Guillén-Nieto, Victoria. 2023. *Hate speech: Linguistic perspectives*. Berlin: De Gruyter. 211 pp. ISBN: 9783110672466

The book by Victoria Guillén-Nieto focuses on hate speech seen through the lens of the combination of various legal and linguistic perspectives which result in several methodologies being called upon to support the analysis of the hate speech phenomenon. The author's starting point is the fact that so far very few significant studies on hate speech in the field of linguistics have been accessible. The general point of view adopted by Guillén-Nieto is that of legal practitioners and linguists who face major difficulties in dealing with language of hatred, particularly with the emergence and rapid evolution of new technologies and social networks. The author develops a linguistic perspective based on data, tools and solutions that linguists may provide to enable legal action to be taken.

The macro-structure of the book consists of a preface and eight chapters. In the Preface the reader will find a detailed review of the bibliography on hate speech accompanied by some general considerations on the current status of research on hate speech in various areas of study. Then, the book is divided into two parts: Legal linguistics (Part I – Chapters 1-4) and Forensic linguistics (Part II – Chapters 5-8). Legal linguistics analyses the doctrinal content of the law and its linguistically-based structure, while forensic linguistics is concerned with helping to establish the facts on which a legal decision is based.

In Chapter 1. *Approaches to the meaning of hate speech*, Guillén-Nieto considers various definitions of hate speech and adopts Wittgenstein's concept of *family resemblance* (2009 [1953]) with the aim of revealing to what extent it can be of use for the researchers in the area of linguistics and law who approach the phenomenon of hate speech. This perspective enables and supports the understanding that hate speech does not have a single meaning but rather several connotations that share certain affinities with each other. Thus it is not possible to identify features that would be shared by all scientific disciplines that deal with hate speech. The aforementioned thesis is proven by the author through Brown's ordinary language analysis (2017).

In the following part of the chapter Guillén-Nieto gives an outline of legal scholarly attempts to define the concept of hate speech and she suggests its division into three categories, namely content-based hate speech, intent-based hate speech and harms-based hate speech. This section provides a diachronic overview of research on hate speech and shows very clearly that it might not be possible to create a single unified definition which could be used both in linguistics and legal studies.

The author's main aim of this part of Chapter 1 is to show how heterogeneous hate speech is, regardless of the discipline that is chosen as the theoretical framework. The author cites a considerable number of studies that prove her thesis, but it shall be acknowledged that this has been a well-known assumption and a starting point for many studies on hate speech, especially in linguistics. The scholars seem to be aware of the complexity and indefiniteness of the phenomenon. At the same time, given the premise the author makes in the preface about combining legal and linguistic perspectives, it might have been useful to place a little more emphasis on aspects related to the latter, as the legal perspective is by far the dominant one here.

The final section of Chapter 1 focuses on approaches to a technical legal definition of hate speech at three levels: international law, common law and civil law (European Union and Member State law). The analysis takes into consideration legal documents such as the Universal Declaration of Human Rights (1948), the International Convention on the Elimination of All Forms of Racial Discrimination (1965), International Covenant on Civil and Political Rights (1966) – for the international law. The common law is represented by Hate Crime Statistic Act, the First Amendment to the Constitution of the United States, Criminal Code of Canada, laws of the United Kingdom, Racial Discrimination Bill (1975) and Racial Vilification Act of Australia. Finally, the author devotes some space to hate speech legislation within the European Union, and particularly she refers to the European Convention on Human Rights, Recommendation No. R (97) 20 of the Committee of Ministers of the Council of Europe to the Member States (1997), the Council Framework Decision (2008) and to Member State law in Germany, France and Spain.

What emerges from the considerations exposed in this chapter is the difference between the common law, on the one hand, and the civil/international law on the other. This concept will be furtherly elaborated in the next chapter. At this point, it is worth noting that this is a particularly insightful section of the book, as the author effectively and clearly shows that the differences in legal systems do have a very significant impact on how difficult it is to create a uniform definition of hate speech. Although the differences in European and American legislation are a matter of common knowledge, only a detailed analysis of the legal acts shows that, from the scholarly perspective, a coherent definition of the phenomenon under discussion will be difficult to reach. Indeed, an utterance which under one legal

system may be considered incitement to hatred, in another system will be protected by the provisions on freedom of expression.

