are portrayed as the “source of evil”, that farmers are portrayed as criminals who exploit workers, and that violence against farmers was “romanticized” by the song.⁷⁷ The EFF said in response that its movement was formed to promote the economic emancipation of black people and in doing so it supports, among others, the expropriation of land without compensation. The EFF said the song should not be taken literally but assessed in light of the

⁷² Equality Act, sec. 10.

⁷³ AfriForum v. Economic Freedom Fighters and Others (EQ 04/2020), ZAGPJHC 599 (2022), para. 3.

⁷⁴ AfriForum and Another v. Malema and Another (Vereniging van Regslui vir Afrikaans as Amicus Curiae) 2011 (12), BCLR 1289 (EqC) (2011).

⁷⁵ AfriForum v. Economic Freedom Fighters and Others (EQ 04/2020), ZAGPJHC 599 (2022), para. 98.

⁷⁶ AfriForum Research Institute, “*Farm Attacks and Farm Murders in South Africa*” (Report, AfriForum Research Institute, 2020), 3.

⁷⁷ AfriForum v. Economic Freedom Fighters and Others (EQ 04/2020), ZAGPJHC 599 (2022), para. 37.

oppressive apartheid system of which the effects remain visible. The word ‘Boers’, argued the EFF, refers to the system of oppression under apartheid, rather than to a specific community.⁷⁸

The Equality Court dismissed the complaint. Justice Molahleni explained his finding, in essence, as follows: the test to be applied is an objective one “that requires a reasonable person test”;⁷⁹ the objective test includes taking into account a reasonable listener having regard to the historical context of the speech; offensive or controversial statements do not equate to hate speech;⁸⁰ the lyrics of the song cannot reasonably be construed to incite hatred; the song should not be taken in its literal meaning;⁸¹ that there was inadequate evidence to link violence on farms to the song; and curtailing the singing of the song would place an undue restriction on freedom of expression since it falls within the scope of ‘political contestations’.⁸²

V. CRITICAL ASSESSMENT OF OLD FLAG AND KILL THE BOER JUDGMENTS

The two aforementioned judgments of the Equality Court of South Africa raise issues that are not only relevant to South Africa but also reflect the ambivalence within international comparative law about the demarcation between freedom of expression and hate speech. Earlier in this paper, the following questions were identified in relation to analysis of the two judgments: (a) when speech can be regarded as hate speech; (b) what test – be it objective or subjective – is applied to ascertain whether speech constitutes hate speech; (c) what evidence is required for a finding to be made; and (d) whether a finding of hate speech is retrospective or prospective.

5.1. When is Speech ‘Hate’ Speech?

There is no universal definition of hate speech, although there is agreement that hate speech is the most extreme form of speech and should not be confused

⁷⁸ AfriForum v. Economic Freedom Fighters and Others, para. 69–71.

⁷⁹ AfriForum v. Economic Freedom Fighters and Others, para. 92.

⁸⁰ Nelson Mandela Trust and Ors v. AfriForum and Ors, EQ 02/18 (2019), para. 97.

⁸¹ Nelson Mandela Trust and Ors v. AfriForum and Ors, para. 107.

⁸² Nelson Mandela Trust and Ors v. AfriForum and Ors, para. 102.

with other forms of dislikeable speech. Various countries, including South Africa, have enacted legislation to define hate speech. The general theme that runs through definitions is that hate speech is not protected by freedom of expression; hate speech includes speech, publication and symbols; and hate speech is speech that is harmful to persons or may promote or propagate hatred, for example, the Equality Act 2000, Section 10(1). Generally, hate speech is understood as any kind of communication in speech, writing, symbol or behavior that attacks or uses pejorative or discriminatory language with reference to a person or a group based on who they are, in other words, based on their religion, ethnicity, nationality, race, color, descent, gender or other identity factors.

In both cases discussed above, the reasoning adopted by the Equality Court is open to criticism. In the *Old Flag* case, the Court concluded that the mere display of South Africa's old flag at a public march against crime was *ipso facto* hate speech. The court did not consider the context in which the old flag was displayed, particularly that the public protest march against crime could be regarded as a political event. The Court also did not take into account that the person who had waved the old flag had not been identified; the event organizers did not promote the old flag; there was no evidence that the display of the old flag was directed at black persons or as some form of rejection of the new democratic order; and the event organizers were not responsible for the display of the old flag. In the *Kill the Boer* case, on the other hand, the Court went to some length to investigate the context within which the song was sung; the absence of evidence to prove that the singing of the song had caused actual violence; and the history and current symbolism of the song. It would seem that the Court's approach in the two cases differed fundamentally. In the *Old Flag* case, the mere display of a symbol was regarded as hate speech; while in the *Kill the Boer* case, the context, history and symbolism of the song were considered and accepted in its defense.

