Speech, Harm and Power

Proposed by Veronika Lassl

There are various ways we can harm each other through speech. Some of these ways are related to structures of oppression and power. Recent literature in social and political philosophy of language has examined these relationships. A considerable portion of this literature is concerned with hate speech and with what is called 'illocutionary silencing'. My proposal is that over the course of this Forum year, we read through central contributions in these debates.

Below are summaries of some of the debates we could read into as part of this topic. I'm happy to place more emphasis on one than another, if there is more interest in one subtopic than another. I'm also happy to broaden the topic to include other debates from social and political philosophy of language (e.g. on the use of generic masculine expressions). The reason they are not included here is because I am less knowledgeable of them than I am of the debates on hate speech and illocutionary silencing.

I. Hate Speech

a. How should we define hate speech?

In legal contexts, examples of hate speech are heterogenous. Nevertheless, attempts have been made to identify unifying features that set hate speech apart from other speech. Brison (1998) offers a definition of hate speech in terms of the types of harms it subjects targets to. Waldron (2014) offers an account where hate speech is defined by the specific way it undermines the social standing and dignity of the target group. Brown (2017a, 2017b) offers reasons to be skeptical about whether a unified definition of hate speech is attainable. For an overview of the various ways one might define hate speech, we could read an interview with Paul Roberts, a legal scholar with philosophical expertise (in Herz and Molnar 2012, 11-36).

b. How does hate speech harm?

Different accounts of the way hate speech harms exist. Some accounts focus on the "immediate" psychological impact on the addressees (e.g. the psychological distress of being on the receiving end of hate speech). Texts which focus on this type of harm are Delgado (1993), Lawrence (1993), and Matsuda (1993). Other accounts (e.g. Waldron (2014)) focus on the structural harm of hate speech, i.e. how it undermines the social standing of the targeted group. Yet other accounts discuss both types of harm (e.g. Langton (2012, 2018), Gelber and McNamara (2016)).

There is some disagreement over whether we can distinguish between consequential and constitutive harms of hate speech. Some have suggested that we can think of hate speech as harming both causally – by spreading certain (bad) beliefs, for instance – and constitutively – in virtue of being the kind of speech act that (when successful) changes the social status of a group for the worse. This distinction between the way hate speech harms can be found in Langton (1993, 2012), Maitra (2012) and McGowan (2009). Kukla (2014) criticizes this distinction. To understand the theoretical background for this distinction (speech act theory), we may want to read Morris (2007).

c. Slurs

Hate speech may be thought to occur in explicit forms or implicit forms. An explicit form of hate speech is the use of slurs. Philosophers of language have been interested in how to explain the semantic and pragmatic profile of slurs with the hope that this can in turn explain how slurs derogate. One common way to think of slurs is to think of them as offensive expressions that have a neutral counterpart (though Ashwell (2016) disagrees with this view). Many philosophers that are happy to regard slurs as expressions that have neutral counterparts have been interested in explaining why the slurs are offensive, while the neutral counterparts are not. Hom (2008) suggests explaining this in terms of ideologies and sets of practices that are responsible for the slurs' derogatory meaning (Jeshion (2013b) objects). Camp (2013) suggests explaining slurs through commitments to specific perspectives on the derogated group. That is, as ways of thinking about a group in which certain affective and evaluative responses appear warranted. Nunberg (2018) objects to Camp and suggests explaining slurs as signals of allegiance to groups with negative views of the target. Jeshion (2013a)

and Richard (2010) propose accounts of slurs which center their ability to express contempt towards the target group. Dummett (1993), Tirell (1999) and Whiting (2008) suggest explaining slurs in terms of the inferences they license from non-offensive content to offensive content. Anderson and Lepore (2013) suggest explaining the offensiveness of slurs with the idea that there are prohibitions in place against their use and that it is offensive to violate these prohibitions (Camp (2018a) objects to this proposal).

