Chapter 9

Introduction

Nicola Henry

The importance of language to our sense of self and our relationships with others makes us especially vulnerable to the injurious potential of speech, particularly "hate speech," which has been the subject of much political and polarizing debate, especially concerning questions of regulation in the form of hate speech or vilification laws. Although there is no universally agreed definition of "hate speech," critical race theorists Matsuda, Lawrence, Delgado, and Crenshaw (2018) define "assaultive speech" as "words that are used as weapons to ambush, terrorize, wound, humiliate, and degrade." The metaphor of physical injury for the harms of speech acts – the "wounds of words" (Matsuda et al., 2018) – was powerfully explored in Toni Morrison's (1993) Nobel Lecture, in which she claimed "Oppressive language does more than represent violence; it *is* violence" (emphasis added). Speech acts then both *cause* harm and *constitute* harm (see Barendt, 2019; Gelber & McNamara, 2016).

Speech harms can be cumulative, long-term, and generational, reinforcing and amplifying discriminatory attitudes and behaviors that treat the "other" as inferior and subordinate, and solidifying existing power relations (Matsuda et al., 2018; see also Calvert, 1997 for a discussion). Jeremy Waldron (2012) argues that hate speech undermines and compromises the dignity "of those at whom it is targeted, both in their own eyes and in the eyes of other members of society" (p. 5). The lived experiences of targets of hate speech attest to these consequential and constitutive harms, such as normalizing discriminatory attitudes and behaviors, as well as "subordination, silencing, fear, victimization, emotional symptoms, restrictions on freedom, lowering of self-esteem, maintenance of power imbalances, and undermining of human dignity" (Gelber & McNamara, 2016, p. 336).

In this section of the Handbook on "speech-based harms," the chapters by Kim Barker and Olga Jurasz, Emma Jane, Benjamin Colliver, Briony Anderson and Mark Wood, and Elina Vaahensalo, each take as their point of departure that

words – in the form of online hate speech, doxxing, and other oppressive forms of online text or speech – cause wounds that are both physical and metaphysical, as well as individual and collective. Jane's chapter on "cyberhate" against women *by women* begins with a discussion of a "vicious attack" against a feminist colleague on Twitter. As Jane explains, online abuse leads to feelings of "hurt, fear, anger, and self-doubt," which are compounded by "despair, disbelief, and betrayal" when the targets discover that the assailants are women because they assume them to be "peers, friends, or allies." One participant in Jane's study described feeling like "she had been hit in the stomach by a baseball bat."

Barker and Jurasz focus their chapter on online text-based sexual abuse, which they define as "written, electronic communication containing threatening and/or disruptive and/or distressing content, such as … textual threats to kill, rape, or otherwise inflict harm on the recipient of such messages." Barker and Jurasz argue that text-based sexual abuse "has silencing effects on women and girls participating online and contributes to the creation of hostile spaces for women." They argue that the legal system has created "hierarchy of harm … in which more credence and gravitas are given to forms of online abuse involving photographic representations of the victim than textual – and frequently very violent – abuse." The resulting harms ranging from psychological to physical to democratic are frequently "shrugged off as less serious than offline," and seen as "'part and parcel' of what happens online."

Likewise, Colliver in his chapter describes a number of interconnected harms relating to transphobic online hate speech. In his study, he examined transphobic discourse in relation to YouTube videos concerning public toilet access and gender neutral toilets. He states that these speech acts reinforce rigid gendered stereotypes and binaries, as well as delegitimizing transpeople by invoking discourses of biological essentialism, functioning to "construct transgender people as a 'scientific absurdity.'"

Anderson and Wood focus their chapter on a newer form of speech harm, "doxxing," defined as "the intentional public release onto the internet of personal information about an individual by a third party, often with the intent to humiliate, threaten, intimidate, or punish the identified individual" (Douglas, 2016, p. 199). They characterize the harms of doxxing as a loss of the target's anonymity, obscurity, and legitimacy, with accompanying mental health effects (e.g., depression, anxiety), and in the case of "organizational doxxing," the loss of competitive advantage.

In addition to exploring the harms of speech acts online, a number of the authors in this section have sought to explore the underlying drivers of oppressive speech acts as a means through which to exert and maintain power over marginalized groups. Anderson and Wood explain that the act of doxxing may or may not involve malicious intention. They use the example of journalists and activists using doxxing to serve the public interest (see also Cheung, this volume), or as a result of carelessness or negligence (e.g., failing to anonymize a source). And yet they also discuss the more hostile forms of doxxing, which involve, for instance, releasing compromising information about a person to prove oneself within an online network of peers, or to extort financial benefits, force individuals or groups of

people to remove themselves from an online forum or platform, or as a form of control or punishment (e.g., in the context of domestic and family violence).

None of the authors in this section of the Handbook treat oppressive speech acts in simplistic, individualistic terms. Colliver, for instance, in his exploration of the ways in which cisgender YouTube commenters construct and position themselves as victims of oppression and political correctness, argues that those who engage in transphobic online hate speech "provide links to one another, and expressly attempt to encourage both recruitment and discussion among like-minded people" (pp. 57-8).

Similarly, Vaahensalo's chapter explores community building through "online othering discourse" by analyzing the cacophony of "antisocial" and "cruel" comments on the Finish forum Suomi24. Vaahensalo argues that "othering" is a result of the social "mechanics of intersubjectivity" and "sociality" that work to reinforce notions of "us" and "them" and is "not always a conscious act of harassment." She notes too that "Othering, whether online or offline, is not always openly hostile or aggressive," but rather can range from "innocent concerns and fears to genuine and open hostility." This is an important point that highlights the insidious nature of less overt forms of discriminatory speech which are often downplayed as harmless expressions of free speech.

Jane too in her chapter examines the complex overlay of internalized misogyny and lateral violence as useful concepts for analyzing woman-to-woman online abuse. She argues that "subjugated peoples may lash out with displaced fury and frustration after internalizing and embodying dominant discourses and ideologies about their own groups." The abuse of women by women, therefore, is a product of a broader, systemic form of interlocking forms of inequality as well as patriarchal oppression. Jane, however, acknowledges that in attributing the causes of cyberhate to structural factors, such as patriarchy, it risks "eliding part or all of subjects' agency and responsibility for their actions." As such, individual traits such as psychopathy, Machiavellianism, and narcissism (or the "Dark Tetrad" as it is also known) must also be considered, alongside the "dark infrastructure" of digital platforms and their facilitation of "outrage and polarization."

References

Barendt, E. (2019). What is the harm of hate speech?. *Ethical Theory and Moral Practice*, 22, 539–553.

Calvert, C. (1997). Hate speech and its harms: A communication theory perspective. *Journal of Communication*, 47(1), 4–19.

Douglas, D. M. (2016). Doxing: A conceptual analysis. *Ethics and Information Technology*, 18(3), 199–210. doi:10.1007/s10676-016-9406-0

Gelber, K., & McNamara, L. (2016). Evidencing the harms of hate speech. *Social Identities*, 22(3), 324–341.

Matsuda, M. J., Lawrence, C. L., Delgado, R., & Crenshaw, K. W. (2018). *Words that wound: Critical race theory, assaultive speech, and the first amendment*. New York, NY: Routledge.

Morrison, T. (1993). Toni Morrison: Nobel lecture. Retrieved from https://www.nobelprize.org/prizes/literature/1993/morrison/lecture/

Waldron, J. (2012). *The harm in hate speech*. Cambridge, MA and London: Harvard University Press.