**Hate speech: from speech acts to economic behaviour**

The “harm in hate speech”, to quote Jeremy Waldron, is multifarious but sometimes difficult to pin down; nonetheless, understanding how hate speech damages people’s lives is a condition for offsetting its effects. The “Pyramid of Hate” model, as propounded by the Anti-Defamation League, represents escalating levels of attitudes and behaviours that start with biased attitudes and can subsequently evolve into acts of bias, systemic discrimination, bias-motivated violence, and even genocide. In this sense, hate speech could be situated on the lowest levels of the pyramid and constitute an actual threat of eventual hate-motivated physical violence. This model, then, identifies the harm in hate speech both with the fear it instils into its victims and the increased probability that such fear will eventually materialize. However, this does not answer the following questions: 1) how do hate words bring about, or contribute to bringing about, hate-motivated actions? 2) Are there other, more direct ways in which hate speech causes damage?

Some have addressed the above questions by applying John L. Austin’s theory of “speech acts” and distinguishing between illocutionary and perlocutionary effects of hate speech. As is well known, the illocutionary acts are the linguistic acts performed by "saying something" (e.g.: she asked me to open the window, she ordered me to open the window, she begged me to open the window, she dared me to open the window, etc.); while the perlocutionary acts are the extra-linguistic effects resulting from an illocutionary act (e.g.: she made me open the window, she made me feel uncomfortable, etc.). The effects of illocutionary acts are purely *conventional*, while those of perlocutionary acts are *material*: the former are *constituted*, while the latter are *caused*. For Austin, among the “felicity conditions” of illocutionary acts (i.e. conditions that make them “felicitous”) we find not only the following of certain conventional procedures, but also that these are performed in appropriate circumstances by the appropriate person: hence, *authority* is required. Practical authority can be either epistemic or practical.

Authors like Rae Langton have focused on the illocutionary effects of hate speech: following Catharine MacKinnon, Langton holds that hate speech can place people in hierarchies (illocutionary effect) thereby strengthening positions of subordination and feelings of superiority/inferiority, and rationalizing or justifying indifference to violence against those considered to be inferior (perlocutionary effect). The authority needed for the illocutionary effect to obtain is both epistemic and practical and, for Langton, it can be purely *de facto*, i.e. informal. It is acquired, among other things, thanks to *accommodation*, i.e. when no one challenges the presupposition that those who practice hate speech have some authority backing their speech act: for instance, that they can actually order some foreigner to leave the country, or that they have epistemic credibility or represent a majority when expressing insulting views on some group. Accommodation can make individuals acquire authority if they simply act as if they already had it and nobody challenged them.

For a number of reasons, however, this account of the effects of hate speech is not fully convincing: among other things, because it underplays the role that the context plays in granting authority to the speakers. There mustalready be social structures of subordination at play, for the hate speech act to be successful and for the speaker to constitutethe illocutionary effects. Authority is not simply conferred by accommodation. Some pre-existing conditions of asymmetries of power, discrimination (racism, sexism, ableism, etc.), inequalities, etc. are needed.

Furthermore, even if we accept that hate speech acts obtain illocutionary effects, it is very difficult to precisely assess the further effects that are caused by illocution, i.e. perlocutionary effects. In the absence of causal consequences, the illocutionary effects remain confined to the world of words and are incapable of doing actual harm to hate speech victims.

Lastly, it is perfectly possible that both the conventional procedures and the ascription of authority required for felicitous speech acts may apply only to restricted groups in society, without actually damaging the standing of hate speech victims as equal members of the polity.

However, a promising attempt to explain of how the illocutionary effects of hate speech can bring about harmful perlocutionary consequences, by causing hate crimes, has been made by Dhammika Dharmapala and Richard H. McAdams, who have provided an economic model based on the fact that potential offenders care about their reputation among like-minded individuals. By applying the “esteem theory” drawn from the economics of social norms, Dharmapala and McAdams show that hate speech *signals* esteem judgments and can actually convey information about the kind of behaviour that, in certain groups, will win fame and appreciation for its perpetrators. In other words, if we assume fame as motivation and esteem as a consumption good, hate speech can make hate crime more alluring among the members of extremist groups. Moreover, Dharmapala and McAdams show that individuals gain “expressive utility” by sending signals via hate speech, but in the presence of “speech costs”, some hate speech may be deterred, thereby creating uncertainty about social attitudes and the expected reputation gains of hate crimes.

A possible extension of the above analysis, based on signaling theory, can be made in order to explain the more *direct* harms of hate speech, which are not done by means of its illocutionary effects, but are an immediate causal consequence of its performance: namely, the damages that are caused to the victims who are exposed to it. Hate speech victims can perceive it as the emission of signals to like-minded people and also as a direct source of information regarding the society in which they are living. Thence, victims are both scared of the possibility of suffering from hate crimes and they are led to think that, in a society that allows hate speech against them, they are not considered of equal value and standing as other members, because their interests are not considered as deserving protection. These “stigmatization effects” cause tangible, concrete harm.