There is also some work beyond the way in which slurs offend. Cepollaro and López de Sa (2022, 2023) have investigated the practice of reclaiming slurs by oppressed groups – for example, the use of the N-word ending on 'a' by BIPoC communities or the use of the F-slur by queer communities. Maitra (2012) and Langton (2018) have discussed interactions between use of hate speech and speaker authority. They discuss the impact of hate speech in situations where it is produced by a speaker with authority or where no bystander objects.

d. Code words and dogwhistles

Implicit forms of hate speech may occur through the use of dogwhistles and code words. They are tools speakers can use to communicate content without explicitly stating it. More specifically, they can be used to express content that is typically only picked up by a specific subset of a general audience. For example, to audiences who are not aware that April 20th is Hitler's birthday and that egg-dumplings were his favorite meal, tweeting "I love egg dumplings" on April 20th may seem like a completely innocuous. However, neo-Nazi audiences may recognize that by tweeting "I love egg dumplings", the tweeter was also communicating that they are sympathetic to Hitler. This covert type of communication allows speakers to communicate content that may be similar to content communicated by explicit hate speech. But it allows the speakers to retain a certain plausible deniability over whether or not they were indeed communicating objectionable content. It is thus a great tool for politicians to hint at certain objectionable content, while it is at the same time harder to hold them accountable for expressing that content. Saul (2018), Tonhauser (2012), Khoo (2017), and O'Donnel (2017) have all developed various frameworks for making sense of how this sneaky type of communication works. For a broader discussion of this type of sneaky communication, we could read Camp (2018b) on insinuating.

II. Illocutionary silencing

A different debate we could dive into as part of the 'Speech, Harm and Authority' topic is the debate around what is called 'illocutionary silencing'. Very basically, the idea here is that speakers can, in some circumstances, fail to perform the action they intended to perform with an utterances, despite being able to produce that utterance. The case that is most commonly discussed in this literature is that of sexual refusal. The suggestion about sexual refusal developed in Langton (1993) is that due to the prominence of favorable rape depictions in pornography, certain conditions obtain that undermine women's ability to perform the illocutionary act of refusal. Specifically, the thought is that even though women succeed in uttering "No", this utterance is not taken by their audience as the refusal it was intended to be. Much debate exists about what it takes for a particular group to be systematically illocutionarily silenced, whether it is plausible that women are silenced in this way and what the consequences are if this view is correct (for instance, if someone utters "No" but is not taken as refusing sexual advances by the hearer, is the hearer committing rape?). Some texts on this topic are: Langton (1993), Hornsby and Langton (1998), Mikkola (2011, 2019), Maitra (2009), Sbisá (2009), Caponetto (2021), Bird (2008).

III. Deadnaming

One last subtopic we could dip into is deadnaming (referring to a trans person by a name they have rejected). I'm only aware of two very recent publications on this issue. Koles (2024) gives a semantic account of deadnames; Lane (2025) criticizes this account and offers an account in terms of linguistic authority instead). But I think there may be interesting similarities and differences between how deadnames and other hate speech work.

Keynotes:

Europe

- Alexander Brown (University of East Anglia)
    - Has published on: skepticism about a unified definition of hate speech
- Rae Langton (Cambridge)
    - Has published on: Hate speech, illocutionary silencing
- Bianca Cepollaro (Vita-Salute San Raffaele University)
    - Has published on: Slurs, Reclaiming slurs
- Elek Lane (University of Vienna)
    - Has published on: Deadnaming
- Judith Tonhauser (University of Stuttgart)
    - Has published on: code words, dogwhistles
- Laura Caponetto (University of Milan)
    - Has published on: Illocutionary silencing

North America

- Elizabeth Camp (Rutgers)
    - Has published on: slurs, insinuation
- Jennifer Saul (University of Waterloo, Canada)
    - Has published on: code words / dogwhistles

References:

Ashwell, Lauren. 2016. "Gendered Slurs." *Social Theory and Practice* 42 (2): 228–39.

Anderson, Luvell, and Ernest Lepore. 2013. "Slurring Words." *Noûs* 47 (1): 25–48.

Bach, Kent. 2018. "Loaded Words: On the Semantics and Pragmatics of Slurs." In *Bad Words: Philosophical Perspectives on Slurs*, edited by David Sosa, 60–76. Oxford: Oxford University Press.

Bird, Alexander. 2002. "Illocutionary Silencing." *Pacific Philosophical Quarterly* 83: 1–15. https://doi.org/10.1111/1468-0114.00137.

Bolinger, Renée Jorgensen. 2017. "The Pragmatics of Slurs." *Noûs* 51 (3): 439–62.