Chapter 2. *Hate Speech as a legal problem* provides an analysis of some problematic aspects of hate speech legislation such as, for example, the coexistence of various different hate speech laws and the consequent lack of consistent jurisprudence on hate speech. It also includes reflections on the dilemma concerning the difficult balance between the right to freedom of expression and the prohibition of incitement to hatred combined with the struggle to find some universal rules that would enable to define which speech acts should be legally banned as those inciting to hatred.

In this chapter, the author succeeds in demonstrating how enormous the challenge is to determine common legal standards that would allow an effective control of hate speech. This will not be an easy goal to achieve since many domestic laws differ in a significant way from what is recommended at international level. What is more, on the basis of considerations included in Chapter 2, one can realise that there will never be a limited repertoire of speech acts and forms of expression that will allow for a direct prohibition or conviction of hate content. This is due to the fact that the communicative context (the speaker's intent and the effect on the recipient) always needs to be considered and this makes it difficult to create standards that are both universal and effective. To make this explicit with the help of terminology derived from linguistic pragmatics, one can say that courts and legal practitioners must not limit themselves to observe only locutionary acts but must also consider the illocutionary and perlocutionary force of an utterance if a legal decision shall be taken. The final section of this chapter gives an outline of how the advent of Internet has affected the dynamics of the phenomenon.

In general, Chapter 2 provides further confirmation of the impossibility of creating a common and universal definition of hate speech. One could conclude that, in the absence of such a technical legal definition, the attempts to create one should be abandoned. Instead, the author invokes again Wittgenstein's concept of family resemblance, according to which hate speech constitutes an umbrella term covering many definitions that can be derived from different legal systems and that will share at least one common element. Such a solution would relieve researchers from trying to arrive at a universal definition of the phenomenon, but on the other hand it would certainly not contribute to the coherence of any research. In that case, each study might be based on a definition of hate speech created *ad hoc*, which is currently, in fact, a fairly common practice. This means that the author in this section describes the factual situation that occurs, giving it a theoretical framework, rather than sets new research standards.

Chapter 3. *The legal reasoning in hate speech court proceedings* provides an analysis of some important legal cases (landmark cases) coming from different periods of time (1949-2019) and different geographical areas (namely the United States and the European Union) in which the motif of hate speech was present. The choice of these particular areas was deliberate as, according to the author, they represent two different approaches to hate speech. Thus, a total of nine cases were subjected to analysis and they come from the United States Supreme Court (Terminiello v. Chicago [1949], Brandenburg v. Ohio [1969], National Socialist Party v. Skokie [1977] and Virginia v. Black [2003]), the United States Court of Appeals for the Armed Forces (United States v. Wilcox [2008]) and the European Court of Human Rights (Jersild v. Denmark [1994], ES v. Austria [2019], Fáber v. Hungary [2012] and A. v. The United Kingdom [2003]).

For a European Union citizen the decisions of the courts in the United States may seem beyond comprehension and vice versa. It becomes very clear why a common and uniform definition of hate speech has not been established so far, as two globally important social elements are competing with each other, namely the right to freedom of expression and other fundamental human rights. In fact, this chapter constitutes a more 'practical' confirmation of the considerations raised in the previous chapters, since it can be clearly seen that both the United States and European Union courts rule on the basis of the situational context in which a contested utterance occurred.

The review of the cases that the author provides in Chapter 3 at the same time constitutes the presentation of the corpus. What is important here is the fact that the cases discussed are analysed as a whole, not being limited to purely linguistic material, although each of them contains verbal content (e.g. the minutes of a party or testimony). This means that the analysis of the corpus in the later part of the volume will be multilevel and the two disciplines, namely law and linguistics, will converge.