5.2. What Test to Apply: Objective or Subjective?

Both judgments reflect the international consensus that for hate speech to be determined, an objective test is required. There is no need to ascertain

the subjective mindset of the person who expressed the speech, albeit that the circumstances within which the speech was expressed is a relevant factor. A statement made in private may be assessed to be different to a statement made at a political rally, compared to a statement made in a scientific forum.

While an objective test is proper, the application thereof in the two judgments raises an important question, namely, how is the objective test of reasonableness of speech applied in practice? In the *Old Flag* case only one aggrieved person gave evidence about the perceptions of the black community of the old flag. That evidence was accepted by the Equality Court as being objectively reasonable. The Court did not accept the evidence that the old flag has a different meaning to different communities; that the old flag represents part of the complex history of South Africa; that the old flag itself reflected the struggle against British imperialism; and that the meaning of the old flag had changed over time. While it cannot be denied that the old flag represents the system of apartheid, it also reflects many other dimensions of South African society. The Court accepted that the old flag, objectively, only has one predominant meaning and that represents hate speech. On the other hand, in the *Kill the Boer* case, the Court rejected evidence of persons who spoke about the harm, fear and anxiety they experience when the song is heard. The Court accepted that it was a pre-democracy liberation song of which the meaning should not be interpreted literally under the current democratic order. In the *Old Flag* case, however, the Court adopted a position that the old flag has a predominant meaning that overshadows any other secondary meaning.

The inconsistency between the two approaches about the application of an objective test raises the question of how, in an ethnically diverse, post-conflict society, does a court ascertain what a reasonable person would perceive and what are the risks, as have been cautioned by the US Supreme Court, of favoritism being shown to one community's sensitivities over those of another community?⁸³ Although one could accept that the 'Kill the Boer' chant was objectively justifiable during the apartheid struggle, the question can be raised if it is equally justifiable

⁸³ R.A.V. v. City of St. Paul, 505 U.S. 377 (1992).

in a democracy where it continues to be used as a political slogan in the context of a country that is plagued by interracial violence and crime? It is also notable that in the *Kill the Boer* case, the Court went to some length to explain the justification of the song in light of the history of its origin; whereas in the *Old Flag* case, the Court heard no evidence from persons who were at the march who may have been offended; the Court discounted that a public march against crime had taken place; and the Court attached no weight to the circumstances within which the old flag was displayed. The judgments highlight that although there is agreement that the test for hate speech is objective, the application of the test may be problematic, particularly in deeply divided ethno-cultural societies. The two judgments leave unanswered the question of how a judge ascertains (other than by his or her personal preference) the depth of hurt, pain or fear experienced by those exposed to hate speech.

5.3. What Evidence is Required for a Finding of Hate Speech to be Made?

Comparative jurisprudence suggests that a wide net is cast to consider whether speech should be restricted, for example the background to the speech (or symbol); the circumstances in which the speech is made; the content of the speech; the forum where the speech is made; the profile or status of the person making the speech; the audience of the speech; the perceptions of people exposed to the speech; and any other relevant factor. The reason for considering all this information is that courts should be slow to restrict freedom of expression. It is only when a court is satisfied that speech exceeds boundaries of reasonableness, that free expression should be restricted. Even if freedom of expression is restricted, the restriction should be proportionate and limited to the specific event and the subject of the complaint.

In the two judgments under discussion, the Equality Court adopted seemingly different approaches to the weighting of evidence. In the *Old Flag* case, the Court only heard evidence from one aggrieved person who had merely heard of the display of the Old Flag on television. The Court accepted the evidence of that witness and attached less weight to submissions and evidence by the organizers of the march about the other meanings of the old flag and the history thereof.

Notably, when the *Old Flag* judgment is compared to the *Kill the Boer* judgment, it appears that in the *Kill the Boer* judgment, the Court went to some length to explain and justify that the song should not be taken to its literal meaning and the history of the song should be considered, whilst in the *Old Flag* judgment the Court rejected those same propositions. Although in international law, a threat of violence or actual violence is not required to prove hate speech, in the *Kill the Boer* case, the Court went to some length to justify its findings on the basis that a causal link between the song and violence had not been established. Both cases highlight the complexity to ascertain hate speech within the context of political rallies and events, and the risk that the opinions of the majority are weighted with greater sympathy than those of the minority.