Brison, Susan. 1998. "The Autonomy Defense of Free Speech." *Ethics* 108 (2): 312–39.

Brown, Alexander. 2017a. "What Is Hate Speech? Part 1: The Myth of Hate." *Law and Philosophy* 36 (4): 419–68.

Brown, Alexander. 2017b. "What Is Hate Speech? Part 2: Family Resemblances." *Law and Philosophy* 36 (5): 561–613.

Camp, Elisabeth. 2013. "Slurring Perspectives." *Analytic Philosophy* 54 (3): 330–49.

Camp, Elisabeth. 2018a. "Slurs as Dual-Act Expressions." In *Bad Words: Philosophical Perspectives on Slurs*, edited by David Sosa, 29–59. Oxford: Oxford University Press.

Camp, Elisabeth. 2018b. "Insinuation, Common Ground, and the Conversational Record." In *New Work on Speech Acts*, edited by Daniel Fogal et al., 40–66. Oxford: Oxford University Press.

Caponetto, Laura. 2021. "A Comprehensive Definition of Illocutionary Silencing." *Topoi* 40: 191–202.

Cepollaro, Bianca. 2015. "In Defense of a Presuppositional Account of Slurs." *Language Sciences* 52: 36–45.

Cepollaro, Bianca. 2020. *Slurs and Thick Terms: When Language Encodes Values*. Philosophy of Language: Connections and Perspectives. Lanham: Rowman & Littlefield.

Cepollaro, Bianca, and Dan López de Sa. 2022. "Who Reclaims Slurs?" *Pacific Philosophical Quarterly*. https://doi.org/10.1111/papq.12403.

Cepollaro, Bianca, and Dan López de Sa. 2023. "The Successes of Reclamation." *Synthese* 202 (205). https://doi.org/10.1007/s11229-023-04393-y.

Delgado, Richard. 1993. "Words That Wound: A Tort Action for Racial Insults, Epithets, and Name Calling." In *Words That Wound: Critical Race Theory, Assaultive Speech, and the First Amendment*, edited by Mari Matsuda, Charles Lawrence III, Richard Delgado, and Kimberlé Crenshaw, 89–110. Boulder, CO: Westview Press.

Dummett, Michael. 1993. *Frege: Philosophy of Language*. Cambridge, MA: Harvard University Press.

Gelber, Katharine, and Luke McNamara. 2016. "Evidencing the Harms of Hate Speech." *Social Identities* 22 (3): 324–41.

Hom, Christopher. 2008. "The Semantics of Racial Epithets." *The Journal of Philosophy* 105: 416–40.

Hornsby, Jennifer, and Rae Langton. 1998. "Free Speech and Illocution." *Legal Theory* 4: 21–37.

Jeshion, Robin. 2013a. "Expressivism and the Offensiveness of Slurs." *Philosophical Perspectives* 27: 231–59.

Jeshion, Robin. 2013b. "Slurs and Stereotypes." *Analytic Philosophy* 54 (3): 314–29.

Khoo, Justin. 2017. "Code Words in Political Discourse." *Philosophical Topics* 45 (2): 33–64.

Koles, Taylor. 2024. "The Semantics of Deadnames." *Philosophical Studies* 181 (4): 715–39.

Kukla, Quill (writing as Rebecca). 2014. "Performative Force, Convention, and Discursive Injustice." *Hypatia* 29 (2): 440–57.

Kukla, Quill. 2018. "Slurs, Interpellation, and Ideology." *The Southern Journal of Philosophy* 56 (S1): 7–32.

Lane, Elek. 2025. "Deadnaming, Taboo, and Linguistic Authority." *Mind* 134 (536): 1015–39.

Langton, Rae. 1993. "Speech Acts and Unspeakable Acts." *Philosophy & Public Affairs* 22 (4): 293–330.

Langton, Rae. 2012. "Beyond Belief: Pragmatics in Hate Speech and Pornography." In *Speech & Harm: Controversies over Free Speech*, edited by Ishani Maitra and Mary Kate McGowan, 72–93. New York: Oxford University Press.