Chapter 4. *Critical discourse analysis* focuses on the exam of surface discourse structures and aims at showing in what way both power and social supremacy may influence the language and how it can be revealed through a linguistic study. It shall be acknowledged that critical discourse analysis is an approach which was subjected to criticism on the grounds of social science and linguistics, among others, due to the lack of a methodology, subjectivism and a more interpretative than analytic character. Nevertheless, the author decided consciously to apply critical discourse analysis with its three theories (the theory of social representations, the theory of ideology and the theory of power as a control) in order to look over one of the landmark cases introduced in Chapter 3 (Brandenburg v. Ohio, 1969). Guillén-Nieto aims at a better understanding of cases in which hate speech appears as a crucial element. Thanks to this approach, one can notice some lexical choices and syntactic strategies which are present in the linguistic material subjected to analysis and which show a clear racist content. The results of this analysis appear even more persuasive in light of the fact that no guilty verdict was ultimately reached, since the Supreme Court reversed the conviction.

Chapter 5. *Register and genre perspectives on hate speech* opens the second part of the book (dedicated to legal linguistics) and exploits the assumption that discourse can be subjected to analysis that focuses on the text register and complex generic forms. The first part of the chapter emphasises the fact that hate speech embedded in a text lasts

over time, since the production of the text and its reception do not necessarily occur in the same spatiotemporal context. According to the author, this premise is of extreme importance since it casts doubts on the legal basis used by the courts in the landmark cases cited in Chapter 3 (the imminence standard) as it relies on the immediate unlawful effects that an incitement to hatred can produce. The possible spatiotemporal gap between the production and reception of the text, on the other hand, means that such effects cannot be predicted and, consequently, the imminence standard would be ineffective.

The second part undertakes the register perspective, in which the author follows Biber and Conrad's (2009) theory that is closely linked to Halliday's 'context of situation' (1978). For this reason, the analysis concentrates on the semiotic functions of tenor, field and mode and argues that four parties participate in a speech event, namely a) the hate-advocating speaker, b) the ingroup, c) the outgroup (the target group) and d) the overhearers.

As a conclusion made on the basis of the results of her analysis, the author observes that it is not possible to establish a concrete and closed repertoire of hateful linguistic features since the speaker is provided with a wide range of channels of communication and forms of expression. Neither will the high-frequency criterion be effective in this case, as hatred can be encapsulated in those linguistic features that are characterised by a low frequency. Hence, the author suggests using the criterion of relevance (collective salience of a linguistic feature) as the indicator of hate register.

A similar situation is encountered in the last part of this chapter, namely in the context of genre perspective, where it is also difficult to investigate hate speech by searching for universal rules since the phenomenon under consideration is not connected to any specific generic form; instead it manifests itself in a multitude of genres such as, for example, protest speeches, demonstrations, posters, leaflets and many others (genre colonisation - Bhatia 2014 [2004]: 100). However, as the author claims, it is possible to identify some common features, thus enabling an analysis from a generic perspective.

Although there is no genre form that was created specifically to spread hateful content, the author suggests that 'hate propaganda' may be considered a super genre to which other forms are subjected. Hate propaganda could be classified as a type of negative propaganda which promotes an ideology that incites prejudice and intolerance towards target groups. Its purpose is to influence emotions, attitudes, behaviour and opinions in favour of the ingroup and against the outgroup. While hate propaganda is not associated with any specific rhetorical form, the subgenres would share the same set of purposes (manifest, disseminate and promote hatred). To illustrate the phenomenon the author presents some textual and visual genre forms typical of racist Nazi propaganda.

The author predicts that it will become increasingly difficult to investigate hate speech, as it will less and less manifest itself explicitly in the form of hateful words or grammar structures (e.g. pronouns), while indirectly transmitted contents will become prevalent, as they easily hinder the automatic detection tools such as, for example, algorithms used by the social media.

Chapter 6. *Speech act theory* is devoted to hate speech seen within the perspective of speech act theory, where, according to the author, hate speech constitutes a 'global-act'. The latter is then supported by minor speech acts. In her analysis Guillén-Nieto takes into consideration Searle's classification of speech acts (1969; 1979: 1–29) and her aim is to determine which speech acts can express hate speech. It results in quite a difficult task as Searle's five categories may overlap and hatred cannot be simply ascribed to be an expressive act. In fact, the author claims that hate speech may be classified as *directives* due to the fact that through incitement the speaker invites the addressee to "do" some unlawful action in the future. However, there is no obligation for the addressee to perform such an action and this is why incitement involves implicitness rather than directness.