5.4. Is a Declaration of Hate Speech Retrospective or Prospective?

Comparative jurisprudence suggests that a declaration of hate speech (or not) applies to a certain incident, in other words, the effect of the finding is retrospective and it does not regulate any future speech of the same nature. The reason being that parliaments may ban or restrict certain forms of speech prospectively, whilst courts are limited in their jurisdiction to a specific event and a particular circumstance.

Whilst in both judgments there is recognition that freedom of expression is an important pillar of a free and democratic society, and that speech should only be restricted to the extent that it is necessary to prevent hate speech, the question that arose in the *Old Flag* case is what the scope of a declaratory order should be. International theory and jurisprudence suggest that unless parliament bans certain speech or insignia, a declaration of hate speech should be directed only at the specific incident. In the *Old Flag* case, however, the Court not only declared that the display of the old flag on 30 October 2017 was hate speech, but the Court also went on to effectively ban the old flag, including the private display of the old flag. The scope of this declaration arguably exceeds the reasonable and proportional restriction that should be placed on freedom of expression. It was arguably disproportionate for the Court to make a declaration about the general display of the old flag, be it in private or in public, as being hate speech.

VI. CONCLUSION

This paper has shown how difficult it is to develop and apply universal norms to determine when freedom of expression should be restricted due to hate speech. Whether speech can be classified as hate speech is largely circumstantial, based on an objective assessment of the speech. It has also been highlighted that to ascertain what a reasonable person's perception is may be unrealistic and possibly highly unreliable, particularly in deeply divided societies. All relevant evidence should be considered by a court in the light of the fundamental importance of freedom of expression to a free and open democratic society. The two judgments discussed highlight the complexity to apply an objective assessment of reasonableness of speech since, particularly in a deeply divided society, the history of the respective communities and their perceptions of history are multidimensional, nuanced and complex. While in India the Supreme Court has upheld the right to certain forms of speech since the Court said it may assist the public to reflect on their past, the two judgments from South Africa seem to place higher value on the experiences and perceptions of one community vis-à-vis another community. In one case, the Equality Court declared that the display of a flag was hate speech; yet in another judgment, the Court declared that a song about killing farmers was not hate speech. The irony is that the South African Parliament had not thought it desirable to ban the display of the old flag specifically due to the sensitivity and complexity of it, yet the Court not only declared the display of the old flag on a specific day as hate speech, the Court also declared the old flag hate speech into the future. Ultimately, both judgments highlight the complexity to determine hate speech; the challenge to weigh up different histories of peoples and communities; the difficulty to separate a literal interpretation of speech from a figurative interpretation; and the intricacy arising from the so-called objective test since a reasonable person test in a deeply divided society may not be as simple to ascertain as theory may suggest. The lack of consistency in comparative jurisprudence reaffirms that the declaration of speech as hate speech should be reserved for the most

extreme forms of speech; it should be proportionate to the speech, including who expressed it, where and when; and any declaration should only be directed at the specific incident and not restrict speech in general.

BIBLIOGRAPHY

AfriForum and Another v. Malema and Another (Vereniging van Regslui vir Afrikaans as Amicus Curiae) 2011 (12), BCLR 1289 (EqC) (2011).

AfriForum v. Economic Freedom Fighters and Others (EQ 04/2020), ZAGPJHC 599 (2022).

AfriForum Research Institute. “*Farm Attacks and Farm Murders in South Africa.*” Report, AfriForum Research Institute, 2020.

Amish Devgan v. Union of India, SCC OnLine SC 994 (2020).

Amnesty International. “Freedom of Expression.” Amnesty International. <https://www.amnesty.org/en/what-we-do/freedom-of-expression/>.

Brown, A. “What Is Hate Speech? Part 1: The Myth of Hate.” *Law and Philosophy* 36 (2017).

Catch the Fire Ministries Inc v. Islamic Council of Victoria Inc, VSCA 284; 15 VR 207; 235 ALR 750 (2006).

De Varennes, F. “Minority Issues: Report of the Special Rapporteur on Minority Issues, Fernand de Varennes.” Report, United Nations, 2021.

De Villiers, B. “Flying a Flag for Freedom of Expression: When Does a Historic Symbol of a Minority Turn into Hate Speech? The Case of the Old Flag of South Africa.” In *Navigating the Unknown – Essays on Selected Case Studies about the Rights of Minorities*, 215–50. Leiden: Brill, 2022.