Langton, Rae. 2018. "The Authority of Hate Speech." In *Oxford Studies in Philosophy of Law*, vol. 3, edited by John Gardner, Leslie Green, and Brian Leiter, 123–52. New York: Oxford University Press.

Lawrence, Charles III. 1993. "If He Hollers Let Him Go: Regulating Racist Speech on Campus." In *Words That Wound: Critical Race Theory, Assaultive Speech, and the First Amendment*, edited by Mari Matsuda, Charles Lawrence III, Richard Delgado, and Kimberlé Crenshaw, 53–88. Boulder, CO: Westview Press.

Maitra, Ishani. 2009. "Silencing Speech." *Canadian Journal of Philosophy* 39 (2): 309–38.

Maitra, Ishani. 2012. "Subordinating Speech." In *Speech & Harm: Controversies over Free Speech*, edited by Ishani Maitra and Mary Kate McGowan, 94–120. New York: Oxford University Press.

Matsuda, Mari. 1993. "Public Response to Racist Speech: Considering the Victim's Story." In *Words That Wound: Critical Race Theory, Assaultive Speech, and the First Amendment*, edited by Mari Matsuda,

Charles Lawrence III, Richard Delgado, and Kimberlé Crenshaw, 17–51. Boulder, CO: Westview Press.

McGowan, Mary Kate. 2009. "Oppressive Speech." *Australasian Journal of Philosophy* 87 (3): 389–407.

McGowan, Mary Kate. 2012. "On 'Whites Only' Signs and Racist Hate Speech: Verbal Acts of Racial Discrimination." In *Speech & Harm: Controversies over Free Speech*, edited by Ishani Maitra and Mary Kate McGowan, 121–47. New York: Oxford University Press.

Mikkola, Mari. 2011. "Illocution, Silencing, and the Act of Refusal." *Pacific Philosophical Quarterly* 92: 415–37.

Mikkola, Mari. 2019. *Pornography: a Philosophical Introduction*. New York: Oxford University Press.

Morris, Michael. 2006. "Austin on speech acts." In *An Introduction to the Philosophy of Language*, 231–247. Cambridge: Cambridge University Press.

Nunberg, Geoff. 2018. "The Social Life of Slurs." In *New Works on Speech Acts*, edited by Daniel W. Harris, Daniel Fogal, and Matt Moss, 237–95. Oxford: Oxford University Press.

O'Donnell, Patrick. 2017. "Generics, Race, and Social Perspectives." *Inquiry: An Interdisciplinary Journal of Philosophy*. Published online January 26, 2017. https://doi.org/10.1080/00207543.2016.1266801.

Parekh, Bhikhu. 2012. "Is There a Case for Banning Hate Speech?" In *The Content and Context of Hate Speech: Rethinking Regulation and Response*, edited by Michael Herz and Peter Molnar, 37–56. Cambridge: Cambridge University Press.

Richard, Mark. 2010. *When Truth Gives Out*. Oxford: Oxford University Press.

Saul, Jennifer M. 2018. "Dogwhistles, Political Manipulation, and Philosophy of Language." In *New Works on Speech Acts*, edited by Daniel W. Harris, Daniel Fogal, and Matt Moss, 360–83. Oxford: Oxford University Press.

Sbisà, Marina. 2009. "Illocution and Silencing." In *Language in Life, and a Life in Language: Jacob Mey—a Festschrift*, edited by Bruce Fraser and Ken Turner, 351–57. Bradford: Emerald.

Tirrell, Lynne. 1999. "Derogatory Terms: Racism, Sexism, and the Inferential Role Theory of Meaning." In *Language and Liberation: Feminism, Philosophy, and Language*, edited by Christina Hendricks and Kelly Oliver, 41–79. Albany: State University of New York Press.

Tonhauser, Judith. 2012. "Diagnosing (Not-)at-Issue Content." *Proceedings of Semantics of Underrepresented Languages in the Americas* 6: 239–54.

West, Caroline. 2012. "Words That Silence? Freedom of Expression and Racist Hate Speech." In *Speech & Harm: Controversies over Free Speech*, edited by Ishani Maitra and Mary Kate McGowan, 222–48. Oxford: Oxford University Press.

Whiting, Daniel. 2008. "Conservatives and Racists: Inferential Role Semantics and Pejoratives." *Philosophia* 36 (3): 375–88.