Then Guillén-Nieto moves on to discuss Austin's idea of explicit and implicit performatives (1962) and Searle's primary and secondary illocutionary acts (1969), as in real world the indirectness in the conversation is dominant over directness and the case of hate speech is not an exception here. In fact, hate content is not conveyed through performative verbs such as *incite*, *advocate*, etc. as it is always performed implicitly.

What emerges from the considerations presented in this chapter is the conclusion that, thanks to the speech act theory, it is possible to prove that the texts of the cases subjected to analysis contain an implicit act of incitement to violence, although some guilty verdicts had been overturned. Another observation to make here is that speech act theory can, indeed, be a tool to investigate hate speech, yet at the same time it presents limitations, as hate speech cannot be narrowed down to a list of specific speech acts due to the complexity of the phenomenon. Similar conclusions have been reached by, for instance, Alfonzetti (2019) and Bazzanella (2020).

Chapter 7. *(Im)politeness theory* provides considerations on the potential application of this approach in the analysis of hate speech. In this section, the author aims to illustrate that hate speech is related to an intentional deviation from polite behaviour operated by the sender. The author proves that the strategies used in hate speech are to be defined as face-threatening (Goffman (1967), Brown and Levinson 1987 [1978]) since they involve emotions traditionally seen as negative (fear, anger, contempt, etc.). Thanks to the linguistic analysis under (im)politeness theory, one can observe that a skilful speaker is able to manipulate pragmatic rules in order to "maintain or enhance communicative concord and comity with the ingroups or communicative discord with the targets" (p. 151). By analysing fragments drawn from the landmark cases, the author sheds light on how the speakers resort to various strategies that enable them to reduce the degree of impoliteness when addressing the ingroup (the audience) and to increase it in communication with the outgroup (the target group).

Chapter 8. *Cognitive pragmatics*, as the title suggests, approaches hate speech from the perspective of cognitive pragmatics relying primarily on the notion of implicature as coined by Grice (1975). In this case, the indispensable premise in order to proceed with the study is the scheme, already seen in the previous chapters, in which hate speech is activated, that is, that of the relationship between the subject (the hate-advocating speaker), the audience (the in-group) and the object (the target group). It is important to emphasise that, within the above-mentioned relationship, the subject and the audience share the same assumptions and attitudes, while the object forms a different interpretive community. The idea of implicature "provides the bridge from what is said to what is meant but not overtly said" (p. 153).

The author proceeds on the basis of the assumption that the speaker in the majority of cases provides information not only explicitly but also indirectly and implicitly, and may do so with the use of both verbal and non-verbal codes such as images or symbols (in the present case, swastika, burning cross and flag). The receiver, on the other hand, in the process of interpretation creates hypotheses in order to trace the speaker's communicative intention. The result of this process is always speculative since one can never be absolutely certain of the speaker's intentions by not having access to the latter's mind.

For the purpose of illustrating the above mentioned considerations, the author subjects to analysis three landmark cases from chapter 3 (Terminiello v. Chicago [1949], Brandenburg v. Ohio [1969], and National Socialist Party v. Skokie [1977]), attempting to reconstruct the ways in which hate-advocating speakers use implicit strategies to convey the message containing the incitement of hatred. The process of reconstruction of the communicative intent by the recipients occurs through the comparison of the implicatures born during the interpretation with their own prior assumptions and experiences, so it is obvious that the audience and the target group will construct quite different hypotheses regarding the same utterance. The author considers this also to be the reason for the different interpretations of cases made by the courts.

In conclusion, the second part of the book (chapters 5-8) explores some aspects of forensic linguistics. It evokes linguistic methods (register and genre perspectives, speech act theory, (im)politeness theory and cognitive pragmatics) which are applied to analyse the corpus consisting of nine landmark cases. It should be noted that the purpose of the analysis is not to determine which of the chosen tools is the most efficient for investigating hate speech. On the contrary, the author rather shows that none of them is able to cover all the aspects of the phenomenon, but this also does not mean that the present study misses its goal. Guillén-Nieto proves that hate speech is not linked to a closed number of genres or to any particular register. It manifests itself through a variety of speech acts, which in themselves are not easy to categorise as their boundaries overlap. It cannot even be reduced to a single sign system, as visual signs and symbols coexist alongside hateful words. Finally, the content that conveys hatred is rarely communicated exclusively in an explicit way, and much more often adopts implicit and indirect forms.