Delcker, J. “Germany’s Balancing Act: Fighting Online Hate While Protecting Free Speech.” Politico, published October 01, 2020. <https://www.politico.eu/article/germany-hate-speech-internet-netzgd-controversial-legislation/>.

- Delport, D. "SA 19de Op Lys van Lande Met Ergste Georganiseerde Misdad." *Netwerk24*, published September 24, 2022. <https://www.netwerk24.com/netwerk24/nuus/misdad/sa-19de-op-lysvan-lande-met-ergste-georganiseerde-misdad-20220923>.
- Duy, I.N. "The Limits to Free Speech on Social Media: On Two Recent Decisions of the Supreme Court of Norway." *Nordic Journal of Human Rights* 38 (2020). "Equality Act. Promotion of Equality and Prevention of Unfair Discrimination Act 4 of 2000." Opened for signature February 2, 2000. <https://www.justice.gov.za/legislation/acts/2000-004.pdf>.
- FA Picture International v. Central Board of Film Certification, AIR 2005 Bom 145 (2005).
- Faurisson v. France Communication, No 550/93 (1996).
- Gelber, K. "Political Culture, Flag Use and Freedom of Speech." *Political Studies* 60 (2012).
- Hagan v. Trustees of the Toowoomba Sports Ground Trust, FCA 1615 (2000).
- Herz, M. and P. Molnar. *The Content and Context of Hate Speech: Rethinking Regulation and Responses*. Cambridge: Cambridge University Press, 2012.
- Hogan Lovells. "The Global Regulation of Online Hate: A Survey of Applicable Laws." Report, Hogan Lovells, 2020.
- Nelson Mandela Trust and Ors v. AfriForum and Ors, EQ 02/18 (2019).
- Perincek v. Switzerland App, No. 27510/08 (ECHR, 2015).
- Portner Tom. "South Africans Hold Black Monday Protests Over Farm Murders." *Newsweek*, published October 30, 2017. <https://www.newsweek.com/blackmonday-south-africans-protesting-murder-white-farmers-696441>.
- Pravasi Bhalai Sangathan v. Union of India and Ors., No. 157 (Supreme Court, 2014).
- Qwelane v. South African Human Rights Commission and Another, ZACC 22 (2021).

R. v. Keegstra, 3 SCR 697 (1990).

Rao, P.V. "After Supreme Court Judgment, We Must Combat Hate Speech at Social and Political Levels." The Leaflet, published December 8, 2020. <https://theleaflet.in/after-supreme-court-judgment-we-must-combat-hate-speech-at-social-and-political-levels/>.

R.A.V. v. City of St. Paul, 505 U.S. 377 (1992).

Refah Partisi (the Welfare Party) and Others v. Turkey App, No. 41340/98 ECTHR 13 (2003).

Rosenfeld, M. "Hate Speech in Constitutional Jurisprudence: A Comparative Analysis." *Cardozo Law Review* 24 (2003).

Sanchez v. France, 45581/15 (2021).

Saskatchewan (Human Rights Commission) v. Whatcott, 2013 SCC 11 (CanLII), 1 SCR 467 (2013).

Schweppe, J. and D. Walsh. *Combating Racism and Xenophobia Through the Criminal Law*. Limerick: Centre of Criminal Justice, 2008.

Snyder v. Phelps, 131 S. Ct. 1207 (US Supreme Court 2011).

Texas v. Johnson, 491 US 397 (1989).

The Federal Constitutional Court, Federal Constitutional Court (First Senate) 15 January 1958 BVerfGE 7, 198 (1958).

The Federal Constitutional Court, Order of the First Senate of 4 November 2009, 1 BvR 2150/08 (2009).

The Prosecutor v. Ferdinand Nahimana Jean-Bosco Barayagwiza Hassan Ngeze, ICTR 99-52-T (2003).

Tontodimamma, A. et. al. "Thirty Years of Research into Hate Speech: Topics of Interest and Their Evolution." *Scientometrics* 126 (2021).

Topidi, K. "Words That Hurt (1): Normative and Institutional Considerations in the Regulation of Hate Speech in Europe." Paper, SSRN, 2019.

UN Strategy Hate Speech. "UN Strategy and Plan of Action on Hate Speech."
Report, Geneva, 2019.

United States v. Schwimmer, 279 U. S. 644, 655 (US Supreme Court, 1929).

Virginia v. Black, 538 U.S. 343 (2003).