Given the aforementioned considerations, it has to be stated that the results of the author's analysis are consistent with the thesis set out at the beginning, that is, that hate speech is such a complex phenomenon that it can be triangulated only as an umbrella term. The situation will only move forward when legal provisions at the international level evolve. Until then, linguistic research on hate speech must rely on definitions that do not capture the phenomenon as a whole. It is in this aspect that the author invoked Wittgenstein's concept of family resemblance.

To sum up, *Hate speech: Linguistic perspectives* by Victoria Guillén-Nieto constitutes a valuable study of the phenomenon of hate speech from the perspective of law and linguistics. The general point of view in the analysis belongs to the field of law, while the methodological tools are drawn from linguistics. The volume incorporates a comprehensive overview of recent research on hate speech and the reference section included in it constitutes an impressive body of knowledge. In addition, the author has managed to demonstrate the initial thesis, that is, the undoubted complexity that the phenomenon of hate speech poses to scholars in the fields of both linguistics and legal studies.

Although the legal perspective is very much in evidence, the volume provides an insightful contribution to the linguistic study of hate speech. However, as any research, it is not free from limitations. In fact, it should be noted that the corpus submitted for analysis (landmark cases) is not linguistically coherent enough to establish universal recommendations or methodologies for linguistic and/or legal research on hate speech. What is more, the methodology applied to the analysis is based on tools that are usually elaborated on the basis of fabricated utterances whose linguistic analysis is not so problematic and produces more precise results than a body of live data. However, this is something the author is aware of. All in all, the volume will certainly provide a reference point for numerous scholars in further research and it shows very clearly what lacunae need to be addressed, especially within the field of law, if hate speech research is to develop effectively.

Alicja Paleta
ORCID: https://orcid.org/0000-0002-5146-6244
Department of Italian Studies
The Jagiellonian University (Cracow, Poland)
alicja.paleta@uj.edu.pl

References

Alfonzetti, Giovanna. (2019). Gli insulti: alcuni criteri di categorizzazione. *Studi linguistici in memoria di Giovanni Tropea*, edited by Salvatore Carmelo Trovato, Alessandria: Edizioni dell'Orso. 67–78.

Austin, John L. (1962). *How to do things with words*. Oxford: Oxford University Press.

Bazzanella, Carla. (2020). Insulti e pragmatica: complessità, contesto, intensità. *Quaderns d'Italià* 25. 11–26.

Bhatia, Vijay K. (2014 [2004]). *Worlds of written discourse. A genre-based view*. London, New Delhi, New York, Sydney: Bloomsbury.

Biber, Douglas & Susan Conrad. (2009). *Register, genre, and style*. Cambridge: Cambridge University Press.

Brown, Alexander. (2017). What is hate speech? Part 2: Family resemblances. *Law and Philosophy* 36 (5). 1–53.

Brown, Penelope & Stephen C. Levinson. (1987 [1978]). *Politeness. Some universals in language use*. Cambridge: Cambridge University Press.

Goffman, Ervin. (1967). *Interaction ritual: Essays on face-to-face interaction*. Chicago: Aldine.

Grice, H. Paul. (1975). Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and Semantics 3, Speech Acts*. New York: Academic Press, 41–58.

Halliday, M.A.K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.

Searle, John R. (1969). *Speech acts*. Cambridge: Cambridge University Press.

Searle, John R. (1979). *Expression and meaning. Studies in the theory of speech acts*. Cambridge. Cambridge University Press.

Wittgenstein, Ludwig. (2009 [1953]). *Philosophical investigations* (4th ed.). P. M. S. Hacker & Joachim Schulte (eds.). G. E. M. Anscombe, P. M. S. Hacker & Joachim Schulte (trans.). Oxford: Blackwell.